การแปลงกราฟีมภาษาอังกฤษเป็นโฟนีมภาษาไทย

นายกฤตนันท์ พิทักษ์ภวัตกุล

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมคอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2555

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของวิทยานิพนธ์ตั้งแต่ปีการศึกษา 2554 ที่ให้บริการในคลังปัญญาจุฬาฯ (CUIR)
เป็นแฟ้มข้อมูลของนิสิตเจ้าของวิทยานิพนธ์ที่ส่งผ่านทางบัณฑิตวิทยาลัย
The abstract and full text of theses from the academic year 2011 in Chulalongkorn University Intellectual Repository(CUIR)
are the thesis authors' files submitted through the Graduate School.

ENGLISH GRAPHEME TO THAI PHONEME CONVERSION

Mr. Kitanan Pitakpawatkul

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Engineering Program in Computer Engineering

Department of Computer Engineering

Faculty of Engineering

Chulalongkorn University

Academic Year 2012

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การแปลงกราฟีมภาษาอังกฤษเป็นโฟนีมภาษาไทย
โดย	นายกฤตนันท์ พิทักษ์ภวัตกุล
สาขาวิชา	วิศวกรรมคอมพิวเตอร์
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	รองศาสตราจารย์ ดร.อติวงศ์ สุชาโต
อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม	ผู้ช่วยศาสตราจารย์ ดร.โปรดปราน บุณยพุกกณะ ดร.ชัย วุฒิวิวัฒน์ชัย

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

…….. คณบดีคณะวิศวกรรมศาสตร์
(รองศาสตราจารย์ ดร.บุญสม เลิศหิรัญวงศ์)

คณะกรรมการสอบวิทยานิพนธ์

……. ประธานกรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.เศรษฐา ปานงาม)

…….. อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(รองศาสตราจารย์ ดร.อติวงศ์ สุชาโต)

……. อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม
(ผู้ช่วยศาสตราจารย์ ดร.โปรดปราน บุณยพุกกณะ)

……. อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม
(ดร.ชัย วุฒิวิวัฒน์ชัย)

……. กรรมการภายนอกมหาวิทยาลัย
(ดร.ชัชวาลย์ หาญสกุลบรรเทิง)

กฤตนันท์ พิทักษ์ภวัตกุล : การแปลงกราฟีมภาษาอังกฤษเป็นโฟนีมภาษาไทย. (ENGLISH GRAPHEME TO THAI PHONEME CONVERSION) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: ร.ศ. ดร.อติวงศ์ สุชาโต, อ. ที่ปรึกษาวิทยานิพนธ์ร่วม: ผ.ศ. ดร.โปรดปราน บุณยพุกกณะ, ด.ร.ชัย วุฒิวิวัฒน์ชัย, 61 หน้า.

ความเป็นธรรมชาติของเสียงสังเคราะห์เป็นเป้าหมายสำคัญของระบบสังเคราะห์เสียง (Text-to-Speech System) ซึ่งกระบวนการสำคัญก็คือการหาคำอ่านของคำที่จะสังเคราะห์เสียง แต่ในระบบสังเคราะห์เสียงภาษาไทยการหาคำอ่านของคำภาษาอังกฤษไม่สามารถหาได้แบบตรงไปตรงมา ดังนั้นเพื่อให้ระบบสังเคราะห์เสียงภาษาไทยสามารถสังเคราะห์เสียงคำภาษาอังกฤษที่อาจจะปรากฏร่วมกับคำภาษาไทยได้นั้น จึงเสนอระบบการหาคำอ่านในหน่วยเสียงภาษาไทยจากคำภาษาอังกฤษ ซึ่งประกอบไปด้วย 3 กระบวนการหลัก คือ โมเดลการทำนายหน่วยเสียง, กระบวนการประมวลผลหลังจากการทำนายหน่วยเสียง และโมเดลการทำนายเสียงวรรณยุกต์ โดยโมเดลการทำนายหน่วยเสียงใช้อัลกอริทึมออนไลน์ดิสคริมิเนทีฟเทรนนิงเฟรมเวิร์ค (Online Discriminative Training Framework) ในกระบวนการเรียนรู้ ร่วมกับค่าลักษณะสำคัญต่าง ๆ ดังนี้ บริบท (Context) , เอ็นแกรม (n-gram), มาร์คอฟออร์เดอร์ (Markov Order), ลิเนียร์เชน (Linear-Chain) และจอยต์เอ็นแกรม (Joint n-gram) ในขณะที่กระบวนการประมวลผลหลังจากการทำนายหน่วยเสียงจะทำการแก้ไขหน่วยเสียงที่ผิดหลักพยางค์ในภาษาไทย และโมเดลการทำนายเสียงวรรณยุกต์อัลกอริทึมต้นไม้ (Classification and Regression Tree: CART) ในกระบวนการเรียนรู้ร่วมกับค่าลักษณะสำคัญต่าง ๆ ที่เกี่ยวข้องกับหลักภาษาศาสตร์ ผลลัพธ์ความแม่นยำเฉลี่ยในระดับพยางค์เท่ากับ 76.03%, ความแม่นยำเฉลี่ยในระดับคำเท่ากับ 53.93% และผลลัพธ์ในแง่ความพึงพอใจพบว่า 90.95% เป็นผลลัพธ์ที่ยอมรับได้

ภาควิชา วิศวกรรมคอมพิวเตอร์ ลายมือชื่อ.....
สาขาวิชา วิศวกรรมคอมพิวเตอร์ ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก.....
ปีการศึกษา 2555 ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์ร่วม.....
ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์ร่วม.....

5570113621 : MAJOR COMPUTER ENGINEERING

KEYWORDS : English to Thai G2P / text-to-speech / Thai phonetization of English words / English syllable / Thai transcription of English word

KITANAN PITAKPAWATKUL :ENGLISH GRAPHEME TO THAI PHONEME CONVERSION. ADVISOR : ASSOC.PROF. ATIWONG SUCHATO, Ph.D., CO-ADVISOR : ASST.PROF. PROADPRAN PUNYABUKKANA, Ph.D., CHAI WUTIWIWATCHAI, Ph.D., 61 pp.

The naturalness of synthesized sound is considered as a primary goal in text-to-speech (TTS) system. The important process, which makes the system reach the purpose, is grapheme-to-phoneme (G2P) conversion. However, in Thai TTS system, transcribing English texts to Thai phones is not a trivial task. This study, therefore, proposes the method of finding Thai transcriptions of English words in order to deal with English words existing among Thai words in modern written language. The system consists of three main components: phone prediction model, post-processing, and tone prediction model. Phone prediction model is constructed by applying the online discriminative training framework as its learning algorithm. The features used in phone prediction model are context, n-gram, markov order, linear-chain and joint n-gram. Post-processing is created for correcting some phone errors which break Thai syllable structure rules. Tone prediction model is built by using classification and regression tree (CART) algorithm with linguistic features. Results show that the average syllable and word accuracy are 76.03% and 53.93% respectively. Moreover, the result of subjective test indicates that 90.95% is acceptable.

Department : Computer Engineering
Field of Study : Computer Engineering
Academic Year : 2012

Student's Signature .....
Advisor's Signature .....
Co-advisor's Signature .....
Co-advisor's Signature.....

กิตติกรรมประกาศ

ข้าพเจ้าขอขอบพระคุณ อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก รองศาสตราจารย์ ดร.อติวงศ์ สุชาโต อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม ผู้ช่วยศาสตราจารย์ ดร.โปรดปราน บุณยพุกกณะ และ ดร.ชัย วุฒิวิวัฒน์ชัย ที่ได้สละเวลาอันมีค่าเพื่อให้ความรู้และคำปรึกษารวมทั้งคำแนะนำต่างๆ ในระหว่างระยะเวลาที่ศึกษานี้ ข้าพเจ้าขอขอบคุณ รองศาสตราจารย์ ดร.วิโรจน์ อรุณมานะกุล, ดร.ชัชวาลย์ หาญสกุลบรรเทิง และ ผู้ช่วยศาสตราจารย์ ดร.เศรษฐา ปานงาม ที่ให้คำแนะนำและแนวคิดในการทำวิทยานิพนธ์นี้ รวมทั้งหน่วยปฏิบัติการวิจัยวิทยาการมนุษยภาษา ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ: NECTEC ที่ให้ข้อมูลสำหรับการทดลองในวิทยานิพนธ์นี้ และภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ที่สนับสนุนทุนการศึกษาตลอดระยะเวลาที่ศึกษา

ข้าพเจ้าขอขอบคุณ คุณสิทธิชัย เจียมพจมาน ที่ได้ให้คำปรึกษาและคำแนะนำที่มีประโยชน์อย่างมากรวมทั้งให้เครื่องมือที่ใช้ในการทดลอง ข้าพเจ้าขอขอบคุณ ครอบครัวและบุคลากรทุกคนของห้องปฏิบัติการระบบภาษาพูด ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ที่ให้ข้อแนะนำและกำลังใจกับข้าพเจ้า ซึ่งส่งผลให้การทำวิทยานิพนธ์ประสบความสำเร็จอย่างดี

สารบัญ

หน้า

บทคัดย่อภาษาไทย..... ง

บทคัดย่อภาษาอังกฤษ..... จ

กิตติกรรมประกาศ..... ฉ

สารบัญ..... ช

สารบัญตาราง..... ฌ

สารบัญภาพ..... ญ

บทที่ 1 บทนำ..... 1

1.1 ความเป็นมาและความสำคัญของปัญหา..... 1

1.2 วัตถุประสงค์ของการวิจัย..... 2

1.3 ขอบเขตของการวิจัย..... 2

1.4 ลำดับขั้นตอนในการเสนอผลการวิจัย..... 2

1.5 ประโยชน์ที่คาดว่าจะได้รับ..... 3

1.6 ผลงานตีพิมพ์จากวิทยานิพนธ์..... 5

1.7 ลำดับการจัดเรียงเนื้อหาในวิทยานิพนธ์..... 5

บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง..... 6

2.1 ทฤษฎีที่เกี่ยวข้อง..... 6

2.1.1. ทฤษฎีเกี่ยวกับภาษาไทย..... 6

2.1.2. หน่วยเสียงสากล, หน่วยเสียงภาษาไทยและหน่วยเสียงภาษาอังกฤษ..... 9

2.2 เอกสารและงานวิจัยที่เกี่ยวข้อง..... 14

2.2.1. งานวิจัยเรื่องการหาคำอ่านในหน่วยเสียงภาษาไทยจากคำภาษาอังกฤษ… 14

2.2.2. งานวิจัยที่เกี่ยวกับระบบสังเคราะห์เสียงสองภาษา..... 15

2.2.3. งานวิจัยที่ใกล้เคียงโดยเกี่ยวกับการหาคำอ่านจากคำในภาษาใดภาษาหนึ่ง 17

2.2.4. งานวิจัยที่เกี่ยวข้องอื่น ๆ..... 19

บทที่ 3 ขั้นตอนการดำเนินงานวิจัย..... 20

3.1 ภาพรวมระบบ..... 20

หน้า

3.2 การปรับแนวและโมเดลทำนายหน่วยเสียง..... 21
3.3 กระบวนการประมวลผลหลังจากการทำนายหน่วยเสียง..... 26
3.4 โมเดลการทำนายเสียงวรรณยุกต์..... 29

บทที่ 4 การทดลอง ผลการทดลองและการวิเคราะห์ผลการทดลอง..... 35
4.1 การทดลอง..... 35
4.1.1. ข้อมูลที่ใช้ในการทดลอง..... 35
4.1.2. เครื่องมือที่ใช้ในการทดลอง..... 35
4.1.3. การประเมินผล..... 37
4.2 ผลการทดลอง..... 38
4.2.1. ผลการทดลองของโมเดลการทำนายหน่วยเสียง..... 38
4.2.2. ผลการทดลองของโมเดลการทำนายเสียงวรรณยุกต์..... 39
4.2.3. ผลการทดลองรวม..... 40
4.2.4. ผลประเมินความพึงพอใจ..... 42
4.3 การวิเคราะห์ผลการทดลอง..... 42
4.3.1. การวิเคราะห์ผลลัพธ์จากโมเดลการทำนายหน่วยเสียง..... 42
4.3.2. การวิเคราะห์ผลลัพธ์จากโมเดลการทำนายเสียงวรรณยุกต์..... 44
4.3.3. การวิเคราะห์ผลลัพธ์รวม..... 45
4.3.4. การวิเคราะห์ผลลัพธ์เทียบกับระบบอ้างอิง..... 48

บทที่ 5 บทสรุปผลการวิจัย และข้อเสนอแนะ..... 51
5.1 สรุปผลการวิจัย..... 51
5.2 ข้อเสนอแนะ..... 52

รายการอ้างอิง..... 53
ภาคผนวก..... 56
ประวัติผู้เขียนวิทยานิพนธ์..... 61

สารบัญตาราง

ตารางที่		หน้า
2.1	ตัวอักษรพยัญชนะไทยเทียบกับหน่วยเสียงไทย.....	7
2.2	การผันเสียงวรรณยุกต์.....	9
2.3	หน่วยเสียงในภาษาไทย.....	11
2.4	หน่วยเสียงภาษาอังกฤษเปรียบเทียบกับหน่วยเสียงสากลและหน่วยเสียงภาษาไทย.....	12
3.1	รูปแบบของค่าลักษณะสำคัญสำหรับโมเดลทำนายหน่วยเสียง.....	25
3.2	ความผิดพลาดที่เกิดขึ้นจากโมเดลทำนายหน่วยเสียง.....	29
3.3	จำนวนพยางค์ของหน่วยเสียงที่แบ่งตามอักขระพยางค์ภาษาอังกฤษและเสียงวรรณยุกต์.....	32
3.4	จำนวนพยางค์ที่เกิดขึ้นจากความสัมพันธ์ระหว่างอักขระพยางค์ถัดไป คำเป็น คำตาย และเสียงวรรณยุกต์.....	33
3.5	ค่าลักษณะสำคัญที่ใช้สำหรับโมเดลการทำนายเสียงวรรณยุกต์.....	34
4.1	ผลลัพธ์ของโมเดลการทำนายหน่วยเสียงก่อนกระบวนการประมวลผลหลังจากการทำนายหน่วยเสียง.....	38
4.2	ผลลัพธ์ของโมเดลการทำนายหน่วยเสียงเมื่อผ่านกระบวนการประมวลผลหลังจากการทำนายหน่วยเสียง.....	39
4.3	ผลลัพธ์ของโมเดลการทำนายเสียงวรรณยุกต์เปรียบเทียบกับระบบอ้างอิง.....	40
4.4	ผลลัพธ์รวมที่ไม่ผ่านกระบวนการประมวลผลหลังจากการทำนายหน่วยเสียง….	41
4.5	ผลลัพธ์รวมที่ผ่านกระบวนการประมวลผลหลังจากการทำนายหน่วยเสียง…….	41
4.6	ผลลัพธ์จากการประเมินความพึงพอใจต่อคำอ่าน.....	42
4.7	ความผิดพลาดแบบเกินแจกแจงตามจำนวนพยางค์ที่เกิน.....	47
4.8	ความผิดพลาดแบบขาดแจกแจงตามจำนวนพยางค์ที่ขาด.....	47

สารบัญรูป

รูปที่		หน้า
1.1	แผนการดำเนินงานวิจัย.....	4
2.1	รายละเอียดหน่วยเสียงสากล.....	10
3.1	ภาพรวมการทำงานของระบบ.....	20
3.2	ตัวอย่างการปรับแนวแบบกลุ่มต่อกลุ่ม.....	22
3.3	กระบวนการในโมเดลการทำนายหน่วยเสียง.....	23
4.1	สัดส่วนของคำที่ผิดพลาดแบบแทนที่ แบ่งตามจำนวนพยางค์ที่ผิดและความยาวของคำ.....	47

บทที่ 1
บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ระบบสังเคราะห์เสียงเป็นการสร้างเสียงสังเคราะห์จากบทความต่าง ๆ ซึ่งมีความสำคัญและนำไปใช้ประโยชน์ในหลายรูปแบบ เช่น การใช้งานเพื่ออำนวยความสะดวกสำหรับผู้พิการทางสายตานั้น เป็นรูปแบบหนึ่งที่เห็นได้ชัดเจน เช่น การนำระบบสังเคราะห์ไปใช้ในการสร้างเครื่องอ่านหน้าจอ [1] ซึ่งจะสามารถอ่านออกเสียงบรรยายข้อความที่อยู่บนหน้าจอคอมพิวเตอร์ให้กับผู้พิการทางสายตาได้ฟังได้ แต่ในกระบวนสังเคราะห์เสียงนั้นประกอบไปด้วยกระบวนย่อย ๆ หลายส่วน ซึ่งกระบวนการหนึ่งที่มีสำคัญก็คือ การหาคำอ่านของคำที่ต้องการสังเคราะห์เสียง โดยทั่วไปแล้วระบบสังเคราะห์เสียงก็จะออกแบบเพื่อรองรับกับคำในภาษาใดภาษาหนึ่งเท่านั้น เช่น ระบบสังเคราะห์เสียงภาษาไทย [2] ซึ่งออกแบบเพื่อรองรับกับคำภาษาไทยเท่านั้น แต่หากพิจารณาบทความในภาษาไทยสมัยใหม่ ก็จะพบว่าบ่อยครั้งที่มักจะมีคำภาษาอังกฤษปรากฏร่วมกับคำภาษาไทย การหาคำอ่านของคำภาษาอังกฤษเหล่านั้นก็ไม่ใช่เรื่องง่าย จึงเกิดปัญหาขึ้นกับกระบวนการสังเคราะห์เสียงที่ไม่สามารถสังเคราะห์เสียงของคำเหล่านั้นได้ การแก้ปัญหาแบบตรงไปตรงมาอาจทำได้โดย การใช้พจนานุกรมคำอ่านที่เตรียมไว้สำหรับคำภาษาอังกฤษที่มีโอกาสจะพบในบทความ แต่ในความจริงอาจใช้งานไม่ได้ประสิทธิภาพเพราะไม่สามารถจัดการกับคำนอกพจนานุกรมได้ การใช้ระบบสังเคราะห์เสียงสำหรับภาษาอังกฤษเพื่อมารองรับกับคำภาษาอังกฤษโดยเฉพาะก็เป็นอีกวิธีหนึ่งที่สามารถแก้ปัญหาได้อย่างตรงไปตรงมา [3] แต่ก็จะพบกับปัญหาของความไม่เป็นธรรมชาติของเสียงสังเคราะห์ ซึ่งเกิดจากความแตกต่างกันระหว่างหน่วยเสียงของสองภาษาซึ่งมีความต่างกัน รวมทั้งสำเนียงและการเน้นพยางค์ที่แตกต่างกันอย่างชัดเจนด้วย ซึ่งความเป็นธรรมชาติของเสียงสังเคราะห์นั้นถือเป็นเป้าหมายหลักของการสังเคราะห์เสียง

