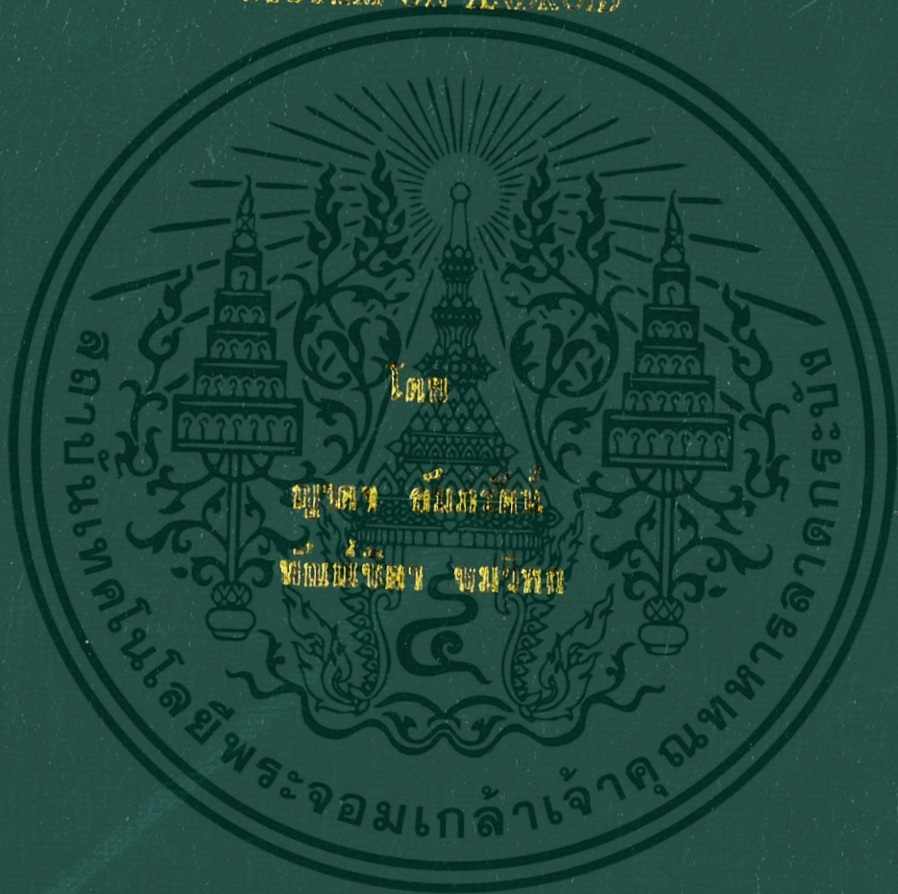


การศึกษานเปรียบเทียบส่วนต่อประสานโปรแกรมประยุกต์สำหรับ
ระบบแปลภาษาด้วยเสียงพูดบนหลายภาษาบนระบบปฏิบัติการแอนดรอยด์

COMPARATIVE STUDY OF APPLICATION PROGRAMMING
INTERFACE FOR MULTILINGUAL SPEECH TRANSLATION
SYSTEM ON ANDROID



วิทยานิพนธ์นี้เป็นลิขสิทธิ์ของมหาวิทยาลัยราชภัฏบุรีรัมย์ โดยสามารถนำข้อมูลบางส่วนไปใช้ในการศึกษาวิจัยได้
แต่ห้ามนำไปเผยแพร่โดยไม่ได้รับอนุญาต

นางสาววิภาดาพร โสภณกุล ศึกษานิเทศก์

คณะศึกษานิเทศก์ มหาวิทยาลัยราชภัฏบุรีรัมย์

มหาวิทยาลัยราชภัฏบุรีรัมย์ วิทยาเขตบุรีรัมย์ ตำบลพิกุลทอง อำเภอเมืองบุรีรัมย์ จังหวัดบุรีรัมย์

ภาคตะวันออกเฉียงเหนือ ปีที่ ๒๕๖๕

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

การศึกษาเปรียบเทียบส่วนต่อประสานโปรแกรมประยุกต์สำหรับ
ระบบแปลภาษาด้วยเสียงพูดแบบหลายภาษาบนระบบปฏิบัติการแอนดรอยด์

COMPARATIVE STUDY OF APPLICATION PROGRAMMING
INTERFACE FOR MULTILINGUAL SPEECH TRANSLATION
SYSTEM ON ANDROID



โดย
ญาดา อัมภรัตน์
พัฒนชิตา ชมวิหก

อาจารย์ที่ปรึกษา
ดร. มานพ พันธุ์โคกกรวด

เลขหมู่.....
เลขทะเบียน...144523
วัน,เดือน,ปี...25...๗๙...2559

600268203
b. 12811816
i.

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรบัณฑิต
สาขาวิชาเทคโนโลยีสารสนเทศ
คณะเทคโนโลยีสารสนเทศ
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ภาคเรียนที่ 2 ปีการศึกษา 2557

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

**COMPARATIVE STUDY OF APPLICATION PROGRAMMING
INTERFACE FOR MULTILINGUAL SPEECH TRANSLATION
SYSTEM ON ANDROID**



**A PROJECT SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENT FOR THE DEGREE OF
BACHELOR OF SCIENCE PROGRAM IN INFORMATION TECHNOLOGY
FACULTY OF INFORMATION TECNOLOGY
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG**

2/2014

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



COPYRIGHT 2015

FACULTY OF INFORMATION TECHNOLOGY

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ใบรับรองปริญญาโท ประจำปีการศึกษา 2557

คณะเทคโนโลยีสารสนเทศ


สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง การศึกษาเปรียบเทียบส่วนต่อประสานโปรแกรมประยุกต์สำหรับระบบ
แปลภาษาด้วยเสียงพูดแบบหลายภาษาบนระบบปฏิบัติการแอนดรอยด์

COMPARATIVE STUDY OF APPLICATION PROGRAMMING
INTERFACE FOR MULTILINGUAL SPEECH TRANSLATION
SYSTEM ON ANDROID

ผู้จัดทำ

1. นางสาวญาดา อัมภรัตน์ รหัสนักศึกษา 54070020
2. นางสาวพัฒนัชชิตา ชมวิหค รหัสนักศึกษา 54070058


.....อาจารย์ที่ปรึกษา

(ดร. มานพ พันธุ์โลกกรวด)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อโครงการ	การศึกษาเปรียบเทียบส่วนต่อประสาน โปรแกรมประยุกต์สำหรับระบบแปลภาษาด้วยเสียงพูดแบบหลายภาษาบนระบบปฏิบัติการแอนดรอยด์		
นักศึกษา	นางสาวณูดา อัมภรัตน์	รหัสนักศึกษา 54070020	
	นางสาวพัฒนชิตา ชมวิหค	รหัสนักศึกษา 54070058	
ปริญญา	วิทยาศาสตรบัณฑิต		
สาขาวิชา	เทคโนโลยีสารสนเทศ		
ปีการศึกษา	2557		
อาจารย์ที่ปรึกษา	ดร. มานพ พันธุ์โคกกรวด		

บทคัดย่อ

ปริญญานิพนธ์ฉบับนี้ได้ศึกษาเปรียบเทียบคุณสมบัติของส่วนต่อประสาน โปรแกรมประยุกต์ในส่วนของ Speech-to-Text, Language Translation และ Text-to-Speech เพื่อเลือก API ที่มีคุณสมบัติเหมาะสมสำหรับแอปพลิเคชันแปลภาษาด้วยเสียงพูดบนระบบปฏิบัติการแอนดรอยด์และนำมาพัฒนาแอปพลิเคชันเพื่อวัดประสิทธิภาพของ API จากการศึกษาพบว่า Google Speech Recognition API, Bing Translator Control API, Google Translator Control API และ Google Text to Speech API มีคุณสมบัติที่เหมาะสมสำหรับแอปพลิเคชัน และเพื่อวัดประสิทธิภาพของ API ดังกล่าว โครงการนี้จึงทำการทดสอบกับกลุ่มผู้ใช้งาน 20 คน โดยทดสอบรับเสียงภาษาไทยและภาษาอังกฤษ ด้วยกลุ่มประโยคตัวอย่าง 30 ประโยค จากการทดสอบพบว่า Google Speech Recognition API มีประสิทธิภาพในการรับเสียงพูดทั้งสองภาษาอยู่ในระดับที่ดีมาก ในส่วนของการแปลภาษาพบว่า Google Translator Control API พบว่ามีความถูกต้องในการแปลดีกว่า Bing Translator Control API ในส่วนการสังเคราะห์เสียงพบว่า ประสิทธิภาพโดยรวมของ Google Text to Speech API อยู่ในระดับที่ดี ทั้งการสังเคราะห์เสียงภาษาอังกฤษและภาษาไทย จากผลการทดสอบชี้ให้เห็นว่าประสิทธิภาพของระบบขึ้นอยู่กับประสิทธิภาพของกระบวนการแปลภาษา

Project Title	Comparative Study of Application Programming Interface for Multilingual Speech Translation System on Android.		
Student	Miss. Yada Amparat	Student ID54070020	
	Miss. Phanchita Chomwihoke	Student ID54070058	
Degree	Bachelor of Science		
Program	Information Technology		
Academic Year	2014		
Advisor	Dr. Manop Phankokkrud		

ABSTRACT

This project proposed comparative study the properties of an application programming interface (API) in the speech-to-text process, language translation process and text-to-speech process. In order to choose API that appropriate for the speech translation system on android then develop an application to measure the performance of the API. In the comparative study, we found the Google Speech Recognition API, Bing Translator Control API, Google Translator Control API and Google Text to Speech API which have the most appropriate for the application. In measurement of the performance of an API, we conduct The experiments with 20 users by having user speak 30 sample sentences. The results shown that Google Speech Recognition API is the most powerful in speech recognition both languages. In the language translation process, the Google Translator Control API can translate more accurate than the Bing Translator Control API. In speech synthesis, we found the efficiency of the Google Text to Speech API is good a speech synthesis in both English and Thai language. In addition, the results indicated that the performance of the system depends on the efficiency of the language translation process.

กิตติกรรมประกาศ

ปริญญาบัตรฉบับนี้ไม่อาจสำเร็จลุล่วงได้ด้วยดีหากขาดความกรุณาจากอาจารย์ที่ปรึกษา
ดร. มานพ พันธุ์โคกกรวด ที่ได้ให้คำแนะนำที่ดีมาโดยตลอด ถึงการปรับปรุงแก้ไขปัญหาต่างๆ
ตลอดมา

ขอขอบคุณ คุณพ่อ คุณแม่ ผู้มีพระคุณสูงสุดที่คอยให้กำลังใจ ให้คำแนะนำในการใช้ชีวิต
ในรั้วมหาวิทยาลัยและให้โอกาสในการศึกษาเล่าเรียน

สุดท้ายนี้ขอขอบคุณคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณ
ทหารลาดกระบังที่เอื้ออำนวยสถานที่ต่างๆ ในการทำปริญญาบัตร รวมถึงขอบคุณเพื่อนๆ คณะ
เทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังทุกคนที่ได้ให้การ
ช่วยเหลือให้คำปรึกษาเพื่อให้ปริญญาบัตรฉบับนี้สำเร็จลุล่วงได้โดยสมบูรณ์



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

หน้า

บทคัดย่อภาษาไทย.....	I
บทคัดย่อภาษาอังกฤษ.....	II
กิตติกรรมประกาศ.....	III
สารบัญ.....	IV
สารบัญตาราง.....	VI
สารบัญรูป.....	VII

บทที่

1. บทนำ.....	1
1.1 ที่มาและความสำคัญ.....	1
1.2 เป้าหมายและวัตถุประสงค์.....	2
1.3 ขอบเขตของงาน.....	2
1.4 วิธีดำเนินงาน.....	3
1.5 ประโยชน์ที่คาดว่าจะได้รับ.....	3
2. ทฤษฎีที่เกี่ยวข้อง.....	5
2.1 ระบบปฏิบัติการแอนดรอยด์.....	5
2.2 Application Programming Interface.....	7
2.3 Speech-to-Text.....	8
2.4 Text-to-Speech.....	9
2.5 การแปลภาษา.....	12
- Google Translate.....	12
- Bing Translator.....	13
- Yandex Translate.....	13
- Fregly Translator.....	14

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
2.6 งานวิจัยที่เกี่ยวข้อง.....	15
2.7 นิยามศัพท์	18
3. การวิเคราะห์และออกแบบระบบ	19
3.1 การศึกษางานระบบเดิม	19
3.2 ปัญหาที่พบในระบบปัจจุบัน	23
3.3 การวิเคราะห์ความต้องการของระบบ.....	23
- ความต้องการที่เป็นหน้าที่หลักของระบบ	23
- ความต้องการของระบบที่ไม่ใช่หน้าที่หลักของระบบ.....	24
3.4 การวิเคราะห์และการออกแบบระบบ	24
- ตารางวิเคราะห์คุณสมบัติ API.....	25
- แผนภาพแสดง Activity Diagram.....	26
4. ผลการดำเนินงาน	27
4.1 หน้าจอแสดงผลแอปพลิเคชัน	27
4.2 การทดสอบประสิทธิภาพของ API	29
4.3 การทดสอบประสิทธิภาพของ API ด้วยแอปพลิเคชันด้วยเสียงพูด.....	36
4.4 ผลรวมแบบสอบถามเพื่อประเมินความพึงพอใจของผู้ใช้แอปพลิเคชัน.....	50
5. บทสรุปและวิเคราะห์ผลการดำเนินงาน	51
5.1 สรุปผล	51
บรรณานุกรม.....	53
ภาคผนวก	
ประวัติผู้เขียน	

สารบัญตาราง

หน้า

ตารางที่

3.1 แสดงการวิเคราะห์คุณสมบัติ API.....	25
4.1 แสดงการทดสอบการรับเสียงพูดภาษาอังกฤษด้วย Google speech recognition.....	29
4.2 แสดงการทดสอบการรับเสียงพูดภาษาไทยด้วย Google speech recognition	30
4.3 แสดงผลการแปลภาษาอังกฤษเป็นภาษาไทยด้วย Bing และ Google.....	32
4.4 แสดงผลการแปลภาษาไทยเป็นภาษาอังกฤษด้วย Bing และ Google.....	34
4.5 แสดงผลการทดสอบการรับเสียงภาษาอังกฤษ.....	37
4.6 แสดงผลการทดสอบการแปลภาษาอังกฤษเป็นภาษาไทย	39
4.7 แสดงผลทดสอบการสังเคราะห์เสียงภาษาไทย.....	41
4.8 แสดงผลจากการทดสอบการรับเสียงภาษาไทย.....	43
4.9 แสดงผลการทดสอบการแปลภาษาไทยเป็นภาษาอังกฤษ.....	44
4.10 แสดงผลการทดสอบการสังเคราะห์เสียงภาษาอังกฤษ	46
4.11 แสดงผลรวมของแบบสอบถามความพึงพอใจ	50

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

หน้า

รูปที่

2.1 แสดงสถาปัตยกรรมของแอนดรอยด์.....	6
2.2 แสดงองค์ประกอบของการประมวลผลทั่วไปของระบบ Text-to-Speech.....	10
2.3 แสดงองค์ประกอบหน้าเว็บไซต์ tts-api.com.....	11
2.4 แสดงองค์ประกอบหน้าเว็บไซต์ www.vocalware.com/index/demo.....	11
2.5 แสดงองค์ประกอบหน้าเว็บไซต์ Google Translate.....	12
2.6 แสดงองค์ประกอบหน้าเว็บไซต์ Bing Translate.....	13
2.7 แสดงองค์ประกอบหน้าเว็บไซต์ Yandex Translate.....	14
2.8 แสดงองค์ประกอบหน้าเว็บไซต์ Frengly Translate.....	15
3.1 แสดงภาพหน้าเว็บไซต์ translate.baidu.co.th.....	19
3.2 แสดงภาพหน้าเว็บไซต์ Dictionary.cambridge.org.....	20
3.3 แสดงภาพหน้าเว็บไซต์ Oxforddictionaries.com.....	20
3.4 แสดงภาพแอปพลิเคชัน THAI DICT.....	21
3.5 แสดงภาพแอปพลิเคชัน แปลภาษา.....	22
3.6 แสดงภาพแอปพลิเคชัน Talking english easy.....	22
3.7 แสดงภาพองค์ประกอบของแอปพลิเคชันแปลภาษาด้วยเสียงพูด.....	24
4.1 แสดงภาพองค์ประกอบโดยรวมของแอปพลิเคชัน.....	27
4.2 แสดงภาพองค์ประกอบโดยรวมของแอปพลิเคชัน.....	28
4.3 กระบวนการทดสอบ.....	37
4.4 แสดงกราฟการรับเสียงภาษาอังกฤษและภาษาไทย.....	48

สารบัญรูป (ต่อ)

หน้า

รูปที่

4.5 แสดงกราฟการแปลภาษาอังกฤษและภาษาไทยของ Bing Translator Control API และ Google Translator Control API	48
4.6 แสดงกราฟการสังเคราะห์เสียงภาษาไทยและภาษาอังกฤษ	49



บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญ

ในปัจจุบันเทคโนโลยีการติดต่อสื่อสารมีการพัฒนาก้าวหน้ามากขึ้น ทุกคนสามารถแลกเปลี่ยนข้อมูลแสดงความคิดเห็นกันได้มากขึ้น และเทคโนโลยีที่เกี่ยวข้องกับการสื่อสารมีการพัฒนาอย่างต่อเนื่อง เพื่อตอบสนองความต้องการของมนุษย์ ในการติดต่อสื่อสารจะต้องอาศัยสื่อกลางเพื่อช่วยในการติดต่อสื่อสาร อาทิเช่น ภาษา เป็นต้น ภาษาถือเป็นสื่อกลางที่สำคัญในการสื่อสารเนื่องจากภาษาเปรียบเสมือนสัญลักษณ์ที่ช่วยให้ผู้รับสารและผู้ส่งสารเข้าใจตรงกัน เทคโนโลยีการสื่อสารได้พัฒนาและมีความก้าวหน้าอย่างรวดเร็ว รวมถึงผลิตภัณฑ์โทรศัพท์เคลื่อนที่สมาร์ทโฟน (Smart Phone) ที่มีรูปแบบการใช้งานที่แตกต่างจากอดีตซึ่งมีความสามารถเพียงแต่ติดต่อพูดคุยกันได้เพียงอย่างเดียวเท่านั้น แต่โทรศัพท์เคลื่อนที่ในปัจจุบันสามารถควบคุมการทำงานด้วยการสัมผัส (Touch Screen) และเชื่อมต่อกับเครือข่ายอินเทอร์เน็ตได้ ในขณะเดียวกันการเปิดกว้างของประเทศทำให้มีโอกาสพบปะติดต่อกับคนที่ไม่ใช่คนในชาติเดียวกันหรือประเทศเดียวกันเป็นไปได้ง่ายขึ้น และถ้าบางประเทศมีภาษาประจำชาติของตัวเองก็จะทำให้การสื่อสารเกิดอุปสรรค อย่างเช่น ในประเทศไทยที่ใช้ภาษาไทยเป็นภาษาประจำชาติ และเมื่อมีการเปิดประชาคมเศรษฐกิจอาเซียน AEC ทำให้ต้องมีการเตรียมพร้อมกับการติดต่อกับอีกหลายๆ ประเทศ ที่เข้ามาดำเนินธุรกิจและทำงานในประเทศ หรือการที่ต้องออกไปดำเนินธุรกิจและทำงานในต่างประเทศ

การเข้าสู่ประชาคมเศรษฐกิจอาเซียน AEC จะทำให้มีโอกาสได้เจอคนต่างชาติที่หลากหลายน่าสนใจและทำให้เกิดอุปสรรคในการสื่อสารเพราะบางทีอาจจะสื่อสารกันคนละภาษา ซึ่งในขณะเดียวกันประเทศของเราก็มีภาษาของตัวเองคือภาษาไทย เหมือนกับอีกหลายๆประเทศที่มีภาษาของตัวเอง แต่ในการสื่อสารกันนั้นจะใช้ภาษาสากลคือ ภาษาอังกฤษ ในความเป็นจริงบุคคลบางกลุ่มอาจจะไม่มีความเชี่ยวชาญในการใช้ภาษาอื่นๆ ที่ไม่ใช่ภาษาประจำชาติของตัวเอง ดังนั้นอาจจะมีปัญหาทางการสื่อสารเกิดขึ้นได้ อาทิเช่น ผู้ที่เดินทางไปท่องเที่ยวในต่างประเทศ หรืออาชีพที่เกี่ยวกับการให้บริการ ที่ต้องทำกิจกรรมหรือแลกเปลี่ยนทัศนคติ การสื่อสารมีความจำเป็นอย่างมากที่จะต้องสื่อสารให้เข้าใจ ไม่ว่าจะเป็น อาชีพพนักงานในร้านอาหาร อาชีพขับแท็กซี่ เป็นต้น การสื่อสารที่ผิดพลาดอาจทำให้เกิดความเข้าใจคลาดเคลื่อนและเสียเวลา ทั้งนี้ในปัจจุบันได้มีแอปพลิเคชันที่เกี่ยวกับการแปลภาษาให้สื่อใช้หลากหลายมากขึ้นในสมาร์ทโฟน (Smart Phone) ซึ่งก็มีรูปแบบในการใช้งานและประสิทธิภาพที่แตกต่างกันออกไป ซึ่งประสิทธิภาพของ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แอปพลิเคชันจะส่งผลกระทบต่อความถูกต้องของผลลัพธ์หรือทำให้ผลลัพธ์คลาดเคลื่อนไปจากความต้องการจริงของผู้ใช้งาน ทำให้การสื่อสารไม่เป็นไปตามที่ผู้ใช้งานต้องการ

จากปัญหาข้างต้นจึงเกิดการศึกษารเปรียบเทียบส่วนต่อประสานโปรแกรมประยุกต์สำหรับระบบแปลภาษาด้วยเสียงพูดขึ้นเพื่อวัดประสิทธิภาพในแต่ละส่วนของระบบแปลภาษาด้วยเสียงพูด รวมทั้งช่วยลดอุปสรรคและลดช่องว่างในการติดต่อสื่อสารดังนั้นผู้ดำเนินโครงการจะนำประโยชน์ของโทรศัพท์เคลื่อนที่สมาร์ทโฟน (Smart Phone) ที่จะพกพากันโดยทั่วไปอยู่แล้ว โดยใช้ความสามารถของซอฟต์แวร์ชุดเครื่องมือหรือที่เรียกอีกอย่างว่า Application Programming Interface (API) ซึ่งจะมีให้เลือกและทดลองใช้อย่างหลากหลาย ที่จะประกอบด้วยชุดคำสั่งที่นักพัฒนาซอฟต์แวร์สามารถเรียกใช้ได้มาใช้ในการพัฒนาแอปพลิเคชันนี้ โดยการเรียกใช้บริการนั้นจะเรียกผ่านอุปกรณ์พกพา (Mobile Device) ที่รองรับระบบปฏิบัติการแอนดรอยด์ (Android) ที่สามารถเชื่อมต่อเข้ากับอินเทอร์เน็ตได้ โดยผู้ดำเนินโครงการนี้จะศึกษาเปรียบเทียบคุณสมบัติ API ที่เกี่ยวข้องกับระบบแปลภาษาด้วยเสียงพูด เพื่อเลือก API ที่มีความเหมาะสมกับสภาพแวดล้อมการใช้งาน และนำ API ในส่วนต่างๆ มาวัดประสิทธิภาพด้วยการพัฒนาแอปพลิเคชันแปลภาษาด้วยเสียงพูด

หลังจากที่ได้ศึกษาเปรียบเทียบและพัฒนาระบบดังกล่าวทำให้ทราบถึงประสิทธิภาพและข้อบกพร่องในแต่ละส่วน รวมถึงประสิทธิภาพโดยรวมของระบบแปลภาษาด้วยเสียงพูด อีกทั้งเป็นประโยชน์กับผู้ที่จะพัฒนาระบบแปลภาษาในรูปแบบอื่นๆ โดยผู้ที่สนใจสามารถศึกษาและเลือกใช้ API ที่เหมาะสำหรับแอปพลิเคชันที่พัฒนาได้ เพื่อลดขั้นตอนในการศึกษาพัฒนารวมทั้งช่วยให้แอปพลิเคชันที่พัฒนามีประสิทธิภาพและช่วยลดช่องว่างในการติดต่อสื่อสารได้ดียิ่งขึ้นได้

1.2 เป้าหมายและวัตถุประสงค์

- 1) เพื่อศึกษาระบบแปลภาษาด้วยเสียงพูดแบบหลายภาษาบนระบบปฏิบัติการแอนดรอยด์
- 2) เพื่อพัฒนาระบบแปลภาษาด้วยเสียงพูดแบบหลายภาษาบนระบบปฏิบัติการแอนดรอยด์
- 3) เพื่อเปรียบเทียบคุณสมบัติของ API (Application programming interface) ในระบบแปลภาษาด้วยเสียงพูด
- 4) เพื่อวัดประสิทธิภาพของ API ที่เลือกใช้กับระบบแปลภาษาด้วยเสียงพูด