ดังนั้นผู้วิจัยจึงมีแนวคิดในการสร้างระบบการหาคำอ่านในหน่วยเสียงภาษาไทยจากคำภาษาอังกฤษ (English to Thai Grapheme-to-Phoneme: English to Thai G2P) เพื่อแก้ปัญหานี้ ซึ่งระบบจะสร้างคำอ่านในหน่วยเสียงภาษาไทยจากคำภาษาอังกฤษ รวมทั้งกำหนดเสียงวรรณยุกต์ด้วย เพื่อทำให้ระบบการสังเคราะห์เสียงภาษาไทยสามารถสร้างเสียงสังเคราะห์ของอักขระหรือคำภาษาอังกฤษในหน่วยเสียงไทยได้อย่างเป็นธรรมชาติมากขึ้น

1.2 วัตถุประสงค์ของการวิจัย

เพื่อเสนอวิธีที่เหมาะสมในการสร้างคำอ่านที่เป็นหน่วยเสียงภาษาไทยจากคำภาษาอังกฤษ

1.3 ขอบเขตของการวิจัย

1. คำที่พิจารณาในระบบจะเป็นคำภาษาอังกฤษเท่านั้น
2. ระบบไม่รองรับคำภาษาอังกฤษที่คำอ่านไม่มีความสัมพันธ์กับรูปอักขระ เช่น คำย่อ
3. ประสิทธิภาพในการประมวลผลจะไม่นับรวมถึงประสิทธิภาพในเชิงเวลา
4. งานวิจัยนี้เสนอเฉพาะวิธีการหาคำอ่านในหน่วยเสียงไทยจากคำภาษาอังกฤษ

1.4 ลำดับขั้นตอนในการเสนอผลการวิจัย

1. ขั้นตอนการศึกษาเบื้องต้น
    1.1. ศึกษางานวิจัยเกี่ยวกับการหาคำอ่านในหน่วยเสียงไทยจากคำภาษาอังกฤษ
    1.2. ศึกษางานวิจัยเกี่ยวกับการหาคำอ่านในหน่วยเสียงภาษาอังกฤษจากคำภาษาอังกฤษ
    1.3. ศึกษางานวิจัยเกี่ยวกับการหาคำอ่านในหน่วยเสียงภาษาไทยจากคำภาษาไทย
    1.4. ศึกษาทฤษฎีเกี่ยวกับหลักภาษาไทย
    1.5. ศึกษาหาค่าลักษณะสำคัญที่มีผลต่อการหาคำอ่าน
    1.6. ศึกษาหาค่าลักษณะสำคัญที่มีผลต่อการผันวรรณยุกต์ในภาษาไทย
    1.7. ศึกษาขั้นตอนวิธีที่เหมาะสมสำหรับการสร้างระบบการหาคำอ่าน
2. ขั้นตอนการออกแบบระบบ
    2.1. ออกแบบการทำงานของระบบ
    2.2. เลือกใช้เครื่องมือในการวิจัยที่เหมาะสม
    2.3. ออกแบบการทดลอง
3. ขั้นตอนการทดลอง
    3.1. ทดสอบการทำงานของระบบ
    3.2. บันทึกผลการทดลอง
    3.3. วิเคราะห์ผลการทดลอง
4. สรุปผลและเรียบเรียงวิทยานิพนธ์

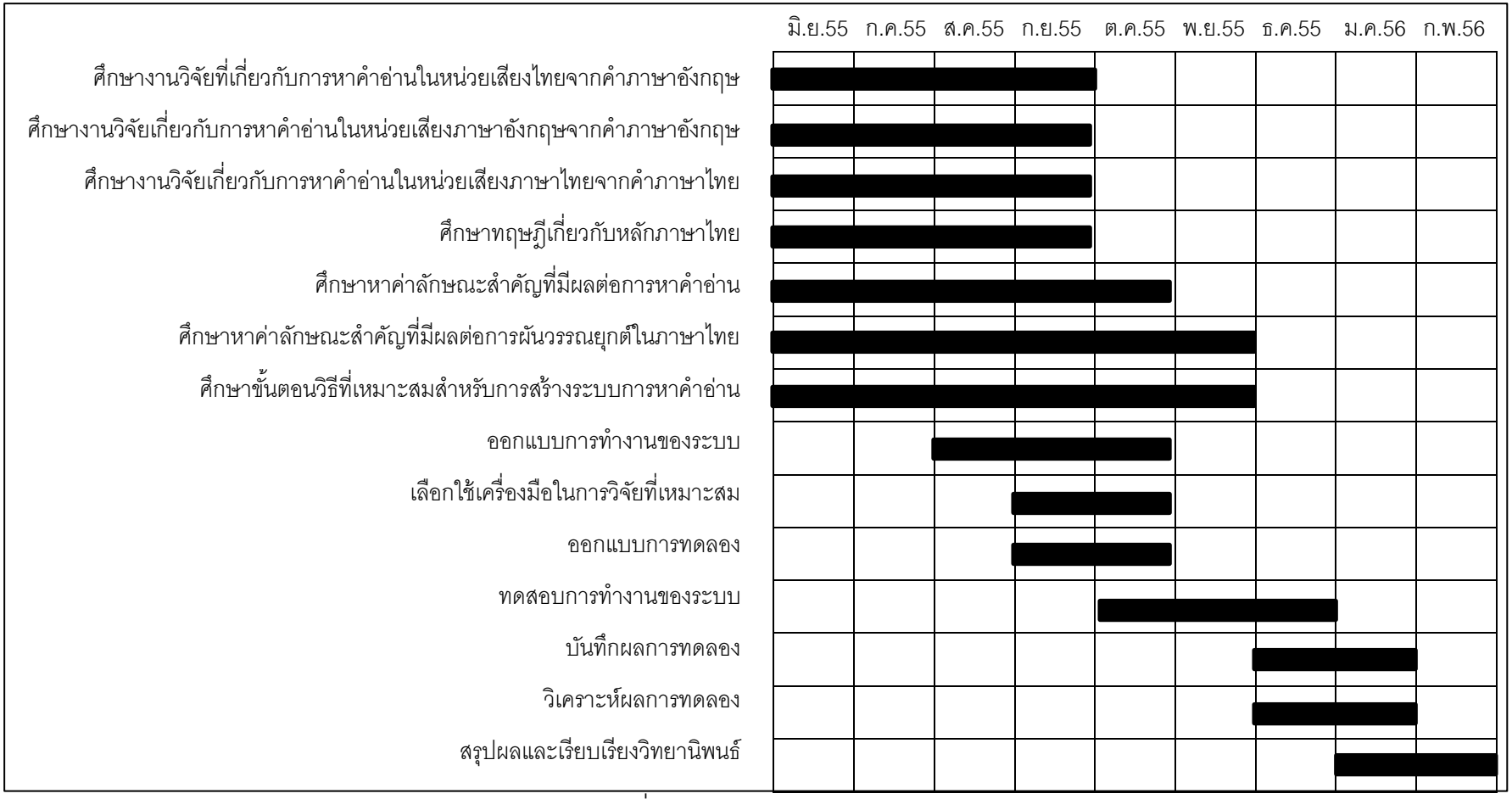
1.5 ประโยชน์ที่คาดว่าจะได้รับ

1. กลุ่มของพยัญชนะสะกดในภาษาไทยและอักขระพยางค์ของคำภาษาอังกฤษ มีส่วนช่วยให้การทำนายเสียงวรรณยุกต์แม่นยำมากขึ้น

2. โมเดลแบบดิสคริมิเนทีฟ คือ ออนไลน์ดิสคริมิเนทีฟเทรนนิงเฟรมเวิร์ค สามารถสร้างโมเดลทำนายหน่วยเสียงที่มีความแม่นยำมากกว่ามากเมื่อเทียบกับโมเดลแบบต้นไม้ซึ่งใช้ในงานวิจัยอ้างอิง

3. กระบวนการประมวลผลหลังจากการทำนายหน่วยเสียง สามารถแก้ไขความผิดพลาดจากโมเดลทำนายหน่วยเสียงได้

4. โมเดลรวมมีความเหมาะสมในการหาคำอ่านในหน่วยเสียงภาษาไทยจากคำภาษาอังกฤษ และสามารถนำไปพัฒนาเข้ากับระบบสังเคราะห์เสียงภาษาไทยต่าง ๆ ได้ เช่น Chula TTS เพื่อให้สามารถสังเคราะห์เสียงคำภาษาอังกฤษได้ และได้เสียงที่มีความเป็นธรรมชาติ



รูปที่ 1.1 แผนการดำเนินงานวิจัย

1.6 ผลงานตีพิมพ์จากวิทยานิพนธ์

ส่วนหนึ่งของวิทยานิพนธ์นี้ได้ตีพิมพ์เป็นบทความทางวิชาการในหัวข้อเรื่อง "A Framework for Thai Phonetization of English words" จัดทำโดย "Kitanan Pitakpawatkul, Atiwong Suchato, Proadpran Punyabukkana" ถูกนำเสนอในงานประชุมวิชาการ "The 16th International Computer Science and Engineering Conference 2012: ICSEC 2012" ณ Garden Cliff Resort & Spa จังหวัดชลบุรี ประเทศไทย ระหว่างวันที่ 17 ตุลาคม 2555 ถึงวันที่ 19 ตุลาคม 2555 และหัวข้อเรื่อง "Thai Phonetization of English Words using English Syllables" จัดทำโดย "Kitanan Pitakpawatkul, Atiwong Suchato, Proadpran Punyabukkana, Chai Wutiwiwatchai" ถูกนำเสนอในงานประชุมวิชาการ "International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology 2013: ECTI-CON 2013" ณ โรงแรมมาริไทม์ ปาร์คแอนด์สปา จังหวัดกระบี่ ประเทศไทย ระหว่างวันที่ 15 พฤษภาคม 2555 ถึงวันที่ 17 พฤษภาคม 2555

1.7 ลำดับการจัดเรียงเนื้อหาในวิทยานิพนธ์

ในวิทยานิพนธ์นี้ได้แบ่งเนื้อหาออกเป็น 5 บท คือ บทที่ 1 บทนำ กล่าวถึง ความเป็นมาและความสำคัญของปัญหา, วัตถุประสงค์ของการวิจัย, ขอบเขตของการวิจัย, ลำดับขั้นตอนในการเสนอผลการวิจัย และผลงานตีพิมพ์จากวิทยานิพนธ์

บทที่ 2 กล่าวถึงทฤษฎีและงานวิจัยที่เกี่ยวข้อง โดยทฤษฎีที่เกี่ยวข้องจะประกอบด้วย ลักษณะสำคัญของภาษาไทย, หน่วยเสียงสากล, หน่วยเสียงภาษาไทยและหน่วยเสียงภาษาอังกฤษ ในส่วนของงานวิจัยที่เกี่ยวข้องจะแบ่งเป็นงานวิจัยที่เกี่ยวข้องกับการหาคำอ่านในหน่วยเสียงภาษาไทยจากคำภาษาอังกฤษ, งานวิจัยที่เกี่ยวข้องกับการสังเคราะห์เสียงสองภาษา และงานงานวิจัยที่เกี่ยวข้องอื่น ๆ

บทที่ 3 กล่าวถึงขั้นตอนการดำเนินงานวิจัย ซึ่งจะประกอบไปด้วย ภาพรวมของระบบ, การปรับแนวและโมเดลทำนายหน่วยเสียง, กระบวนการประมวลผลหลังจากการทำนายหน่วยเสียงและโมเดลการทำนายเสียงวรรณยุกต์

บทที่ 4 กล่าวถึงการทดลอง ผลการทดลองและการอภิปรายผล ซึ่งจะประกอบไปด้วย ข้อมูลที่ใช้ในการทดลอง, เครื่องมือที่ใช้ในการทดลอง, การประเมินผล, ผลการทดลองและการอภิปรายผล

บทที่ 5 กล่าวถึงบทสรุปผลการวิจัย และข้อเสนอแนะต่าง ๆ

บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง

ทฤษฎีที่เกี่ยวข้องสามารถแบ่งได้ออกได้เป็นสองบทย่อย บทย่อยที่ 1 จะกล่าวถึงทฤษฎีที่เกี่ยวข้องกับลักษณะต่าง ๆ ที่สำคัญของภาษาไทย บทย่อยที่ 2 จะกล่าวถึงสัทอักษรสากล หน่วยเสียงภาษาไทย และหน่วยเสียงภาษาอังกฤษ

2.1.1 ทฤษฎีเกี่ยวกับภาษาไทย

2.1.1.1 พยัญชนะ

ตัวอักษรพยัญชนะในภาษาไทยมี 44 ตัวอักษร คือ ก ถึง ฮ โดยสามารถแบ่งตัวอักษรพยัญชนะออกได้เป็นสามกลุ่ม คือ อักษรสูง อักษรกลาง และอักษรต่ำ ตามการผันวรรณยุกต์ที่แตกต่างกัน [4] เช่น อักษรสูงผันได้สามเสียง คือ เสียงวรรณยุกต์เอก เสียงโท และจัตวา อักษรกลางผันได้ครบทั้งห้าเสียง ส่วนอักษรต่ำผันได้สามเสียงคือ เสียงสามัญ เสียงโท เสียงจัตวา เป็นต้น โดยรายละเอียดเป็นดังนี้

อักษรสูงมี 11 ตัว ได้แก่ ข ฃ ฉ ฐ ถ ผ ฝ ส ศ ษ ห
อักษรกลางมี 9 ตัว ได้แก่ ก จ ฎ ฏ ด ต บ ป อ
อักษรต่ำมี 24 ตัว ได้แก่ ค ฅ ฆ ง ช ซ ฌ ญ ฑ ฒ ท ธ น พ ฟ ภ ม ย ร ล ว ฬ ฮ

โดยอักษรต่ำยังสามารถแบ่งได้เป็นสองกลุ่ม คือ อักษรต่ำเดี่ยว และ อักษรต่ำคู่ อักษรต่ำคู่คืออักษรต่ำที่มีอักษรสูงเป็นเสียงคู่ มี 14 ตัว ได้แก่ ค ฅ ฆ (คู่กับ ข ฃ) ช ฌ (คู่กับ ฉ) ซ (คู่กับ ศ ษ ส) ฑ ฒ ท ธ (คู่กับ ฐ ถ) พ ภ (คู่กับ ผ) ฟ (คู่กับ ฝ) ฮ (คู่กับ ห) อักษรต่ำเดี่ยวคืออักษรต่ำที่เหลืออีก 10 ตัว ได้แก่ ง ญ ณ น ม ย ร ล ว ฬ

ซึ่งพยัญชนะทั้ง 44 ตัวสามารถแบ่งเป็นเสียงพยัญชนะได้ทั้งสิ้น 21 เสียงหรือหน่วยเสียง ตารางที่ 2.1 แสดงพยัญชนะที่สัมพันธ์กับหน่วยเสียงนั้น ๆ โดยหน่วยเสียงจะถูกกำหนดด้วยตัวอักษรภาษาอังกฤษ

ตารางที่ 2.1 ตัวอักษรพยัญชนะไทยเทียบกับหน่วยเสียงไทย [4]

ตัวอักษรพยัญชนะ	**หน่วยเสียง**
ก	k
ค ข ฆ ค ฆ	kh
ง	ng
จ	c
ช ฉ ฌ	ch
ซ ศ ษ ส	s
ด ฎ	d
ต ฏ	t
ท ธ ถ ฐ ฑ ฒ	th
น ณ	n
บ	b
ป	p
พ ผ ภ	ph
ฟ ฝ	f
ม	m
ย ญ	j
ร	r
ล ฬ	l
ว	w
ฮ ห	h
อ	z

แต่หากจะแบ่งตัวอักษรพยัญชนะตามการออกเสียงของพยัญชนะสะกด ก็จะแบ่งออกเป็น 8 มาตรา คือ มาตราแม่กก แม่กง แม่กด แม่กน แม่กบ แม่กม แม่เกย และแม่เกอว ดังนี้

แม่กก ก ข ฆ ค ศ ฆ

แม่กง ง

แม่กด จ ฉ ช ซ ฌ ฎ ฏ ฐ ฑ ฒ ด ต ถ ท ธ ศ ษ ส

แม่กน ญ ณ น ร ล ฬ

แม่กบ บ ป พ ภ ผ ฝ ฟ
แม่กม ม
แม่เกย ย
แม่เกอว ว

2.1.1.2 **สระ**

สำหรับเสียงสระ เสียงสระในภาษาไทยมีทั้งหมด 24 เสียง โดยแบ่งเป็นเสียงสระเดี่ยว 18 เสียง และเสียงสระผสม 6 เสียง ดังนี้

สระเดี่ยว ได้แก่ อะ อา อิ อี อึ อื อุ อู เอะ เอ แอะ แอ โอะ โอ เอาะ ออ เออะ เออ
สระผสม ได้แก่ เอียะ เอีย เอือะ เอือ อัวะ อัว

หรือเมื่อจับคู่ความสั้นความยาวของเสียงสระจะได้คู่สระดังนี้ (สั้น-ยาว)

อะ-อา, อิ-อี, อึ-อื, เอะ-เอ, แอะ-แอ, โอะ-โอ, เอาะ-ออ, เออะ-เออ, เอียะ-เอีย, เอือะ-เอือ, อัวะ-อัว

2.1.1.3 **คำเป็นคำตาย**

การแยกคำเป็นคำตายนั้นเป็นประโยชน์ในการอธิบายในเรื่องการผันวรรณยุกต์ โดยหลักในการแยกคำเป็นคำตายดังนี้

หากคำนั้นสะกดในมาตราแม่ กก กด กบ คำนั้นเป็นคำตาย
หากคำนั้นสะกดในมาตราแม่ กง กน กม เกย เกอว คำนั้นเป็นคำเป็น
หากคำนั้นไม่มีตัวสะกด ถ้าเป็นสระเสียงสั้น คำนั้นเป็นคำตาย แต่หากเป็นสระเสียงยาว คำนั้นเป็นคำเป็น

2.1.1.4 **การผันวรรณยุกต์ในภาษาไทย**

การผันวรรณยุกต์ในภาษาไทย ถูกกำหนดจากองค์ประกอบหลัก ๆ 4 ส่วน คือ กลุ่มของพยัญชนะต้น, คำเป็นคำตาย, ความสั้นยาวของสระ และ อักษรนำ ซึ่งสามารถสรุปได้ตามรายละเอียดในตารางที่ 2.2

ตารางที่ 2.2 การผันเสียงวรรณยุกต์ [4]

				สามัญ	เอก	โท	ตรี	จัตวา
อักษรกลาง	คำเป็น			กา	ก่า	ก้า	ก๊า	ก๋า
	คำตาย			-	กะ	ก้ะ	ก๊ะ	ก๋ะ
อักษรสูง	คำเป็น			-	ข่า	ข้า	-	ขา
	คำตาย			-	ขาก	ข้าก	-	-
อักษรต่ำคู่	คำเป็น			คา		ค่า	ค้า	
					(ข่า)	(ข้า)		(ขา)
	คำตาย	สระสั้น		-		ค่ะ	คะ	ค๋ะ
					(ขะ)	(ข้ะ)		
		สระยาว		-		คาก	ค้าก	ค๋าก
					(ขาก)	(ข้าก)		
อักษรต่ำเดี่ยว	คำเป็น			งา		ง่า	ง้า	
			สูงนำ		หง่า	หง้า		หงา
			กลางนำ	อยา	อย่า	อย้า	อย๊า	อย๋า
	คำตาย	สระสั้น		-		ง่ะ	งะ	ง๋ะ
			สูงนำ		หงะ	หง้ะ		
			กลางนำ	-	อยะ	อย้ะ	อย๊ะ	อย๋ะ
		สระยาว		-		งาก	ง้าก	ง๋าก
			สูงนำ		หงาก	หง้าก		
			กลางนำ	-	อยาก	อย้าก	อย๊าก	อย๋าก

2.1.2. **หน่วยเสียงสากล, หน่วยเสียงภาษาไทยและหน่วยเสียงภาษาอังกฤษ**

สัทอักษรสากล (International Phonetic Alphabet: IPA) [5] คือ สัญลักษณ์ที่ถูกกำหนดขึ้นมาเพื่อแทนหน่วยเสียงที่มนุษย์สามารถเปล่งเสียงออกมาได้ โดยแบ่งออกตามตำแหน่งการออกเสียง ลักษณะการออกเสียง เป็นต้น ซึ่งเป็นหน่วยเสียงที่ครอบคลุมกับเสียงในทุก ๆ ภาษา สัญลักษณ์ต่าง ๆ ที่กำหนดขึ้นใช้อักษรโรมันเป็นหลัก ร่วมกับอักษรกรีกและอักษรอื่น ๆ ที่ประดิษฐ์ขึ้นมาใหม่ รายละเอียดแสดงในรูปที่ 2.1

the international phonetic alphabet (2005)

consonants (pulmonic)

	LABIAL		CORONAL				DORSAL				RADICAL		LARYNGEAL
	Bilabial	Labio-dental	Dental	Alveolar	Palato-alveolar	Retroflex	Alveolo-palatal	Palatal	Velar	Uvular	Pharyngeal	Epi-glottal	Glottal
Nasal	m	ɱ	n			ɳ		ɲ	ŋ	ɴ			
Plosive	p b		t d			ʈ ɖ	c ɟ		k g	q ɢ		ʡ	ʔ
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ɕ ʑ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	ʜ ʢ	h ɦ
Approximant		ʋ		ɹ		ɻ		j	ɰ				
Tap, flap		ⱱ		ɾ		ɽ							
Trill	ʙ			r						ʀ		ʀ	
Lateral fricative				ɬ ɮ									
Lateral approximant				l		ɭ		ʎ	ʟ				
Lateral flap				ɺ									

Where symbols appear in pairs, the one to the right represents a modally voiced consonant, except for murmured *fi*.
Shaded areas denote articulations judged to be impossible. Light grey letters are unofficial extensions of the IPA.

consonants (non-pulmonic)

clicks		implosives		ejectives	
ʘ	Bilabial fricated	ɓ	Bilabial	ʼ	*examples:*
ǀ	Laminal alveolar fricated ("dental")	ɗ	Dental or alveolar	pʼ	Bilabial
ǃ	Apical (post)alveolar abrupt ("retroflex")	ᶑ	Retroflex	tʼ	Dental or alveolar
‼	Subapical retroflex	ʄ	Palatal	kʼ	Velar
ǂ	Laminal postalveolar abrupt ("palatal")	ɠ	Velar	tɬʼ	Lateral affricate
ǁ	Lateral alveolar fricated ("lateral")	ʛ	Uvular	sʼ	Alveolar fricative

consonants (co-articulated)

ʍ	Voiceless labialized velar approximant
w	Voiced labialized velar approximant
ɥ	Voiced labialized palatal approximant
ʃ͡ʒ	Simultaneous *x* and ʃ (existance disputed)
t͡ʃ	Affricates and double articulations
g͡b	may be joined by a tie bar

vowels

	Front	Near front	Central	Near back	Back
Close	i • y		ɨ • ʉ		ɯ • u
Near close		ɪ • ʏ		ʊ •	
Close mid	e • ø		ɘ • ɵ		ɤ • o
Mid			ə		
Open mid	ɛ • œ		ɜ • ɞ		ʌ • ɔ
Near open		æ	ɐ		
Open		a • ɶ			ɑ • ɒ

Vowels at right & left of bullets are rounded & unrounded.

suprasegmentals

ˈ	Primary stress	ˌ	Secondary stress
ˌfoʊnəˈtɪʃən			
eː	Long	eˑ	Half-long
e	Short	ĕ	Extra-short
.	Syllable break	‿	Linking (no break)

intonation

|	Minor (foot) break	‖	Major (intonation) break
↗	Global rise	↘	Global fall

(tone)

level tones		*contour tones (e.g.)*	
˥	Top	˩˥	Rising
˦	High	˥˩	Falling
˧	Mid	˧˥	High rising
˨	Low	˨˩	Low rising
˩	Bottom	˥˧	High falling

tone terracing

↑	Upstep	˧˩	Low falling
↓	Downstep	˩˧	Peaking
		˩˥˩	Dipping

diacritics

Diacritics may be moved to fit a letter, as n̥ or ɳ̊. Other letters may be used as diacritics of phonetic detail:
t^{s} (fricative release), $b^{ɦ}$ (breathy voice), $m^{ʔ}$ (glottalized), ^{ə} (epenthetic schwa), o^{w} (off-glide), $u^{β}$ (compressed).

SYLLABICITY & RELEASES		PHONATION		PRIMARY ARTICULATION		SECONDARY ARTICULATION			
n̩ ɹ̩	Syllabic	n̥ d̥	Voiceless *or* Slack voice	t̪ b̪	Dental	tʷ dʷ	Labialized	ɔ̹ χ̹	More rounded
e̯ o̯	Non-syllabic	s̬ d̬	Modal voice *or* Stiff voice	t̺ d̺	Apical	tʲ dʲ	Palatalized	ɔ̜ χ̜ʷ	Less rounded
tʰ ʰt	(Pre)aspirated	n̤ a̤	Breathy voice	t̻ d̻	Laminal	tˠ dˠ	Velarized	ẽ z̃	Nasalized
dⁿ	Nasal release	n̰ a̰	Creaky voice	u̟ t̟	Advanced	tˤ dˤ	Pharyngealized	ɚ ɝ	Rhoticity
dˡ	Lateral release	n̼ a̼	Strident	i̠ t̠	Retracted	ɬ ʒ	Velarized or pharyngealized	e̝ o̝	Advanced tongue root
t̚	No audible release	n̪ d̪	Linguolabial	ä j̈	Centralized	ŭ	Mid-centralized	e̞ o̞	Retracted tongue root
e̬ β̬	Lowered (β̬ is a bilabial approximant)			e̝ ɹ̝	Raised (ɹ̝ is a voiced alveolar non-sibilant fricative, r̝ a fricative trill)				

brackets: //morphophonemic// /phonemic/ [phonetic] ⟨orthographic⟩

รูปที่ 2.1 รายละเอียดหน่วยเสียงสากล (คัดลอกจาก [5])

สำหรับภาษาไทย คำอ่านจะถูกแบ่งออกเป็นหน่วยย่อย คือ พยางค์ ซึ่งสามารถเขียนแทนได้ในรูปแบบของ /C_i V C_f ^{T}/ โดยที่ C_i คือเสียงพยัญชนะต้น, V คือเสียงสระ, C_f คือ เสียงพยัญชนะสะกด และ T คือเสียงวรรณยุกต์ โดยเสียงพยัญชนะต้นอาจเป็นได้ทั้งเสียงพยัญชนะเดี่ยว (Single) และเสียงพยัญชนะประสม (Cluster) รวมทั้งเสียงสระก็อาจเป็นได้ทั้ง เสียงสระเดี่ยว (Single) และ เสียงสระประสม (Diphthong) ส่วนเสียงวรรณยุกต์มี 5 ระดับ แสดงด้วยตัวเลขตั้งแต่ 0 ถึง 4 คือ เสียงสามัญ (Mid Tone: 0), เสียงเอก (Low Tone: 1), เสียงโท (Falling Tone: 2), เสียงตรี (High Tone, 3) และ เสียงจัตวา (Rising Tone: 4) ตารางที่ 2.3 แสดงหน่วยเสียงภาษาไทยทั้งหมด

ตารางที่ 2.3 หน่วยเสียงในภาษาไทย [6]

ชนิด		**สัญลักษณ์**
เสียงพยัญชนะต้น	เดี่ยว	p, t, c, k, z, ph, th, ch, kh, b, d, m, n, ng, l, r, f, s, h, w, j
	ประสม	pr, phr, tr, kr, khr, pl, phl, thr, kl, khl, kw, khw, br, bl, fr, fl, dr
เสียงสระ	เดี่ยว	a, aa, i, ii, v, vv, u, uu, e, ee, x, xx, o, oo, @, @@, q, qq
	ประสม	ia, iia, va, vva, ua, uua
เสียงพยัญชนะสะกด		p^, t^, k^, n^, m^, ng^, j^, w^, f^, l^, s^, ch^
วรรณยุกต์		0, 1, 2, 3, 4

สำหรับภาษาอังกฤษ คำอ่านจะถูกแทนด้วยหน่วยเสียงในภาษาอังกฤษ แต่จะไม่มีการแบ่งเป็นพยางค์เหมือนอย่างในภาษาไทย ในงานนี้ผู้วิจัยอ้างอิงตาม CMU Pronounciation Dictionary: CMUDICT [7] ซึ่งเป็นพจนานุกรมคำอ่านภาษาอังกฤษ ซึ่งมีมากกว่า 125,000 คำ มีการแทนคำอ่านด้วยหน่วยเสียงในภาษาอังกฤษ 39 หน่วยเสียงและการเน้นพยางค์ (Stress) อีก 3 ระดับ โดยตารางที่ 2.4 แสดงหน่วยเสียงภาษาอังกฤษที่อ้างอิงจาก CMUDICT เปรียบเทียบกับหน่วยเสียงสากลและหน่วยเสียงภาษาไทย

ตารางที่ 2.4 หน่วยเสียงภาษาอังกฤษเปรียบเทียบกับหน่วยเสียงสากลและหน่วยเสียงภาษาไทย [6]

สัทอักษรอังกฤษ	**สัทอักษรสากล**	**สัทอักษรไทย**
AA	ɔ	@
AE	æ	x
AH	ə	v
AO	ɔː	@@
AW	au	a w^
AY	ai	a j^
EH	e	e
ER	ɜ	q
EY	eː	ee
IH	I	i
IY	iː	ii
OW	oː	oo
OY	ɔi	@ j^
UH	ʊ	u
UW	u	uu
B	b	b
CH	ʧ	ch
D	d	d
DH	ð	th
F	f	f
G	g	k
HH	h	h
JH	dʒ	c
K	k	k, kh
L	l	l
M	m	m
N	n	n

สัทอักษรอังกฤษ	สัทอักษรสากล	สัทอักษรไทย
NG	ŋ	ng
P	p	p, ph
R	r	r
S	s	s
SH	ʃ	ch
T	t	t, th
TH	θ	th
V	v	w
W	w	w
Y	j	j
Z	z	s
ZH	ʒ	ch

2.2 เอกสารและงานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องจะแบ่งออกได้เป็น 4 หมวดหลัก ๆ คือ 1.งานวิจัยที่เกี่ยวข้องโดยตรงกับการหาคำอ่านในหน่วยเสียงภาษาไทยจากคำภาษาอังกฤษ 2.งานวิจัยที่เกี่ยวกับระบบสังเคราะห์เสียงสองภาษา (Bilingual Text-to-Speech) 3.งานวิจัยที่ใกล้เคียงโดยเกี่ยวกับหาคำอ่านจากคำในภาษาใดภาษาหนึ่งโดยเฉพาะ เช่น การหาคำอ่านไทยจากคำภาษาไทย (Thai Grapheme-to-Phoneme: Thai G2P) และการหาคำอ่านอังกฤษจากคำภาษาอังกฤษ (English Grapheme-to-Phoneme: English G2P) และ 4. งานวิจัยที่เกี่ยวข้องอื่น ๆ

2.2.1. งานวิจัยเรื่องการหาคำอ่านในหน่วยเสียงภาษาไทยจากคำภาษาอังกฤษ

งานวิจัยในเรื่องการหาคำอ่านในหน่วยเสียงภาษาไทยจากคำภาษาอังกฤษ เริ่มต้นจากการใช้วิธีกฎเกณฑ์ (Rule Based) เป็นวิธีแรก [8] โดยได้ทำการสร้างกฎเกณฑ์แบบง่ายในการสร้างคำอ่านขึ้นมาจากตัวอักษรของคำภาษาอังกฤษนั้น ซึ่งได้ประยุกต์วิธีการนี้มาจากงานวิจัยที่ทำการหาคำอ่านอังกฤษจากคำภาษาอังกฤษ [9] โดยกฎถูกสร้างจากข้อมูลที่เตรียมไว้และสร้างเฉพาะกฎการแปลงเสียงสระเท่านั้น และในส่วนอื่น ๆ ก็ใช้ความรู้ทางภาษาศาสตร์เพื่อแปลงแบบตรงไปตรงมา เพราะเสียงสระมีความกำกวมมากกว่าเสียงพยัญชนะ สำหรับเสียงวรรณยุกต์ก็ใช้กฎเกณฑ์ทางภาษาศาสตร์ในการทำนาย โดยดูจากคำเป็นคำตายและจำนวนพยางค์ เป็นต้น การวัดผลใช้การวัดผลในแง่ความพึงพอใจซึ่งมีสามระดับ คือ “poor”, “fair” และ “good” โดยผลลัพธ์ที่ถือว่ายอมรับได้คืออยู่ในระดับ “fair” หรือ “good” ซึ่งรวมแล้วมีทั้งหมด 44% เป็นผลลัพธ์ที่ยอมรับได้ โดยวิธีนี้มีข้อเสียคือกฎเกณฑ์ไม่สามารถจัดการกับความกำกวมของคำได้ดี และการสร้างกฎก็ทำเพียงเฉพาะกับเสียงสระเท่านั้น ซึ่งเสียงพยัญชนะเองก็มีความกำกวมเช่นกัน ดังนั้นการใช้กฎเกณฑ์ที่ตายตัวในการกำหนดเสียงของพยัญชนะและเสียงวรรณยุกต์จึงทำให้เกิดความผิดพลาดได้ง่าย

เพื่อพัฒนาให้การหาคำอ่านมีความแม่นยำมากขึ้น ในงานวิจัยต่อมา [6] ซึ่งเป็นงานวิจัยอ้างอิง (Baseline) จึงได้พัฒนาโดยนำโมเดลต้นไม้ (Classification and Regression Tree: CART) มาใช้ในการทำนายหน่วยเสียง และเสียงวรรณยุกต์ โดยสร้างเป็นสองโมเดลแยกจากกันและใช้ลักษณะสำคัญ (Feature) สำหรับโมเดลทำนายหน่วยเสียง คือ บริบทของอักษรข้างเคียง

และ บริบทของหน่วยเสียงภาษาอังกฤษข้างเคียง ซึ่งได้มาจาก CMUDICT โดยกระบวนที่สำคัญที่จะต้องทำก่อนการสร้างโมเดลการทำนายหน่วยเสียงได้นั้นก็คือ กระบวนการปรับแนว (Alignment Process) ระหว่างอักษรภาษาอังกฤษ, หน่วยเสียงภาษาอังกฤษ และหน่วยเสียงอ่านภาษาไทย เพื่อที่จะทำให้เกิดความสัมพันธ์ระหว่างอักษรและหน่วยเสียงและสามารถนำไปใช้ในการสร้างโมเดลต้นไม้ต่อไปได้ โดยงานวิจัยนี้ได้นำเทคนิคไดนามิกไทม์วอร์ปปิง (Dynamic time warping: DTW) มาประยุกต์ในกระบวนการปรับแนว ในส่วนของโมเดลการทำนายเสียงวรรณยุกต์ ค่าลักษณะสำคัญที่ใช้สำหรับโมเดลการทำนายเสียงวรรณยุกต์จะเกี่ยวข้องกับหลักภาษาศาสตร์ภาษาไทย ได้แก่ เสียงพยัญชนะต้น, เสียงสระ, เสียงพยัญชนะสะกด, กลุ่มพยัญชนะต้น (สูง กลาง ต่ำ), กลุ่มสระ (สั้น ยาว), จำนวนพยางค์ในคำ, ประเภทของพยางค์ (คำเป็น คำตาย), ตำแหน่งของพยางค์, เสียงวรรณยุกต์ก่อนหน้า ซึ่งประสิทธิภาพรวมของระบบสามารถทำนายคำอ่านได้ความแม่นยำในระดับพยางค์เท่ากับ 53.3% และความแม่นยำในระดับคำเท่ากับ 24.5% รวมทั้งการประเมินผลในในแง่ความพึงพอใจในลักษณะเดียวกันกับในงานวิจัยก่อนหน้า [8] ซึ่งได้ผลลัพธ์ 62% เป็นผลลัพธ์ที่ยอมรับได้ เนื่องจากผลลัพธ์สุดท้ายเกิดจากผลลัพธ์จากทั้งสองโมเดลรวมกัน ดังนั้นหากโมเดลใดโมเดลหนึ่งให้ผลลัพธ์ที่ผิดพลาดก็จะทำให้ผลลัพธ์รวมผิดทันที ข้อผิดพลาดที่เกิดขึ้นบ่อยจากวิธีนี้ก็คือความกำกวมในแยกระหว่างเสียงพยัญชนะต้นและพยัญชนะสะกด รวมทั้งความสั้นยาวของเสียงสระ จึงทำให้ความแม่นยำไม่สูง รวมทั้งค่าลักษณะสำคัญที่มีผลทำให้โมเดลมีความแม่นยำมากขึ้นก็คือ คำอ่านในหน่วยเสียงภาษาอังกฤษของคำนั้น ซึ่งคำภาษาอังกฤษที่ไม่อยู่ในพจนานุกรมก็จะไม่สามารถหาคำอ่านได้และทำให้ความแม่นยำลดลงมาก

2.2.2. **งานวิจัยที่เกี่ยวกับระบบสังเคราะห์เสียงสองภาษา**

ปัญหาที่เกิดขึ้นจากคำภาษาอังกฤษที่ปรากฏร่วมกับคำภาษาไทยในระบบการสังเคราะห์เสียงนั้น สามารถแก้ไขได้ในอีกลักษณะที่แตกต่างจากการหาคำอ่านในหน่วยเสียงภาษาไทย คือการนำระบบสังเคราะห์เสียงภาษาอังกฤษเข้ามารองรับโดยตรงกับคำภาษาอังกฤษ แต่วิธีการนี้จะทำให้เกิดความแตกต่างของระดับการเน้นพยางค์ในเสียงสังเคราะห์ ซึ่งอาจทำให้เสียงไม่เป็นธรรมชาติ ดังนั้นจึงมีงานวิจัยที่เกี่ยวข้อง [3] เสนอวิธีในการปรับระดับการเน้นพยางค์ระหว่างเสียงสังเคราะห์สองภาษาให้มีความราบเรียบเป็นธรรมชาติมากขึ้น โดยได้นำเสนอการใช้เทคนิคเอช