1.3 ขอบเขตของงาน

- 1) เพื่อศึกษาค้นคว้าและพัฒนาระบบแปลภาษาด้วยเสียงพูดในระบบปฏิบัติการ Android

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- 2) เพื่อเปรียบเทียบคุณสมบัติของ API ในส่วน Speech-to-Text, Translation Language และ Text-to-Speech
- 3) เพื่อเลือก API ที่มีคุณสมบัติเหมาะสมกับแอปพลิเคชัน
- 4) แอปพลิเคชันพัฒนาโดยใช้ API เป็นส่วนประกอบในขั้นตอนต่างๆ
- 5) แอปพลิเคชันสามารถรับเสียงพูดได้และแปลงออกเป็นภาษาอื่นได้
- 6) แอปพลิเคชันสามารถเลือกภาษาที่ต้องการแปลได้
- 7) แอปพลิเคชันประมวลผลการทำงานแบบออนไลน์
- 8) แอปพลิเคชันสามารถแปลงข้อความเป็นเสียงพูดได้
- 9) แอปพลิเคชันสามารถแปลงเสียงพูดเป็นข้อความได้
- 10) แอปพลิเคชันสามารถเลือก API สำหรับแปลภาษาได้

1.4 วิธีการดำเนินงาน

- 1) ศึกษาสถาปัตยกรรมของ Android Application และการพัฒนา Application บนแอนดรอยด์
- 2) ศึกษาทฤษฎีที่เกี่ยวข้อง
- 3) ศึกษาวิธีการเขียนและเครื่องมือที่ใช้ในการพัฒนาระบบ
 - ศึกษาการพัฒนาซอฟต์แวร์ที่ใช้ในระบบปฏิบัติการ Android Application
 - ศึกษาการเขียนโปรแกรมที่เรียกใช้ API
- 4) วิเคราะห์และออกแบบระบบให้สอดคล้องกับวิธีการที่ศึกษา
 - กำหนดรายละเอียดของโปรแกรม
 - ออกแบบ User Interface ของโปรแกรม
- 5) พัฒนาระบบตามที่ได้วิเคราะห์และออกแบบ
- 6) ทำการตรวจสอบระบบ (ทดลองใช้) และปรับปรุงแก้ไขระบบ
- 7) ประเมินผลระบบ ทำคู่มือการใช้งานและนำระบบไปใช้งานจริง

1.5 ประโยชน์ที่คาดว่าจะได้รับ

- 1) ช่วยอำนวยความสะดวกให้กับบุคคลที่ไม่มีความเชี่ยวชาญในภาษาที่ต้องการสื่อสาร ตัวอย่างเช่นนักท่องเที่ยว อาชีพให้บริการ แท็กซี่ เป็นต้น
- 2) สามารถเข้าถึงการใช้งานได้ง่าย สะดวกรวดเร็วเพราะเรียกใช้งานผ่านอุปกรณ์พกพา (Mobile Device)
- 3) แอปพลิเคชันสามารถทำให้การสื่อสารมีความเข้าใจมากขึ้น และลดความผิดพลาดในการสื่อสาร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- 4) แอปพลิเคชันใช้ในการเรียนรู้ภาษาสำหรับผู้สนใจ
- 5) แอปพลิเคชันสะดวกสำหรับบุคคลที่มีปัญหาในเรื่องสายตาเพราะสามารถรับเสียงพูดได้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

ทฤษฎีพื้นฐานที่เกี่ยวข้อง

เนื่องจากโครงการนี้เป็นการพัฒนาแอปพลิเคชันที่ใช้งานบนระบบปฏิบัติการแอนดรอยด์ และนำประโยชน์ของ API ที่เปิดให้บริการโดยทั่วไปมาใช้ในขั้นตอนต่างๆของการพัฒนาโปรแกรม ไม่ว่าจะเป็น Speech-to-Text, Translation Language และ Text-to-Speech โดยใช้ภาษา JAVA ในการสร้างและพัฒนาแอปพลิเคชันดังกล่าว ดังนั้นจึงต้องทำการศึกษาเกี่ยวกับทฤษฎีพื้นฐาน เพื่อที่จะสามารถเข้าใจและง่ายต่อการนำไปพัฒนาโครงการ โดยในบทที่ 2 จะกล่าวถึงส่วนต่างๆ ที่ได้กล่าวไปแล้วข้างต้นและบทวิจัยที่เกี่ยวข้องกับโครงการ

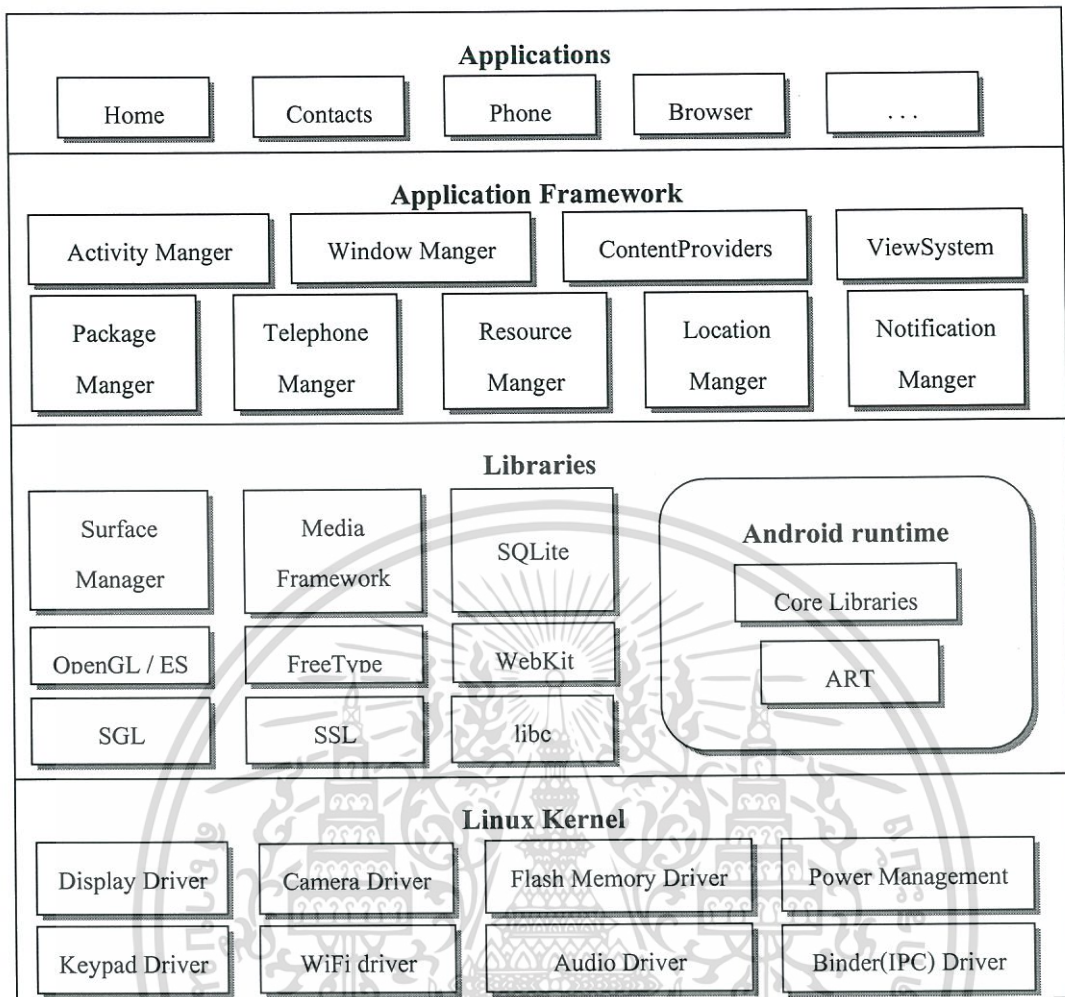
2.1 ระบบปฏิบัติการแอนดรอยด์

Google Android เป็นชื่อเรียกอย่างเป็นทางการของระบบปฏิบัติการแอนดรอยด์ โดยมีบริษัท Google โดยเปิดให้นักพัฒนาสามารถนำรหัสต้นฉบับไปพัฒนาปรับแต่งได้อย่างอิสระ ทำให้แอนดรอยด์มีผู้เข้าร่วมพัฒนาเป็นจำนวนมาก และพัฒนาไปได้อย่างรวดเร็ว

แอนดรอยด์เป็นซอฟต์แวร์ที่มีโครงสร้างแบบเรียงทับซ้อนหรือแบบ Stack ซึ่งรวมเอาระบบปฏิบัติการ Middleware และแอปพลิเคชันที่สำคัญเข้าไว้ด้วยกัน เพื่อใช้สำหรับทำงานบนอุปกรณ์พกพาเคลื่อนที่เช่น โทรศัพท์มือถือและแท็บเล็ต เป็นต้น

การทำงานของแอนดรอยด์มีพื้นฐานอยู่บนระบบ Linux Kernel ซึ่งใช้ Android Software Development Kit (SDK) เป็นเครื่องมือสำหรับการพัฒนาแอปพลิเคชันบนระบบปฏิบัติการแอนดรอยด์และใช้ภาษา Java ในการพัฒนา [1]

สถาปัตยกรรมของแอนดรอยด์ (Android Architecture) นั้นถูกแบ่งออกเป็นลำดับชั้นออกเป็น 4 ชั้นหลักดังในรูปที่ 2.1



รูปที่ 2.1 แสดงสถาปัตยกรรมของแอนดรอยด์

เป็นการแสดงสถาปัตยกรรมของแอนดรอยด์ซึ่งประกอบด้วยส่วนสำคัญหลัก 5 ส่วนคือ Applications, Application framework, Libraries, Android runtime และ Linux kernels [2]

1) **Applications** เป็นชั้นบนสุดของโครงสร้าง Android ซึ่งเป็นส่วนของแอปพลิเคชันที่พัฒนาขึ้นมาใช้งาน เช่น แอปพลิเคชันรับส่งอีเมล, แอปพลิเคชันโทรศัพท์ (Phone Dial) และแอปพลิเคชันเว็บเบราว์เซอร์ (Web Browser) เป็นต้น ทั้งนี้โปรแกรมในชั้นแอปพลิเคชันนั้นจะอยู่ในรูปแบบของไฟล์ .apk ซึ่งโดยทั่วไปแล้วจะอยู่ในไดเรกทอรี data/app ของโทรศัพท์ ซึ่งแอปพลิเคชันทั้งหมดในส่วนนี้จะเขียนด้วยภาษา JAVA

2) **Application framework** ในชั้นนี้นักพัฒนาสามารถเข้าเรียกใช้งาน เป็นตัวกลางช่วยจัดการ ซึ่งถูกพัฒนาโดยใช้ภาษา JAVA ผ่าน API โดยแอนดรอยด์ได้ออกแบบไว้เพื่อลดความซ้ำซ้อนในการใช้งาน application component จะช่วยอำนวยความสะดวกในการจัดการในการเข้าถึงแหล่งข้อมูลต่างๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3) **Libraries** เป็นชั้นที่แอนดรอยด์ได้รวบรวมกลุ่มของไลบรารีต่างๆ ที่สำคัญและมีความจำเป็นเอาไว้มากมาย เพื่ออำนวยความสะดวกให้กับนักพัฒนาและง่ายต่อการพัฒนาโปรแกรมซึ่งไลบรารีจะถูกเขียนด้วยภาษา C/C++ ที่สามารถใช้งานผ่าน API ของ Framework ที่ซึ่งระบบปฏิบัติการแอนดรอยด์ได้จัดไว้ให้

4) **Android runtime** เป็นชั้นย่อยอยู่ใน Libraries มี 2 ส่วน

- ART เป็น runtime ตัวใหม่ที่ Google เอามาแทน Dalvik ซึ่งจะมีใช้งานใน android 4.4 KitKat ขึ้นไป ART ใช้การทำงานแบบ Ahead-Of-Time (AOT) คือ ตอนที่ติดตั้งจะแปลงโค้ดจากภาษาระดับสูงไปภาษาเครื่องครั้งเดียว จากนั้นจะใช้โค้ดส่วนที่แปลงแล้ว run ทำให้เร็วขึ้นเนื่องจากการแปลงตอนที่ลงครั้งแรกครั้งเดียวเท่านั้น

- Core Java Library ส่วนนี้เป็นไลบรารีมาตรฐานซึ่งรวบรวมคำสั่งที่ถูกเขียนมาจากภาษา Java

5) **Linux kernels** ระบบแอนดรอยด์นั้นถูกสร้างบนพื้นฐานของระบบปฏิบัติการ Linux โดยส่วนมากแล้วจะเกี่ยวข้องกับฮาร์ดแวร์โดยตรงเช่นการจัดการหน่วยความจำ, การจัดการโปรเซสและการเชื่อมต่อเครือข่าย

2.2 Application Programming Interface

Application Programming Interface (API) คือ ชุดคำสั่งหรือข้อกำหนด ที่ถูกออกแบบมาเพื่อทำหน้าที่เป็นตัวกลางในการเชื่อมต่อการทำงานระหว่างแอปพลิเคชันหรือเชื่อมการทำงานเข้ากับระบบปฏิบัติการ [3] การที่แอปพลิเคชันจะเชื่อมต่อการทำงานกับระบบปฏิบัติการได้นั้น จะต้องใช้ API เป็นตัวเชื่อม ซึ่งถ้าหากไม่มีการเปิดเผย API ของระบบปฏิบัติการออกมา จะทำให้ผู้พัฒนาไม่สามารถพัฒนาแอปพลิเคชันให้ทำงานเข้ากับระบบปฏิบัติการได้

ในปัจจุบันมี API มากมายเพื่อให้ นักพัฒนาสามารถนำไปใช้ในการพัฒนาแอปพลิเคชันได้ อาทิ YouTube API, Amazon Product Advertising API, Google API และ Bing API เป็นต้น ซึ่งในแอปพลิเคชันนี้ผู้พัฒนาได้เลือกศึกษาและใช้ API ที่เกี่ยวกับการแปลภาษาดังนี้

- Speech-to-text API คือ กระบวนการของการแปลงเสียงพูดของมนุษย์ให้เป็นข้อความลักษณะการใช้งานคือ ต้องพูดผ่านอุปกรณ์เสริมอย่างเช่น ไมโครโฟน เป็นต้น
- Translator API คือ การถ่ายทอดข้อความจากภาษาหนึ่งไปยังอีกภาษาหนึ่งโดยรักษารูปแบบและความหมายในข้อความเดิมไว้อย่างครบถ้วน
- Text-to-speech API คือ กระบวนการของการแปลงข้อความให้ออกมาเป็นเสียงพูดของมนุษย์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประโยชน์ของ Application Programming Interface

- 1) สามารถรับและส่งข้อมูลข้าม Server ได้
- 2) ไม่จำเป็นต้องเข้าหน้าเว็บหลัก ก็มีข้อมูลของเว็บหลัก จากเว็บที่ดึง API
- 3) เป็นตัวกลางที่ทำให้โปรแกรมประยุกต์เชื่อมต่อกับโปรแกรมประยุกต์อื่นหรือเชื่อมการทำงานเข้ากับระบบปฏิบัติการ
- 4) อำนวยความสะดวกให้กับนักพัฒนามากขึ้น ไม่ต้องเขียนระบบหรือฟังก์ชันในส่วนนั้นเอง เป็นการลดความยุ่งยากและซับซ้อน

2.3 Speech-to-Text

Speech-to-Text (STT) หรือ Speech Recognition คือการที่คอมพิวเตอร์สามารถรับรู้เสียงของมนุษย์ได้โดยอัตโนมัติเป็น Software driver ที่แปลงสัญญาณเสียงเป็นสัญญาณ digital และส่งคำที่แปลงได้ออกมาเป็นข้อความให้โปรแกรมทำงาน [3] Recognizers ส่วนมากจะสนับสนุนคำพูดที่มีความต่อเนื่อง ทำให้สามารถพูดได้อย่างเป็นธรรมชาติผ่านไมโครโฟนเข้าไปด้วยความเร็วที่ใช้ในการสนทนากันตามปกติซึ่ง Recognizers ในปัจจุบันนี้จะสนับสนุนการทำงาน 2 แบบคือ

- ผู้ใช้สามารถป้อนข้อมูลให้คอมพิวเตอร์โดยตรง เช่น การอ่านบันทึก, จดหมาย หรือ e-mail ให้โปรแกรมพิมพ์ตาม เป็นต้น

- ผู้ใช้สั่งให้โปรแกรมเริ่มการทำงานโดยการพูดคำสั่งหรือการถามคำถาม

Speech Recognition วิเคราะห์และควบคุมการทำงานระบบด้วยเสียงโดยจะพยายามทำความเข้าใจในชุดข้อมูลเสียงที่ได้รับด้วยการตรวจสอบโครงสร้างและรูปแบบของเสียงว่าถูกต้องตามที่กำหนดไว้หรือไม่และจะดำเนินการตามเงื่อนไขของโปรแกรมที่ได้กำหนดไว้

เทคโนโลยีเหล่านี้ช่วยให้สามารถนำระบบรู้จำเสียงพูดไปใช้งานจริงได้อย่างสะดวกและเกิดประโยชน์ต่อการทำงานและการดำรงชีวิตเป็นอย่างมาก เทคโนโลยีเหล่านี้ได้แก่ Distributed Speech Recognition (DSR) [3]

ในขณะที่ Mobile Computing มีความแพร่หลายมากขึ้น แอปพลิเคชันบนอุปกรณ์อิเล็กทรอนิกส์แบบพกพาเหล่านี้ อาจอาศัยระบบการรู้จำเสียงช่วยในการรับข้อมูลเข้าจากผู้ใช้งานซึ่งทำให้ผู้ใช้ได้รับความสะดวกในการป้อนข้อมูลในกรณีที่ผู้ใช้พิมพ์ไม่ถนัด แต่เนื่องจากการประมวลผลการรู้จำเสียงต้องใช้ความสามารถของ CPU และ Memory สูงมากเพื่อให้ได้ผลลัพธ์รวดเร็วและแม่นยำ การใช้งานระบบการรู้จำเสียงบนอุปกรณ์อิเล็กทรอนิกส์แบบพกพาซึ่งมักจะมีสมรรถนะในการทำงานจำกัดจึงเป็นเรื่องที่ยากมาก ด้วยเหตุผลนี้เอง จึงมีการออกแบบให้การรู้จำเสียงทำงานโดยวิเคราะห์คลื่นเสียงบนอุปกรณ์อิเล็กทรอนิกส์แบบพกพา แล้วส่งข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ลักษณะเฉพาะของคลื่นเสียงไปยังเซิร์ฟเวอร์เพื่อประมวลผลการรู้จำเสียง แล้วจึงส่งผลกลับไปยังอุปกรณ์อิเล็กทรอนิกส์แบบพกพา วิธีการนี้เรียกว่า Distributed Speech Recognition (DSR)

API ที่นำมาพิจารณาในขั้นตอนของ Speech-to-Text คือ

- 1) Google Speech Recognition
- 2) Bing Speech Recognition Control

ทั้ง Google Speech Recognition และ Bing Speech Recognition เป็น API ที่มีลักษณะการแปลงสัญญาณเสียงพูดให้เป็นข้อความเขียน โดยการรับข้อมูลเสียงจากไมโครโฟนและส่งข้อมูลเสียงไปยังบริการเว็บสำหรับการวิเคราะห์และจากนั้นจะส่งกลับมามีความคำพูดของผู้พูดเป็นข้อความ

2.4 Text-to-Speech

Text-to-Speech (TTS) หรือ Speech Synthesis คือกระบวนการของการแปลงข้อความ (Text) ให้เป็นเสียงภาษาพูดของมนุษย์ในการใช้งานส่วนใหญ่จะต้องใช้งานร่วมกับเทคโนโลยีด้านการประมวลผลภาษา (Language Processing Technology) ทำให้ได้เทคโนโลยีสังเคราะห์เสียงจากข้อความ Text-to-Speech Synthesis (TTS) ที่สามารถนำไปประยุกต์ใช้ได้ เป็นกระบวนการที่มีการแยกย่อยคำออกได้เป็นส่วนที่เล็กที่สุดของเสียงในภาษาพูดโดยมีการวิเคราะห์ในเรื่องของการออกเสียงสูงต่ำการเว้นวรรคและสร้าง Output ที่เป็น Digital audio ออกมา สามารถแสดงกระบวนการสังเคราะห์เสียงพูด Text-to-Speech ดังรูปภาพที่ 2.2

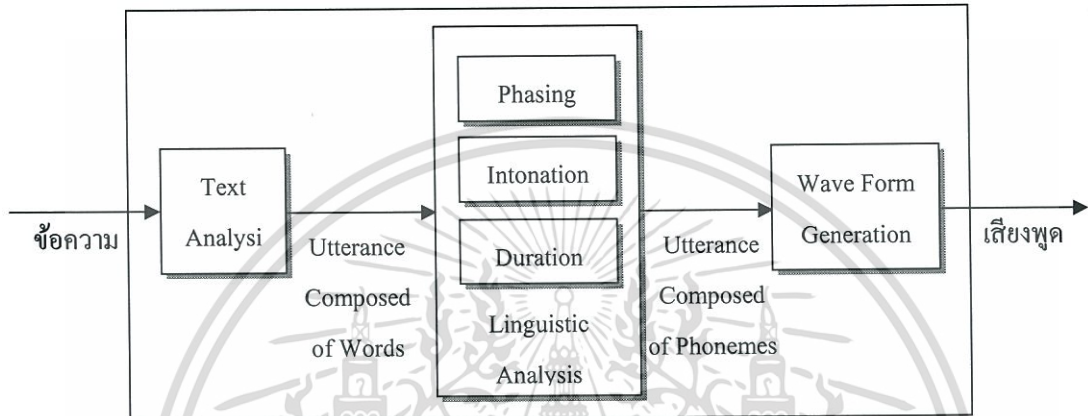
สามารถแบ่งประเภทของ Text-to-Speech engine ได้จากวิธีที่ใช้ในการแปลงส่วนที่เล็กที่สุดของเสียงในภาษาพูด เป็นเสียงที่สามารถได้ยินได้ทั้งหมด 3 ประเภท ดังนี้

1) Concatenated Word ถึงแม้ว่าวิธี Concatenated Word จะไม่ใช่กระบวนการสังเคราะห์เสียงอย่างแท้จริงแต่ก็เป็นวิธีที่ใช้กันมากในระบบ Text-to-Speech ทั่วๆไปซึ่ง engine ที่ใช้วิธีนี้ผู้ออกแบบโปรแกรมจะทำการบันทึกคำพูดและคำเฉพาะลงในฐานข้อมูล เมื่อมีการเรียกใช้โปรแกรมตัว engine จะทำการดึงเสียงที่ได้บันทึกไว้จากฐานข้อมูลแล้วนำมาต่อกันและออกเสียงเป็นประโยคหรือคำพูด

2) Synthesis Text-to-Speech engine จะใช้กระบวนการสังเคราะห์เสียงเพื่อทำการสร้างเสียงที่คล้ายเสียงพูดของมนุษย์ และการใช้ตัวกรองเพื่อจำลองเสียงที่ออกมาจากลำคอ, ช่องปาก, ลักษณะของริมฝีปากและตำแหน่งของลิ้น เสียงที่ถูกสร้างขึ้นโดยเทคโนโลยีการสังเคราะห์เสียงปัจจุบันมีแนวโน้มที่จะได้เสียงใกล้เคียงกับเสียงมนุษย์น้อยกว่าเสียงที่ผลิตโดยวิธี Diphone Concatenation แต่ก็มีความเป็นไปได้ที่จะทำให้มีคุณภาพเสียงดีขึ้นโดยการเปลี่ยน Parameter บางตัว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3) Diphone Concatenation Text-to-Speech engine จะใช้การปะติดปะต่อเชื่อมส่วนของ digital-audio สั้นๆ เข้าด้วยกันและทำให้รอยต่อระหว่างเสียงมีความราบรื่นเพื่อให้เสียงที่ได้มีความต่อเนื่องมากขึ้น การเตรียมส่วนต่างๆ ของเสียงจะได้มาจากการบันทึกเสียงของมนุษย์ในหลายรูปแบบและจะต้องพยายามระบุนการเริ่มต้นและการสิ้นสุดของพยางค์เสียงถึงแม้ว่าเทคนิคนี้สามารถสร้างเสียงที่มีความเหมือนจริงมากขึ้น แต่ก็ต้องคำนึงถึงจำนวนของงานเพื่อสร้างเสียงใหม่และต้องเป็นเสียงที่ไม่จำกัดเพราะว่าพยางค์เสียงจะขึ้นอยู่กับภาษาที่ผู้พูดใช้เท่านั้น



รูปที่ 2.2 แสดงองค์ประกอบของการประมวลผลทั่วไปของระบบ Text-to-Speech

- ขั้นตอนในการสังเคราะห์เสียงพูด

ประกอบด้วย 2 ขั้นตอน ดังนี้

1) ส่วนต้น (Front-end) ประกอบด้วยงานหลัก 2 งานคือ

- การเปลี่ยนข้อความเข้าที่อยู่ในรูปสัญลักษณ์และตัวเลขให้อยู่ในรูปของคำศัพท์ ซึ่งขั้นตอนนี้มักถูกเรียกว่า Text Normalization, Pre-Processing หรือ Tokenization