เอ็มเอ็มอินเทอร์โพเลชัน (HMM Interpolation) ในการปรับระดับการเน้นพยางค์ โดยใช้การปรับแนวระหว่างคำอ่านของสองภาษา (Cross-language Phone Alignment) และ การเชื่อมความสัมพันธ์สถานะระหว่างเสียงสองภาษา (HMM State Mapping) ผลลัพธ์ที่ได้นั้นประเมินผลในแง่ความพึงพอใจในหลายลักษณะ ซึ่งสรุปได้ว่า หากคำภาษาอังกฤษเป็นคำที่ภาษาไทยมีการยืมมาใช้ (Loan Word) ผู้ทดสอบพอใจกับการออกเสียงแบบไทยมากกว่า แต่หากไม่ใช่ ผู้ทดสอบจะพอใจกับการออกเสียงแบบภาษาอังกฤษมากกว่า และเมื่อสรุปความพึงพอใจในแง่มุมของอาชีพของผู้ทดสอบ อาชีพนักเรียนพอใจกับการฟังแบบไทยมากที่สุด วิศวกรพอใจกับการฟังทั้งสองแบบ และ นักภาษาศาสตร์พอใจกับการฟังแบบอังกฤษมากที่สุด ซึ่งก็จะเห็นได้ว่า ถึงแม้การปรับการเน้นพยางค์ระหว่างสองภาษาแล้วนั้น ก็ยังพบว่ามีคนบางกลุ่มที่ยังพอใจกับการฟังการออกเสียงแบบภาษาไทยอยู่

นอกจากในงานภาษาไทย การแก้ไขปัญหาการเกิดคำมากกว่าหนึ่งภาษา ก็ปรากฏในภาษาจีนด้วยเช่นกัน การปรากฏของคำภาษาอังกฤษขึ้นในบทความภาษาจีนซึ่งโดยมากก็จะเป็นชื่อเฉพาะ ทำให้ระบบสังเคราะห์เสียงภาษาจีนต้องออกแบบเพื่อรองรับการอ่านภาษาอังกฤษด้วย โดยงานวิจัย [10] ได้ออกแบบระบบสังเคราะห์เสียงที่พยายามสร้างเสียงให้เกิดความเป็นธรรมชาติที่สุด โดยแทนที่จะใช้ระบบสังเคราะห์เสียงสองภาษาแยกจากกัน งานวิจัยนี้ทำการสร้างระบบการหาคำอ่านของคำภาษาจีนและคำภาษาอังกฤษแยกจากกัน จากนั้นประโยคที่จะถูกสังเคราะห์เสียงจะถูกนำไปหาคำอ่านตามภาษาของคำนั้น แล้วนำคำอ่านมาเรียงกันตามลำดับคำในประโยคจริง จากนั้นก็ทำการเลือกเสียงที่ได้ทำการเรียนรู้มาแล้ว มาสร้างเป็นเสียงสังเคราะห์ต่อเนื่องกัน เพื่อให้เกิดความเป็นธรรมชาติงานวิจัยนี้ใช้ผู้ทดลองบันทึกเสียงเป็นคนที่อ่านออกเสียงทั้งภาษาจีนและภาษาอังกฤษได้อย่างดี และทำการเก็บเสียงทั้งหน่วยเสียงภาษาจีนและภาษาอังกฤษ

2.2.3. **งานวิจัยที่ใกล้เคียงโดยเกี่ยวกับการหาคำอ่านจากคำในภาษาใดภาษาหนึ่ง**

นอกเหนือจากงานวิจัยที่ได้กล่าวมาแล้วนั้น ผู้วิจัยได้ศึกษางานวิจัยที่เกี่ยวข้องกับการหาคำอ่านไทยจากคำภาษาไทย และ การหาคำอ่านอังกฤษจากคำภาษาอังกฤษ เพื่อที่จะศึกษาเทคนิคต่าง ๆ ในงานเหล่านี้นำมาใช้ประโยชน์ในการสร้างระบบการหาคำอ่านในหน่วยเสียงภาษาไทยจากคำภาษาอังกฤษ ซึ่งเป็นงานที่ใกล้เคียงกัน โดยสามารถแบ่งงานวิจัยออกได้เป็นหลาย ๆ ประเภท ตามลักษณะวิธีกระบวนการสร้าง

2.2.3.1 **พจนานุกรมคำอ่าน** (Dictionary Based)

วิธีการใช้พจนานุกรมคำอ่าน เป็นวิธีที่ตรงไปตรงมาและไม่ซับซ้อนจึงมีหลายงานวิจัยที่เสนอการหาคำอ่านภาษาไทยจากคำภาษาไทยด้วยวิธีการนี้ [2], [11] แต่ไม่สามารถจัดการกับคำที่ไม่มีในพจนานุกรมได้ จึงนิยมใช้ร่วมกับวิธีอื่นเพื่อให้สามารถจัดการกับคำนอกพจนานุกรมได้ด้วย อย่างเช่นงานวิจัยที่ใช้พจนานุกรมคำอ่านภาษาไทยซึ่งมีจำนวนคำประมาณ 25,000 คำ เป็นข้อมูลหลัก ในการหาคำอ่านของคำต่าง ๆ [11] แต่ก็จะเกิดปัญหาเมื่อเจอกับคำที่ไม่อยู่ในพจนุกรม ดังนั้นในงานวิจัยนี้จึงใช้กฎเกณฑ์ มาช่วยจัดการกับคำที่ไม่มีในพจนุกรม การใช้พจนานุกรมคำอ่านให้ได้ประสิทธิภาพ จำเป็นต้องเตรียมพจนานุกรมที่มีคำครอบคลุมคำที่จะเจอทั้งหมด ซึ่งในความเป็นจริงไม่สามารถทำได้ เพราะมีคำใหม่ ๆ เกิดขึ้นอยู่ตลอดเวลา

2.2.3.2 **กฎเกณฑ์** (Rule Based)

การสร้างกฎเพื่อใช้สำหรับการหาคำอ่านเป็นวิธีที่มีข้อดี คือ ง่ายและประมวลผลได้รวดเร็ว แต่ก็มีข้อเสียคือเมื่อเจอคำที่มีความกำกวมที่สามารถอ่านออกเสียงได้หลายลักษณะตามบริบทของอักษร จะไม่สามารถจัดการได้อย่างมีประสิทธิภาพ ตัวอย่างเช่น มีงานวิจัยที่สร้างกฎเกณฑ์ขึ้นจากการจำรูปแบบการแปลงจากอักขระอังกฤษเป็นคำอ่านภาษาอังกฤษ [9] โดยใช้ข้อมูลที่เตรียมไว้ และสร้างเป็นตารางสำหรับไว้ค้นหา โดยจะมีกฎเกณฑ์ที่เป็นค่าพื้นฐานอยู่แล้วที่ได้มาจากความรู้ทางภาษาศาสตร์ จากนั้นก็ทำการสร้างกฎจากข้อมูลที่มีการปรับแนวระหว่างอักขระและหน่วยเสียงแล้ว

2.2.3.3 **โมเดลสถิติ** (Statistical Model)

การใช้โมเดลทางสถิติ เป็นวิธีที่ได้รับความนิยมมากกว่าวิธีอื่น ๆ โดยมีตั้งแต่การใช้โมเดลต้นไม้ (Classification and Regression Tree) กับทั้งการทำคำอ่านไทยจากคำภาษาอังกฤษ, การหาคำอ่านอังกฤษจากคำภาษาอังกฤษและการหาคำอ่านไทยจากคำภาษาไทย [6], [12], [13] ฮิดเดนมาร์คอฟโมเดล (Hidden Markov Models: HMM) [14], คอนดิชันนอลแรนดอมฟิลด์ (Conditional Random Fields: CRFs) [15] และออนไลน์ดิสคริมิเนทีฟเทรนนิงเฟรมเวิร์ค (Online Discriminative Training Framework) [16], [17] เป็นต้น โดยโมเดลต้นไม้เป็นโมเดลที่ได้รับความนิยมมากในช่วงเริ่มต้น เพราะประมวลผลเร็วและไม่ซับซ้อน โดยมีงานวิจัยที่ได้นำโมเดลต้นไม้มาเป็นโมเดลในการแปลงรูปอักขระเป็นคำอ่านในหลาย ๆ ภาษา ตั้งแต่ภาษาอังกฤษ, ฝรั่งเศษ และเยอรมัน [12] โดยใช้เทคนิคไดนามิกไทม์วอร์ปปิง มาช่วยในการปรับแนวระหว่างอักขระและคำอ่านในข้อมูลที่จะนำมาทำการเรียนรู้ จากนั้นจึงใช้โมเดลต้นไม้มาทำการเรียนรู้โดยใช้ลักษณะสำคัญ คือ บริบทรอบ ของอักษรที่กำลังพิจารณา โดยจะใช้สามหน่วยทางด้านซ้ายและสามหน่วยทางด้านขวา ในส่วนของการกำหนดการเน้นพยางค์ (Stress Assignment) จะใช้กฎเกณฑ์เข้ามาช่วยการในการทำนาย โดยจะมีเพียงสองระดับ คือ เน้น และ ไม่เน้น ซึ่งในงานวิจัยของการหาคำอ่านภาษาไทยก็มีการนำวิธีนี้มาประยุกต์เช่นกัน [13]

นอกจากโมเดลต้นไม้ที่ได้รับความนิยมแล้ว ยังมีงานวิจัยที่ทำการหาคำอ่านอังกฤษจากคำภาษาอังกฤษอีกจำนวนหนึ่งที่ใช้โมเดลใหม่ ๆ อย่างเช่น คอนดิชันนอลแรนดอมฟิลด์และดิสคริมิเนทีฟเทรนนิงเฟรมเวิร์ค ซึ่งเป็นโมเดลที่เป็นลักษณะแบบดิสคริมิเนทิฟ (Discriminative model) ซึ่งมีความเหมาะสมกับการงานในลักษณะของ การเปลี่ยนแปรสายอักขระ (String Transducer) เช่น การแปลงอักขระเป็นคำอ่าน (Grapheme-to-Phoneme) และ การสะกดคำภาษาอังกฤษเป็นอักษรไทย (English to Thai Transliteration) มากกว่าโมเดลแบบอื่น ๆ โดย งานวิจัยที่ใช้ดิสคริมิเนทีฟเทรนนิงเฟรมเวิร์ค [16], [17] สำหรับการหาคำอ่านอังกฤษจากคำภาษาอังกฤษ ได้ใช้ลักษณะสำคัญหลาย ๆ อย่างที่มีผลต่อการทำนาย นอกเหนือจากบริบทข้าง ๆ (Context) แต่ยังมีลักษณะสำคัญอื่น ๆ คือ เอ็นแกรม (N-gram), ลิเนียร์เชน (Linear-chain) และ จอยนต์เอ็นแกรม (Joint n-gram) เป็นต้น ซึ่งให้ผลลัพธ์ที่มีประสิทธิภาพมากในงานวิจัยการหาคำอ่านภาษาอังกฤษและสามารถนำมาปรับประยุกต์ใช้กับงานกับระบบการหาคำอ่านในหน่วยเสียงภาษาไทยจากคำภาษาอังกฤษได้

2.2.4. งานวิจัยที่เกี่ยวข้องอื่น ๆ

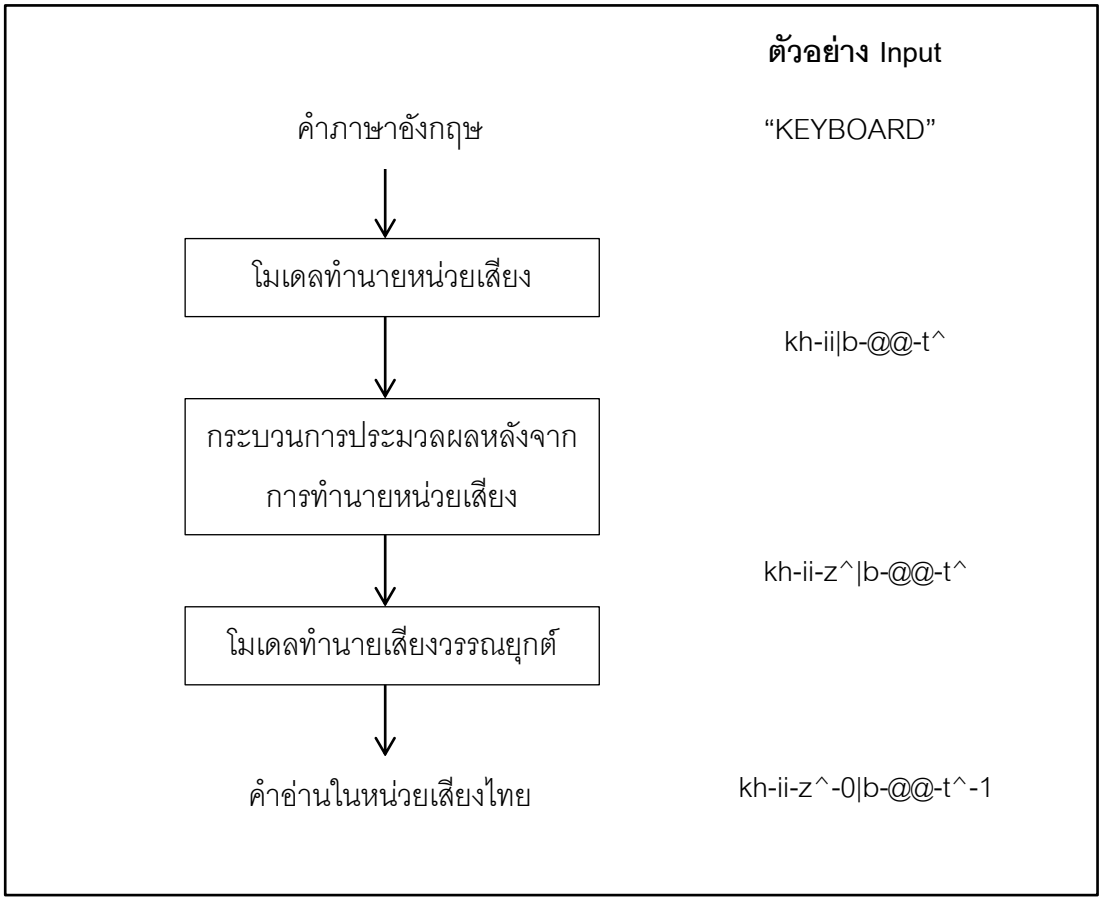
งานวิจัยอื่น ๆ ที่เกี่ยวข้องก็คืองานวิจัยที่สร้างคำทับศัพท์ภาษาจีนจากคำภาษาอังกฤษ [18] โดยศึกษาเปรียบเทียบถึงความสามารถของการแบ่งพยางค์ภาษาอังกฤษ ที่จะช่วยให้ผลลัพธ์การสร้างคำทับศัพท์มีความถูกต้องแม่นยำมากขึ้นหรือไม่ โดยใช้โมเดลทางสถิติ คือ โมเดลที่ใช้ในการแปลภาษาในระดับวลี (Phrase-Based Statistical Machine Translation) สำหรับการสร้างโมเดลทำนายคำทับศัพท์ ซึ่งพบว่าหากสร้างโมเดลการแบ่งพยางค์อังกฤษแยกออกจากโมเดลการทำนายคำทับศัพท์ เพื่อช่วยในการแบ่งพยางค์ก่อนที่จะทำนายคำทับศัพท์ ผลลัพธ์มีความแม่นยำมากกว่าการใช้โมเดลการแปลภาษาในระดับวลีในการทำนายคำทับศัพท์เพียงโมเดลเดียว โมเดลการแบ่งพยางค์สร้างจากอัลกอริทึมคอนดิชันนอลแรนดอมฟิลด์ โดยใช้ค่าลักษณะสำคัญเอ็นแกรมของตัวอักษร งานวิจัยนี้แสดงให้เห็นถึงการใช้ประโยชน์จากพยางค์ภาษาอังกฤษในการทำนายคำทับศัพท์ในภาษาจีน ซึ่งในงานวิจัยการหาคำอ่านของคำภาษาอังกฤษในหน่วยเสียงภาษาไทยเองก็มีความใกล้เคียงกันเพราะมีแนวคิดในลักษณะเดียวกัน คือ การแปลงจากอักขระในภาษาหนึ่งไปเป็นอักขระหรือหน่วยเสียงในอีกภาษาหนึ่ง ข้อมูลพยางค์ของภาษาต้นทางจึงมีความสำคัญอย่างมากในการทำนายอักขระหรือหน่วยเสียงในอีกภาษา ซึ่งหากพิจารณาในงานการหาคำอ่านภาษาไทยจากคำภาษาอังกฤษแล้วนั้น พยางค์ภาษาอังกฤษก็น่าจะมีส่วนช่วยในการทำนายเสียงวรรณยุกต์ไทยด้วยเช่นกัน ดังนั้นในวิทยานิพนธ์นี้จึงได้ทำการทดลองโดยใช้ข้อมูลจากพยางค์ภาษาอังกฤษกับโมเดลการทำนายเสียงวรรณยุกต์ แต่มีสิ่งหนึ่งที่น่าสนใจก็คือ การแบ่งพยางค์ภาษาอังกฤษโดยทั่วไปกับการแบ่งพยางค์ภาษาอังกฤษเพื่อใช้ในการทำนายข้ามภาษา ก็มีความแตกต่างกัน เพราะการแบ่งพยางค์ภาษาอังกฤษโดยทั่วไปจะแบ่งตามการอ่านของเจ้าของภาษา แต่สำหรับงานการแปลงข้ามภาษาการแบ่งพยางค์อาจมีความแตกต่างกันเพราะลักษณะการอ่านของผู้อ่านที่ไม่ใช่เจ้าของภาษาก็อาจจะอ่านออกเสียงคนละแบบ และมีจำนวนพยางค์ไม่เท่ากันได้ ดังนั้นหากต้องใช้การแบ่งพยางค์ภาษาอังกฤษเพื่อช่วยในการทำนายอักขระหรือหน่วยเสียงไทย ก็จำเป็นที่จะต้องสร้างโมเดลการแบ่งพยางค์ภาษาอังกฤษขึ้นมาใหม่ ตามการอ่านของคนไทย ไม่สามารถนำโมเดลการแบ่งพยางค์ภาษาอังกฤษของเจ้าของภาษามาใช้ได้โดยตรง

บทที่ 3
ขั้นตอนการดำเนินงานวิจัย

ในบทนี้จะกล่าวถึงภาพรวมและองค์ประกอบของระบบที่เสนอ รวมทั้งรายละเอียดในแต่ละองค์ประกอบย่อย ตั้งแต่ โมเดลการทำนายหน่วยเสียง (Phone Prediction Model), กระบวนการประมวลผลภายหลังจากการทำนายหน่วยเสียง (Post-processing) และโมเดลการทำนายเสียงวรรณยุกต์ (Tone Prediction Model) ซึ่งจะกล่าวถึงรายละเอียดของค่าลักษณะสำคัญหลักที่นำเสนอ ซึ่งก็คือ มาตราของพยัญชนะสะกด และ พยางค์ภาษาอังกฤษ

3.1 ภาพรวมของระบบ

กระบวนการสร้างคำอ่านจากคำภาษาอังกฤษ มีทั้งหมด 3 องค์ประกอบหลัก คือ โมเดลการทำนายหน่วยเสียง, กระบวนการประมวลผลหลังจากการทำนายหน่วยเสียง และ โมเดลการทำนายเสียงวรรณยุกต์



รูปที่ 3.1 ภาพรวมการทำงานของระบบ

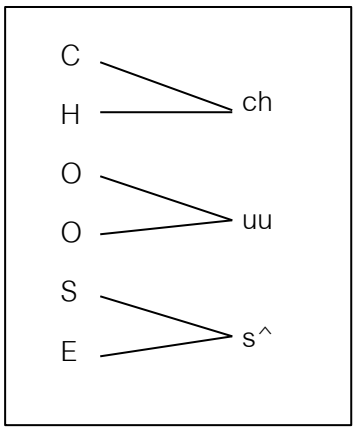
กระบวนการสร้างคำอ่านเริ่มจากการรับข้อความภาษาอังกฤษเพื่อทำการทำนายหน่วยเสียงคำอ่านจากโมเดลการทำนายหน่วยเสียง ผลลัพธ์ที่ได้จากกระบวนนี้มีโอกาสผิดพลาดแบบที่สามารถเห็นได้ชัดเจน เช่น ผิดหลักของพยางค์ในภาษาไทยที่มีลักษณะของพยัญชนะต้นที่ต้องตามด้วยสระและตามด้วยพยัญชนะสะกด หากเป็นเช่นนี้อาจจะสามารถแก้ไขให้ถูกต้องได้ในกระบวนการประมวลผลหลังจากการทำนายหน่วยเสียง หลังจากนั้นก็จะทำการทำนายเสียงวรรณยุกต์ให้กับทุก ๆ พยางค์ในคำนั้น เพื่อให้ได้คำอ่านที่สมบูรณ์เป็นผลลัพธ์ของระบบ ดังแสดงในรูปที่ 3.1 การหาคำอ่านของคำภาษาอังกฤษ “KEYBOARD” โมเดลทำนายหน่วยเสียงจะทำนายคำอ่านได้ผลลัพธ์คือ “kh-ii|b-@@-t^” ซึ่งจะเห็นว่าพยางค์แรกนั้นขาดพยัญชนะสะกดซึ่งผิดหลักของพยางค์ ดังนั้น ในกระบวนการประมวลผลหลังจากการทำนายหน่วยเสียงจะแก้ไขเพื่อให้ถูกต้องเป็นไปตามหลักของพยางค์และจะได้ผลลัพธ์คือ “kh-ii-z^|b-@@-t^” หลังจากนั้นก็จะทำนายเสียงวรรณยุกต์ให้กับทั้งสองพยางค์ด้วยโมเดลทำนายเสียงวรรณยุกต์และได้เป็นคำอ่านที่สมบูรณ์ คือ “kh-ii-z^-0|b-@@-t^-1”

3.2 การปรับแนวและโมเดลทำนายหน่วยเสียง

3.2.1. กระบวนการปรับแนว

ก่อนการสร้างโมเดลทำนายหน่วยเสียงนั้น จำเป็นจะต้องทำการปรับแนวหรือการจับคู่ระหว่างอักขระอังกฤษและหน่วยเสียงไทยก่อน เพื่อกำหนดความสัมพันธ์กันระหว่างอักขระภาษาอังกฤษกับหน่วยเสียงไทย ผู้วิจัยเลือกใช้การปรับแนวแบบกลุ่มต่อกลุ่ม (Many-to-Many Alignment) [19], [20] ซึ่งคือการเชื่อมโยงระหว่างอักขระที่มากกว่าเท่ากับหนึ่งอักขระไปยังหน่วยเสียงที่มากกว่าเท่ากับหนึ่งหน่วยเสียงได้ ซึ่งสามารถแสดงความเชื่อมโยงกันได้เหมาะสมและให้ผลลัพธ์ที่ดีกว่าการปรับแนวแบบหนึ่งต่อหนึ่ง (One-to-One Alignment) [12], [21] ซึ่งกำหนดให้อักขระหนึ่งอักขระต้องเชื่อมโยงกับหน่วยเสียงเดียวเท่านั้น เพราะเนื่องจากจำนวนอักขระในคำภาษาอังกฤษกับหน่วยเสียงไทยมีโอกาสไม่เท่ากันจึงทำให้ต้องมีการกำหนดหน่วยเสียงว่าง (Null) ขึ้นมาเพื่อเชื่อมโยงกับอักขระที่ไม่สามารถจับคู่ได้ เป็นต้น ซึ่งโดยทั่วไปการปรับแนวแบบหนึ่งต่อหนึ่งมักจะใช้เทคนิคไดนามิกไทม์วอร์ปปิงในการเรียนรู้ สำหรับการปรับแนวแบบกลุ่มต่อกลุ่มที่ผู้วิจัยเลือกใช้ [20] อาศัยเทคนิคเอ็กซ์เปคเตชันแม็กซิไมซ์เซชัน (Expectation Maximization : EM) ในการเรียนรู้ ซึ่งจะคำนวณคะแนนเพื่อหาคู่ที่เหมาะสมที่สุดโดยใช้ ฟอร์เวิร์ดเมนีทูเมนีอัลกอริทึม (Forward-many2many Algorithm) และ แบ็กเวิร์ดเมนีทูเมนีอัลกอริทึม (Backward-many2many Algorithm)

ซึ่งในงานวิจัย [20] ได้ทำการทดลองเพื่อเปรียบเทียบความสามารถระหว่างการปรับแนวทั้งสองแบบ พบว่าการปรับแนวแบบกลุ่มต่อกลุ่มทำให้โมเดลการทำนายหน่วยเสียงมีความแม่นยำกว่าอย่างเห็นได้ชัด ซึ่งแสดงถึงความเหมาะสมของการปรับแนวแบบกลุ่มต่อกลุ่มกับงานการหาคำอ่าน ดังแสดงในรูปที่ 3.2 คือการปรับแนวระหว่างอักขระอังกฤษและหน่วยเสียงไทยแบบกลุ่มต่อกลุ่มของคำว่า “CHOOSE”



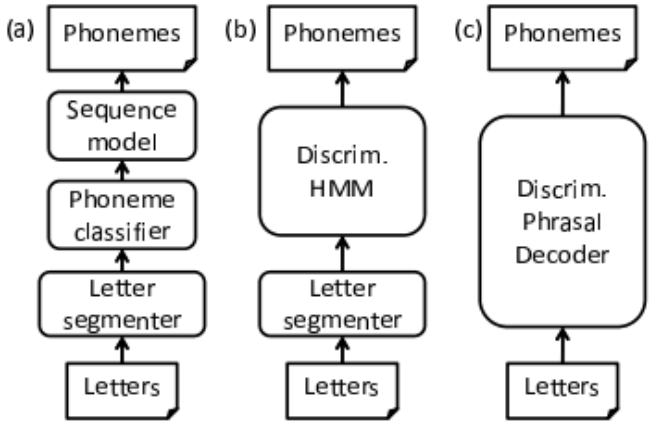
รูปที่ 3.2 ตัวอย่างการปรับแนวแบบกลุ่มต่อกลุ่ม

3.2.2. โมเดลทำนายหน่วยเสียง

กระบวนการสร้างโมเดลการทำนายหน่วยเสียงเริ่มต้นจากการเตรียมข้อมูลที่ได้ทำการปรับแนวแล้วนำมาทำการเรียนรู้จากโมเดลทางสถิติ ซึ่งผู้วิจัยเลือกใช้โมเดลออนไลน์ดิสคริมิเนทีฟเทรนนิงเฟรมเวิร์ค (Online Discriminative Training Framework) [16] ในการสร้างโมเดลทำนายหน่วยเสียง เนื่องจากเป็นโมเดลที่เหมาะสมกับปัญหาการแปลงอักขระเป็นคำอ่าน มากกว่าโมเดลอื่น ๆ และเป็นโมเดลที่ให้ผลลัพธ์ที่ดีที่สุดในงานภาษาอังกฤษ นอกจากนั้นค่าลักษณะสำคัญที่ใช้ในการทำนายหน่วยเสียงพิจารณาทั้งอักขระและหน่วยเสียงไปพร้อม ๆ กัน ต่างจากโมเดลแบบคลาสสิฟิเคชัน (Classification Model) เช่น โมเดลต้นไม้ ที่พิจารณาค่าลักษณะสำคัญจากเฉพาะแต่อักขระเพื่อทำนายหน่วยเสียงเท่านั้น

โดยทั่วไปกระบวนการทำนายหน่วยเสียงจะมี 3 ขั้นตอนหลักที่แยกจากกัน ก็คือ การแบ่งอักขระ (Letter Segmentation), การทำนายหน่วยเสียง (Phoneme Classification) และ การลำดับหน่วยเสียง (Sequence Modeling) ซึ่งคำที่จะถูกทำนายหน่วยเสียงจะถูกทำการแบ่งส่วนอักขระเป็นส่วน ๆ ก่อน จากนั้นจึงทำนายหน่วยเสียงในแต่ละส่วน และทำการรวมหน่วยเสียงเป็นผลลัพธ์ ซึ่งกระบวนการเหล่านี้ทำแยกออกจากกันทำให้เกิดความผิดพลาดแบบต่อเนื่องจากโมเดล

ก่อนหน้าส่งผลไปยังโมเดลต่อไป ดังนั้นออนไลน์ดิสคริมิเนทีฟเทรนนิงเฟรมเวิร์คได้ทำการรวมกระบวนการทั้ง 3 เข้าเป็นโมเดลเดียว นอกจากจะทำให้มีความแม่นยำเพิ่มขึ้น ยังลดความยุ่งยากด้วย ดังแสดงในรูปที่ 3.3 คอลัมน์ a แสดงกระบวนย่อยทั้งหมดซึ่งประกอบไปด้วย 3 กระบวนหลักตามที่ได้กล่าวข้างต้น และในคอลัมน์ b แสดงการรวมกระบวนการทำนายหน่วยเสียงและการลำดับหน่วยเสียงเข้าเป็นโมเดลเดียวกัน โดยใช้โมเดลเพอร์เซปตรอนเอชเอ็มเอ็ม (The Perceptron HMM Model) [23] ส่วนในคอลัมน์ c แสดงถึงการรวมทั้งสามกระบวนการเข้าเป็นโมเดลเดียวกัน โดยประยุกต์ใช้เฟรเซิลดีโคดเดอร์ (Phrasal Decoder) [22] ในการแบ่งส่วนอักขระ



รูปที่ 3.3 กระบวนการในโมเดลการทำนายหน่วยเสียง (คัดลอกจาก [16])

ซึ่งจะสามารถสรุปขั้นตอนการสร้างโมเดลได้เป็น 3 กระบวนการหลัก ๆ ดังนี้

1. การแบ่งส่วนของคำภาษาอังกฤษให้เป็นส่วนย่อย ๆ ซึ่งในกระบวนการเรียนรู้ (Training Process) นั้นการแบ่งส่วนจะถูกทำเสร็จสิ้นตั้งแต่กระบวนการปรับแนวแล้วที่จะแบ่งอักษรอังกฤษเพื่อจับคู่กับหน่วยเสียง แต่ในขณะที่ใช้งานโมเดลเพื่อทำนายหน่วยเสียงการแบ่งส่วนของคำภาษาอังกฤษที่ต้องการทำนายคำอ่านนั้นจะถูกแบ่งด้วย เฟรเซิลดีโคดเดอร์ ซึ่งจะทำการแบ่งส่วนคำภาษาอังกฤษเพื่อใช้ทำนายโดยอัตโนมัติ ทำให้ไม่จำเป็นต้องนำคำภาษาอังกฤษมาแบ่งส่วนก่อนเหมือนดังในกระบวนเรียนรู้ โดยเฟรเซิลดีโคดเดอร์นั้นคือโมเดลที่ใช้ในการแปลภาษาในระดับวลี (Phrase-Based Statistical Machine Translation) ซึ่งเมื่อเปรียบเทียบกับงานการหาคำอ่านก็สามารถมองได้ว่าเป็นการแปลหน่วยอักขระไปเป็นหน่วยเสียง ซึ่งด้วยเหตุนี้เฟรเซิลดีโคดเดอร์จึงเหมาะสมกับการประยุกต์ใช้ในกระบวนการแบ่งอักขระออกเป็นส่วน ๆ และทำการค้นหารูปแบบที่เหมาะสมที่สุด หรือก็คือการสร้างรูปแบบที่เป็นไปได้ทั้งหมดเพื่อคำนวณคะแนนและเลือกรูปแบบที่ได้คะแนนมากที่สุดเป็นคำตอบ

2. ทำการสร้างโมเดลจากการให้คะแนนกับแต่ละคู่อักขระกับหน่วยเสียงด้วยการใช้ค่าลักษณะสำคัญต่าง ๆ ดังนี้ บริบท (Context), เอ็นแกรม (N-gram), ลิเนียร์เชน (Linear-chain), มาร์คอฟออร์เดอร์ (Markov Order) และ จอยนต์เอ็นแกรม (Joint N-gram) โดยค่าลักษณะสำคัญต่าง ๆ เหล่านี้ สามารถแสดงได้ในรูปแบบลำดับ ดังในตารางที่ 3.1 โดย x_i คือ สายอักขระย่อย (Substring) ของอักษรอังกฤษที่ถูกแบ่งส่วน, y_i ก็คือสายอักขระย่อยของหน่วยเสียง และ c คือ ค่าความกว้างของหน่วยบริบทที่พิจารณา (Window Length) ผู้วิจัยเลือกใช้ค่าพารามิเตอร์ต่าง ๆ ดังนี้ ค่าความกว้างของหน่วยบริบทที่พิจารณาเท่ากับ 5 หน่วย, เอ็นแกรม 11 หน่วย, มาร์คอฟออร์เดอร์ 1 หน่วย, จอยนต์เอ็นแกรม 4 หน่วย ซึ่งค่าเหล่านี้เป็นค่าที่ให้ผลลัพธ์ที่ดีที่สุดในงานวิจัยภาษาอังกฤษ โดยได้มีการทดลองในงานวิจัย [16], [17]

3. ทำการปรับค่าน้ำหนักในทุก ๆ รอบของการเรียนรู้ เพื่อให้โมเดลมีความแม่นยำมากขึ้น โดยใช้อัลกอริทึมมาร์จินอินฟิวส์รีแล็กซ์ แทนที่จะใช้อัลกอริทึมเดอะเพอร์เซปตรอนที่อยู่ในโมเดลอยู่แล้วมาเป็นอัลกอริทึมในการปรับค่าน้ำหนัก เพราะเนื่องจากอัลกอริทึมเดอะเพอร์เซปตรอนเป็นอัลกอริทึมที่ตรงไปตรงมา ในแต่ละรอบของการวนเรียนรู้นั้นจะใช้ผลลัพธ์ที่ถูกต้องจากครั้งก่อนในการนำมาปรับค่าน้ำหนัก (Weight) เท่านั้น ดังนั้นหากในแต่ละรอบของการวนเรียนรู้นั้นได้ผลลัพธ์ที่ถูกต้องบ่อย ๆ ก็จะไม่มีการปรับค่าน้ำหนักเกิดขึ้นเลย ด้วยเหตุนี้อาจทำให้โมเดลไม่ได้คำนึงถึงผลลัพธ์ที่ผิดที่ควรจะถูกปรับค่าน้ำหนักด้วยเช่นกัน ในขณะที่อัลกอริทึมมาร์จินอินฟิวส์รีแล็กซ์ (Margin Infused Relaxed Algorithm: MIRA) [24] มีความเหมาะสมและละเอียดกว่า การปรับค่าน้ำหนักของอัลกอริทึมมาร์จินอินฟิวส์รีแล็กซ์จะอาศัยผลลัพธ์ของการวนเรียนรู้ที่ดีที่สุดเป็นจำนวนเอ็นผลลัพธ์ (n-best) ซึ่งก็เพื่อให้มีการปรับค่าน้ำหนักของหน่วยเสียงที่ผิดให้มีความแตกต่างจากหน่วยเสียงที่ถูกมากขึ้น เพื่อลดโอกาสที่โมเดลจะทำนายผิด นอกเหนือจากการปรับค่าน้ำหนักกับเฉพาะหน่วยเสียงที่ถูกเท่านั้น ซึ่งค่าเอ็นที่เหมาะสมและให้ผลลัพธ์ที่ดีจากการทดลองในงานวิจัย [16] ก็คือ 10

หลังจากโมเดลถูกสร้างจากกระบวนการเรียนรู้จนเสร็จสิ้นแล้ว การทำนายหน่วยเสียงก็จะใช้ค่าลักษณะสำคัญต่าง ๆ ที่ได้กล่าวไปนี้ มาคำนวณคะแนนและหาหน่วยเสียงที่ได้คะแนนมากที่สุดเป็นผลลัพธ์ต่อไป โดยจะใช้เฟรเซิลดีโคดเดอร์ ในการแบ่งส่วนอักขระของคำภาษาอังกฤษที่ต้องการจะทำนายหน่วยเสียงและคำนวณคะแนน เพื่อหารูปแบบอักขระและหน่วยเสียงที่ให้คะแนนมากที่สุดเป็นคำตอบ

ตารางที่ 3.1 รูปแบบของค่าลักษณะสำคัญสำหรับโมเดลทำนายหน่วยเสียง [17]

ค่าลักษณะสำคัญ	**รูปแบบ**
บริบท (Context)	$x_{i-c}\ y_i$ … $x_{i+c}\ y_i$ $x_{i-c}\ x_{i-c+1}\ y_i$ … $x_{i+c-1}\ x_{i+c}\ y_i$ …… x_{i-c} … $x_{i+c}\ y_i$
มาร์คอฟออร์เดอร์ (Markov order)	$y_{i-1}\ y_i$
ลิเนียร์เชน (Linear-Chain)	$x_{i-c}\ y_{i-1}\ y_i$ … $x_{i+c}\ y_{i-1}\ y_i$ $x_{i-c}\ x_{i-c+1}\ y_{i-1}\ y_i$ … $x_{i+c-1}\ x_{i+c}\ y_{i-1}\ y_i$ …… x_{i-c} … $x_{i+c}\ y_{i-1}\ y_i$
จอยนต์เอ็นแกรม (Joint n-gram)	$x_{i+1-n}\ y_{i+1-n}\ x_i\ y_i$ … $x_{i-1}\ y_{i-1}\ x_i\ y_i$ $x_{i+1-n}\ y_{i+1-n}\ x_{i+2-n}\ y_{i+2-n}\ x_i\ y_i$ … $x_{i-2}\ y_{i-2}\ x_{i-1}\ y_{i-1}\ x_i\ y_i$ …… $x_{i+1-n}\ y_{i+1-n}$ … $x_{i-1}\ y_{i-1}\ x_i\ y_i$