- การนำคำศัพท์แต่ละคำมาแยกและทำสัญลักษณ์ว่าแต่ละคำนั้นเป็นพยางค์หรือประโยค ขั้นตอนนี้เรียกว่ากระบวนการ Text-to-Phoneme หรือ Grapheme-to-Phoneme จากนั้นก็นำผลลัพธ์ที่ได้มาเพิ่มข้อมูลเกี่ยวกับเสียงพูด แล้วส่งต่อไปให้กระบวนการต่อไป

2) ส่วนปลาย (Back-end) ทำหน้าที่รับประโยคที่มีเสียงกำกับ (phoneme) แล้วนำมาสังเคราะห์เป็นเสียงพูด

ตัวอย่างเว็บไซต์ที่เป็น Text-to-Speech API หรือ Speech Synthesis

รูปที่ 2.3 แสดงองค์ประกอบหน้าเว็บไซต์ tts-api.com

ทางเว็บ <http://tts-api.com> ได้เปิดให้ใช้บริการฟรี โดยพิมพ์ข้อความที่ต้องการอยากรู้ ว่าข้อความนั้นอ่านออกเสียงว่าอย่างไร แล้วกด Convert from text to speech (MP3) หลังจากนั้นจะได้ยินเสียงของข้อความนั้น นอกจากนั้นยังมีลิงค์ URL ตัวอย่างให้อีกด้วย ดังรูปที่ 2.3

รูปที่ 2.4 แสดงองค์ประกอบหน้าเว็บไซต์ www.vocalware.com/index/demo

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

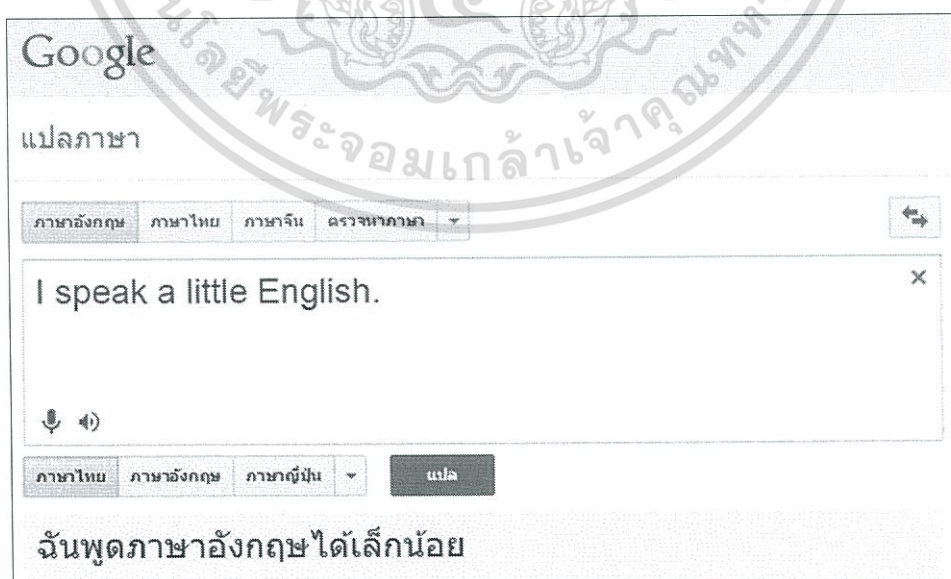
เว็บไซต์ Vocalware เป็นเว็บที่เปิดให้ใช้บริการฟรีผ่านอินเทอร์เน็ต โดยผู้ใช้งานสามารถเลือกภาษาที่จะให้โปรแกรมออกเสียงให้เป็นตัวอย่างได้ โดยการพิมพ์ข้อความของภาษานั้นๆ โปรแกรมจะจำกัดจำนวนตัวอักษรไว้ที่ 150 ตัวอักษรดังรูปที่ 2.4

2.5 การแปลภาษา

การแปลภาษา คือ การถ่ายทอดข้อความจากภาษาหนึ่งไปยังอีกภาษาหนึ่ง โดยรักษารูปแบบและความหมายในข้อความเดิมไว้อย่างครบถ้วน [4] ในปัจจุบันการแปลภาษาสามารถใช้งานในรูปแบบออนไลน์ได้ นอกจากนี้ในบางเว็บไซต์ที่ให้บริการแปลภาษายังมีการเปิดให้ใช้ API ให้ผู้พัฒนานำไปใช้กันอีกด้วย อาทิเช่น

1) **Google Translate** ในส่วนของ Android version Google Translate เป็นโปรแกรมที่สามารถดาวน์โหลดได้ฟรีสำหรับผู้ใช้งานระบบปฏิบัติการบนอุปกรณ์อิเล็กทรอนิกส์แบบพกพาที่สร้างโดย Google สำหรับรุ่นแรกที่เปิดตัวให้ใช้ในการทำงานจะเป็นการแปลภาษาสำหรับ Android มีเพียงแค่สองตัวเลือกหลัก คือ SMS translate และ History

Google Translate API เป็นบริการที่เปิดให้ใช้บริการ ซึ่งสามารถแปลภาษาได้หลายภาษาให้บริการโดย Google Inc. ในการแปลข้อความให้เป็นลายลักษณ์อักษรจากภาษาหนึ่งไปยังอีกหลายๆ ภาษา Google Translate มีส่วนติดต่อผู้ใช้รูปแบบเว็บไซต์ (web Interface) สำหรับติดต่อกับผู้ใช้งานบนอุปกรณ์อิเล็กทรอนิกส์แบบพกพาสำหรับ Android และ API ที่นักพัฒนาซอฟต์แวร์สามารถใช้งานได้เพื่อสร้างส่วนขยายในแอปพลิเคชัน โปรแกรมหรือซอฟต์แวร์อื่นๆ

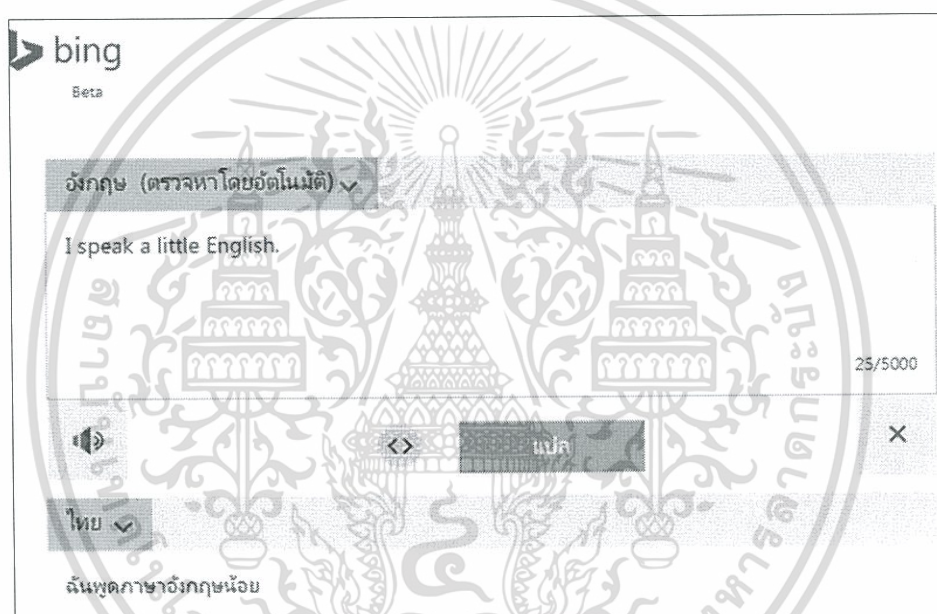


รูปที่ 2.5 แสดงองค์ประกอบหน้าเว็บไซต์ Google Translate

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เว็บไซต์ Google Translate จะมีหลายภาษาให้เลือกในการที่จะแปลข้อความของภาษาหนึ่งไปยังอีกภาษาหนึ่งยกตัวอย่างเช่น คำว่า I speak a little English. เลือกภาษาปลายทางเป็นภาษาไทย เมื่อพิมพ์ภาษาต้นทางคือภาษาอังกฤษลงในกรอบสี่เหลี่ยมด้านบนดังรูป 2.5 แล้วกดแปลความหมาย จะปรากฏในสี่เหลี่ยมด้านล่าง นอกจากนี้ยังสามารถแก้คำที่ผิดบริบทหรือความหมายได้อีกด้วย

2) **Bing Translator** เป็นเว็บไซต์แปลภาษาที่ให้บริการโดยไมโครซอฟท์ ซึ่งเปิดให้บริการทั้ง Windows Store และ Bing Translator Control API ใช้สำหรับแปลข้อความภาษาต่างๆ โดยใช้เครื่องมือ Microsoft Translator API อีกชั้นหนึ่งซึ่งการใช้บริการ Bing Translator Control API ต้องสมัคร App ID ของ Bing เสียก่อน ซึ่งในการให้บริการ API ตัวฟรีจะสามารถแปลได้ 2,000 บทความต่อเดือนซึ่งแต่ละบทความต้องน้อยกว่า 1,000 ตัวอักษร

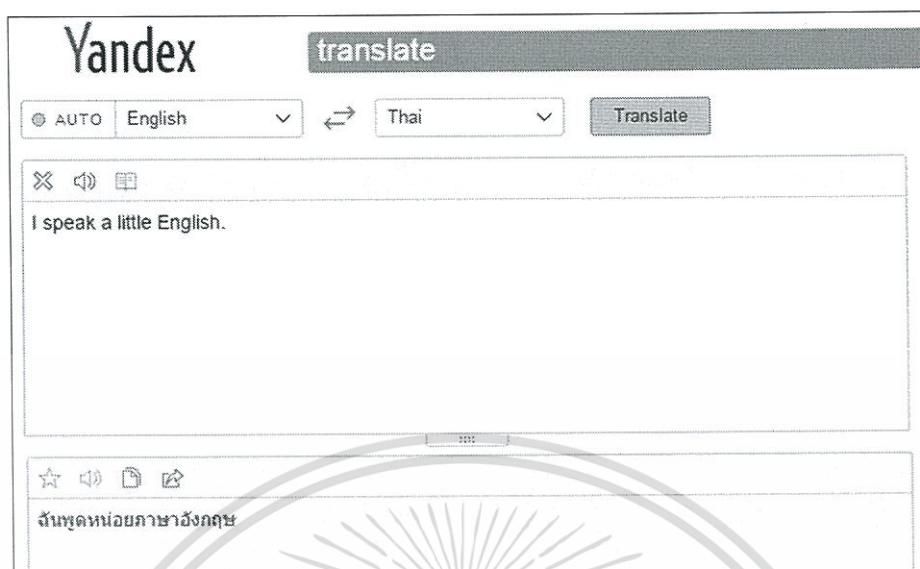


รูปที่ 2.6 แสดงองค์ประกอบหน้าเว็บไซต์ Bing Translate

เว็บไซต์ Bing Translate มีลักษณะการทำงานคล้ายๆ กับเว็บไซต์ Google Translate แต่ Bing Translate ไม่สามารถแก้ไขบริบทหรือข้อความ คำศัพท์ที่แปลมาได้ ทำได้เพียงการเลื่อนเมาส์เพื่อดูความหมายของคำศัพท์แต่ละคำที่อยู่ในประโยคได้ ดังรูป 2.6

3) **Yandex Translate** เป็นเว็บไซต์เสิร์ชเอนจินที่ได้รับความนิยมมากในประเทศรัสเซีย Yandex ได้รับการจัดอันดับเป็นเว็บไซต์แปลภาษาอันดับ 1 ในประเทศรัสเซีย โดย Yandex ให้บริการพจนานุกรมและการแปลภาษาด้วยการใช้เทคโนโลยีจากบริษัทซอฟต์แวร์ภาษา www.ABBYY.com และเว็บไซต์ Mail.ru สำหรับแปลภาษาและพจนานุกรมภาษา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

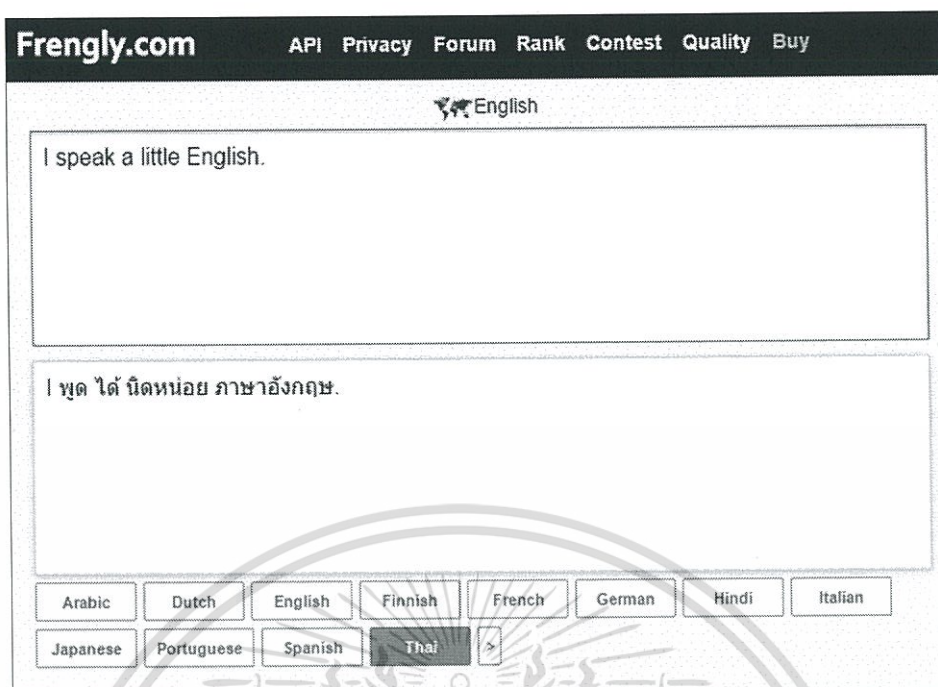


รูปที่ 2.7 แสดงองค์ประกอบหน้าเว็บไซต์ Yandex Translate

เว็บไซต์ Yandex Translate เนื่องจากเป็นเว็บไซต์ของประเทศรัสเซียการแปลข้อความออกมาเป็นภาษาไทยอาจจะคลาดเคลื่อนมากกว่าเว็บที่นิยมใช้ในวงกว้างเช่น Google และ Bing ดังรูป 2.7 ที่ Yandex Translate แปลออกมาเป็นภาษาไทยว่า ฉันพูดน้อยภาษาอังกฤษ จากประโยคภาษาอังกฤษ I speak a little English.

4) **Frengly Translator** เป็นเว็บให้บริการแปลภาษาข้อความจากภาษาหนึ่งไปยังอีกหลายๆ ภาษาซึ่งสามารถแปลออกมาได้ถึง 24 ภาษาโดยการพิมพ์ข้อความเพียงครั้งเดียว Frengly ง่ายต่อการใช้เพียงแค่พิมพ์ข้อความที่ต้องการจะแปลในกล่องข้อความและคลิกภาษาที่ได้รับการแปลเป็นภาษาอัตโนมัติตรวจพบและแปลได้อย่างรวดเร็วเพราะมันจะขึ้นอยู่กับผลการแปลผลของ Google นอกจากนี้ Frengly เป็นบริการแปลฟรีที่อยู่บนพื้นฐานของ Google Ajax ภาษา API มีลักษณะการใช้งานเหมือน Google Translator แต่ใช้งานได้ง่ายกว่าในการเลือกภาษาปลายทางเพราะไม่ต้องกดเลื่อนลงเพื่อเลือกภาษาซึ่งในเว็บของ Frengly ก็จะมี Frengly Translation API ให้บริการด้วยเช่นกัน

เว็บไซต์ Frengly Translate การแปลภาษาที่เห็นได้ดังรูป 2.8 ยังมีความคลาดเคลื่อนโดยที่เว็บไม่แปลประธานที่เป็น I และเว็บแปลตามคำศัพท์ตามประโยคโดยที่ไม่มีการเรียงใหม่ให้เป็นประโยค



รูปที่ 2.8 แสดงองค์ประกอบหน้าเว็บไซต์ Frengly Translate

ประโยชน์ของ Translate API

- 1) รวดเร็ว เนื่องจากไม่จำเป็นต้องอัปเดตไฟล์ฯลฯ ข้อความที่ผู้ใช้ต้องการแปลจะถูกส่งโดยตรงจากระบบของผู้ใช้และระบบจะได้รับข้อความที่แปลกลับมาอีกครั้ง
- 2) ใช้งานง่าย ทันทีที่ข้อความที่ต้องการถูกส่งไปแปลโดยอัตโนมัติ เมื่อดำเนินการแปลเสร็จสมบูรณ์แล้วข้อความที่แปลก็จะถูกส่งกลับมาที่ระบบโดยอัตโนมัติเช่นกัน

2.6 งานวิจัยที่เกี่ยวข้อง

2.6.1) Speech-to-Text

M. Tomalin และทีมงาน [10] ได้ศึกษาและพัฒนาระบบแปลภาษาจากเสียงพูดภาษาแมนดารินเป็นข้อความภาษาอังกฤษ และประเมินประสิทธิภาพการทำงานในช่วงการถอดความเสียงพูดและการแปลภาษา การทำงานของระบบนี้แบ่งเป็น 3 องค์ประกอบหลัก ได้แก่ ส่วน Speech-to-Text (STT) คือ การแปลงเสียงพูดภาษาแมนดารินเป็นข้อความภาษาแมนดาริน, ส่วน Integration คือ การแบ่งประโยคให้เป็นคำ รวมถึงพวกตัวเลขและส่วน Statistical Machine Translation (SMT) คือ การแปลงข้อความภาษาแมนดารินไปเป็นข้อความภาษาอังกฤษ พวกเขาพัฒนาระบบแปลภาษาขึ้นมา 2 ระบบจากนั้นพวกเขาทำการเปรียบเทียบประสิทธิภาพของสองระบบ โดยใช้วิธี Character Error Rate (CER) สำหรับประเมินประสิทธิภาพของ STT และใช้วิธี Translation Edit Rate (TER) และ BLEU metrics เพื่อวัดประสิทธิภาพการทำงานของทั้งระบบ ผลจากการประเมินประสิทธิภาพทำให้ออกเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

พวกเขาพบว่า SMT มีความถูกต้องมากขึ้นเมื่อใช้เทคนิคการแบ่งส่วนเพื่อป้อนประโยคที่มีรูปแบบเหมือนกัน

Ahmad A. M. และทีมงาน [11] ได้ศึกษาการออกแบบและการทำงานของระบบการจำแนกเสียงพูดด้วยวิธีการ Hidden Markov Model โดยพวกเขาเน้นที่การจำแนกเสียงพูดในตัวเลขภาษาอังกฤษตั้งแต่ 0-9 ดังนั้นพวกเขาจึงพัฒนา 2 โมดูล ได้แก่ การจำแนกเสียงแยกคำและการจำแนกเสียงแบบต่อเนื่อง จากนั้นพวกเขาประเมินผลใน multi-speaker mode และ speaker-independent mode ของแต่ละโมดูล ในสภาพแวดล้อมที่ปราศจากเสียงรบกวนและมีเสียงรบกวน ซึ่งผลการประเมินพบว่า การจำแนกเสียงใน multi-speaker mode มีอัตราที่ดีกว่า speaker-independent mode แต่นอกจากนั้นความถูกต้องของการจำแนกเสียงยังขึ้นอยู่กับปัจจัยอื่นๆอีก เช่น ระดับความแตกต่างของเสียงโดยรอบในช่วงการบันทึกและคุณภาพของไมโครโฟนที่ใช้ในการรับเสียงพูด เป็นต้น ซึ่งปัจจัยเหล่านี้ถ้าสามารถควบคุมได้จะทำให้ความถูกต้องของการจำแนกเสียงดีขึ้น

2.6.2) Text -to-Speech

NtsakoBaloyi [12] ได้ศึกษาระบบแปลงข้อความที่เป็นเสียงสำหรับภาษา Xitsonga ซึ่งเป็นหนึ่งในภาษาพื้นเมืองของทวีปแอฟริกา เนื่องจากการสำรวจประชากรพื้นเมืองในแอฟริกาใต้พบว่ามีเพียง 4.5 ล้านคนที่ศึกษาอยู่ในเกรด 12 ขึ้นไป และมีประชากรถึง 18 ล้านคนที่ต่ำกว่าเกรด 12 ดังนั้นกระบวนการพัฒนาระบบแปลงข้อความที่เป็นเสียงในขั้นตอนแรกจะนำไปสู่การสร้างข้อมูลและความสามารถในการเข้าถึงเทคโนโลยี รวมทั้งความง่ายในการใช้งานสำหรับบุคคลที่มีความสามารถในการอ่านออกเขียนที่แตกต่างกัน งานวิจัยนี้ให้ความสำคัญกับการสังเคราะห์เสียงให้มีความเป็นธรรมชาติ มีความถูกต้อง และความยืดหยุ่นดังนั้นในงานวิจัยนี้ได้เลือกใช้วิธี Hidden Markov models ที่มีความสามารถในการให้ผลลัพธ์ที่มีความเข้าใจและมีเสียงที่สังเคราะห์ให้มีความเป็นธรรมชาติ รวมทั้งง่ายต่อการจับค่าคุณสมบัติทางสถิติของเสียง ซึ่งผลลัพธ์ที่ได้จากงานวิจัยนี้ จะวิเคราะห์ใน 3 ส่วนด้วยกัน ได้แก่ ความเข้าใจ, ความเป็นธรรมชาติ และความพอใจ จากการทดลองพบว่าผู้ฟังสามารถเข้าใจได้โดยไม่ต้องใช้ความพยายามในการฟังให้เข้าใจ ความเป็นธรรมชาติของเสียงอยู่ในระดับปานกลาง และผู้ฟังมีความพึงพอใจในเสียงที่ได้รับค่อนข้างดี และในภาพรวมของระบบอยู่ในเกณฑ์ที่ดีจากการศึกษางานวิจัยนี้ ทำให้ทราบว่าวิธีการสังเคราะห์เสียงด้วย HMM นั้นจะให้ความเป็นธรรมชาติและความถูกต้องของเสียงมาก ซึ่งถือเป็นจุดสำคัญของการสังเคราะห์เสียงในด้านของการแปลภาษา

Dilshan De Silvac และทีมงาน [13] ได้ศึกษาเป็นงานวิจัยเกี่ยวกับเครื่องมือแปลภาษาจากภาษาสิงหลภาษาเป็นภาษาอังกฤษ ภาษาสิงหลเป็นภาษาพื้นเมืองของประเทศศรีลังกา คนส่วนใหญ่ในประเทศศรีลังกาจะใช้ภาษานี้ในการสื่อสาร แต่ในปัจจุบันภาษาอังกฤษเป็นภาษาที่มีบทบาทสำคัญในประเทศศรีลังกา โดยเฉพาะในกลุ่มคนทำงาน ที่มีทักษะ แต่ขาดทักษะด้านภาษาอังกฤษ จึงเกิดอุปสรรคในการทำงาน รวมทั้งยังไม่มีระบบแปลภาษาสิงหลเป็นภาษาอังกฤษ ผู้วิจัยจึงได้

พัฒนาระบบนี้ขึ้น ในการพัฒนาจะให้ความสำคัญกับความถูกต้องตามหลักไวยากรณ์, ความสอดคล้อง, ความเร็ว, ความถูกต้องและแม่นยำ จึงเลือกใช้เทคนิค Rule-based วิธีนี้จะได้คุณภาพสูงในการแปลภาษารวมถึงความถูกต้องแม่นยำ จากผลการทดสอบพบว่าระบบให้ผลลัพธ์ที่ดี สามารถช่วยลดอุปสรรคการใช้ภาษาอังกฤษ และช่วยให้เกิดการเรียนรู้ภาษาอังกฤษได้ง่ายขึ้นและส่วนของการสังเคราะห์เสียง

Ahmad A. M. Abushariah และทีมงาน [14] ได้ศึกษางานวิจัยเกี่ยวกับการออกแบบและการดำเนินการของระบบรู้จำเสียงตัวเลขภาษาอังกฤษโดยใช้ Mat lab งานวิจัยนี้ใช้หลักของ Hidden Markov model ซึ่งเป็นวิธีที่ให้ผลลัพธ์ที่มีความน่าเชื่อถือสูง ระบบรู้จำเสียงใช้เทคนิค Mel Frequency Cepstral Coefficients (MFCC) เพื่อแปลงคลื่นเสียงไปเป็นกลุ่มของเวกเตอร์ งานวิจัยนี้จะให้ความสำคัญกับตัวเลขภาษาอังกฤษซึ่งจะขึ้นอยู่กับโครงสร้างการแยกคำ มีการพัฒนาทั้งหมด 2 โมดูล ได้แก่ การรู้จำเสียงพูดแบบคำโดดและการรู้จำเสียงพูดต่อเนื่อง โดยจะนำทั้ง 2 โมดูลมาทดลองในสภาพแวดล้อมที่มีเสียงรบกวนและไม่มีเสียงรบกวน ในโหมด multi-speaker และ speaker-independent ผลลัพธ์ที่ได้อัตราการรู้จำเสียงของ multi-speaker มีความแม่นยำสูงกว่า speaker-independent ในทั้งสองสภาพแวดล้อม ซึ่งความแม่นยำของกระบวนการรับรู้เสียงขึ้นอยู่กับหลายๆปัจจัย เช่น ความแตกต่างของสภาพแวดล้อม, คุณภาพของตัวรับเสียง เป็นต้น จากการศึกษา งานวิจัยนี้ทราบถึงเทคนิคที่ใช้ในการรู้จำเสียง และปัจจัยที่ส่งผลกระทบต่อกระบวนการรับรู้เสียง