3.3 กระบวนการประมวลผลหลังจากการทำนายหน่วยเสียง

เนื่องจากผลลัพธ์จากโมเดลการทำนายหน่วยเสียงนั้น มีโอกาสผิดพลาดซึ่งอาจเกิดขึ้นเฉพาะเจาะจงกับโมเดลที่เลือกใช้และพิจารณาว่าเป็นจุดบกพร่องของโมเดลได้ ซึ่งความผิดพลาดเหล่านี้มักผิดหลักของพยางค์ภาษาไทย ที่ทุก ๆ พยางค์จะประกอบไปด้วย เสียงพยัญชนะต้น เสียงสระ เสียงพยัญชนะสะกด ตามลำดับ ดังนั้นผู้วิจัยจึงทำการวิเคราะห์ถึงความผิดพลาดทั้งหมดที่เกิดจากโมเดลการทำนายหน่วยเสียง เพื่อเป็นแนวทางในการสร้างกฎเพื่อแก้ไขความผิดพลาดเหล่านี้ต่อไป

ผู้วิจัยได้จำแนกความผิดพลาดจากโมเดลทำนายหน่วยเสียง โดยใช้ข้อมูลจากพจนานุกรมคำอ่านภาษาไทยจากคำภาษาอังกฤษ [6] ในการทดสอบ ซึ่งมีทั้งหมด 8,283 คำ แบ่งข้อมูล 90% เป็นข้อมูลที่ใช้ในการเรียนรู้ และข้อมูลส่วนที่เหลืออีก 10% เป็นข้อมูลในการทดสอบ ซึ่งมีทั้งหมด 828 คำ ในจำนวนคำทดสอบทั้งหมดนี้ มีคำที่ผิดพลาดอยู่ทั้งหมด 301 คำ และในคำที่ผิดพลาดนี้มีตำแหน่งที่ผิดพลาดรวมแล้วทั้งสิ้น 499 ตำแหน่ง โดยสามารถแบ่งได้เป็นสองหมวดหลัก ๆ คือ ความผิดพลาดที่มีโอกาสในการแก้ไขได้ (Correctable Error) และ ความผิดพลาดที่ไม่มีโอกาสในการแก้ไขได้ (Un-Correctable Error) รายละเอียดของความผิดพลาดแสดงใน ตารางที่ 3.2 ความผิดพลาดส่วนใหญ่เป็นความผิดพลาดที่ไม่มีโอกาสในการแก้ไขได้ มีจำนวน 87.37% ของทั้งหมด แต่สิ่งที่น่าสนใจคือความผิดพลาดที่มีโอกาสแก้ไขได้ มีจำนวน 12.63% โดยสามารถแบ่งได้เป็น 6 ลักษณะ ดังนี้

1. ตกเสียงพยัญชนะต้น

คำอ่านของคำภาษาอังกฤษ “ALLOWABLE” คือ “z-a-z^|l-aa-w^|z-ee-z^|b-q-l^” แต่โมเดลทำนายหน่วยเสียงทำนายได้ผลลัพธ์คือ “z-a-z^|l-aa-w^|ee-z^|b-q-l^” ซึ่งจะเห็นว่าตกเสียงพยัญชนะต้น “z” ที่พยางค์ที่สาม กรณีแบบนี้สามารถแก้ไขได้ด้วยกฎการเติมเสียงพยัญชนะต้น “z” ระหว่างเสียงพยัญชนะสะกดที่ติดกับเสียงสระ แต่กฎนี้ไม่อาจแก้ไขได้ทุกกรณีของการตกเสียงพยัญชนะต้น เพราะเสียงพยัญชนะต้นมีทั้งหมด 38 เสียง ซึ่งผู้วิจัยวิเคราะห์จากข้อผิดพลาดแล้วพบว่า เกิดการตกเสียง “z” บ่อยที่สุด ดังนั้นจึงเลือกที่แก้ไขด้วยการเติมเสียง “z” ในกรณีนี้

2. ตกเสียงสระ

คำอ่านของคำภาษาอังกฤษ “ALPHABETIC” คือ “z-a-l^|f-aa-z^|b-ee-z^|th-i-k^” แต่โมเดลทำนายหน่วยเสียงทำนายได้ผลลัพธ์คือ “z-l^|f-aa-z^|b-ee-z^|th-i-k^” ซึ่งตกเสียงสระ “a”

ในพยางค์แรก กรณีแบบนี้สามารถแก้ไขได้ด้วยการเติมเสียงสระ “a” ระหว่างเสียงพยัญชนะต้นที่ติดกับเสียงพยัญชนะสะกด แต่กฎนี้ไม่อาจแก้ไขได้ทุก ๆ กรณีของการตกเสียงสระเช่นเดียวกันกับการตกเสียงพยัญชนะต้น เพราะเสียงสระมีทั้งหมดถึง 24 เสียง ผู้วิจัยได้เลือกการเติมเสียง “a” หรือเสียง “อะ” เนื่องจากเป็นเสียงสระพื้นฐานที่สามารถพบได้บ่อย ถึงแม้การเติมเสียง “a” จะไม่ถูกเสมอไป แต่ก็อย่างน้อยการเติมเสียงสระ “a” ก็จะทำให้พยางค์นั้นสมบูรณ์ขึ้น และจะไม่ทำให้ผลลัพธ์ผิดมากไปกว่าเดิมอย่างแน่นอน

3. ตกเสียงพยัญชนะสะกด

คำอ่านของคำภาษาอังกฤษ “NORWAY” คือ “n-@@-z^|w-ee-z^” แต่โมเดลทำนายหน่วยเสียงทำนายได้ผลลัพธ์คือ “n-@@|w-ee-z^” ซึ่งจะเห็นว่าตกเสียงพยัญชนะสะกด “z^” ที่พยางค์ที่แรก กรณีแบบนี้สามารถแก้ไขได้ด้วยการเติมเสียงพยัญชนะสะกด “z^” ระหว่างเสียงสระที่ติดกับเสียงพยัญชนะต้น แต่กฎนี้ไม่อาจแก้ไขได้ทุกกรณีของการตกเสียงพยัญชนะสะกดเช่นเดียวกันกับการตกเสียงพยัญชนะต้นและการตกเสียงสระ เพราะเสียงพยัญชนะสะกดมีทั้งหมด 12 เสียง ซึ่งผู้วิจัยวิเคราะห์จากข้อผิดพลาดแล้วพบว่า เกิดการตกเสียง “z^” มากที่สุด ดังนั้นจึงเลือกแก้ไขด้วยการเติมเสียง “z^”

4. เสียงพยัญชนะต้นสองเสียงติดกัน

คำอ่านของคำภาษาอังกฤษ “WILSON” คือ “w-i-l^|s-a-n” แต่โมเดลทำนายหน่วยเสียงทำนายได้ผลลัพธ์คือ “w-i-l^|l-s-a-n” ซึ่งจะเห็นว่าเกิดเสียงพยัญชนะต้นเกินมาและติดกันที่พยางค์ที่สอง ซึ่งผิดต่อหลักพยางค์ในภาษาไทย ดังนั้น การแก้ไขในกรณีนี้ คือ การตัดเสียงพยัญชนะต้นตัวใดตัวหนึ่งออก เพื่อให้ถูกต้องตามหลักพยางค์ในภาษาไทย จากการวิเคราะห์ข้อผิดพลาดของผู้วิจัย พบว่าพยัญชนะต้นสองเสียงที่ติดกัน มักจะมีเสียงใดเสียงหนึ่งที่เป็นคำตอบที่ถูก ดังนั้นการเลือกเสียงใดเสียงหนึ่งเป็นคำตอบก็มีโอกาสถูกหรือผิดเท่า ๆ กัน ผู้วิจัยได้เลือกเสียงพยัญชนะต้นเสียงที่สองเป็นคำตอบ และตัดเสียงแรกทิ้งไป การแก้ไขด้วยวิธีนี้ก็ไม่สามารถแก้ไขได้ทั้งหมดเพราะก็มีโอกาสที่ทั้งสองเสียงไม่มีเสียงใดเป็นเสียงที่ถูกเลย แต่ไม่ว่าอย่างไรก็จะไม่ทำให้ผลลัพธ์แย่ลงไปกว่าเดิม นอกจากนั้นหากพิจารณาถึงการเกิดเสียงพยัญชนะต้นสองตัวติดกันอาจเกิดได้ในอีกลักษณะ เช่น “w-s-a-n” เกิดการหายไปของเสียงสระและเสียงพยัญชนะสะกดของพยางค์แรก การแก้ไขกับกรณีนี้คือการเติมเสียงสระและเสียงพยัญชนะสะกดระหว่างเสียงพยัญชนะต้นสองตัวที่ติดกัน แต่กรณีแบบนี้ผู้วิจัยไม่พบเกิดขึ้นในข้อมูล

5. เสียงสระสองเสียงติดกัน

คำอ่านของคำภาษาอังกฤษ “ADIEU” คือ “z-a-z^|d-uu-z^” แต่โมเดลทำนายหน่วยเสียงทำนาย “z-a-z^|d-iia-uu-z^” ซึ่งจะเห็นว่าเกิดเสียงสระติดกันที่พยางค์ที่สอง ซึ่งผิดต่อหลักพยางค์ในภาษาไทย ดังนั้น การแก้ไขในกรณีนี้ คือ การตัดเสียงสระตัวใดตัวหนึ่งออก เพื่อให้ถูกต้องตามหลักพยางค์ในภาษาไทย จากการวิเคราะห์ข้อผิดพลาดของผู้วิจัย พบว่าสระสองเสียงที่ติดกันมักจะมีเสียงใดเสียงหนึ่งที่เป็นคำตอบที่ถูก ดังนั้นการเลือกเสียงใดเสียงหนึ่งเป็นคำตอบก็มีโอกาสถูกหรือผิดเท่า ๆ กัน ผู้วิจัยได้เลือกเสียงสระเสียงที่สองเป็นคำตอบ และตัดเสียงแรกทิ้งไป เช่นเดียวกันกับกรณีเสียงพยัญชนะต้นติดกัน

6. เสียงพยัญชนะสะกดสองเสียงติดกัน

คำอ่านของคำภาษาอังกฤษ “PROGRAMMING” คือ “pr-oo-z^|kr-x-m^|m-i-ng^” แต่โมเดลทำนายหน่วยเสียงทำนายได้ผลลัพธ์คือ “pr-oo-z^|kr-x-m^-m^|m-i-ng^” ซึ่งจะเห็นว่าเกิดเสียงพยัญชนะสะกดติดกันที่พยางค์ที่สอง ซึ่งผิดต่อหลักพยางค์ในภาษาไทย ดังนั้น การแก้ไขในกรณีนี้ คือ การตัดเสียงพยัญชนะสะกดตัวใดตัวหนึ่งออก เพื่อให้ถูกต้องตามหลักพยางค์ในภาษาไทย จากการวิเคราะห์ข้อผิดพลาดของผู้วิจัย พบว่าพยัญชนะสะกดสองเสียงที่ติดกัน มักจะมีเสียงใดเสียงหนึ่งที่เป็นคำตอบที่ถูก ผู้วิจัยได้เลือกเสียงที่สองเป็นคำตอบ เช่นเดียวกันกับกรณีเสียงพยัญชนะต้นติดกันและเสียงสระติดกันตามที่ได้กล่าวไว้ข้างต้น

จากการสังเกตข้อผิดพลาดเหล่านี้ทำให้ผู้วิจัยสร้างกฎขึ้นมาเพื่อจัดการกับกรณีเหล่านี้ ซึ่งสามารถสรุปได้ดังนี้

1. เติมเสียงพยัญชนะต้น “z” ระหว่างเสียงพยัญชนะสะกดที่ติดกับเสียงสระ
2. เติมเสียงสระ “a” ระหว่างเสียงพยัญชนะต้นที่ติดกับเสียงพยัญชนะสะกด
3. เติมเสียงพยัญชนะสะกด “z^” ระหว่างเสียงสระที่ติดกับเสียงพยัญชนะต้น
4. เลือกเสียงที่สองเป็นคำตอบ เมื่อเจอเสียงพยัญชนะต้นติดกัน หรือ เสียงสระติดกัน หรือ เสียงพยัญชนะสะกดติดกัน

โดยกฎทั้งหมดที่กำหนดขึ้นจะทำให้ผลลัพธ์มีความถูกต้องตามหลักการพยางค์ได้ ซึ่งอาจจะไม่สามารถทำให้ผลลัพธ์ถูกต้องได้ทุกกรณี แต่ก็จะไม่ทำให้ผลลัพธ์ด้อยกว่าเดิม

ตารางที่ 3.2 ความผิดพลาดที่เกิดขึ้นจากโมเดลทำนายหน่วยเสียง

ชนิดของความผิดพลาด	**รายละเอียด**	**จำนวน (%)**
ความผิดพลาดที่มีโอกาสแก้ไขได้	ตกเสียงพยัญชนะต้น	5.61
	ตกเสียงสระ	0.60
	ตกเสียงพยัญชนะสะกด	3.01
	เสียงพยัญชนะต้นสองเสียงติดกัน	1.20
	เสียงสระสองเสียงติดกัน	1.40
	เสียงพยัญชนะสะกดสองเสียงติดกัน	0.80
ความผิดพลาดที่ไม่มีโอกาสแก้ไขได้	ทำนายเสียงพยัญชนะต้นผิด	11.02
	ทำนายเสียงสระผิด	47.29
	ทำนายเสียงพยัญชนะสะกดผิด	22.65
	ทำนายพยางค์เกิน	4.41
	ทำนายพยางค์ขาด	2.00

3.4 โมเดลการทำนายเสียงวรรณยุกต์

ภาษาไทยเป็นภาษาที่มีเสียงวรรณยุกต์ ดังนั้นการหาคำอ่านของคำภาษาอังกฤษก็จำเป็นจะต้องกำหนดเสียงวรรณยุกต์ให้กับทุก ๆ พยางค์ด้วย แต่การกำหนดเสียงวรรณยุกต์ไม่ใช่เรื่องง่ายและตรงไปตรงมา เพราะการทำนายเสียงวรรณยุกต์จากหน่วยเสียงที่ได้มาจากโมเดลการทำนายหน่วยเสียงนั้นไม่เหมือนกับการทำนายเสียงวรรณยุกต์จากรูปอักขระภาษาไทยที่มีหลักการชัดเจน ดังนั้นค่าลักษณะสำคัญที่เกี่ยวข้องกับหลักภาษาไทยจึงเป็นสิ่งที่สำคัญที่สามารถช่วยให้การทำนายเสียงวรรณยุกต์มีความแม่นยำ ในงานวิจัยอ้างอิง [6] ได้มีการเสนอค่าลักษณะสำคัญ 9 อย่างคือ เสียงพยัญชนะต้น, เสียงสระ, เสียงพยัญชนะสะกด, กลุ่มของพยัญชนะต้น (สูง กลาง ต่ำ), ความยาวสระ (สั้น ยาว), จำนวนพยางค์ในคำ, คำเป็นคำตาย, ตำแหน่งของพยางค์ในคำ และ เสียงวรรณยุกต์ของพยางค์ก่อนหน้า ผู้วิจัยได้ทำการศึกษาถึงค่าลักษณะสำคัญที่มีผลต่อการทำนายวรรณยุกต์เพื่อพัฒนาให้โมเดลการทำนายเสียงวรรณยุกต์มีความแม่นยำมากขึ้น โดยเสนอสองค่าลักษณะสำคัญเพิ่มเติม คือ มาตราสะกด และ พยางค์ภาษาอังกฤษ

3.4.1. **มาตราสะกด**

ในภาษาไทยมีมาตราสะกดทั้งหมด 8 มาตรา ซึ่งไม่ว่าจะเป็นอักษรใด ก็จะสามารถแบ่งตามการออกเสียงได้เป็น 8 มาตรา คือ มาตราแม่กก แม่กง แม่กด แม่กน แม่กบ แม่กม แม่เกย และแม่เกอว แต่เมื่อพิจารณาเสียงพยัญชนะสะกดที่ใช้ในการสังเคราะห์เสียงกลับมีทั้งหมด 12 เสียง ที่เป็นเช่นนั้นก็เพราะว่ามีการกำหนดเสียงพยัญชนะสะกดที่เป็นเสียงจากคำต่างประเทศเพิ่มเติมขึ้นมาอีก 4 เสียง คือ “f^”, “l^”, “s^” และ “ch^” ซึ่งเมื่อพิจารณาแล้วจะพบว่า เสียง “s^”, “ch^” สามารถจัดเป็นกลุ่มเดียวกันกับมาตราแม่กดได้ เสียง “f^” สามารถจัดอยู่ในกลุ่มเดียวกันกับมาตราแม่กบได้ และ เสียง “l^” สามารถจัดอยู่ในกลุ่มเดียวกันกับมาตราแม่กน เพราะออกเสียงในลักษณะเดียวกันและไม่ได้ส่งผลต่อเสียงวรรณยุกต์ เช่น คำว่า “JUST” และ “JUDGE” มีคำอ่านคือ “c-a-s^-3” และ “c-a-t^-3” ตามลำดับ ทั้งสองคำมีเสียงวรรณยุกต์เดียวกันและอ่านในลักษณะเดียวกัน ต่างกันตรงที่มีเสียงเอสในภาษาอังกฤษที่ข้างท้ายของพยางค์ นอกจากนั้นในข้อมูลที่นำมาใช้ในการทดลอง [6] พบว่ามีเสียงพยัญชนะสะกดที่เพิ่มเติมเสียงเอสเข้าไปกับเสียงพยัญชนะสะกดดั้งเดิม เช่น “zs^”, “ps^”, “ts^”, “ks^”, “ns^”, “ms^”, “ngs^”, “js^” และ “ws^” เป็นต้น เพราะเนื่องจากคำภาษาอังกฤษที่ลงท้ายด้วยเสียงเอส ไม่สามารถเขียนเป็นหน่วยเสียงไทยในพยางค์เดียวได้ ดังนั้นจึงมีการเพิ่มหน่วยเสียงเหล่านี้เข้าไป แต่ในการใช้งานจริงระบบสังเคราะห์เสียงที่ออกแบบมาจะต้องรองรับหน่วยเสียงเหล่านี้ด้วย ซึ่งจะเห็นว่าหน่วยเสียงเหล่านี้ไม่มีผลต่อการทำนายเสียงวรรณยุกต์เพราะเป็นเพียงเสียงเอสที่ลงท้ายของพยางค์เท่านั้น และโอกาสที่จะพบหน่วยเสียงเหล่านี้ก็มีน้อยมากเมื่อเทียบกับหน่วยเสียงเดียวกันที่ไม่มีเสียงเอส การแยกพิจารณาอาจทำให้ความแม่นยำในการทายเสียงวรรณยุกต์ของพยางค์ที่ลงท้ายด้วยเสียงเอสไม่แม่นยำ เนื่องจากข้อมูลในการเรียนรู้อาจไม่พอ จึงควรจัดหน่วยเสียงเหล่านี้เข้าไปเป็นกลุ่มเดียวกันกับมาตราสะกดของหน่วยเสียงดั้งเดิม ดังนั้นผู้วิจัยจึงเสนอมาตราสะกดในภาษาไทย เป็นค่าลักษณะสำคัญหนึ่งในการทดลอง

3.4.2. **พยางค์ภาษาอังกฤษ**

ค่าลักษณะสำคัญทั้งหมดที่กล่าวมาก่อนหน้านี้ ทั้งในงานวิจัยก่อนหน้าและมาตราสะกดต่างเป็นการใช้ประโยชน์จากการพิจารณาเฉพาะหน่วยเสียงเท่านั้น ซึ่งทำให้เกิดข้อจำกัดหลายอย่าง เช่น คำภาษาอังกฤษหลาย ๆ คำมีการออกเสียงเหมือนกัน แต่เสียงวรรณยุกต์ต่างกัน หากใช้เพียงหน่วยเสียงในการกำหนดค่าลักษณะสำคัญแล้ว ก็จะมองว่าคำทั้งสองเป็นคำที่ออกเสียงวรรณยุกต์เดียวกัน ทั้งที่ถ้าหากมองที่อักขระภาษาอังกฤษของพยางค์เหล่านั้นก็จะเห็นว่ามีความแตกต่างกัน เช่น คำว่า “COMPLEMENT” และ “BUSINESSMEN” คำอ่านของทั้งพยางค์

“MENT” และ “MEN” ต่างเหมือนกัน คือ “m-e-n^” แต่เสียงวรรณยุกต์ของทั้งสองพยางค์กลับต่างกัน คือ “MENT” มีเสียงวรรณยุกต์เสียงตรี แต่ “MEN” มีเสียงวรรณยุกต์สามัญ ซึ่งหากใช้ค่าลักษณะสำคัญที่กล่าวถึงก่อนหน้านี้ เช่น เสียงพยัญชนะต้น, เสียงสระ, เสียงพยัญชนะสะกด, คำเป็นคำตาย, กลุ่มพยัญชนะต้น และความยาวสระ ก็จะเห็นว่าได้ค่าที่เหมือนกัน และส่งผลให้อาจไม่สามารถแยกความต่างของเสียงวรรณยุกต์ได้ ดังนั้นผู้วิจัยจึงได้ทำการวิเคราะห์จากข้อมูลพจนานุกรมคำอ่านที่ใช้ในการทดลอง ที่มีจำนวนคำทั้งหมด 8,283 คำ เพื่อศึกษาดูตัวอย่างของพยางค์ที่เกิดในลักษณะเช่นนี้ รายละเอียดแสดงในตารางที่ 3.3 พบว่าคำอ่านที่มีหน่วยเสียง “m-e-n^” มีทั้งสิ้น 145 พยางค์ และสามารถแบ่งตามอักขระพยางค์ได้ 3 รูปแบบ คือ “MENT”, “MEN” และ “MEND” ซึ่ง 99% ของอักขระพยางค์ “MENT” เป็นเสียงวรรณยุกต์ตรี, 88% ของอักขระพยางค์ “MEN” เป็นเสียงวรรณยุกต์สามัญ และอักขระพยางค์ “MEND” ทุกพยางค์เป็นเสียงวรรณยุกต์สามัญ ในขณะที่คำอ่านที่มีหน่วยเสียง “t-ii-z^” มีทั้งสิ้น 37 พยางค์ สามารถแบ่งตามอักขระพยางค์ได้ 5 รูปแบบ คือ “TY”, “TIES”, “TEE”, “TI” และ “TE” ซึ่งทุกอักขระพยางค์ถูกกำหนดด้วยเสียงวรรณยุกต์เดียว เช่นเดียวกันกับคำอ่านในหน่วยเสียง “th-ii-z^” ที่มีทั้งหมด 168 พยางค์ สามารถแบ่งตามอักขระพยางค์ได้ 8 รูปแบบ ซึ่งมีเพียงรูปแบบเดียว คือ “TI” ที่ถูกกำหนดด้วยเสียงวรรณยุกต์สองเสียง ในขณะที่อักขระพยางค์ที่เหลือทั้งหมดถูกกำหนดด้วยเสียงวรรณยุกต์เดียว ซึ่งจะเห็นว่าอักขระพยางค์ของคำภาษาอังกฤษมีความสามารถในการช่วยแยกเสียงวรรณยุกต์ได้ดีมาก

ตารางที่ 3.3 จำนวนพยางค์ของหน่วยเสียงที่แบ่งตามอักขระพยางค์ภาษาอังกฤษและเสียงวรรณยุกต์

หน่วยเสียง	**อักขระพยางค์**	**เสียงวรรณยุกต์**	**จำนวน (พยางค์)**
m-e-n^	MENT	3	98
		0	1
	MEN	0	38
		3	3
		2	2
	MEND	0	3
t-ii-z^	TY	2	30
	TIES	2	1
	TEE	0	3
	TI	0	2
	TE	0	1
th-ii-z^	TY	2	95
	TIES	2	15
	THY	2	7
	TIE	2	2
	TE	0	24
	THE	0	15
	TEE	0	5
	TI	0	3
		1	2

นอกจากอักขระพยางค์ของคำภาษาอังกฤษที่พิจารณาอยู่แล้วนั้น อักขระพยางค์ของพยางค์ก่อนหน้าและพยางค์ต่อไปก็มีส่วนช่วยในการกำหนดเสียงวรรณยุกต์เช่นกัน โดยผู้วิจัยได้ทำการวิเคราะห์จากข้อมูลในพจนานุกรมคำอ่าน และพบว่าเมื่อพิจารณาอักขระพยางค์ของพยางค์ภาษาอังกฤษพยางค์ถัดไปร่วมกับคำเป็นคำตาย จะพอสามารถแบ่งแยกเสียงวรรณยุกต์ได้ดังแสดงในตารางที่ 3.4 เช่น หากอักขระพยางค์ถัดไปคือ “TION” และพยางค์ที่พิจารณาอยู่เป็นคำตายแล้ว 98% จะเป็นเสียงวรรณยุกต์ตรี ตัวอย่างเช่นคำว่า “TRANSACTION” มีคำอ่านคือ “thr-aa-n^-0|s-x-k^-3|ch-a-n^-2” ที่พยางค์ที่สามเป็นคำตายและอักขระพยางค์ถัดไปคือ “TION” เป็นต้น ในขณะที่ถ้าหากพยางค์ที่พิจารณาเป็นคำเป็นแล้ว 100% ของพยางค์เหล่านั้น จะเป็นเสียงวรรณยุกต์สามัญทั้งหมด เช่น คำว่า “SOLUTION” มีคำอ่านคือ “s-oo-z^-0|l-uu-z^-0|ch-a-n^-2” ที่พยางค์ที่สองเป็นคำเป็น และอักขระพยางค์ถัดไปคือ “TION” ในขณะที่ในตัวอย่างที่สองเมื่อพิจารณาอักขระพยางค์ถัดไป “BLE” ร่วมกับ คำเป็นคำตาย ก็จะสามารถแยกเสียงวรรณยุกต์ได้ค่อนข้างมีประสิทธิภาพ ในลักษณะเดียวกันกับตัวอย่างแรก

ตารางที่ 3.4 จำนวนพยางค์ที่เกิดขึ้นจากความสัมพันธ์ระหว่างอักขระพยางค์ถัดไป คำเป็นคำตาย และเสียงวรรณยุกต์

อักขระพยางค์ของพยางค์ถัดไป	**คำเป็นคำตายของพยางค์ที่พิจารณา**	**เสียงวรรณยุกต์**	**จำนวน (พยางค์)**
TION	คำตาย	3	113
		1	3
	คำเป็น	0	347
BLE	คำตาย	1	21
		3	11
	คำเป็น	0	104

จากข้อมูลการวิเคราะห์ที่แสดงให้เห็นถึงความสามารถของค่าลักษณะสำคัญ คือ มาตราพยัญชนะสะกดและอักขระพยางค์ภาษาอังกฤษ ผู้วิจัยจึงนำค่าลักษณะสำคัญทั้งสองอย่างนี้ร่วมกับค่าลักษณะสำคัญที่ได้ถูกนำเสนอในงานวิจัยอ้างอิง [6] ทดลองร่วมกันเพื่อสร้างโมเดลการทำนายเสียงวรรณยุกต์ ซึ่งสรุปค่าลักษณะสำคัญที่ใช้ทั้งหมดได้ดังรายละเอียดในตารางที่ 3.5 สำหรับอัลกอริทึมในการเรียนรู้ ผู้วิจัยใช้โมเดลต้นไม้เป็นอัลกอริทึมสำหรับการเรียนรู้ เนื่องจากเป็นโมเดลที่เหมาะสมและให้ผลลัพธ์ที่ดีกับงานประเภทนี้ และใช้เวลาในการประมวลผลเร็ว

ตารางที่ 3.5 ค่าลักษณะสำคัญที่ใช้สำหรับโมเดลการทำนายเสียงวรรณยุกต์

ค่าลักษณะสำคัญ
เสียงพยัญชนะต้นของพยางค์ที่กำลังพิจารณา
เสียงสระของพยางค์ที่กำลังพิจารณา
เสียงพยัญชนะสะกดของพยางค์ที่กำลังพิจารณา
กลุ่มของพยัญชนะต้นของพยางค์ที่กำลังพิจารณา (สูง กลาง ต่ำ)
ความยาวของเสียงสระของพยางค์ที่กำลังพิจารณา (สั้น ยาว)
มาตราสะกดของพยางค์ที่กำลังพิจารณา
จำนวนพยางค์ของคำที่กำลังพิจารณา
ตำแหน่งของพยางค์ที่กำลังพิจารณา (หน้า กลาง หลัง)
คำเป็นคำตายของพยางค์ที่กำลังพิจารณา
เสียงวรรณยุกต์ของพยางค์ก่อนหน้า
อักขระพยางค์ภาษาอังกฤษของพยางค์ที่กำลังพิจารณา
อักขระพยางค์ภาษาอังกฤษของพยางค์ก่อนหน้า
อักขระพยางค์ภาษาอังกฤษของพยางค์ถัดไป

บทที่ 4

การทดลอง ผลการทดลองและการวิเคราะห์ผลการทดลอง

4.1 การทดลอง

ในส่วนของการทดลองจะกล่าวถึงข้อมูลที่ใช้ในการทดลอง, เครื่องมือที่ใช้ในการทดลอง และวิธีการประเมินผล

4.1.1. ข้อมูลที่ใช้ในการทดลอง

ข้อมูลที่ใช้ในการทดลองในวิทยานิพนธ์นี้ คือ พจนานุกรมคำอ่านในหน่วยเสียงภาษาไทยจากคำภาษาอังกฤษ ซึ่งเป็นข้อมูลชุดเดียวกันกับงานวิจัยอ้างอิง [6] จำนวนคำทั้งหมด 8,283 คำ ซึ่งคำทั้งหมดถูกกำกับเสียงวรรณยุกต์ให้ด้วย สำหรับพยางค์ภาษาอังกฤษผู้วิจัยได้ทำการแบ่งพยางค์เหล่านั้นเอง คำอ่านในพจนานุกรมทั้งหมดมาจากการเลือกโดยผู้ประเมิน 3 คน โดยเลือกคำอ่านที่เป็นเสียงส่วนมาก การทดลองแบ่งข้อมูลเป็นสองส่วน ส่วนแรกคือข้อมูลสำหรับกระบวนการเรียนรู้ 90% และส่วนที่สองคือข้อมูลสำหรับการประเมิน 10% โดยการเลือกคำโดยเว้นไปทีละ 10 คำ เพื่อที่จะได้ข้อมูลสำหรับการประเมินผลที่กระจายตัวอักษรอย่างเท่า ๆ กัน จากนั้นได้นำเทคนิคการแบ่งส่วนข้อมูลเพื่อสลับการเป็นข้อมูลสำหรับการเรียนรู้และข้อมูลสำหรับการประเมินผล (K-Fold Cross Validation) ผู้ทดลองได้แบ่งส่วนข้อมูลเป็น 10 ส่วนเท่า ๆ กัน โดยส่วนแรกเกิดจากการเลือกคำที่หนึ่งแล้วเว้นไปทีละ 10 คำ ข้อมูลส่วนที่สองจะเลือกจากคำที่สองแล้วเว้นไปทีละ 10 คำเช่นกัน ทำเช่นนี้ไปจนครบได้ครบ 10 ชุดข้อมูล การทดลองจะทำ 10 ครั้ง โดยเลือกชุดข้อมูล 1 ใน 10 มาเป็นข้อมูลสำหรับทดสอบ และ 9 ส่วนที่เหลือไว้สำหรับการเรียนรู้ ผลการทดลองจะเป็นค่าเฉลี่ยของทั้ง 10 ครั้งการทดลอง ซึ่งจะทำให้ผลลัพธ์มีความยุติธรรมและน่าเชื่อถือ

4.1.2. เครื่องมือที่ใช้ในการทดลอง

เครื่องมือที่ใช้ในการทดลองสามารถแบ่งเป็น 4 หมวด คือ เครื่องมือสำหรับการปรับแนวระหว่างอักขระอังกฤษและหน่วยเสียงไทย, เครื่องมือสำหรับสร้างโมเดลการทำนายหน่วยเสียง, เครื่องมือสำหรับการสร้างโมเดลการทำนายเสียงวรรณยุกต์ และ เครื่องมืออื่น ๆ

4.1.2.1 **เครื่องมือสำหรับการปรับแนวระหว่างอักขระอังกฤษและหน่วยเสียงไทย**

ผู้วิจัยใช้เครื่องมือ “Many-to-Many alignment model: m2m-aligner” [20] ในการปรับแนวระหว่างอักขระอังกฤษและหน่วยเสียงไทยแบบกลุ่มต่อกลุ่ม เพราะเป็นการปรับแนวที่เหมาะสมและให้ผลลัพธ์ที่ดีที่สุดสำหรับนำไปใช้ในกระบวนการสร้างโมเดลการทำนายหน่วยเสียง โดยการปรับแนวแบบกลุ่มต่อกลุ่มจะแสดงความสัมพันธ์ระหว่างอักขระอังกฤษและหน่วยเสียงไทยได้ดีกว่าแบบอื่น ๆ m2m-aligner ได้ถูกสร้างมาจากอัลกอริทึมเอ็กซ์เปคเตชันแม็กซิไมซ์เซชัน ซึ่งได้กล่าวในหัวข้อกระบวนการปรับแนวในบทที่ 3 เนื่องจากการจับคู่เป็นแบบกลุ่มต่อกลุ่มดังนั้นจึงต้องตั้งค่าพารามิเตอร์สองค่า คือ จำนวนอักขระมากสุดที่จะจับคู่กับหน่วยเสียงได้ และจำนวนหน่วยเสียงมากสุดที่จะจับคู่กับอักขระได้ ซึ่งผู้วิจัยตั้งค่าทั้งสองอยู่ที่ 2 หน่วยเป็นค่าที่เหมาะสม

4.1.2.2 **เครื่องมือสำหรับการสร้างโมเดลการทำนายหน่วยเสียง**

ผู้วิจัยใช้เครื่องมือ “DirecTL+” [17] ซึ่งเป็นเครื่องมือเดียวสำหรับการสร้างออนไลน์ดิสคริมิเนทิฟเฟรมเวิค ในการเรียนรู้และสร้างโมเดลการทำนายหน่วยเสียง ซึ่งได้มีการอธิบายถึงออนไลน์ดิสคริมิเนทิฟเฟรมเวิคในหัวข้อโมเดลการทำนายหน่วยเสียงในบทที่ 3 โดยผู้ที่คิดค้นเฟรมเวิคนี้ได้ทำการสร้างเครื่องมือนี้ขึ้นมาสำหรับการสร้างโมเดลนี้โดยเฉพาะ ในงานวิจัยนี้ผู้วิจัยใช้ค่าพารามิเตอร์สำหรับแต่ละค่าลักษณะสำคัญทั้งหมดดังนี้ ขนาดของบริบทที่พิจารณา 5 หน่วย, เอ็นแกรม 11 หน่วย, มาคอฟออร์เดอร์ 1 หน่วย, จอยนต์เอ็นแกรม 4 หน่วย และ เลือกใช้ค่าลักษณะสำคัญลิเนียร์เชนในการเรียนรู้ด้วย ซึ่งค่าพารามิเตอร์เหล่านี้เป็นค่าที่ให้ผลลัพธ์มีความแม่นยำที่สูง ซึ่งอ้างอิงจากการทดลอง [16], [17] ซึ่งรายงานว่าหากปรับค่าพารามิเตอร์ต่าง ๆ ให้มากกว่านี้ ความแม่นยำที่ได้อาจจะน้อยลงหรือเพิ่มขึ้นเพียงเล็กน้อยเท่านั้น

4.1.2.3 **เครื่องมือสำหรับการสร้างโมเดลการทำนายเสียงวรรณยุกต์**

ผู้วิจัยใช้เครื่องมือ “Wagon Cart Tool” [25] ซึ่งเป็นเครื่องมือที่ใช้อัลกอริทึมต้นไม้ในการทำนายเสียงวรรณยุกต์ เนื่องจากเป็นเครื่องมือที่ประมวลผลเร็ว และสามารถปรับค่าพารามิเตอร์ต่าง ๆ ได้มากมาย โดยมีการใช้งานเครื่องมือนี้ในงานลักษณะการประมวลผลข้อความอย่างแพร่หลาย

4.1.2.4 **เครื่องมืออื่น ๆ ที่ใช้ในการทดลอง**

นอกเหนือจากเครื่องมือต่าง ๆ ที่ได้กล่าวไปข้างต้นแล้ว บางกระบวนการจำเป็นต้องทำแยกต่างหาก เช่น การดึงค่าลักษณะสำคัญจากข้อมูล เพื่อที่จะนำไปใช้เรียนรู้ในการสร้างโมเดล

การทำนายเสียงวรรณยุกต์ ผู้วิจัยใช้เครื่องมือ “Visual Studio 2008” ในการเขียนโปรแกรมสำหรับกระบวนนี้และกระบวนที่จำเป็นอื่น ๆ

4.1.3. การประเมินผล

ในการประเมินผลระบบนั้นจะวัดผลจากความแม่นยำในการทำนายหน่วยเสียงจากโมเดลการทำนายหน่วยเสียง และความแม่นยำในการทำนายเสียงวรรณยุกต์จากโมเดลการทำนายเสียงวรรณยุกต์ รวมทั้งความแม่นยำรวมเมื่อทำนายทั้งหน่วยเสียงและเสียงวรรณยุกต์ต่อเนื่องกันจากทั้งสองโมเดล ผู้วิจัยใช้ตัววัดผล 3 แบบ คือ ความแม่นยำระดับหน่วยเสียง (Phoneme Accuracy), ความแม่นยำระดับพยางค์ (Syllable Accuracy), และความแม่นยำระดับคำ (Word Accuracy) โดยรายละเอียดการคำนวณของทั้ง 3 ตัววัดใช้รูปแบบเดียวกัน ดังที่แสดงในสมการ (1) เพียงแต่แตกต่างกันที่ระดับของหน่วยว่าเป็น หน่วยเสียง, พยางค์ หรือคำ

$$\text{ความแม่นยำ} = \frac{(N-S-D-I)}{N} \times 100\ (\%) \qquad (1)$$

โดยที่

N คือ จำนวนหน่วยเสียงหรือพยางค์หรือคำทั้งหมด

S คือ จำนวนหน่วยเสียงหรือพยางค์หรือคำที่ผิดพลาดแบบแทนที่ (Substitution Error)

D คือ จำนวนหน่วยเสียงหรือพยางค์หรือคำที่ผิดพลาดแบบหายไป (Deletion Error)