2.6.3) Speech-to-Speech

Kei Hashimoto และทีมงาน [15] ได้ศึกษาวิเคราะห์ผลลัพธ์ของเครื่องแปลภาษาและระบบการแปลภาษาด้วยเสียงพูด ระบบแปลภาษาด้วยเสียงพูดประกอบด้วย 3 องค์ประกอบหลักคือ การรู้จำเสียง (speech recognition), เครื่องมือแปลภาษา (machine translation) และการสังเคราะห์เสียง (speech synthesis) ซึ่งในงานวิจัยนี้พวกเขามุ่งเน้นที่เครื่องมือแปลภาษาและการสังเคราะห์เสียง รวมไปถึงการประเมินผลเกี่ยวกับการวิเคราะห์ผลลัพธ์ในแต่ละองค์ประกอบที่ได้กล่าวข้างต้น พวกเขาประเมินผลในระบบแปลภาษาด้วยเสียงพูดจากภาษาฟินแลนด์เป็นภาษาอังกฤษ ในการประเมินผลพวกเขาใช้เครื่องแปลภาษาแบบ HiFST และใช้ Hidden Markov Model สำหรับการสังเคราะห์เสียง โดยพวกเขาจะเลือกผลลัพธ์การแปลประโยคที่ดีที่สุด 5 อันดับแรกจากเครื่องมือแปลภาษาและนำมาสังเคราะห์เสียง โดยใช้วิธี N-gram สำหรับประเมินความเป็นธรรมชาติและความเข้าใจของเสียงที่สังเคราะห์ ผลลัพธ์ที่ได้คือเสียงสังเคราะห์มีความเป็นธรรมชาติและสามารถเข้าใจความหมายของประโยคที่สังเคราะห์เป็นเสียงพูดได้ ทำให้พวกเขาพบว่าความเป็นธรรมชาติของเสียงและความเข้าใจความหมายของประโยคจากการฟังเสียงจะดีหรือไม่ดีขึ้นอยู่กับผลลัพธ์ที่ได้จากการแปลภาษาโดยตรง เนื่องจากผลลัพธ์ขององค์ประกอบหนึ่งจะเป็นอินพุตขององค์ประกอบถัดไป ดังนั้นความผิดพลาดขององค์ประกอบก่อนหน้าจะส่งผลกระทบต่อประสิทธิภาพการทำงานขององค์ประกอบถัดไป

Seung Yun และทีมงาน [16] ได้ศึกษาถึงประสิทธิภาพของ Genietalk ซึ่งเป็นระบบแปลภาษาด้วยเสียงพูดจากภาษาเกาหลีเป็นภาษาอังกฤษ, ภาษาเกาหลีเป็นภาษาญี่ปุ่น และภาษาเกาหลีเป็นภาษาจีน ซึ่งระบบแปลภาษาด้วยเสียงพูดนี้ใช้งานผ่านอินเทอร์เน็ต โดยพวกเขาจะสำรวจความต้องการจากกลุ่มตัวอย่างเพื่อนำมาออกแบบแอปพลิเคชัน ทำให้ในการศึกษาครั้งนี้พวกเขาจะสร้างฐานข้อมูลของภาษาและฐานข้อมูลเสียงพูดของสถานการณ์ที่จะใช้ระบบการแปลภาษาด้วยเสียงพูด นอกจากนี้พวกเขาก็พิจารณาการออกแบบ UI ของระบบแปลภาษาด้วยเสียงพูด ความยากง่ายในการใช้งาน รวมถึงการลดลงของข้อผิดพลาดในกระบวนการแปลภาษา ผลที่ได้จากการศึกษาทำให้พวกเขาพบว่าผู้ใช้พอใจแอปพลิเคชันและชี้ให้เห็นความก้าวหน้าเพื่อเพิ่มความพึงพอใจของผู้ใช้ผ่านคุณสมบัติเพิ่มเติม เช่น การค้นหาสำหรับผลลัพธ์การแปลอื่นๆ

2.7 นิยามศัพท์

1. ระบบแปลภาษาด้วยเสียงพูดแบบหลายภาษา คือ ระบบสำหรับแปลประโยคหรือข้อความจากภาษาแรกเป็นภาษาอื่นๆมากกว่า 1 ภาษา
2. ส่วนต่อประสานโปรแกรมประยุกต์ คือ ชุดคำสั่งหรือข้อกำหนด ที่ถูกออกแบบมาเพื่อทำหน้าที่เป็นตัวกลางในการเชื่อมต่อการทำงานระหว่างแอปพลิเคชันหรือเชื่อมการทำงานเข้ากับระบบปฏิบัติการ
3. Speech-to-Speech คือ การแปลงเสียงพูดภาษาแรก เป็นเสียงพูดของภาษาอื่นๆ
4. Speech-to-Text คือ การแปลงเสียงพูดให้เป็นตัวอักษร
5. Text-to-Speech คือ การแปลงข้อความให้เป็นเสียงพูด
6. Language Translation คือ การถ่ายทอดข้อความจากภาษาหนึ่งไปยังอีกภาษาหนึ่งโดยรักษารูปแบบและความหมายในข้อความเดิมไว้อย่างครบถ้วน

บทที่ 3

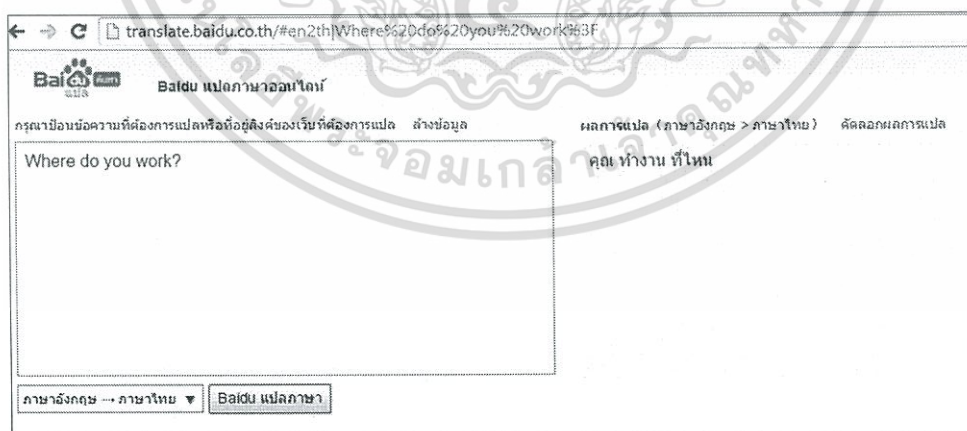
การวิเคราะห์และออกแบบระบบ

ในบทนี้จะกล่าวถึงการวิเคราะห์และออกแบบระบบ คือ การหาความต้องการของระบบว่าต้องการอะไร และการนำความต้องการของระบบมาเขียนเป็นแบบแผนในการสร้างระบบ รวมถึงวิธีการที่ใช้ในการสร้างระบบขึ้นมาใหม่ นอกจากนั้นแล้วการวิเคราะห์ระบบที่มีอยู่ยังช่วยในการแก้ไขระบบเดิมให้มีประสิทธิภาพดียิ่งขึ้น รวมถึงการออกแบบแอปพลิเคชันและฟังก์ชันการใช้งานของแอปพลิเคชันให้มีความสะดวกและง่ายต่อการใช้งาน นอกจากนั้นแล้วในบทนี้ยังมีการเปรียบเทียบและการเลือกใช้ API ที่ได้นำมาใช้ในส่วนต่างๆ ของการพัฒนาแอปพลิเคชันอีกด้วย

3.1 การศึกษางานระบบเดิม

การแปลภาษาในแบบเดิมจะทำได้จากการเปิดหนังสือดิกชันนารี, เว็บไซต์ออนไลน์ ดิกชันนารี, Talking Dict หรือทางอื่นยกตัวอย่างเช่น แอปพลิเคชันแปลภาษาในสมาร์ทโฟน ซึ่งในการรับข้อมูลที่ต้องการแปลแบบเดิมคือยังเป็นการรับในรูปแบบการพิมพ์ และจะแปลคำศัพท์หรือประโยคที่ผู้ใช้งานที่ต้องการออกมาจากนั้นถึงจะกดฟังคำศัพท์หรือประโยคนั้นๆ ได้ และผู้จัดทำได้ยกเอาตัวอย่าง เว็บไซต์แปลภาษาและแอปพลิเคชันแปลภาษาดังต่อไปนี้

1) เว็บไซต์แปลภาษาและเว็บไซต์ดิกชันนารี ตัวอย่างเว็บแปลภาษาและเว็บดิกชันนารีที่เปิดให้บริการฟรีในปัจจุบัน อาทิเช่น



รูปที่ 3.1 แสดงภาพหน้าเว็บไซต์ translate.baidu.co.th

จากรูปที่ 3.1 เว็บแปลภาษาของ translate.baidu.co.th เป็นเว็บไซต์สัญชาติจีนที่ได้รับความนิยมอย่างกว้างขวางในปัจจุบันให้บริการเพียง ภาษาจีน, อังกฤษ, ญี่ปุ่น, ไทย เท่านั้นให้บริการใน

รูปแบบออนไลน์ผ่านเว็บไซต์ นอกจากนี้ยังมีบริการของ Baidu API in Android อีกด้วย
เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อใช้ในการศึกษาเท่านั้น ไม่อนุญาตให้เผยแพร่ไปโดยไม่ขออนุญาต
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Cambridge Dictionaries Online
The most popular online dictionary and thesaurus for learners of English

dictionary Search! British English ✓ American English Business English More...

English definition of "dictionary"
dictionary
noun [C] UK ɪˈdɪkʃənəri / US ɪˈdɪ-/er.ri

1 a book that contains a list of words in alphabetical order and that explains their meanings, or gives a word for them in another language; a similar product for use on a computer:
a French-English/English-French dictionary
a bilingual/monolingual dictionary
To check how a word is spelled, look it up in a dictionary.

2 a book that gives information about a particular subject, in which the entries (= words or phrases) are given in alphabetical order:
a biographical/science dictionary
a dictionary of quotations

(Definition of dictionary from the Cambridge Advanced Learners Dictionary & Thesaurus © Cambridge University Press)
What is the pronunciation of dictionary?

Browse related topics
You are looking at an entry to do with **Books: reference books**, but you might be interested in these topics from the **Media & publishing** category:

Add Cambridge dictionaries to your browser to your website

Definitions of "dictionary" in other dictionaries

American English | Learner's

Translations of "dictionary"

in Spanish | in Arabic | in Catalan | in Chinese (Simplified) | in Chinese (Traditional) | in Italian | in Japanese | in Korean | in Polish | in Portuguese | in Russian | in Turkish | in French | in German

SMART Thesaurus: Books: reference books

"dictionary": synonyms and related words:

almanac annual atlas bible coursebook directory gazetteer **grammar** guide **guidebook** handbook **manual** peerage

รูปที่ 3.2 แสดงภาพหน้าเว็บไซต์ Dictionary.cambridge.org

จากรูป 3.2 เว็บไซต์ดิกชันนารีของ Cambridge จะให้บริการออนไลน์โดยจะเปรียบเสมือนดิกชันนารีไม่สามารถแปลทั้งประโยคได้ ข้อดีคือ มีหมวดหมู่ภาษาให้เลือกหลากหลายและมี Dictionary API Cambridge ให้บริการ นอกจากนี้ยังให้บริการผ่านแอปพลิเคชันทั้งระบบปฏิบัติการแอนดรอยด์และระบบปฏิบัติการ iOS โดยชื่อว่า Cambridge ADVANCED Learner's

Oxford Dictionaries
Language matters

ABOUT | HELP | SUBSCRIBE | LOGIN

ENGLISH

dictionary Dictionary Synonyms Grammar Explore Get premium

dictionary
Line breaks: dictionary
Pronunciation: /ˈdɪkʃən(ə)rɪ ˈdɪ /

Definition of dictionary in English:
NOUN (plural dictionaries)

1 A **BOOK** or electronic resource that lists the words of a language (typically in alphabetical order) and gives their meaning, or gives the equivalent words in a different language, often also providing information about pronunciation, origin, and usage:
'I'll look up 'love' in the dictionary'
'the website gives access to an online dictionary'
[AS MODIFIER: *'the dictionary definition of 'smile''*

Get more from Oxford Dictionaries
Subscribe to remove adverts and access premium resources
[Find out more >](#)

Top searches

1 experience 6 knowledge
2 practice 7 provide

รูปที่ 3.3 แสดงภาพหน้าเว็บไซต์ Oxforddictionaries.com

จากรูปที่ 3.3 เว็บไซต์ดิกชันนารีของ Oxford จะให้บริการออนไลน์โดยจะเปรียบเสมือนดิกชันนารีไม่สามารถเลือกแปลทั้งประโยคได้ แปลได้เป็นคำๆ เท่านั้น ข้อดีคือ มีภาษาและคำศัพท์เอกสารเป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ที่หลากหลาย นอกจากนี้ยังสามารถเลือกคีย์บอร์ดให้เป็นภาษาอารบิกและละตินเพื่อค้นหาคำศัพท์ได้อีกด้วย รวมถึงเว็บไซต์ Oxford ยังให้บริการ Oxford Dictionaries API เพื่อนักพัฒนาสามารถนำไปใช้ได้อีกด้วย

2) แอปพลิเคชันแปลภาษา ตัวอย่างของแอปพลิเคชันแปลภาษาที่เปิดให้ดาวน์โหลดเพื่อให้บริการในอุปกรณ์พกพาเคลื่อนที่ อาทิเช่น



รูปที่ 3.4 แสดงภาพแอปพลิเคชัน THAI DICT

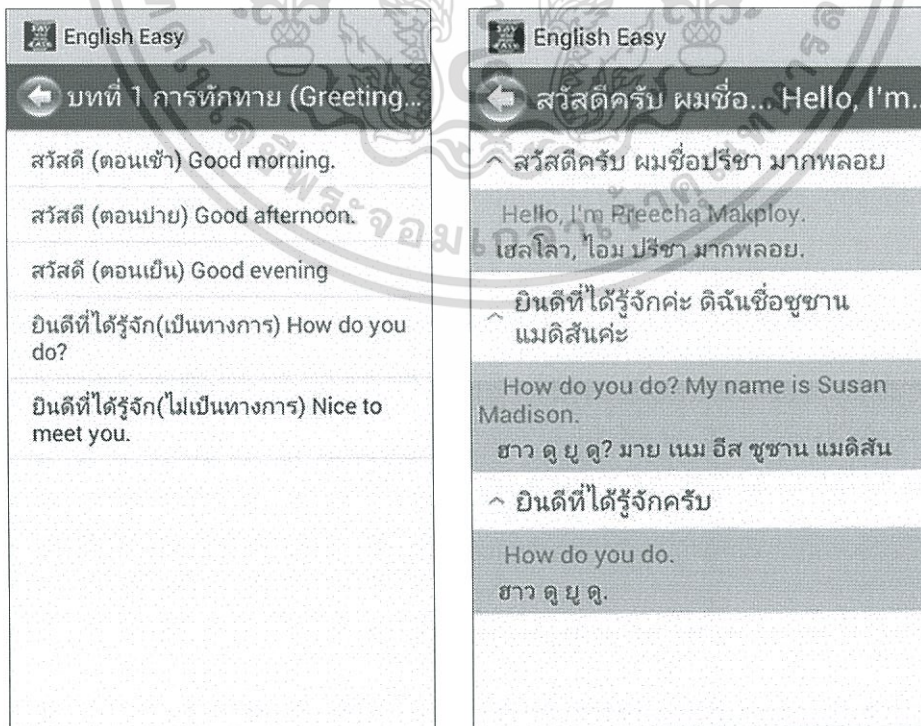
จากรูปที่ 3.4 แอปพลิเคชันแปลภาษา THAI DICT จะเป็นดิกชันนารีในระบบปฏิบัติการแอนดรอยด์ที่ผู้ใช้งานสามารถโหลดมาใช้ทำงานในเครื่องได้เลย ข้อดีคือ สามารถใช้งานได้โดยไม่ต้องผ่านอินเทอร์เน็ต บอกชนิดของคำ ความหมาย คำเหมือนและคำใกล้เคียงสามารถกดเพื่อฟังการออกเสียงของคำนั้นๆได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.5 แสดงภาพแอปพลิเคชัน แปลภาษา

จากรูปที่ 3.5 แอปพลิเคชันแปลภาษา รองรับการแปลทั้งประโยคหรือเป็นคำได้ เลือกแปลได้แค่ 2 ภาษาเท่านั้นคือภาษาไทยและภาษาอังกฤษ แอปพลิเคชันรับข้อมูลในรูปแบบการพิมพ์ ข้อดีคือ แอปพลิเคชันมีเสียงพูดสามารถใช้ฝึกในการออกเสียงได้ และยังแปลเปรียบเทียบกันระหว่าง Google translate และ Bing translate สามารถกดคัดลอกคำแปลโดยที่สัมผัสหน้าจอค้างไว้



รูปที่ 3.6 แสดงภาพแอปพลิเคชัน Talking english easy

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อใช้ประโยชน์ในการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปภาพ 3.6 เป็นแอปพลิเคชัน Talking english easy ที่นำบทสนทนาพื้นฐานที่ใช้สื่อสารในชีวิตประจำวันให้ผู้ใช้ได้ฝึกพูดและฟังเสียงอ่าน โดยมีประโยคไทย และภาษาอังกฤษ พร้อมคำอ่านโดยแบ่งตามหมวดหมู่ 22 บท มีมากกว่า 300 ประโยค จุดอ่อนของแอปพลิเคชันคือ แม้จะมีคำอ่านมาให้แต่ไม่มีเสียงพูดให้เป็นตัวอย่างและถ้าไม่มีประโยคที่ต้องการก็จะไม่สามารถนำมาใช้งานได้

3.2 ปัญหาที่พบในระบบปัจจุบัน

ในเว็บดิกชันนารี, แปลภาษาและแอปพลิเคชันดิกชันนารี, แอปพลิเคชันแปลภาษา ส่วนมากยังใช้งานได้ไม่สะดวกมากนัก การแปลภาษาหรือแปลคำศัพท์ยังเป็นการรับข้อมูลในรูปแบบการพิมพ์ ซึ่งอาจจะไม่สะดวกสำหรับผู้ที่มีปัญหาในเรื่องสายตาหรือการแปลในเว็บซึ่งในขณะนั้นผู้ใช้งานอาจจะเดินทางไปนอกสถานที่ซึ่งไม่ได้พกคอมพิวเตอร์ไปด้วยหรือต่ออินเทอร์เน็ตไม่ได้ ดังตัวอย่างเช่น เว็บไซด์แปลภาษาและแอปพลิเคชันที่ได้ยกตัวอย่างไปข้างต้น เว็บไซด์แปลภาษาและเว็บไซด์ดิกชันนารีดังรูป 3.1-3.3 จำเป็นจะต้องมีคอมพิวเตอร์หรือแท็บเล็ตที่สามารถเชื่อมต่ออินเทอร์เน็ตได้ในการใช้งานส่วนแอปพลิเคชัน THAI DICT ดังรูป 3.4 แอปพลิเคชันการแปลภาษาดังรูป 3.5 จะไม่สามารถรับเสียงพูดเข้าไปได้ผู้ใช้งานจะต้องป้อนข้อมูลผ่านการพิมพ์ และแอปพลิเคชัน Talking english easy ดังรูป 3.6 เป็นแอปพลิเคชันที่กำหนดควลีและประโยคมาไว้ให้ในระบบเลยโดยจะเขียนคำอ่านประกอบมาด้วยแต่ไม่สามารถให้ออกเสียงเป็นตัวอย่างได้ จุดอ่อนของระบบคือไม่สามารถออกเสียงเป็นตัวอย่างให้กับผู้ใช้งานได้และถ้าไม่มีประโยคที่ต้องการใช้งานก็จะไม่สามารถนำมาใช้งานได้

3.3 การวิเคราะห์ความต้องการของระบบ

ก่อนการสร้างระบบใหม่เราจำเป็นต้องทำการวิเคราะห์ความต้องการของระบบ เพื่อที่จะกำหนดความต้องการว่ามีอะไรบ้าง โดยความต้องการที่เรากำหนดขึ้นมาให้นำมาใช้ประโยชน์ได้ทั้งในมุมมองของผู้ใช้งานและของเจ้าของงาน

3.3.1) ความต้องการที่เป็นหน้าที่หลักของระบบ (Functional Requirement) คือ Requirement หรือสิ่งที่ระบบควรจะทำหน้าที่หลักของระบบที่จะต้องทำ

- ระบบสามารถแปลภาษาได้อย่างถูกต้อง
- ระบบสามารถใช้งานออนไลน์
- ระบบสามารถรับเสียงพูดของผู้ใช้งานได้
- ระบบสามารถแปลงเสียงพูดเป็นข้อความได้
- ระบบสามารถแปลงข้อความเป็นเสียงพูดได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

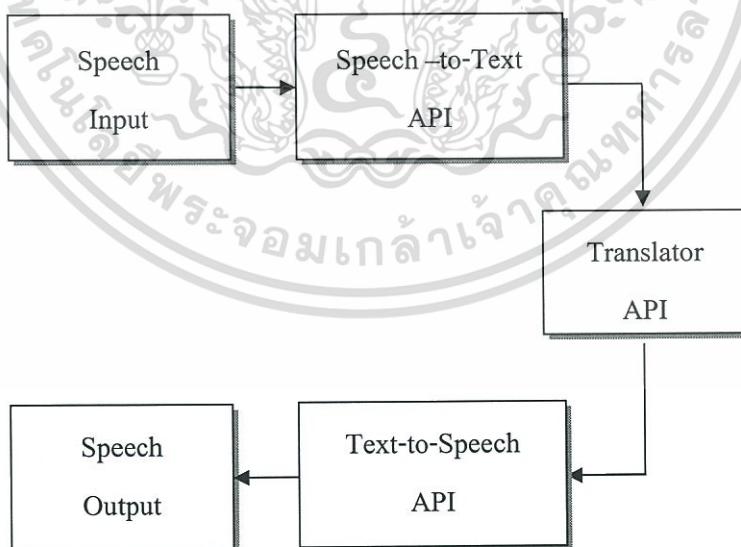
3.3.2) ความต้องการของระบบที่ไม่ใช่หน้าที่หลักของระบบ (Non-Functional Requirement) คือ Requirement อื่นๆ ที่ไม่ใช่หน้าที่หลักๆที่ต้องทำแต่เป็นคุณสมบัติอื่นๆ ที่เราอยากได้จากระบบ

- ผู้ใช้สามารถใช้งานได้อย่างสะดวกและมีความรู้สึกสมเหตุสมผลในการใช้งาน
- ระบบมีความเสถียรภาพในการใช้งาน
- ระบบทำให้ผู้ใช้รู้สึกว่ามีความง่ายต่อความเข้าใจในการใช้งาน
- ระบบมีรูปแบบสวยงามน่าสนใจในการใช้งาน
- ระบบมีรูปแบบเหมาะสมในการใช้งาน

3.4 การวิเคราะห์และออกแบบระบบ

ผู้จัดทำมีความต้องการที่จะให้ระบบแปลภาษาบนระบบปฏิบัติการแอนดรอยด์สามารถรับเสียงพูดและการแปลภาษาที่มีความถูกต้องในการใช้งานมีประโยชน์สำหรับบุคคลทั่วไปและสามารถใช้งานได้สะดวก เพราะในปัจจุบัน โทรศัพท์เคลื่อนที่แบบสมาร์ตโฟน หรืออุปกรณ์จำพวกแท็บเล็ต มีการใช้งานกันอย่างแพร่หลายผู้จัดทำจึงมีแนวคิดที่จะลดช่องว่างในการสื่อสารโดยการพัฒนาแอปพลิเคชันแปลภาษาขึ้นมา โดยที่ระบบจะสามารถใช้งานได้แบบออนไลน์

ดังนั้นผู้จัดทำจึงได้ออกแบบองค์ประกอบของแอปพลิเคชันแปลภาษาด้วยเสียงพูด โดยจะมีรายละเอียดดังรูปที่ 3.7



รูปที่ 3.7 แสดงภาพองค์ประกอบของแอปพลิเคชันแปลภาษาด้วยเสียงพูด

การทำงานของแอปพลิเคชันจะเป็นไปตามรูปที่ 3.7 ภาพรวมของแอปพลิเคชันแปลภาษาเอกด้วยเสียงพูดบนระบบปฏิบัติการแอนดรอยด์โดยเริ่มจากผู้ใช้ป้อนข้อมูลด้วยเสียงพูด แอปพลิเคชันจะทำการแปลภาษาจากคำพูดที่ป้อนเข้ามาเป็นข้อความ และทำการแปลข้อความนั้นเป็นภาษาที่ต้องการ และทำการนำข้อความที่แปลแล้วไปใช้

จะเปลี่ยนจากเสียงพูดเป็นข้อความจากนั้นในขั้นตอนต่อมาระบบจะแปลข้อความเป็นภาษาปลายทางที่ผู้ใช้งานเลือกและแปลงข้อความเป็นเสียงพูดให้ผู้ใช้งานในขั้นตอนสุดท้าย

3.4.1) ตารางวิเคราะห์คุณสมบัติของ API จะทำการวิเคราะห์ API ที่เลือกมาในแต่ละส่วน เพื่อพัฒนาแอปพลิเคชันแปลภาษาด้วยเสียงพูด โดยการพิจารณาความเหมาะสมกับการใช้งานและสภาพแวดล้อมการใช้งาน จะทำการวิเคราะห์ทั้งส่วน Speech-to-text, Translation และ Text-to-Speech

ตารางที่ 3.1 แสดงการวิเคราะห์คุณสมบัติ API

ส่วนที่นำมาวิเคราะห์	Bing	Google	Yandex
Speech-to-text			
- สามารถรองรับเสียงภาษาไทยได้	x	✓	-
- สามารถรองรับเสียงภาษาอังกฤษได้	✓	✓	-
- รองรับการพัฒนาบนระบบปฏิบัติการแอนดรอยด์	✓	✓	-
Translation			
- สามารถรองรับการแปลภาษาไทยได้	✓	✓	x
- สามารถรองรับการแปลภาษาอังกฤษได้	✓	✓	✓
- สามารถรองรับการแปลภาษาจีน, ญี่ปุ่น, เกาหลี	✓	✓	✓
- รองรับการพัฒนาบนระบบปฏิบัติการแอนดรอยด์	✓	✓	✓
Text-to-Speech			
- สามารถสังเคราะห์เสียงภาษาไทย	x	✓	-
- สามารถสังเคราะห์เสียงภาษาอังกฤษ	✓	✓	-
- สามารถรองรับการสังเคราะห์เสียงภาษาจีน, ญี่ปุ่น, เกาหลี	✓	✓	-
- รองรับการพัฒนาบนระบบปฏิบัติการแอนดรอยด์	✓	✓	-

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ส่วนที่ 1 Speech-to-Text

การเลือกใช้ API จากการวิเคราะห์ Speech Recognition API ที่ได้ศึกษามาทั้ง Google Speech Recognition และ Bing Speech Recognition Control พบว่า Bing สามารถใช้ได้แค่ 7 ภาษาเท่านั้น German (Germany), English (Unites States), English (Great Britain), Spanish (Spain), French (France), Italian (Italy), Chinese (China) ซึ่งไม่มีในส่วนภาษาไทย ในงานวิจัยนี้จึงเลือกใช้บริการของ Google Speech Recognition เนื่องด้วยรองรับการรับเสียงพูดของภาษาไทย ดังตารางที่ 3.1

ส่วนที่ 2 Translators

จากการศึกษาและค้นคว้าคุณสมบัติของ Google Translate API, Bing Translator Control API, Yandex Translate API ดังตารางที่ 3.1 มี API สามารถรองรับภาษาไทยได้ทั้ง Google Translate API, Bing Translator API งานวิจัยชิ้นนี้จึงได้เลือกใช้ API ทั้ง 2 ในขั้นตอนแปลภาษา โดยการนำมาทดสอบประสิทธิภาพ ด้วยการพัฒนาแอปพลิเคชันเพื่อเปรียบเทียบทั้ง 2 API ว่าแต่ละ API มีประสิทธิภาพมากน้อยเพียงใด

ส่วนที่ 3 Text-to-speech

การเลือกใช้ API จากการวิเคราะห์ที่ได้ศึกษาค้นคว้ามาทั้ง Google Text to Speech API และ Bing Text to Speech API พบว่า Bing รองรับเพียงบางภาษา เช่น English (US), French (FR), German (DE), Japanese (JA), Korean (KR), Chinese (TW) เป็นต้น ซึ่งไม่รองรับภาษาไทย ส่วนของ Google นั้นรองรับได้หลากหลายภาษารวมถึงภาษาไทยที่จำเป็นต้องมีในระบบดังตารางที่ 3.1 ดังนั้นในงานวิจัยชิ้นนี้จึงเลือก Google Text to Speech API มาใช้ในส่วนนี้

3.4.2) แผนภาพแสดง Activity Diagram

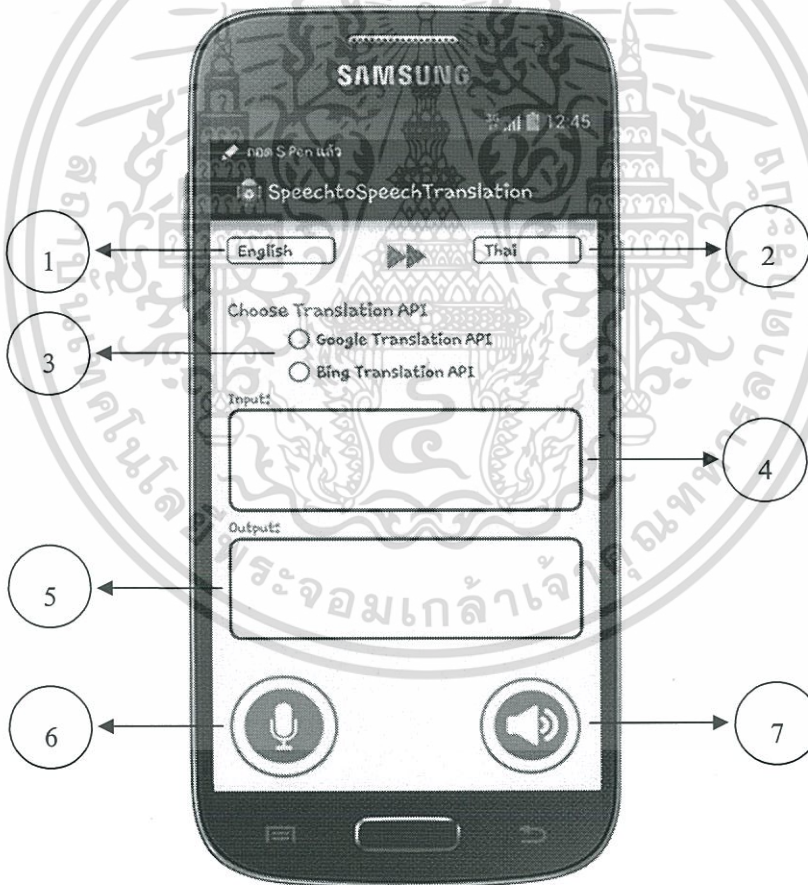
- 1) แผนภาพกิจกรรมการทำงานของระบบแปลภาษาด้วยเสียงพูดทั้งระบบ (ตารางที่ ก.1)
- 2) แสดงแผนภาพกิจกรรมการเลือกภาษา (ตารางที่ ก.2)
- 3) แสดงแผนภาพกิจกรรมการรับเสียงพูด (ตารางที่ ก.3)
- 4) แสดงแผนภาพกิจกรรมการแปลภาษา (ตารางที่ ก.4)
- 5) แสดงแผนภาพกิจกรรมการฟังข้อความเสียง (ตารางที่ ก.5)
- 6) แสดงแผนภาพกิจกรรมการเลือก API (ตารางที่ ก.6)

บทที่ 4

ผลการดำเนินงาน

ในบทนี้เป็นการนำเสนอผลการดำเนินการของการพัฒนาแอปพลิเคชันแปลภาษาด้วยเสียงพูดบนระบบปฏิบัติการแอนดรอยด์ตามที่ได้วิเคราะห์และออกแบบไว้ โดยเนื้อหาจะกล่าวถึงการทำงานต่างๆ ของแอปพลิเคชันและการทดสอบการทำงานในขั้นต่อต่างๆ ของระบบไม่ว่าจะเป็นขั้นตอนในการรับเสียง การแปลความหมาย การสังเคราะห์เสียงและการประเมินการใช้งานระบบ

4.1 หน้าจอแสดงผลแอปพลิเคชัน



รูปที่ 4.1 แสดงภาพองค์ประกอบโดยรวมของแอปพลิเคชัน

จากรูปภาพที่ 4.1 จะแสดงให้เห็นถึงหน้าแอปพลิเคชันทั้งหมดโดยแบ่งออกเป็น 7

หมายเลข

เอกสารนี้เป็น 1 คือ เลือกภาษาในการรับเสียง เพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- 2 คือ เลือกภาษาที่ต้องการแปล
- 3 คือ เลือก API ที่ผู้ใช้ต้องการใช้ในการแปลภาษา
- 4 คือ กรอบแสดงข้อความที่รับเข้ามาโดยการพูดของผู้ใช้งาน ในที่นี้ถ้ายังไม่มีการรับเสียงเข้ามาในส่วนนี้จะป็นหน้าจอว่าง
- 5 คือ กรอบแสดงข้อความของภาษาที่แปลเสร็จแล้ว
- 6 คือ ปุ่มที่ใช้ในการรับเสียงของผู้พูด โดยการใช้นั้นผู้ใช้งานต้องกดสัมผัสหน้าจอเพื่อรับเสียงและปล่อยเมื่อสิ้นสุดประโยค
- 7 คือ ปุ่มที่ใช้กดฟังเสียงเพื่อใช้เป็นตัวอย่างในการออกเสียงภาษาที่แปลเสร็จสิ้นแล้ว โดยการกดสัมผัสที่ปุ่มแล้วปล่อย



รูปที่ 4.2 แสดงภาพองค์ประกอบ โดยรวมของแอปพลิเคชัน

จากรูปที่ 4.2 เมื่อแอปพลิเคชัน ไม่สามารถรับเสียงพูดได้ในกรณีที่พูดไม่ชัดหรือกดปุ่มผิดพลาดในส่วนของหน้าจอจะขึ้นข้อความว่า No speech input และในกรณีที่สมาร์ทโฟน ไม่มีการเชื่อมต่ออินเทอร์เน็ตเมื่อกดปุ่มเพื่อรับเสียงหน้าจอจะขึ้นข้อความว่า Network error เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์การใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2 การทดสอบประสิทธิภาพของ API

การทดสอบประสิทธิภาพของตัว API ในส่วนของ Speech-to-Text และ Translation ก่อนที่จะนำ API ไปพัฒนาแอปพลิเคชันแปลภาษาด้วยเสียงพูดแบบหลายภาษาบนระบบปฏิบัติการแอนดรอยด์ เพื่อทดสอบประสิทธิภาพโดยรวมของระบบทั้งหมด

4.2.1) การทดสอบการรับเสียงพูดภาษาอังกฤษด้วย Google speech recognition ทดสอบด้วยการรับเสียงจากคลิปเสียงข้อความของเว็บไซต์ www.newsinlevels.com ซึ่งผลการทดสอบดังแสดงในตาราง

ตารางที่ 4.1 แสดงการทดสอบการรับเสียงพูดภาษาอังกฤษด้วย Google speech recognition

ประโยค	ผลลัพธ์
1. It may not have the same worldwide fame as the Olympic Games, but the Pan American Games certainly knows how to do an opening ceremony.	it may not have the same worldwide fame as the Olympic Games but the Pan American Games certainly knows how to do an opening ceremony
2. She came to America speaking only her native language	She came to America speaking only her native language
3. The parachutes help them go down slowly	parachutes help them go down slowly
4. People held the world paper airplane championships in Austria.	People help the world paper airplane championships in Austria
5. Christmas is celebrated all over the world, as a religious holiday or as a time of celebration by Christians and non-Christians alike.	Christmas is celebrated all over the world, as a religious holiday or as a time of celebration by Christians and non-Christians alike.
6. A black bear and her cubs chased a group of tourists on a bridge in the park.	black bear and her cubs chased a group of tourists on a bridge in the car
7. She's been programmed to speak Japanese but can also be fine-tuned to speak in Chinese and Korean, too.	She's been programmed to speak Japanese but can also be fine tuned to speak in Chinese and Korean too
8. The flamingo, which is entirely black, was feeding with the other pink flamingos on the banks of the lake on Wednesday afternoon.	The flamingos which is entirely black eating with the other pink flamingos on the banks of the lake on Wednesday afternoon.

ตารางที่ 4.1 (ต่อ) แสดงการทดสอบการรับเสียงพูดด้วย Google speech recognition

ประโยค	ผลลัพธ์
9. The company will begin testing the vehicles on public roads in the coming months, but it turns out they will have steering wheels and brakes, something they had not envisioned a year ago.	the company will begin testing the vehicles on public roads in the coming months, but it turns out they will have steering wheels and brakes something they had not envisioned a year ago.
10. The men working in the northwest of the country discovered a huge manta ray which local media have said weighs just over a ton.	the men working in the northwest of the country discovered a huge manta ray which local media have said ways just over a ton

จากตารางที่ 4.1 การทดสอบรับเสียงพูดภาษาอังกฤษ พบว่าจากการรับเสียง 10 ประโยค สามารถรับเสียงถูกต้องได้ 6 ประโยค และผิด 4 ประโยค โดยพิจารณาความถูกต้องจากการพิจารณาผลลัพธ์ที่ได้ตรงกับประโยคที่ใช้ทดสอบ ซึ่งประโยคที่รับเสียงผิดนั้นเกิดจากคำที่ออกเสียงคล้ายกัน เช่น feeding กับ eating, park กับ car และ held กับ help เป็นต้น และผิดพลาดจากการไม่สามารถรับเสียงคำนำหน้าคำนามได้ เช่น a, an และ the เป็นต้น

4.2.2) การทดสอบการรับเสียงพูดภาษาไทยด้วย Google speech recognition ทดสอบด้วยการรับเสียงจากการคลิปลีเสียงประโยคตัวอย่าง ซึ่งผลการทดสอบดังแสดงในตาราง ตารางที่ 4.2 แสดงการทดสอบการรับเสียงพูดภาษาไทยด้วย Google speech recognition

ประโยค	ผลลัพธ์
1. ภาษาเป็นที่มาของวัฒนธรรมและวัฒนธรรมก็เป็นที่มาของภาษาเช่นกัน	ภาษาเป็นที่มาของวัฒนธรรมและวัฒนธรรมก็เป็นที่มาของภาษาเช่นกัน
2. วัฒนธรรมคือสิ่งที่มนุษย์เปลี่ยนแปลงปรับปรุง หรือผลิตขึ้นเพื่อความเจริญอกงามในวิถีชีวิต	วัฒนธรรมคือสิ่งที่มนุษย์เปลี่ยนแปลงปรับปรุงหรือผลิตขึ้นเพื่อความเจริญอกงามในวิถีชีวิต
3. การสื่อสารข้อมูลเป็นการถ่ายทอดความรู้ ความคิดความรู้สึจากคนหนึ่งไปสู่อีกคนหนึ่ง โดยอาศัยสื่อหรือเครื่องมือต่างๆเป็นช่องทางในการสื่อสาร	การสื่อสารข้อมูลเป็นการถ่ายทอดความรู้ ความคิดความรู้สึจากคนหนึ่งไปสู่อีกคนหนึ่ง โดยอาศัยสื่อหรือเครื่องมือต่างๆเป็นช่องทางในการสื่อสาร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.2 (ต่อ) แสดงการทดสอบการรับเสียงพูดภาษาไทยด้วย Google speech recognition

ประโยค	ผลลัพธ์
4. สิ่งแวดล้อม หมายถึงสิ่งที่เกิดขึ้นเองตามธรรมชาติและสิ่งที่มีมนุษย์สร้างขึ้น	สิ่งแวดล้อมหมายถึงสิ่งที่เกิดขึ้นเองตามธรรมชาติและสิ่งที่มีมนุษย์สร้างขึ้น
5. ปัญหาสิ่งแวดล้อม สามารถเกิดขึ้นทุกพื้นที่ โดยเฉพาะอย่างยิ่งพื้นที่ ที่ประชากรอาศัยอยู่นานแน่น	ปัญหาสิ่งแวดล้อมสามารถเกิดขึ้นทุกพื้นที่ โดยเฉพาะอย่างยิ่งพื้นที่ ที่ประชากรอาศัยอยู่นานแน่น
6. AEC เกี่ยวข้องกับการรวมตัวทางเศรษฐกิจของสมาชิกอาเซียนซึ่งครอบคลุมการเปิดเสรีและอำนวยความสะดวกในการเคลื่อนย้ายทุน, การเคลื่อนย้ายแรงงาน, การปรับพิธีการศุลกากร, การกำหนดมาตรฐานสินค้าและนโยบายเศรษฐกิจให้สอดคล้องกันระหว่างประเทศสมาชิกอาเซียน	AEC เกี่ยวข้องกับการรวมตัวทางเศรษฐกิจของสมาชิกอาเซียนซึ่งครอบคลุมการเปิดเสรีและอำนวยความสะดวกในการเคลื่อนย้ายทุนการเคลื่อนย้ายแรงงานการปรับพิธีการศุลกากรการกำหนดมาตรฐานสินค้าและนโยบายเศรษฐกิจให้สอดคล้องกันระหว่างประเทศสมาชิกอาเซียน
7. ปลาฉลามเป็นสัตว์น้ำทะเลที่เป็นที่รู้จักกันอย่างแพร่หลายเพราะรูปร่างที่แปลกตาและ ความไม่มีพิษภัยของมัน	ปลาฉลามเป็นสัตว์น้ำทะเลที่เป็นที่รู้จักกันอย่างแพร่หลายเพราะรูปร่างที่แตกต่างและความไม่มีพิษภัยของมัน
8. ฝนแล้ง เป็นภัยธรรมชาติซึ่งเกิดจากฝนไม่ตกตามฤดูกาล ส่งผลให้เกิดการขาดแคลนน้ำซึ่งมีผลกระทบต่อ และก่อให้เกิดความเสียหายมากมายต่อมนุษย์ทั้งทางตรงและทางอ้อม	ฝนแล้ง เป็นภัยธรรมชาติซึ่งเกิดจากฝนไม่ตกตามฤดูกาล ส่งผลให้เกิดการขาดแคลนน้ำ ผลกระทบและก่อให้เกิดความเสียหายมากมายต่อมนุษย์ทั้งทางตรงและทางอ้อม
9. โทรศัพท์มือถือเป็นเครื่องมือสื่อสารที่ทำให้ประชาชนเข้าถึงข้อมูลข่าวสารได้อย่างสะดวก เพราะสามารถพกพาได้ และให้บริการได้เกือบทุกพื้นที่	โทรศัพท์มือถือเป็นเครื่องมือสื่อสารที่ทำให้ประชาชนเข้าถึงข้อมูลข่าวสารได้อย่างสะดวก เพราะสามารถพกพาได้และให้บริการได้เกือบทุกพื้นที่
10. ตอนนี้โทรศัพท์มือถือได้กลายมาเป็นส่วนหนึ่งของชีวิตประจำวันและมีใช้กันอย่างแพร่หลาย	ตอนนี้โทรศัพท์มือถือได้กลายมาเป็นส่วนหนึ่งของชีวิตประจำวันและมีใช้กันอย่างแพร่หลาย

จากตารางที่ 4.2 การทดสอบรับเสียงพูดภาษาไทย พบว่าจากการรับเสียง 10 ประโยค สามารถรับเสียงถูกต้องได้ 8 ประโยค และผิด 2 ประโยค โดยพิจารณาความถูกต้องจากการพิจารณาผลลัพธ์ที่ได้ตรงกับประโยคที่ใช้ทดสอบ ซึ่งประโยคที่รับเสียงผิดนั้นเกิดจากคำที่ออกเสียงใกล้เคียงเกินไป และการรับเสียงได้ไม่ครบทั้งประโยค เช่น ฝนแล้ง เป็นภัยธรรมชาติซึ่งเกิดจากฝนไม่ตกตาม รค่า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ฤดูกาล ส่งผลให้เกิดการขาดแคลนน้ำซึ่งมีผลกระทบ และก่อให้เกิดความเสียหายมากมายต่อมนุษย์ ทั้งทางตรงและทางอ้อม รับประทานเป็น ผืนเลี้ยง เป็นภัยธรรมชาติซึ่งเกิดจากฝนไม่ตกตามฤดูกาลส่งผลให้เกิดการขาดแคลนน้ำผลกระทบและก่อให้เกิดความเสียหายมากมายต่อมนุษย์ทั้งทางตรงและทางอ้อม เป็นต้น

4.2.3) การทดสอบประสิทธิภาพการแปลภาษาอังกฤษเป็นภาษาไทยของ Bing และ Google

ในการทดสอบการแปลจะใช้ประโยคชุดเดียวกับประโยคที่ใช้ในการรับเสียงภาษาอังกฤษ

ตารางที่ 4.3 แสดงผลการแปลภาษาอังกฤษเป็นภาษาไทยด้วย Bing และ Google

ประโยค	แปลด้วย Bing	แปลด้วย Google
1. It may not have the same worldwide fame as the Olympic Games, but the Pan American Games certainly knows how to do an opening ceremony.	มันอาจไม่มีชื่อเสียงทั่วโลก เหมือนกับเป็น โอลิมปิก แต่แพนอเมริกันเกมส์แน่นอนรู้วิธีการทำพิธีเปิดการ	มันอาจจะไม่ได้มีชื่อเสียงไปทั่วโลกเช่นเดียวกับการแข่งขันกีฬาโอลิมปิก แต่แพนอเมริกันเกมส์อย่างแน่นอนรู้วิธีการทำพิธีเปิด
2. She came to America speaking only her native language	เธอมาถึงอเมริกาพูดเฉพาะภาษาพื้นเมืองของเธอ	เธอมาถึงอเมริกาเพียงการพูดภาษาบ้านเกิดของเธอ
3. The parachutes help them go down slowly	Parachutes ช่วยให้พวกเขาลงไปอย่างช้า ๆ	ร่มชูชีพช่วยให้พวกเขาลงไปอย่างช้า ๆ
4. People held the world paper aeroplane championships in Austria.	คนถือกระดาษเครื่องบินชิงแชมป์โลกที่ประเทศออสเตรีย	คนที่จัดขึ้นชิงแชมป์เครื่องบินกระดาษโลกในออสเตรีย
5. Christmas is celebrated all over the world, as a religious holiday or as a time of celebration by Christians and non-Christians alike.	คริสต์มาสคือการเฉลิมฉลองทั่วโลก เป็นวันหยุดทางศาสนา หรือ เป็นระยะเวลาแห่งการเฉลิมฉลองโดยคริสเตียนและไม่ใช่ผู้ที่นับถือศาสนาเหมือนกัน	มีการเฉลิมฉลองคริสต์มาสทั่วโลกที่เป็นวันหยุดทางศาสนาหรือเป็นเวลาของการเฉลิมฉลองโดยชาวคริสต์และไม่ใช่คริสเตียนเหมือนกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.3 (ต่อ) แสดงผลการแปลภาษาอังกฤษเป็นภาษาไทยด้วย Bing และ Google

ประโยค	แปลด้วย Bing	แปลด้วย Google
6. A black bear and her cubs chased a group of tourists on a bridge in the park.	หมีดำและลูกของเธอคนนั้นได้ไล่กลุ่มนักท่องเที่ยวบนสะพานในสวน	หมีสีดำและลูกของเธอไล่ล่ากลุ่มของนักท่องเที่ยวบนสะพานในสวนสาธารณะ
7. She's been programmed to speak Japanese but can also be fine-tuned to speak in Chinese and Korean, too.	เธอถูกตั้งโปรแกรมพูดภาษาญี่ปุ่น แต่ยังมี fine-tuned พูดในจีนและเกาหลีเกินไป	เธอถูกตั้งโปรแกรมให้พูดภาษาญี่ปุ่น แต่ยังสามารถปรับแต่งที่จะพูดในภาษาจีนและภาษาเกาหลีมากขึ้นไป
8. The flamingo, which is entirely black, was feeding with the other pink flamingos on the banks of the lake on Wednesday afternoon.	ฟลามิงโกซึ่งเป็นสีดำทั้งหมดให้อาหารกับอื่น ๆ flamingos ชมพุกริมทะเลสาบในตอนบ่ายวันพุธ	นกกระเรียนที่เป็นสีดำทั้งหมดได้รับการให้อาหารด้วยฟลามิงโกสีชมพูอื่น ๆ ที่อยู่บนฝั่งของทะเลสาบในตอนบ่ายวันพุธ
9. The company will begin testing the vehicles on public roads in the coming months, but it turns out they will have steering wheels and brakes, something they had not envisioned a year ago.	บริษัทจะเริ่มการทดสอบรถบนถนนสาธารณะในเดือนมา แต่มันเปิดออกจะมีพวงมาลัยล้อและเบรค สิ่งที่พวกเขาไม่มีจินตนาการปีที่ผ่านมา	บริษัท จะเริ่มต้นการทดสอบยานพาหนะบนถนนสาธารณะในเดือนที่ผ่านมา แต่มันกลับกลายเป็นพวกเขาจะมีพวงมาลัยและเบรคสิ่งที่พวกเขาไม่เคยจินตนาการปีที่ผ่านมา
10. The men working in the northwest of the country discovered a huge manta ray which local media have said weighs just over a ton.	คนที่ทำงานในตะวันตกเฉียงเหนือของประเทศพบปลากระเบนแมนตามากกว่า สี่ห้องถิ่นที่มีน้ำหนักเพียงกว่าตัน	คนทำงานในทิศตะวันตกเฉียงเหนือของประเทศพบกระเบนราหูขนาดใหญ่ที่สี่ห้องถิ่นได้กล่าวว่ามีน้ำหนักเพียงตัน

จากตารางที่ 4.3 พบว่าการแปลภาษา จากภาษาอังกฤษเป็นภาษาไทย ด้วย Google translation มีความถูกต้องในการสื่อความหมายมากกว่าการแปลภาษาด้วย Bing translator และความเหมาะสมของการใช้คำของ Bing translatorจะน้อยกว่า Google translation เช่น การแปลว่า

ประโยค *She's been programmed to speak Japanese but can also be fine-tuned to speak in Chinese and Korean, too.*

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Bing translator จะแปลได้ว่า เธอถูกตั้ง โปรแกรมพูดภาษาญี่ปุ่น แต่ยังมี fine-tuned พูดในจีนและเกาหลี เกินไป

Google translator จะแปลได้ว่า เธอถูกตั้ง โปรแกรมให้พูดภาษาญี่ปุ่น แต่ยังสามารถปรับแต่งที่จะพูดในภาษาจีนและภาษาเกาหลีมากเกินไป

จะเห็นได้ว่าในคำว่า but can also be fine-tuned การแปลด้วย Bing จะแปลว่า *แต่ยังมี fine-tuned* ส่วน Google จะแปลว่า *แต่ยังสามารถปรับแต่ง* ซึ่งการเลือกใช้คำที่เหมาะสมจะทำให้ประโยคสามารถสื่อความหมายได้ถูกต้อง เข้าใจได้มากกว่า

4.2.4) การทดสอบประสิทธิภาพการแปลภาษาไทยเป็นภาษาอังกฤษของ Bing และ Google ในการทดสอบการแปลจะใช้ประโยคชุดเดียวกับประโยคที่ใช้ในการรับเสียงภาษาไทย

ตารางที่ 4.4 แสดงผลการแปลภาษาไทยเป็นภาษาอังกฤษด้วย Bing และ Google

ประโยค	แปลด้วย Bing	แปลด้วย Google
1. ภาษาเป็นที่มาของ วัฒนธรรมและวัฒนธรรมก็ เป็นที่มาของภาษาเช่นกัน	Language is the source of the culture, and culture is the origin of the language as well.	Language is a source of culture, and culture is the origin of language as well.
2. วัฒนธรรมคือสิ่งที่มนุษย์ เปลี่ยนแปลง ปรับปรุง หรือ ผิดดีขึ้นเพื่อความเจริญอก งามในวิถีชีวิต	Culture is something that human beings change. Update or Antiaging prosperity flourish in life	Human culture is what changes or improvements made to the growth in life.
3. การสื่อสารข้อมูลเป็น การถ่ายทอดความรู้ ความคิดความรู้สึจากคน หนึ่งไปสู่อีกคนหนึ่ง โดย อาศัยสื่อหรือเครื่องมือ ต่างๆเป็นช่องทางในการ สื่อสาร	Data communication is to convey knowledge from one person feel to one another through the media or tools such as a communication channel.	Communications data is transmitted knowledge, thoughts and feelings from one person to another through the media or in the channels of communication.
4. สิ่งแวดล้อม หมายถึงสิ่ง ที่เกิดขึ้นเองตามธรรมชาติ และสิ่งที่มนุษย์สร้างขึ้น	Environment refers to what happens naturally and man-made things.	Environment refers to naturally occurring and man-made.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.4 (ต่อ) แสดงผลการแปลภาษาไทยเป็นภาษาอังกฤษด้วย Bing และ Google

ประโยค	แปลด้วย Bing	แปลด้วย Google
5. ปัญหา สิ่งแวดล้อม สามารถเกิดขึ้น ทุกพื้นที่ โดยเฉพาะอย่างยิ่ง พื้นที่ ที่ประชากรอาศัยอยู่ หนาแน่น	Environmental problems can occur in all areas, especially areas where dense living population.	Environmental problems can occur in all areas. Especially densely populated areas.
6. AEC เกี่ยวข้องกับการ รวมตัวทางเศรษฐกิจของ สมาชิกอาเซียนซึ่งครอบคลุม การเปิดเสรีและอำนวยความสะดวก สะดวกในการเคลื่อนย้ายทุน, การเคลื่อนย้ายแรงงาน, การ ปรับพิธีการศุลกากร, การ กำหนดมาตรฐานสินค้าและ นโยบายเศรษฐกิจให้ สอดคล้องกันระหว่าง ประเทศสมาชิกอาเซียน	AEC is related to the economic integration of ASEAN members. Which covers the liberalization and facilitation of movement of capital, labor mobility, improving customs brokerage, product standardization and consistent economic policy between the ASEAN member countries.	AEC related to the economic integration of ASEAN. Which covers liberalization and facilitating the movement of capital, labor mobility, the customs, standards and harmonizing economic policies among member countries.
7. ปลาฉลามเป็นสัตว์น้ำทะเลที่ เป็นที่รู้จักกันอย่างแพร่หลาย เพราะรูปร่างที่แปลกตาและ ความไม่มีพิษภัยของมัน	Starfish is a sea creature that is widely known because of the unusual shape, and there is no danger of its venom.	Starfish and sea creatures that are widely known because of its shape and quaint. No danger of it
8 ฝนแล้ง เป็นภัยธรรมชาติ ซึ่งเกิดจากฝนไม่ตกตาม ฤดูกาล ส่งผลให้เกิดการขาด แคลนน้ำซึ่งมีผลกระทบ และ ก่อให้เกิดความเสียหาย มากมายต่อมนุษย์ทั้งทางตรง และทางอ้อม	Drought is a natural disaster, which caused the rain does not fall in season. Resulted in water shortages, which have affected and caused a lot of damage against human beings, both direct and indirect.	Drought is a natural disaster caused by seasonal rain. Resulting in a water shortage which has affected. And causing extensive damage to humans both directly and indirectly.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.4 (ต่อ) แสดงผลการแปลภาษาไทยเป็นภาษาอังกฤษด้วย Bing และ Google

ประโยค	แปลด้วย Bing	แปลด้วย Google
9. โทรศัพท์มือถือเป็นเครื่องมือสื่อสารที่ทำให้ประชาชนเข้าถึงข้อมูลข่าวสารได้อย่างสะดวกเพราะสามารถพกพาได้และให้บริการได้เกือบทุกพื้นที่	A mobile phone is a communication tool that allows people to access the information easily. It can carry and serve almost any area.	The mobile phone is a communication tool that allows people to access information easily. Because portability And serve almost every area.
10. ตอนนี้โทรศัพท์มือถือได้กลายมาเป็นส่วนหนึ่งของชีวิตประจำวันและมีใช้กันอย่างแพร่หลาย	Now, the mobile phone has become a part of everyday life and is widely used.	Now mobile phones have become part of everyday life and is widely used.

จากตารางที่ 4.4 พบว่าการแปลภาษาไทย-ภาษาอังกฤษ ด้วย Google translation ผลลัพธ์ที่ได้มีความถูกต้องและใช้คำได้เหมาะสมกว่า Bing Translator เช่น การสื่อสารข้อมูลเป็นการถ่ายทอดความรู้ ความคิดความรู้สึกจากคนหนึ่ง ไปสู่อีกคนหนึ่ง โดยอาศัยสื่อหรือเครื่องมือต่างๆ เป็นช่องทางในการสื่อสาร

Bing translator จะแปลได้ว่า Data communication is to convey knowledge from one person feel to one another through the media or tools such as a communication channel.

Google translator จะแปลได้ว่า Communications data is transmitted knowledge, thoughts and feelings from one person to another through the media or in the channels of communication.

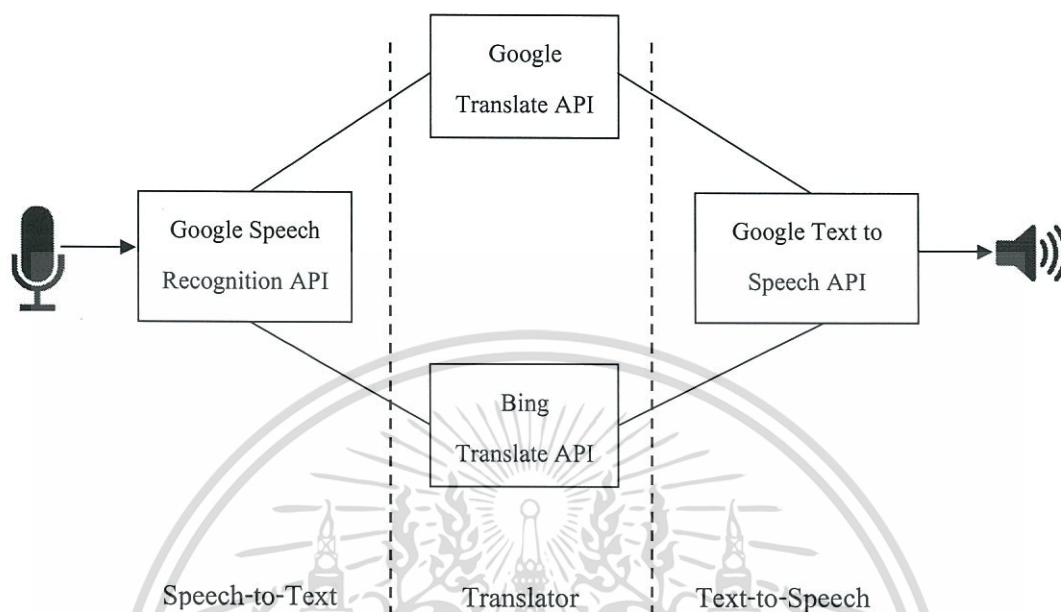
จะเห็นว่า Bing ใช้คำว่า *convey knowledge* แต่ Google ใช้คำว่า *transmitted knowledge* ซึ่งเป็นคำที่เหมาะสมกับประโยค และให้ความหมายที่ถูกต้องมากกว่า

4.3 การทดสอบประสิทธิภาพของ API ด้วยแอปพลิเคชันแปลภาษาด้วยเสียงพูด

การทดสอบ API ที่ได้เลือกมาจากการศึกษาเปรียบเทียบคุณสมบัติที่มีความเหมาะสมกับแอปพลิเคชัน จะทดสอบโดยการพัฒนาแอปพลิเคชันแปลภาษาด้วยเสียงพูดบนระบบปฏิบัติการแอนดรอยด์

ในการทดสอบประสิทธิภาพของ API ผู้จัดทำทดสอบด้วยการรับเสียงเป็นภาษาอังกฤษและภาษาไทยเนื่องจากเป็นภาษาที่นิยมใช้มากที่สุดในชีวิตประจำวัน แต่ในแอปพลิเคชันที่เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

พัฒนาขึ้นมาจะสามารถเลือกให้แปลได้หลายภาษา (ไทย, อังกฤษ, เกาหลี, จีน, ญี่ปุ่น) โดยในการทดสอบจะเป็นการรับเสียง ภาษาอังกฤษแปลเป็นภาษาไทยและภาษาไทยแปลเป็นภาษาอังกฤษ กระบวนการทดสอบแอปพลิเคชันแปลภาษาด้วยเสียงพูด ดังรูปที่ 4.3



รูปที่ 4.3 กระบวนการทดสอบ

4.3.1) ทดสอบแอปพลิเคชันโดยการรับเสียงภาษาอังกฤษแปลเป็นภาษาไทย โดยใช้ผู้ประเมิน 10 คน (แบ่งเป็นผู้เชี่ยวชาญ 3 คน คนทั่วไป 7 คน) ด้วยประโยคตัวอย่าง 30 ประโยค โดยในขั้นตอน Translation จะทดสอบทั้งแบบที่ใช้ API ของ Bing และ API ของ Google เพื่อเปรียบเทียบถึงประสิทธิภาพของ API โดยแบ่งกลุ่มประโยคตามการออกเสียงสูง-ต่ำโดยการให้ผู้ประเมิน ทดสอบการทำงานของแต่ละชั้นลงในแบบทดสอบโดยมีตัวอย่างแบบทดสอบ (ตารางที่ ข.1)

- การทดสอบรับเสียงภาษาอังกฤษโดยใช้ Google Speech Recognition API

เนื่องจากผู้จัดทำได้แบ่งการทดสอบออกเป็น 2 ครั้ง เพื่อเปรียบเทียบประสิทธิภาพการแปลภาษาโดยใช้ Bing และ Google เพราะฉะนั้นในส่วนของ Speech-to-Text จะถูกทดสอบทั้งหมด 20 ครั้ง และจะนำมาสรุปผลด้วยการพิจารณาจากผลลัพธ์ที่ได้จากผู้ประเมินดังตาราง

ตารางที่ 4.5 แสดงผลการทดสอบการรับเสียงภาษาอังกฤษ

ประโยค	การรับเสียง (จำนวนครั้งที่ถูก)
1. Good morning.	18
2. How are you?	20

เอกสารนี้เป็นเอกสารสงวนลิขสิทธิ์สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาติให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.5 (ต่อ) แสดงผลการทดสอบการรับเสียงภาษาอังกฤษ

ประโยค	การรับเสียง (จำนวนครั้งที่ถูก)
3. How tall are you?	20
4. What's your name?	20
5. How old are you?	19
6. I like to read.	18
7. Open the door, please.	18
8. Are you angry at me?	19
9. What are you doing?	17
10. May I help you?	19
11. How much is it?	20
12. Can I pay by credit card?	19
13. What do you want?	18
14. I'm going to play football.	19
15. I speak a little English.	17
16. Where are we going?	18
17. Do you like to travel?	18
18. It's a nice day.	20
19. Could you take a photo for me?	15
20. I going to the airport.	17
21. See you again.	19
22. Will you wake me up at 6 o'clock?	14
23. Where do you work?	18
24. Can you sing?	17
25. What time is it?	20
26. What have you been up to?	20
27. Did you buy anything?	20
28. How long will you be staying?	16
29. Do you need a taxi?	19
30. I'd like to go see a movie.	18

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากตารางที่ 4.5 แสดงผลการทดสอบรับเสียงพูดภาษาอังกฤษด้วย Google Speech Recognition API จากกลุ่มผู้ประเมิน 10 คน โดยแต่ละคนจะออกเสียงประโยคละ 2 ครั้ง เนื่องจากต้องทดสอบการแปลภาษาด้วย Bing กับ Google เพราะฉะนั้นผลรวมแต่ละประโยคจะต้องถูกออกเสียงทั้งหมด 20 ครั้งจากการทดสอบพบว่า การรับเสียงพูดประโยคบอกเล่าและประโยคคำถามส่วนใหญ่รับเสียงได้อย่างถูกต้อง มีเพียงบางประโยคที่รับเสียงผิดพลาด 4-5 ครั้งจาก 20 ครั้ง ซึ่งเป็นประโยคคำถามประเภทขึ้นต้นด้วยกริยาช่วย ดังนั้นจากการทดสอบทำให้ทราบว่า Google Speech Recognition API มีความถูกต้องในการรับเสียงพูดภาษาอังกฤษร้อยละ 91.66 แสดงให้เห็นว่าประสิทธิภาพในการรับเสียงพูดภาษาอังกฤษอยู่ในระดับที่ดีมาก

- ทดสอบการแปลภาษาอังกฤษเป็นภาษาไทย โดยใช้ Bing Translator Control API และ Google Translator Control API จากผลการทดสอบการแปลในแต่ละประโยคของผู้ประเมินทั้ง 10 คนจะนำมาสรุปผลด้วยการพิจารณาจากผลลัพธ์ที่ได้จากผู้ทดสอบดังตาราง

ตารางที่ 4.6 แสดงผลการทดสอบการแปลภาษาอังกฤษเป็นภาษาไทย

ประโยค	การแปล (ครั้ง)	
	Bing	Google
1. Good morning.	9	9
2. How are you?	10	6
3. How tall are you?	2	5
4. What's your name?	10	5
5. How old are you?	9	10
6. I like to read.	6	9
7. Open the door, please.	2	5
8. Are you angry at me?	2	3
9. What are you doing?	9	10
10. May I help you?	10	9
11. How much is it?	9	10
12. Can I pay by credit card?	10	9
13. What do you want?	8	10
14. I'm going to play football.	9	10

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.6 (ต่อ) แสดงผลการทดสอบการแปลภาษาอังกฤษเป็นภาษาไทย

ประโยค	การแปล (ครั้ง)	
	Bing	Google
15. I speak a little English.	6	10
16. Where are we going?	9	9
17. Do you like to travel?	4	6
18. It's a nice day.	9	10
19. Could you take a photo for me?	6	5
20. I going to the airport.	10	7
21. See you again.	9	5
22. Will you wake me up at 6 o'clock?	0	5
23. Where do you work?	2	10
24. Can you sing?	8	8
25. What time is it?	10	7
26. What have you been up to?	5	2
27. Did you buy anything?	1	1
28. How long will you be staying?	2	8
29. Do you need a taxi?	8	6
30. I'd like to go see a movie.	9	10

จากตารางที่ 4.6 แสดงผลการทดสอบแปลภาษาอังกฤษเป็นภาษาไทยด้วย Bing Translator Control API และ Google Translator Control API จากกลุ่มผู้ทดสอบ 10 คน พบว่า Bing มีความถูกต้องในการแปลภาษาร้อยละ 67.6 และ Google มีความถูกต้องในการแปลภาษาร้อยละ 73 แสดงให้เห็นว่า Google มีประสิทธิภาพการแปลภาษาดีกว่า Bing

- ทดสอบการสังเคราะห์เสียงภาษาไทยจากการแปลภาษาด้วย Google Text to Speech API โดยพิจารณาจากความเข้าใจและความเป็นธรรมชาติของเสียงด้วยวิธีการ Mean Opinion Score (MOS) [17] หมายถึง ผู้ทดสอบเข้าใจการสื่อความหมายของประโยค มากกว่าการเข้าใจว่าประโยคนั้นประกอบด้วยคำไหนบ้างและให้คะแนนจากความพึงพอใจ เป็นระดับคือ 5 = ดีเยี่ยม, 4 = ดี, 3 = พอใช้, 2 = ปรับปรุง, 1 = ไม่มี

เนื่องจากผู้จัดทำได้แบ่งการทดสอบออกเป็น 2 ครั้ง เพื่อเทียบประสิทธิภาพการแปลภาษาโดยใช้ Bing และ Google เพราะฉะนั้นในส่วนของ Text-to-Speech จึงได้แยกเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

พิจารณาความเข้าใจและความเป็นธรรมชาติของเสียงโดยพิจารณาการสังเคราะห์เสียงจากข้อความที่แปลด้วย Bing และ Google เพื่อนำมาสรุปผลในแต่ละประโยคดังตารางตารางที่ 4.7 แสดงผลทดสอบการสังเคราะห์เสียงภาษาไทย

ประโยค	ความเข้าใจและความเป็นธรรมชาติของเสียง (ร้อยละ)	
	แปลโดย Bing	แปลโดย Google
1. Good morning.	88	92
2. How are you?	90	82
3. How tall are you?	38	64
4. What's your name?	94	58
5. How old are you?	84	92
6. I like to read.	60	82
7. Open the door, please.	32	58
8. Are you angry at me?	36	60
9. What are you doing?	88	90
10. May I help you?	86	96
11. How much is it?	68	92
12. Can I pay by credit card?	90	92
13. What do you want?	86	90
14. I'm going to play football.	74	98
15. I speak a little English.	52	98
16. Where are we going?	80	84
17. Do you like to travel?	60	68
18. It's a nice day.	70	94
19. Could you take a photo for me?	44	52
20. I going to the airport.	90	76
21. See you again.	88	70
22. Will you wake me up at 6 o'clock?	26	52
23. Where do you work?	52	68
24. Can you sing?	70	80

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.7 (ต่อ) แสดงผลทดสอบการสังเคราะห์เสียงภาษาไทย

ประโยค	ความเข้าใจและความเป็นธรรมชาติของเสียง (ร้อยละ)	
	แปลโดย Bing	แปลโดย Google
25. What time is it?	92	66
26. What have you been up to?	62	32
27. Did you buy anything?	58	48
28. How long will you be staying?	36	70
29. Do you need a taxi?	80	84
30. I'd like to go see a movie.	78	96

จากตารางที่ 4.7 แสดงผลการทดสอบการสังเคราะห์เสียงภาษาไทยด้วย Google Text to Speech API พบว่าผลการทดสอบการสังเคราะห์เสียงที่แปลด้วย Bing Translator Control API และ Google Translator Control API ผู้ทดสอบมีความพึงพอใจอยู่ที่ร้อยละ 68.40 และ 76.13 ตามลำดับ ซึ่งแสดงให้เห็นว่าการสังเคราะห์เสียงภาษาไทยที่แปลจาก Google มีความเข้าใจและสื่อความหมายได้ดีกว่าการสังเคราะห์เสียงที่ได้จากการแปลด้วย Bing จากผลการทดสอบทำให้ประสิทธิภาพของ Google Text to Speech API ในการสังเคราะห์เสียงภาษาไทยโดยรวมอยู่ในระดับปานกลาง

4.3.2) ทดสอบแอปพลิเคชันโดยการรับเสียงภาษาไทยแปลเป็นภาษาอังกฤษ โดยใช้ผู้ประเมิน 10 คน (ในส่วนของรับเสียงภาษาไทยจะใช้คนทั่วไป 10 คน) ด้วยประโยคตัวอย่าง 30 ประโยค โดยในขั้นตอน Translation จะทดสอบทั้งแบบที่ใช้ API ของ Bing และ API ของ Google โดยจะเป็นประโยคที่แปลมาจากประโยคที่ใช้ทดสอบในการรับเสียงภาษาอังกฤษโดยการให้ผู้ประเมิน ทดสอบการทำงานของแต่ละขั้นลงในกระดาษแบบทดสอบ (ตารางที่ ข.2)

- การทดสอบรับเสียงภาษาไทยโดยใช้ Google Speech Recognition API

เนื่องจากผู้จัดทำได้แบ่งการทดสอบออกเป็น 2 ครั้ง เพื่อเปรียบเทียบประสิทธิภาพการแปลภาษาโดยใช้ Bing และ Google เพราะฉะนั้นในส่วนของ Speech-to-Text จะถูกทดสอบทั้งหมด 20 ครั้ง และจะนำมาสรุปผลด้วยการพิจารณาจากผลลัพธ์ที่ได้จากผู้ประเมินดังตาราง

ตารางที่ 4.8 แสดงผลจากการทดสอบการรับเสียงภาษาไทย

ประโยค	การรับเสียง (จำนวนครั้งที่รับถูก)
1. สวัสดิ์	19
2. คุณสบายดีไหม	20
3. คุณสูงเท่าไร	20
4. คุณชื่ออะไร	20
5. คุณอายุเท่าไร	20
6. ฉันชอบอ่านหนังสือ	16
7. กรุณาเปิดประตูด้วย	17
8. คุณโกรธฉันหรือเปล่า	16
9. คุณทำอะไรอยู่	19
10. ให้นินช่วยไหม	20
11. ราคาเท่าไร	18
12. จ่ายด้วยบัตรเครดิตได้ไหม	19
13. คุณต้องการอะไร	19
14. ฉันจะเล่นฟุตบอล	19
15. ฉันพูดภาษาอังกฤษได้นิดหน่อย	19
16. เราจะไปไหนกัน	19
17. คุณชอบท่องเที่ยวไหม	18
18. วันนี้อากาศดีจังนะ	20
19. คุณช่วยถ่ายรูปให้ฉันหน่อยได้ไหม	19
20. ฉันจะไปลงสนามบิน	19
21. แล้วพบกันใหม่	20
22. ปลุกฉันตอน 6 โมงได้ไหม	19
23. คุณทำงานที่ไหน	18
24. คุณร้องเพลงเป็นไหม	8
25. ก็โมงแล้ว	20
26. คุณทำอะไรบ้างช่วงนี้	18
27. คุณได้ชื่ออะไรไหม	16
28. คุณอยู่นานแค่ไหน	19

เอกสารนี้เป็นเอกสารลิขสิทธิ์ของมหาวิทยาลัยราชภัฏวไลยอลงกรณ์ในพระบรมราชูปถัมภ์ ใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.8 (ต่อ) แสดงผลจากการทดสอบการรับเสียงภาษาไทย

ประโยค	การรับเสียง (ครั้ง)
29. คุณต้องการแท็กซี่หรือไม่	19
30. ฉันอยากไป ดูหนัง	20

จากตารางที่ 4.8 แสดงผลการทดสอบรับเสียงพูดภาษาไทยด้วย Google Speech Recognition API จากกลุ่มผู้ประเมิน 10 คน โดยผู้ประเมินแต่ละคนจะออกเสียงประโยคละ 2 ครั้ง เนื่องจากต้องทดสอบการแปลภาษาด้วย Bing กับ Google เพราะฉะนั้นผลรวมแต่ละประโยคจะต้องถูกออกเสียงทั้งหมด 20 ครั้ง จากการทดสอบพบว่าส่วนใหญ่สามารถรับเสียงพูดภาษาไทยได้ถูกต้อง มีเพียงประโยค "คุณร้องเพลงเป็นไหม" ที่รับเสียงผิดถึง 12 ครั้ง จาก 20 ครั้ง ดังนั้นจากผลการทดสอบทำให้ทราบว่า Google Speech Recognition API มีความถูกต้องในการรับเสียงพูดภาษาไทยร้อยละ 92.16 แสดงให้เห็นว่าประสิทธิภาพในการรับเสียงพูดภาษาไทยอยู่ในระดับที่ดีมาก

- ทดสอบการแปลภาษาไทยเป็นภาษาอังกฤษ โดยใช้ Bing Translator Control API และ Google Translator Control API จากผลการทดสอบการแปลในแต่ละประโยคของผู้ทดสอบทั้ง 10 คนจะนำมาสรุปผลด้วยการพิจารณาจากผลลัพธ์ที่ได้จากผู้ทดสอบดังตาราง ตารางที่ 4.9 แสดงผลการทดสอบการแปลภาษาไทยเป็นภาษาอังกฤษ

ประโยค	การแปลภาษา (ครั้ง)	
	Bing	Google
1. สวัสดี	10	10
2. คุณสบายดีไหม	10	9
3. คุณสูงเท่าไร	6	9
4. คุณชื่ออะไร	10	5
5. คุณอายุเท่าไร	10	10
6. ฉันชอบอ่านหนังสือ	6	9
7. กรุณาเปิดประตูด้วย	7	10
8. คุณโกรธฉันหรือเปล่า	6	10
9. คุณทำอะไรอยู่	10	9
10. ให้ฉันช่วยไหม	6	10
11. ราคาเท่าไร	9	9

ตารางที่ 4.9 (ต่อ) แสดงผลการทดสอบการแปลภาษาไทยเป็นภาษาอังกฤษ

ประโยค	การแปลภาษา (ครั้ง)	
	Bing	Google
12. จ่ายด้วยบัตรเครดิตได้ไหม	9	10
13. คุณต้องการอะไร	6	10
14. ฉันจะเล่นฟุตบอล	8	2
15. ฉันพูดภาษาอังกฤษได้นิดหน่อย	8	10
16. เราจะไปไหนกัน	4	3
17. คุณชอบท่องเที่ยวไหม	9	0
18. วันนี้อากาศดีจังนะ	9	6
19. คุณช่วยถ่ายรูปให้ฉันหน่อยได้ไหม	2	6
20. ฉันจะไปลงสนามบิน	10	9
21. แล้วพบกันใหม่	8	10
22. ปลุกฉันตอน 6 โมงได้ไหม	2	5
23. คุณทำงานที่ไหน	9	9
24. คุณร้องเพลงเป็นไหม	3	0
25. ก็โมงแล้ว	10	10
26. คุณทำอะไรบ้างช่วงนี้	6	10
27. คุณได้ซื้ออะไรไหม	6	3
28. คุณอยู่นานแค่ไหน	6	9
29. คุณต้องการแท็กซี่หรือไม่	10	9
30. ฉันอยากไป ดูหนัง	9	10

จากตารางที่ 4.9 แสดงผลการทดสอบการแปลภาษาไทยเป็นภาษาอังกฤษด้วย Bing Translator Control API และ Google Translator Control API จากกลุ่มผู้ทดสอบ 10 คน พบว่าทั้ง 2 API มีความถูกต้องในการแปลใกล้เคียงกัน คือ Bing มีความถูกต้องในการแปลภาษาร้อยละ 74.66 และ Google มีความถูกต้องในการแปลภาษาร้อยละ 77 เป็นผลให้ประสิทธิภาพในการแปลของ Bing Translator Control API และ Google Translator Control API อยู่ในระดับที่ดี คือ มีความถูกต้อง และสามารถรักษาความหมายของข้อความต้นทางได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ทดสอบการสังเคราะห์เสียงภาษาอังกฤษจากการแปลภาษาด้วย Google Text to Speech API โดยพิจารณาจากความเข้าใจและความเป็นธรรมชาติของเสียงด้วยวิธีการ Mean Opinion Score (MOS) [17] หมายถึง ผู้ประเมินเข้าใจการสื่อความหมายของประโยคมากกว่าการเข้าใจว่าประโยคนั้นประกอบด้วยคำไหนบ้างและให้คะแนนจากความพึงพอใจเป็นระดับคือ 5 = ดีเยี่ยม, 4 = ดี, 3 = พอใช้, 2 = ปรับปรุง, 1 = ไม่ดี

เนื่องจากผู้จัดทำได้แบ่งการทดสอบออกเป็น 2 ครั้ง เพื่อเปรียบเทียบประสิทธิภาพการแปลภาษาโดยใช้ Bing และ Google เพราะฉะนั้นในส่วนของ Text-to-Speech จึงได้แยกพิจารณาความเข้าใจและความเป็นธรรมชาติของเสียงโดยพิจารณาการสังเคราะห์เสียงจากการแปลทั้ง Bing และ Google เพื่อนำมาสรุปผลในแต่ละประโยคดังตาราง

ตารางที่ 4.10 แสดงผลการทดสอบการสังเคราะห์เสียงภาษาอังกฤษ

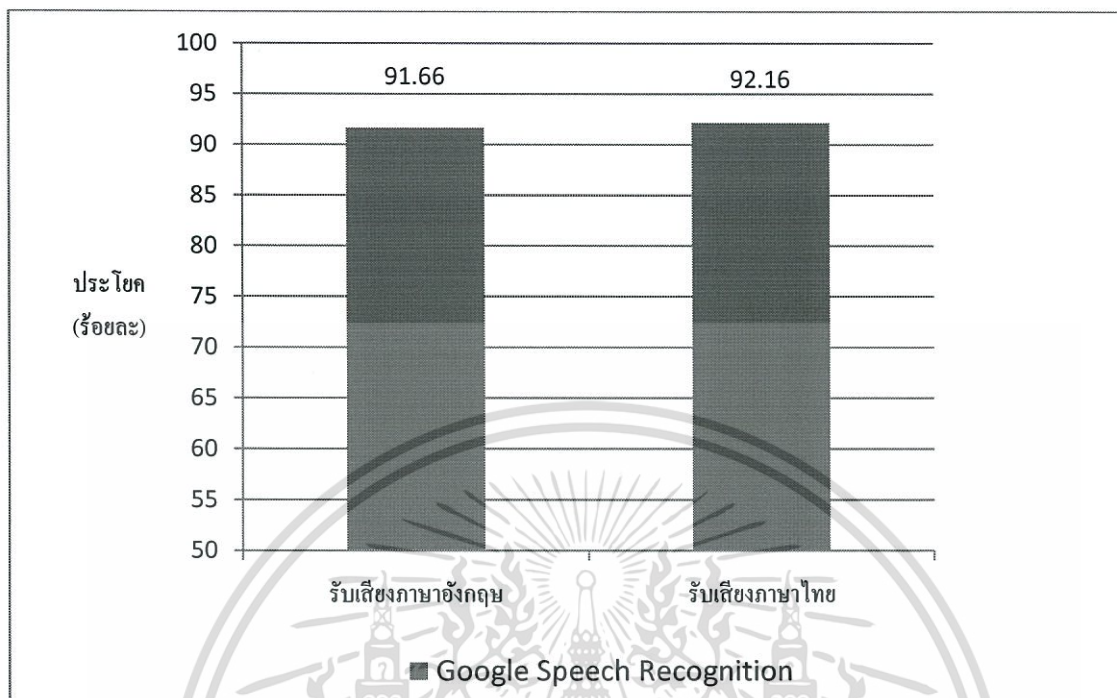
ประโยค	ความเข้าใจและความเป็นธรรมชาติของเสียง (ร้อยละ)	
	แปลโดย Bing	แปลโดย Google
1. สวัสดี	84	100
2. คุณสบายดีไหม	76	86
3. คุณสูงเท่าไร	48	100
4. คุณชื่ออะไร	98	96
5. คุณอายุเท่าไร	100	94
6. ฉันชอบอ่านหนังสือ	78	96
7. กรุณาเปิดประตูด้วย	80	96
8. คุณโกรธฉันหรือเปล่า	68	92
9. คุณทำอะไรอยู่	92	90
10. ให้ฉันช่วยไหม	64	92
11. ราคาเท่าไร	70	88
12. จ่ายด้วยบัตรเครดิตได้ไหม	84	84
13. คุณต้องการอะไร	74	92
14. ฉันจะเล่นฟุตบอล	82	80
15. ฉันพูดภาษาอังกฤษได้นิดหน่อย	82	90
16. เราจะไปไหนกัน	62	58
17. คุณชอบท่องเที่ยวไหม	86	48
18. วันนี้อากาศดีจังนะ	76	82

ตารางที่ 4.10 (ต่อ) แสดงผลการทดสอบการสังเคราะห์เสียงภาษาอังกฤษ

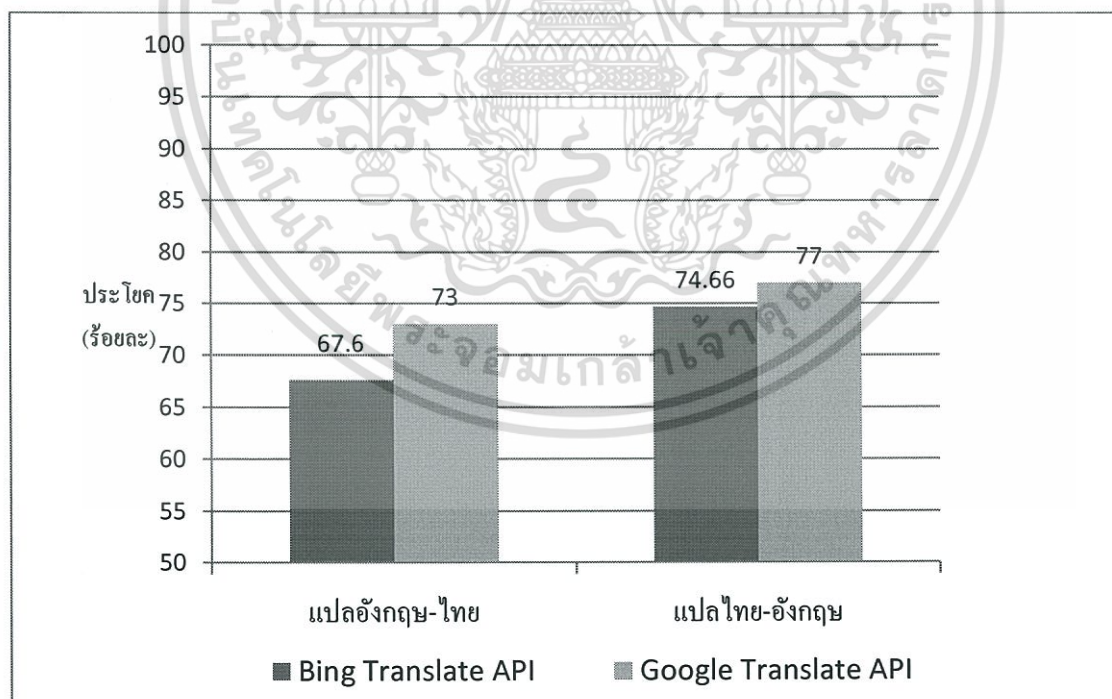
ประโยค	ความเข้าใจและความเป็นธรรมชาติของเสียง (ร้อยละ)	
	แปลโดย Bing	แปลโดย Google
19. คุณช่วยถ่ายรูปให้ฉันหน่อยได้ไหม	66	42
20. ฉันจะไปลงสนามบิน	92	90
21. แล้วพบกันใหม่	82	92
22. ปลุกฉันตอน 6 โมงได้ไหม	54	30
23. คุณทำงานที่ไหน	74	90
24. คุณร้องเพลงเป็นไหม	46	28
25. ก็โมงแล้ว	92	98
26. คุณทำอะไรบ้างช่วงนี้	50	84
27. คุณได้ซื้ออะไรไหม	70	48
28. คุณอยู่นานแค่ไหน	72	74
29. คุณต้องการแท็กซี่หรือไม่	90	70
30. ฉันอยากไป ดูหนัง	82	96

จากตารางที่ 4.10 แสดงผลการทดสอบการสังเคราะห์เสียงภาษาอังกฤษด้วย Google Text to Speech API พบว่าผลการทดสอบการสังเคราะห์เสียงที่แปลด้วย Bing Translator Control API และ Google Translator Control API ผู้ทดสอบมีความพึงพอใจอยู่ที่ร้อยละ 75.80 และร้อยละ 80.20 ตามลำดับ ซึ่งแสดงให้เห็นว่า การสังเคราะห์เสียงภาษาอังกฤษที่แปลจาก Google มีความเข้าใจและสื่อความหมายได้ดีกว่าการสังเคราะห์ที่แปลจาก Bing จากผลการทดสอบทำให้ประสิทธิภาพของ Google Text to Speech API ในการสังเคราะห์เสียงภาษาอังกฤษโดยรวมอยู่ในระดับดี

จากผลการทดสอบตารางที่ 4.5 - 4.10 สามารถสรุปผลเป็นกราฟได้ดังนี้

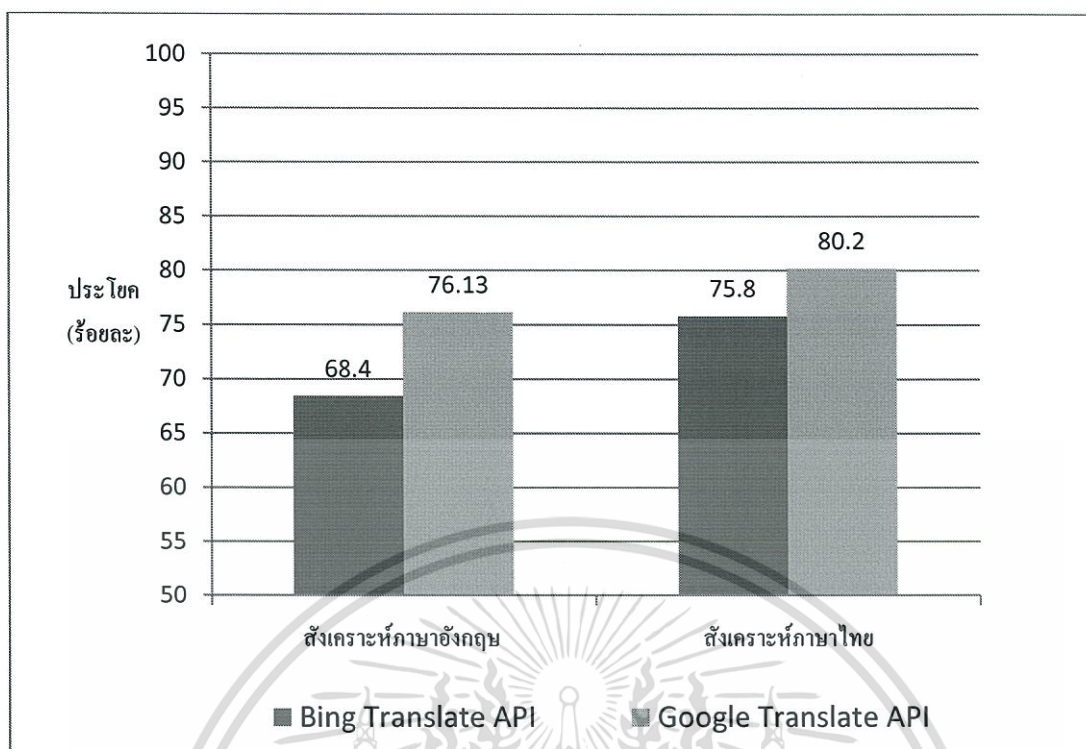


รูปที่ 4.4 แสดงกราฟการรับเสียงภาษาอังกฤษและภาษาไทยด้วย Google Speech Recognition



รูปที่ 4.5 แสดงกราฟการแปลภาษาอังกฤษและภาษาไทยของ Bing Translator Control API และ Google Translator Control API

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.6 แสดงกราฟการสังเคราะห์เสียงภาษาไทยและภาษาอังกฤษด้วย Google Speech Recognition

จากรูปที่ 4.4 แสดงให้เห็นว่าระบบมีประสิทธิภาพในการรับเสียงภาษาอังกฤษได้ดีกว่าภาษาไทยเพียงเล็กน้อยเท่านั้น เนื่องจากผลต่างการรับเสียงภาษาอังกฤษและภาษาไทยต่างกันเพียงร้อยละ 0.5 ดังนั้นทำให้ระบบมีประสิทธิภาพที่ดีทั้งในการรับเสียงภาษาอังกฤษและภาษาไทย ในส่วนของการแปลภาษา ระบบที่เลือกใช้การแปลด้วย Google Translator Control API จะมีประสิทธิภาพที่ดีกว่าการแปลด้วย Bing Translator Control API ทั้งภาษาอังกฤษและภาษาไทย โดยเฉพาะในการแปลภาษาอังกฤษจะเห็นได้ว่า Google มีประสิทธิภาพดีกว่า Bing มาก เนื่องจากมีผลต่างกันถึงร้อยละ 5.4 ดังแสดงในรูปที่ 4.5 ในส่วนการสังเคราะห์เสียงประสิทธิภาพของระบบจะขึ้นอยู่กับส่วนการแปลภาษา กล่าวคือถ้าในส่วนการแปลภาษาเลือกแปลด้วย Google จะส่งผลให้ประสิทธิภาพในการสังเคราะห์เสียงดีทั้งในภาษาอังกฤษและภาษาไทย ดังแสดงในรูปที่ 4.6 จากผลการทดสอบข้างต้นพบว่า ส่วนการรับเสียงของระบบแปลภาษาด้วยเสียงพูดนั้นมีประสิทธิภาพที่ดีมาก แสดงว่าผลจากการแปลภาษาจะต้องได้ผลลัพธ์ที่ดีหรือเทียบเท่าการรับเสียง แต่ผลที่ได้กลับมีความถูกต้องน้อยกว่ากระบวนการรับเสียงเป็นอย่างมาก และส่งผลให้ความเข้าใจในการสังเคราะห์เสียงต่ำลง จึงชี้ให้เห็นว่าประสิทธิภาพของระบบจะเพิ่มขึ้นหรือลดลงเป็นผลมาจากกระบวนการแปลภาษา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.4 ผลรวมของแบบสอบถามเพื่อประเมินความพึงพอใจของผู้ใช้แอปพลิเคชัน

เมื่อผู้ประเมินทำการทดสอบแอปพลิเคชันเรียบร้อยแล้ว จะมีแบบสอบถามเพื่อวัดความพึงพอใจของการใช้งานแอปพลิเคชัน โดยผู้ประเมินให้คะแนนความพึงพอใจในแต่ละหัวข้อ ระดับความพึงพอใจสูงสุดคือ 5 จนถึงต่ำสุดคือ 1 (ตารางที่ ข.3)

จากการประเมินผลความพึงพอใจจากผู้ทดสอบตอบแบบสอบถามทั้งหมด 20 คน

- เพศหญิง 16 คน, เพศชาย 4

- เคยใช้แอปพลิเคชันแปลภาษา 14 คน, ไม่เคยใช้แอปพลิเคชันแปลภาษา 6 คน

ตารางที่ 4.11 แสดงผลรวมของแบบสอบถามความพึงพอใจ

รายการประเมิน	ระดับความพึงพอใจ (ร้อยละ)
1.ความสามารถของแอปพลิเคชัน ในด้านการรับเสียง	78
2.ความสามารถของแอปพลิเคชันในด้านการออกเสียง	84
3.ความสามารถของแอปพลิเคชันในด้านการแปลภาษา	74
4.ความรวดเร็วในด้านการประมวลผล	86
5.ความง่ายต่อการใช้งานของแอปพลิเคชัน	96
6.ความเหมาะสมในการเลือกใช้ขนาดของตัวอักษรในแอปพลิเคชัน	92
7.ความเหมาะสมของการออกแบบแอปพลิเคชันสำหรับแปลภาษา	92
8.ความเหมาะสมในการปฏิสัมพันธ์โต้ตอบกับผู้ใช้	84
9.ประโยชน์ที่ได้จากแอปพลิเคชัน	94
10.ประสิทธิภาพโดยรวมของแอปพลิเคชัน	90
ผลรวมเฉลี่ย	79.1

จากแบบสอบถามเพื่อประเมินความพึงพอใจของการใช้งานแอปพลิเคชันที่ใช้ในการแปลภาษาด้วยเสียงพูดบนระบบปฏิบัติการแอนดรอยด์ โดยผู้ประเมินทั้งหมด 20 คน มีรายการประเมิน 10 หัวข้อ ผลเฉลี่ยที่ผู้ใช้งานประเมินความพึงพอใจอยู่ในระดับดีคือข้อ 1 กับ 3 คือในด้านการรับเสียงและการแปลภาษา รายการประเมินที่ได้ค่าเฉลี่ยสูงสุด 96% คือความง่ายต่อการใช้งานของแอปพลิเคชัน ส่วนข้ออื่นๆผลเฉลี่ยรวมความพึงพอใจอยู่ในระดับดีมาก นอกจากนี้ผลเฉลี่ยรวมทั้งหมดของแอปพลิเคชันคือ 79.1% ผู้ใช้มีความพึงพอใจในระดับดี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อความที่แปลด้วย Google สามารถสื่อความหมายของประโยคให้ผู้ฟังเข้าใจได้ดีถึงร้อยละ 76.13 สำหรับภาษาอังกฤษและร้อยละ 80.20 สำหรับภาษาไทยซึ่งมากกว่าการสังเคราะห์เสียงข้อความที่แปลด้วย Bing เนื่องจากการสังเคราะห์เสียงข้อความที่แปลด้วย Bing ให้ผลลัพธ์เพียงร้อยละ 68.40 ในภาษาอังกฤษและร้อยละ 75.80 ในภาษาไทยเท่านั้น ดังนั้นประสิทธิภาพของการสังเคราะห์เสียงจะขึ้นอยู่กับกระบวนการแปลภาษา และจากผลการทดสอบทำให้ทราบว่าประสิทธิภาพของระบบจะเพิ่มขึ้นหรือลดลงเป็นผลมาจากกระบวนการแปลภาษามากกว่ากระบวนการรับเสียง เนื่องจากในกระบวนการรับเสียงมีประสิทธิภาพที่ดี สามารถรับเสียงได้ถูกต้องทั้งภาษาไทยและภาษาอังกฤษ แต่ในกระบวนการแปลภาษาประสิทธิภาพที่ได้กลับลดลง จึงส่งผลให้ความเข้าใจในการสังเคราะห์เสียงต่ำลง นอกจากนี้ยังพบว่าสภาพแวดล้อมที่ใช้งาน, ตำแหน่งการพูดและปฏิสัมพันธ์การใช้งานระหว่างระบบกับผู้ใช้มีผลต่อประสิทธิภาพโดยรวมของทั้งระบบเช่นกัน

จากการศึกษาดังกล่าวผู้จัดทำคาดหวังว่าจะสามารถเป็นประโยชน์สำหรับบุคคลที่สนใจศึกษาแอปพลิเคชันที่เกี่ยวข้องกับ Speech-to-Text, Translation Language และ Text-to-Speech เนื่องจากผู้จัดทำได้ศึกษาเปรียบเทียบประสิทธิภาพและคุณสมบัติของ API ในแต่ละกระบวนการที่กล่าวมา ทำให้ผู้ที่สนใจสามารถศึกษาและเลือกใช้ API ที่เหมาะสำหรับแอปพลิเคชันที่พัฒนาได้เพื่อลดขั้นตอนในการศึกษาพัฒนา รวมทั้งช่วยให้แอปพลิเคชันที่พัฒนามีประสิทธิภาพยิ่งขึ้นได้

บรรณานุกรม

- [1] “สถาปัตยกรรมของแอนดรอยด์”[Online] Available
<http://kadroidz.blogspot.com/2012/03/android-architecture.html> 2012.
- [2] “**ANDROID TUTORIAL**”[Online] Available :
http://www.tutorialspoint.com/android/android_tutorial.pdf
- [3] “การรู้จำเสียงพูด”[Online] Available : <http://www.amivoicethai.com/ami-engine/#basic>
- [4] “การแปลคืออะไร”[Online] Available : <http://www.l3nr.org/posts/2025>
- [5] “**Spelling Checker Framework.**”[Online] Available :
<http://developer.android.com/guide/topics/text/spell-checker-framework.html>
- [6] “พจนานุกรม”[Online] Available : <http://th.wikipedia.org/wiki/พจนานุกรม>
- [7] “**Dictionary API.**”[Online] Available : <http://www.dictionaryapi.com/>
- [8] “**The Oxford Dictionaries API.**”[Online] Available :
<http://developer.oxforddictionaries.com/>
- [9] “**The Cambridge Dictionaries.**”[Online] Available : <http://dictionary-api.cambridge.org/>
- [10] Tomalin M., Gales M.J.F., Liu X.A., Sim K.C., Sinha R., Wang L., Woodland P.C. และ Yu K., “**Improving Speech Transcription for Mandarin-English Translation.**”, IEEE International Conference on Acoustics, Speech and Signal Processing, pp.IV-97-IV-100, April 2007
- [11] Abushariah A.A.M., Gunawan T.S., Khalifa O.O. และ Abushariah M.A.M., “**English digits speech recognition system based on Hidden Markov Models.**”, International Conference on Computer and Communication Engineering, pp.1-5, May 2010
- [12] NtsakoBaloyi. “**A Text-to-speech Synthesis System for Xitsonga Using Hidden Markov Models.**”, University of Limpopo, 2012
- [13] De Silva D., Alahakoon A., Udayangani I., Kumara V., Kolonnage, D., Perera H. และ ThelijjagodaS., “**Sinhala to English Language Translator.**”, 4th International Conference on Information and Automation for Sustainability, pp.419-424, December 2008

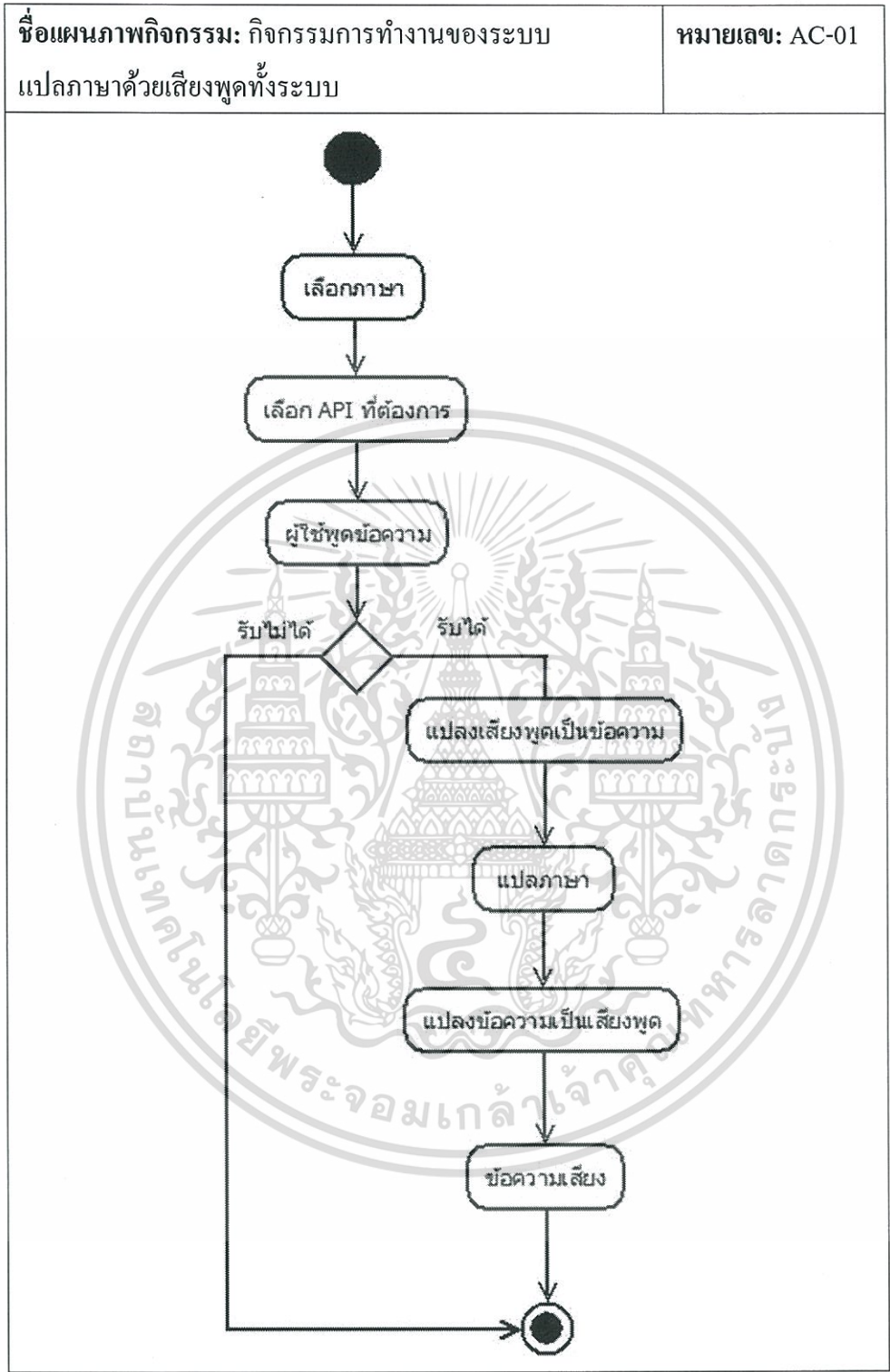
บรรณานุกรม (ต่อ)

- [14] Abushariah AA M, Gunawan, T.S., Khalifa, O.O. และ Abushariah, M. A M, **“English digits speech recognition system based on Hidden Markov Models.”**, International Conference on Computer and Communication Engineering, pp.1-5, May 2010
- [15] Hashimoto K., Yamagishi J., Byrne W., King S. and Tokuda K., **“An analysis of machine translation and speech synthesis in speech-to-speech translation system.”**, IEEE International Conference on Acoustics, Speech and Signal Processing, pp.5108-5111, May 2011
- [16] Seung Yun, Young-Jik Lee and Sang-Hun Kim, **“Multilingual speech-to-speech translation system for mobile consumer devices.”**, IEEE Transactions on Consumer Electronics, pp.508-516, August 2014
- [17] Z. Orhan, Z. Gormez, **“A concatenative Turkish text-to-speech system and evaluation process,”** International Conference on Electrical and Electronics Engineering, 2009, pp.II-106-II-110.



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.1 แสดงแผนภาพกิจกรรมการทำงานของระบบแปลภาษาด้วยเสียงพูดทั้งระบบ



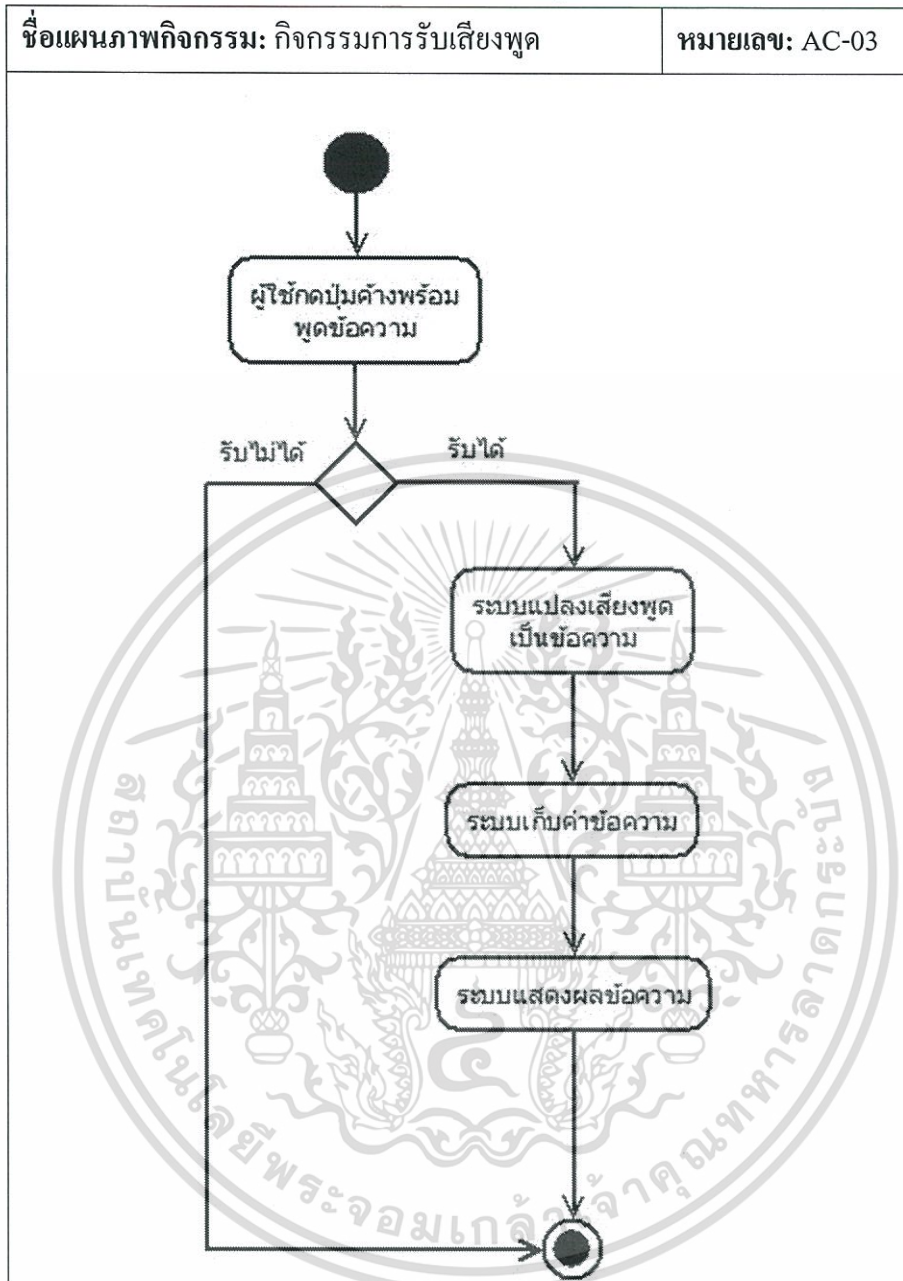
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.2 แสดงแผนภาพกิจกรรมการเลือกภาษา



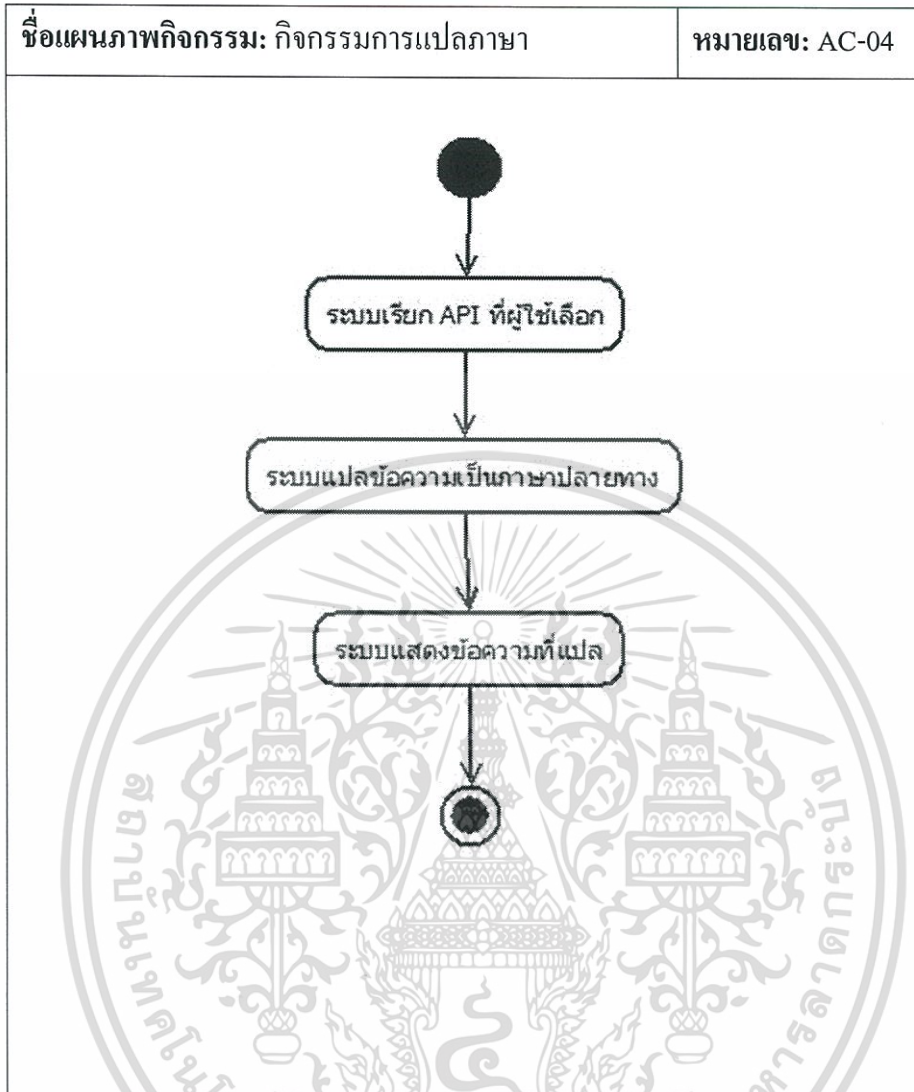
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.3 แสดงแผนภาพกิจกรรมการรับเสียงพูด



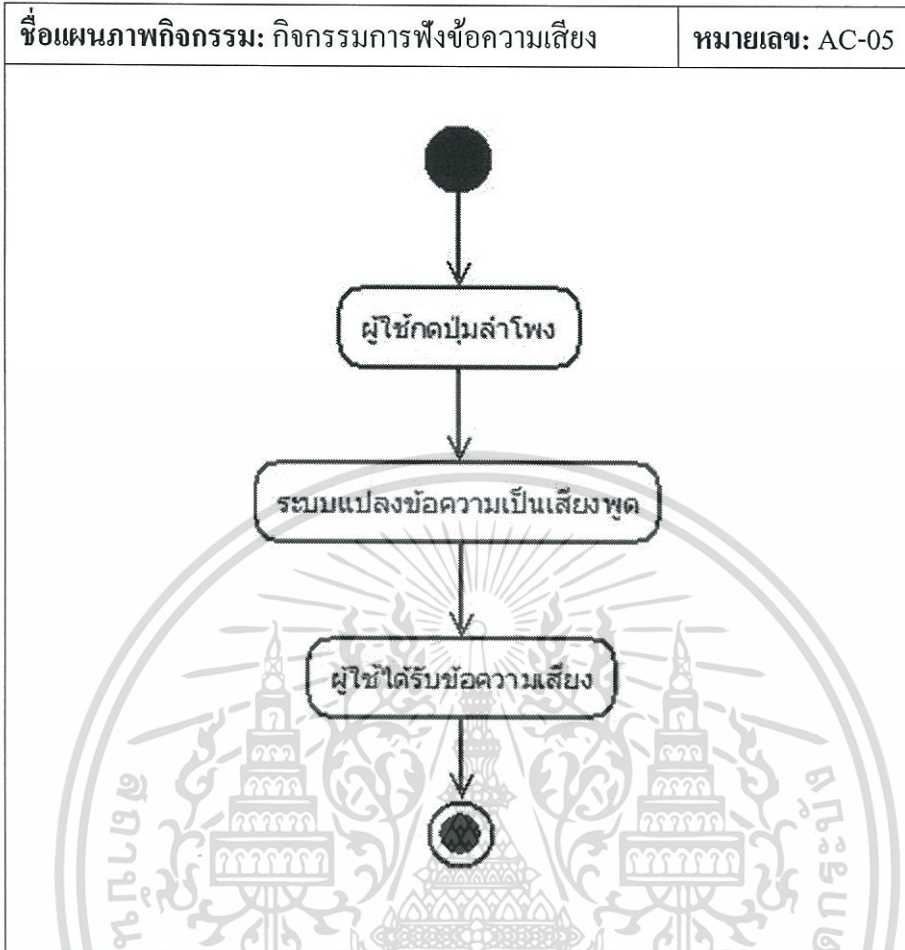
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.4 แสดงแผนภาพกิจกรรมการแปลภาษา



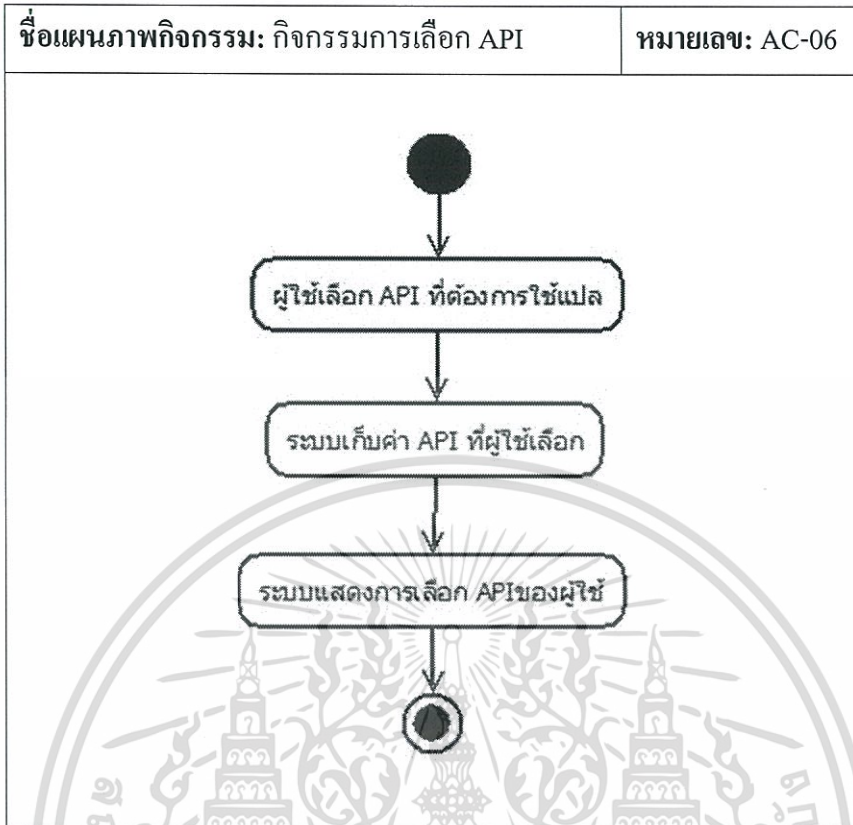
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.5 แสดงแผนภาพกิจกรรมการฟังข้อความเสียง



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.6 แสดงแผนภาพกิจกรรมการเลือก API



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาคผนวก ข
ตารางแบบประเมินแอปพลิเคชันแปลภาษาด้วยเสียงพูดแบบหลายภาษา
บนระบบปฏิบัติการแอนดรอยด์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.1 แสดงตัวอย่างแบบประเมินการรับเสียงภาษาอังกฤษแปลเป็นภาษาไทย

ประโยค	กระบวนการแปลภาษาด้วยเสียงพูด		
	Speech -to-Text	Translation	Text -to-Speech
1. Good morning. (สวัสดี)			
2. How are you? (คุณสบายดีไหม)			
3. How tall are you? (คุณสูงเท่าไร)			
4. What's your name? (คุณชื่ออะไร)			
5. How old are you? (คุณอายุเท่าไร)			
6. I like to read. (ฉันชอบอ่านหนังสือ)			
7. Open the door, please. (กรุณาเปิดประตูด้วย)			
8. Are you angry at me? (คุณโกรธฉันอยู่หรือเปล่า)			
9. What are you doing? (คุณทำอะไรอยู่)			
10. May I help you? (ให้ฉันช่วยไหม)			
11. How much is it? (ราคาเท่าไร)			
12. Can I pay by credit card? (จ่ายด้วยบัตรเครดิตได้ไหม)			
13. What do you want? (คุณต้องการอะไร)			
14. I'm going to play football. (ฉันจะเล่นฟุตบอล)			
15. I speak a little English. (ฉันพูดภาษาอังกฤษได้นิดหน่อย)			
16. Where are we going? (เราจะไปไหนกัน)			
17. Do you like to travel? (คุณชอบเที่ยวไหม)			
18. It's a nice day. (วันนี้อากาศดีจังนะ)			
19. Could you take a photo for me? (คุณช่วยถ่ายรูปให้ฉันหน่อยได้ไหม)			
20. I going to the airport. (ฉันจะไปสนามบิน)			
21. See you again. (แล้วพบกันใหม่)			
22. Will you wake me up at 6 o'clock? (ปลุกฉันตอน 6 โมงได้ไหม)			

ตารางที่ ข.1 (ต่อ) แสดงตัวอย่างแบบประเมินการรับเสียงภาษาอังกฤษแปลเป็นภาษาไทย

ประโยค	กระบวนการแปลภาษาด้วยเสียงพูด		
	Speech -to-Text	Translation	Text -to-Speech
23. Where do you work? (คุณทำงานที่ไหน)			
24. Can you sing? (คุณร้องเพลงเป็นไหม)			
25. What time is it? (กี่โมงแล้ว)			
26. What have you been up to? (คุณทำอะไรบ้าง ช่วงนี้)			
27. Did you buy anything? (คุณได้ซื้ออะไร ใหม่)			
28. How long will you be staying? (คุณอยู่นาน แค่ไหน)			
29. Do you need a taxi? (คุณต้องการแท็กซี่ ไหม)			
30. I'd like to go see a movie. (ฉันอยากไป ดู หนัง)			

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.2 แสดงตัวอย่างแบบประเมินการรับเสียงภาษาไทยแปลเป็นภาษาอังกฤษ

ประโยค	กระบวนการแปลภาษาด้วยเสียงพูด		
	Speech -to-Text	Translation	Text -to-Speech
1. สวัสดี (Good morning.)			
2. คุณสบายดีไหม (How are you?)			
3. คุณสูงเท่าไร (How tall are you?)			
4. คุณชื่ออะไร (What's your name?)			
5. คุณอายุเท่าไร (How old are you?)			
6. ฉันชอบอ่านหนังสือ (I like to read.)			
7. กรุณาเปิดประตูด้วย (Open the door, please.)			
8. คุณโกรธฉันหรือเปล่า (Are you angry at me?)			
9. คุณทำอะไรอยู่ (What are you doing?)			
10. ให้ฉันช่วยไหม (May I help you?)			
11. ราคาเท่าไร (How much is it?)			
12. จ่ายด้วยบัตรเครดิตได้ไหม (Can I pay by credit card?)			
13. คุณต้องการอะไร (What do you want?)			
14. ฉันจะเล่นฟุตบอล (I'm going to play football.)			
15. ฉันพูดภาษาอังกฤษได้นิดหน่อย (I speak a little English.)			
16. เราจะไปไหนกัน (Where are we going?)			
17. คุณชอบท่องเที่ยวไหม (Do you like to travel?)			
18. วันนี้อากาศดีจังนะ (It's a nice day.)			
19. คุณช่วยถ่ายรูปให้ฉันหน่อยได้ไหม (Could you take a photo for me?)			
20. ฉันจะไปลงสนามบิน (I going to the airport.)			

ตารางที่ ข.2 (ต่อ) แสดงตัวอย่างแบบประเมินการรับเสียงภาษาไทยแปลเป็นภาษาอังกฤษ

ประโยค	กระบวนการแปลภาษาด้วยเสียงพูด		
	Speech -to-Text	Translation	Text -to-Speech
21. แล้วพบกันใหม่ (See you again.)			
22. ปลุกฉันตอน 6 โมงได้ไหม (Will you wake me up at 6 o'clock?)			
23. คุณทำงานที่ไหน (Where do you work?)			
24. คุณร้องเพลงเป็นไหม (Can you sing?)			
25. กี่โมงแล้ว (What time is it?)			
26. คุณทำอะไรบ้างช่วงนี้ (What have you been up to?)			
27. คุณได้ซื้ออะไรไหม (Did you buy anything?)			
28. คุณอยู่นานแค่ไหน (How long will you be staying?)			
29. คุณต้องการแท็กซี่หรือไม่ (Do you need a taxi?)			
30. ฉันอยากไปดูหนัง (I'd like to go see a movie.)			

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

**แบบสอบถามเพื่อประเมินความพึงพอใจของผู้ใช้แอปพลิเคชันแปลภาษาด้วยเสียงพูดแบบ
หลายภาษาบนระบบปฏิบัติการแอนดรอยด์**

I. โปรดทำเครื่องหมาย ลงในช่อง และกรอกข้อมูลที่ตรงกับความเป็นจริง

- เพศ

ชาย หญิง

- เคยใช้บริการแอปพลิเคชันระบบในการแปลภาษาหรือไม่

เคย ไม่เคย

II. โปรดทำเครื่องหมาย ลงในช่องในแบบสอบถามที่ตรงกับระดับความคิดเห็นของท่านมากที่สุด โดยตัวเลขของระดับความพึงพอใจแต่ละด้านมีความหมายดังนี้

5 หมายถึง ความพึงพอใจในระดับมากที่สุด

4 หมายถึง ความพึงพอใจในระดับมาก

3 หมายถึง ความพึงพอใจในระดับปานกลาง

2 หมายถึง ความพึงพอใจในระดับน้อย

1 หมายถึง ความพึงพอใจในระดับน้อยที่สุด

ตารางที่ ข.3 แบบสอบถามเพื่อประเมินความพึงพอใจของผู้ใช้แอปพลิเคชัน

รายการประเมิน	ระดับความพึงพอใจ				
	5	4	3	2	1
- ความสามารถของแอปพลิเคชันในด้านการรับเสียง					
- ความสามารถของแอปพลิเคชันในด้านการออกเสียง					
- ความสามารถของแอปพลิเคชันในด้านการแปลภาษา					
- ความรวดเร็วในด้านการประมวลผล					
- ความง่ายต่อการใช้งานของแอปพลิเคชัน					
- ความเหมาะสมในการเลือกใช้นาฬิกาของตัวอักษรในแอปพลิเคชัน					
- ความเหมาะสมของการออกแบบแอปพลิเคชันสำหรับแปลภาษา					
- ความเหมาะสมในการปฏิสัมพันธ์ได้ตอบกับผู้ใช้					
- ประโยชน์ที่ได้จากแอปพลิเคชัน					
- ประสิทธิภาพโดยรวมของแอปพลิเคชัน					

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

ชื่อ นางสาวญาดา อัมภรัตน์
วัน เดือน ปี เกิด 2 สิงหาคม 2535
ที่อยู่ เลขที่ 38 หมู่ 1 ตำบล ดูน อำเภอ กันทรารมณ จังหวัด ศรีสะเกษ 33130
ประวัติการศึกษา วิทยาศาสตร์บัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ชื่อ นางสาวพัฒนชีดา ชมวิหก
วัน เดือน ปี เกิด 11 ธันวาคม 2535
ที่อยู่ เลขที่ 81/15 ตำบล บางกระสอ อำเภอ เมือง จังหวัด นนทบุรี 11000
ประวัติการศึกษา วิทยาศาสตร์บัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้