I คือ จำนวนหน่วยเสียงหรือพยางค์หรือคำที่ผิดพลาดแบบเกิน (Insertion Error)

นอกจากการวัดผลในเชิงความแม่นยำแล้ว ยังมีการวัดผลในเชิงความพึงพอใจ (Subjective Test) ด้วย เนื่องจากจุดประสงค์ของการหาคำอ่าน ก็เพื่อที่จะนำไปใช้ในงานสังเคราะห์เสียง ดังนั้นผู้วิจัยจึงวัดผลจากความพึงพอใจต่อคำอ่านเหล่านี้ ด้วยการแบ่งเป็นสามระดับ 1.ดี (Good) 2. พอใช้ (Fair) 3.ไม่ดี (Poor) โดยใช้หลักการประเมินเดียวกันกับงานวิจัย [6], [8] ที่กำหนดว่าหากคำอ่านที่ทำนายนั้นตรงกับคำอ่านที่ผู้ประเมินคิดไว้ถือว่าอยู่ระดับดี แต่หากคำอ่านที่ทำนายออกมานั้นไม่ตรงกับที่ผู้ประเมินคิดไว้แต่เป็นคำอ่านที่อาจอ่านออกเสียงเช่นนี้ได้หรือยอมรับได้จะถือว่าอยู่ในระดับพอใช้ และคำอ่านที่ไม่น่าจะอ่านเช่นนี้หรือยอมรับไม่ได้ถือว่าอยู่ในระดับไม่ดี โดยผลลัพธ์ที่ถือว่ายอมรับได้คือผลลัพธ์ที่ประเมินอยู่ในระดับพอใช้หรือดี คำที่เลือกมาประเมินนั้นเลือกมาจากคำภาษาอังกฤษที่พบในบทความปัจจุบันบ่อย ๆ ทั้งหมด 84 คำ โดยมีการกระจายตัวอักษรตั้งแต่ A ถึง Z และกระจายความยาวของพยางค์ตั้งแต่ 1 ถึง 5 พยางค์ รายละเอียดของคำทั้งหมดและคำอ่านที่ทำนายจากระบบ แสดงในภาคผนวก ก

4.2 ผลการทดลอง

ผลการทดลองแบ่งออกเป็น 4 ส่วน คือ ผลการทดลองของโมเดลการทำนายหน่วยเสียง, ผลการทดลองของโมเดลการทำนายเสียงวรรณยุกต์, ผลการทดลองรวม และผลประเมินความพึงพอใจ

4.2.1. ผลการทดลองของโมเดลการทำนายหน่วยเสียง

ผลการทดลองของโมเดลการทำนายหน่วยเสียง จะรายงานทั้งผลลัพธ์ก่อนกระบวนการประมวลผลหลังจากการทำนายหน่วยเสียงซึ่งได้ผลลัพธ์ค่าเฉลี่ยความแม่นยำระดับคำที่ 62.87% และผลลัพธ์หลังกระบวนการประมวลผลหลังจากการทำนายหน่วยเสียงซึ่งได้ผลลัพธ์ค่าเฉลี่ยความแม่นยำระดับคำที่ 64.23% รายละเอียดของผลการทดลองของทุกโฟลด์ (Fold) แสดงในตารางที่ 4.1 และ 4.2 ตามลำดับ ซึ่งเมื่อเทียบกับระบบอ้างอิง [6] ซึ่งใช้ข้อมูลชุดเดียวกันและแบ่งข้อมูลสำหรับเรียนรู้และทดสอบเป็นสัดส่วนเท่ากันโดยใช้โมเดลต้นไม้เป็นอัลกอริทึมในการเรียนรู้รายงานผลลัพธ์ความแม่นยำเฉลี่ยระดับคำที่ 32.20%

ตารางที่ 4.1 ผลลัพธ์ของโมเดลการทำนายหน่วยเสียงก่อนกระบวนการประมวลผลหลังจากการทำนายหน่วยเสียง

ชุดข้อมูล	**ความแม่นยำระดับหน่วยเสียง (%)**	**ความแม่นยำระดับพยางค์ (%)**	**ความแม่นยำระดับคำ (%)**
โฟลด์ 1	90.55	80.51	63.65
โฟลด์ 2	90.69	80.61	63.57
โฟลด์ 3	90.48	79.52	62.24
โฟลด์ 4	90.79	80.58	64.90
โฟลด์ 5	90.89	81.27	65.94
โฟลด์ 6	90.83	80.09	62.92
โฟลด์ 7	91.33	80.06	61.71
โฟลด์ 8	90.13	79.05	61.47
โฟลด์ 9	89.95	78.67	60.63
โฟลด์ 10	90.56	79.70	61.71
ค่าเฉลี่ย	**90.62**	**80.01**	**62.87**

ตารางที่ 4.2 ผลลัพธ์ของโมเดลการทำนายหน่วยเสียงเมื่อผ่านกระบวนการประมวลผลหลังจากการทำนายหน่วยเสียง

ชุดข้อมูล	**ความแม่นยำระดับหน่วยเสียง (%)**	**ความแม่นยำระดับพยางค์ (%)**	**ความแม่นยำระดับคำ (%)**
โฟลด์ 1	90.84	81.71	64.86
โฟลด์ 2	90.85	81.71	64.66
โฟลด์ 3	90.78	80.93	63.69
โฟลด์ 4	90.91	81.49	65.62
โฟลด์ 5	91.19	82.37	66.91
โฟลด์ 6	91.14	81.46	64.37
โฟลด์ 7	91.72	81.55	63.53
โฟลด์ 8	90.55	80.79	63.16
โฟลด์ 9	90.28	80.03	62.32
โฟลด์ 10	90.92	80.95	63.16
ค่าเฉลี่ย	**90.92**	**81.30**	**64.23**

4.2.2. ผลการทดลองของโมเดลการทำนายเสียงวรรณยุกต์

ผลการทดลองของโมเดลการทำนายเสียงวรรณยุกต์จะวัดความแม่นยำในระดับพยางค์และระดับคำ ซึ่งต่างจากโมเดลการทำนายหน่วยเสียงที่จะมีการวัดระดับหน่วยเสียงด้วย เพราะการทำนายเสียงวรรณยุกต์จะทำนายทีละพยางค์ ซึ่งต่างจากโมเดลการทำนายหน่วยเสียงซึ่งทำนายหน่วยเล็กที่สุดคือหน่วยเสียง โดยผลลัพธ์ของโมเดลทำนายเสียงวรรณยุกต์ ได้ความแม่นยำระดับพยางค์เฉลี่ย 92.73% และความแม่นยำระดับคำ 82.09% ซึ่งเพื่อเปรียบเทียบกับผลลัพธ์จากระบบอ้างอิง [6] ผู้วิจัยได้ทำการพัฒนาโมเดลต้นไม้โดยใช้ค่าลักษณะสำคัญตามระบบอ้างอิงเพื่อทดลองเปรียบเทียบผลลัพธ์กับข้อมูลที่เหมือนกันทุกประการ ได้ความแม่นยำเฉลี่ยระดับพยางค์ 92.13% และความแม่นยำเฉลี่ยระดับคำ 80.72% รายละเอียดผลการทดลองของทุกโฟลด์แสดงเปรียบเทียบกับระบบอ้างอิงในตารางที่ 4.3

ตารางที่ 4.3 ผลลัพธ์ของโมเดลการทำนายเสียงวรรณยุกต์เปรียบเทียบกับระบบอ้างอิง

ชุดข้อมูล	**ความแม่นยำระดับพยางค์ของโมเดลจากงานวิจัยอ้างอิง (%)**	**ความแม่นยำระดับคำของโมเดลจากงานวิจัยอ้างอิง (%)**	**ความแม่นยำระดับพยางค์ของโมเดลใหม่ (%)**	**ความแม่นยำระดับคำของโมเดลใหม่ (%)**
โฟลด์ 1	91.76	79.71	92.23	81.04
โฟลด์ 2	91.93	80.22	92.65	82.15
โฟลด์ 3	91.93	80.58	92.73	82.27
โฟลด์ 4	93.07	83.35	93.50	83.96
โฟลด์ 5	91.83	80.19	93.20	83.70
โฟลด์ 6	92.77	82.00	92.96	82.13
โฟลด์ 7	92.71	82.25	92.76	82.25
โฟลด์ 8	92.28	81.04	92.98	82.49
โฟลด์ 9	91.33	78.62	91.84	79.47
โฟลด์ 10	91.66	79.23	92.45	81.40
ค่าเฉลี่ย	**92.13**	**80.72**	**92.73**	**82.09**

4.2.3. **ผลการทดลองรวม**

ผลการทดลองรวมคือการนำผลลัพธ์ที่เป็นหน่วยเสียงจากโมเดลการทำนายหน่วยเสียง ไปทำนายเสียงวรรณยุกต์ต่อไปเพื่อให้เป็นคำอ่านที่สมบูรณ์ โดยที่พยางค์หรือคำที่จะถูกพิจารณาว่าถูกต้องนั้น จะต้องถูกทั้งพยางค์หรือทั้งคำ โดยถูกทั้งหน่วยเสียงและเสียงวรรณยุกต์ เนื่องจากกระบวนการประมวลผลหลังจากการทำนายหน่วยเสียงสามารถมองได้ว่าเป็นการแก้ไขข้อผิดพลาดที่เฉพาะเจาะจงกับโมเดลเดียว ดังนั้นจึงได้ทำการทดลองโดยแบ่งเป็น 2 โมเดล คือ โมเดลรวมที่ไม่ผ่านกระบวนการประมวลผลหลังจากการทำนายหน่วยเสียง และ โมเดลรวมที่ผ่านกระบวนการประมวลผลหลังจากการทำนายหน่วยเสียง เพื่อให้สามารถแสดงการเปรียบเทียบกับระบบอ้างอิงได้อย่างชัดเจนมากขึ้น โดยโมเดลแรกได้ผลลัพธ์ความแม่นยำระดับพยางค์เฉลี่ย 74.68% และความแม่นยำระดับคำเฉลี่ย 52.78% ดังแสดงในตารางที่ 4.4 ซึ่งเมื่อเทียบกับระบบอ้างอิง [6] รายงานผลความแม่นยำเฉลี่ยระดับพยางค์ 53.30% และความแม่นยำเฉลี่ยระดับคำ 24.50%

ตารางที่ 4.4 ผลลัพธ์รวมที่ไม่ผ่านกระบวนการประมวลผลหลังจากการทำนายหน่วยเสียง

ชุดข้อมูล	**ความแม่นยำระดับพยางค์ (%)**	**ความแม่นยำระดับคำ (%)**
โฟลด์ 1	74.94	53.62
โฟลด์ 2	75.02	52.83
โฟลด์ 3	74.28	51.99
โฟลด์ 4	75.80	55.97
โฟลด์ 5	76.24	56.28
โฟลด์ 6	75.12	52.29
โฟลด์ 7	74.68	52.05
โฟลด์ 8	74.20	52.54
โฟลด์ 9	72.57	49.88
โฟลด์ 10	73.96	50.36
ค่าเฉลี่ย	**74.68**	**52.78**

ตารางที่ 4.5 ผลลัพธ์รวมที่ผ่านกระบวนการประมวลผลหลังจากการทำนายหน่วยเสียง

ชุดข้อมูล	**ความแม่นยำระดับพยางค์ (%)**	**ความแม่นยำระดับคำ (%)**
โฟลด์ 1	76.08	54.71
โฟลด์ 2	76.07	53.68
โฟลด์ 3	75.70	53.32
โฟลด์ 4	76.61	56.69
โฟลด์ 5	77.47	57.25
โฟลด์ 6	76.34	53.50
โฟลด์ 7	76.09	53.38
โฟลด์ 8	75.75	54.11
โฟลด์ 9	74.96	51.09
โฟลด์ 10	75.25	51.57
ค่าเฉลี่ย	**76.03**	**53.93**

4.2.4. ผลประเมินความพึงพอใจ

จากผู้ประเมินทั้งหมด 5 คน ซึ่งเป็นนิสิตปริญญาโท ทั้งชายและหญิง ได้ทำการประเมินความพึงพอใจต่อคำอ่านไทยจากคำภาษาอังกฤษทั้งหมด 84 คำ ที่คัดเลือกมาจากคำภาษาอังกฤษที่พบบ่อยในบทความปัจจุบัน โดยแบ่งเป็นสามระดับคือ ดี พอใช้และแย่ โดยได้สรุปผลการประเมินได้ดังตารางที่ 4.6 ผลลัพธ์เฉลี่ยของการประเมินในระดับดีเท่ากับ 78.81%, ผลลัพธ์เฉลี่ยของการประเมินในระดับพอใช้เท่ากับ 12.14% และผลลัพธ์เฉลี่ยของการประเมินในระดับไม่ดีเท่ากับ 9.05% รวมเป็นผลลัพธ์ที่ยอมรับได้คือระดับดีและพอใช้รวมกันเท่ากับ 90.95%

ตารางที่ 4.6 ผลลัพธ์จากการประเมินความพึงพอใจต่อคำอ่าน

ลำดับผู้ประเมิน	ไม่ดี (คำ)	พอใช้ (คำ)	ดี (คำ)
1	9	13	62
2	10	11	63
3	9	8	67
4	3	9	72
5	7	10	67

4.3 การวิเคราะห์ผลการทดลอง

ในส่วนนี้จะกล่าวถึงการวิเคราะห์ผลลัพธ์ที่เกิดขึ้น จากโมเดลการทำนายหน่วยเสียงซึ่งจะรวมไปถึงกระบวนการประมวลผลหลังจากการทำนายหน่วยเสียงด้วย และโมเดลการทำนายเสียงวรรณยุกต์ซึ่งจะกล่าวถึงความสามารถของค่าลักษณะสำคัญที่เพิ่มเติมเข้าไป รวมทั้งการวิเคราะห์ผลลัพธ์รวมเมื่อทำการทำนายทั้งหน่วยเสียงและเสียงวรรณยุกต์ และการวิเคราะห์ผลลัพธ์เทียบกับระบบอ้างอิง

4.3.1. การวิเคราะห์ผลลัพธ์จากโมเดลการทำนายหน่วยเสียง

ผู้วิจัยได้ทำการวิเคราะห์ผลลัพธ์จากโมเดลการทำนายหน่วยเสียง จากการทดลองโฟลด์ 1 ซึ่งเมื่อพิจารณาความผิดพลาดที่เกิดขึ้นจากโมเดลการทำนายหน่วยเสียงนั้น จะสามารถแบ่งได้เป็น 2 ลักษณะหลัก ๆ ก็คือ ความผิดพลาดที่มีโอกาสแก้ไขได้ ซึ่งมีสัดส่วนประมาณ 12.62% และความผิดพลาดที่ไม่มีโอกาสแก้ไขได้ที่มีสัดส่วนประมาณ 87.38% ดังแสดงในตารางที่ 3.2 ในบทที่ 3

ความผิดพลาดที่ไม่มีโอกาสแก้ไขได้มีทั้งหมด 5 ลักษณะ คือ 1. ทำนายเสียงพยัญชนะต้นผิด 2. ทำนายเสียงสระผิด 3. ทำนายเสียงพยัญชนะสะกดผิด 4. ทำนายพยางค์เกิน และ 5. ทำนายพยางค์ขาด โดยความผิดพลาดที่เกิดขึ้นมากที่สุดจากโมเดลทำนายหน่วยเสียง ก็คือการทำนายเสียงสระผิดพลาด ซึ่งมีสัดส่วนถึงเกือบครึ่งหนึ่งของความผิดพลาดทั้งหมด โดยภายในความผิดพลาดที่เกิดจากการทำนายเสียงสระผิดนั้น เป็นการทำนายผิดในความสั้นยาวของสระประมาณ 25% ของความผิดพลาดจากการทำนายเสียงสระผิดซึ่งเป็นสัดส่วนที่ไม่มากและเป็นการแก้ไขจุดผิดพลาดจากโมเดลเดิมในงานวิจัยอ้างอิงได้ดีขึ้นที่กล่าวไว้ว่าความผิดพลาดส่วนใหญ่เกิดขึ้นเกิดจากความสั้นยาวของเสียงสระ ซึ่งหากพิจาณาความผิดพลาดจากความสั้นยาวของสระกับความผิดพลาดอื่น ๆ ก็ถือว่าเป็นความผิดพลาดที่ไม่ส่งผลกระทบต่อผลลัพธ์คำอ่านมากนักในเชิงการนำไปสังเคราะห์เสียงจริง เพราะความสั้นยาวของสระอาจไม่แตกต่างกันมากและมีโอกาสที่จะยอมรับได้ นอกจากนี้ความผิดพลาดที่เกิดขึ้นมากรองลงมา ก็คือการทำนายเสียงพยัญชนะสะกดผิดที่มีสัดส่วนประมาณ 22% ของความผิดพลาดทั้งหมด ซึ่งเป็นความผิดพลาดที่แก้ไขได้ยาก เพราะเกิดจากความกำกวมของอักขระภาษาอังกฤษที่มีความไม่แน่นอนในรูปแบบของการอ่าน จึงทำให้เกิดความกำกวมในเสียงพยัญชนะสะกด เช่นจากตัวอย่างของข้อมูลที่ใช้ในการทดสอบ คำว่า “ATTRITION” มีคำอ่านคือ “z-x-t^-3|thr-i-z^-3|ch-a-n^-3” แต่ผลลัพธ์จากโมเดลการทำนายหน่วยเสียงได้ผลลัพธ์คือ “z-x-t^-3|thr-i-t^-3|ch-a-n^-3” ที่พยางค์ที่สองโมเดลทำนายเสียงพยัญชนะสะกดผิดเพราะ อักขระภาษาอังกฤษมีความกำกวมในการอ่าน บางครั้งตัวอักขระพยัญชนะก็เป็นทั้งเสียงพยัญชนะสะกดและเสียงพยัญชนะต้นของพยางค์ถัดไป ในขณะที่บางครั้งก็ไม่ใช่ การแก้ไขความผิดพลาดนี้อาจจะทำได้จากการเพิ่มขนาดของข้อมูลที่มาทำการเรียนรู้ให้มากขึ้น และความผิดพลาดที่เกิดบ่อยเป็นอันดับที่สาม ก็คือการทำนายเสียงพยัญชนะต้นผิดที่มีสัดส่วนประมาณ 11% ของความผิดพลาดทั้งหมด ซึ่งในจำนวนนี้เป็นความผิดพลาดระหว่างเสียงพยัญชนะต้นที่มีความสัมพันธ์กันถึงประมาณ 50% ของความผิดพลาดในการทำนายเสียงพยัญชนะต้นผิด เช่น ความผิดพลาดระหว่างเสียง “t”-“th”, “c”-“ch”, “p”-“ph”, “k”-“kh” และ “tr”-“thr” ซึ่งเกิดขึ้นเพราะอักขระภาษาอังกฤษหลาย ๆ อักขระสามารถอ่านออกเสียงได้หลายเสียงขึ้นอยู่กับแต่ละคำ เช่น “K” สามารถอ่านออกเสียงได้สองเสียงที่พบบ่อยก็คือเสียง “k” และ “kh” จึงทำให้เกิดความกำกวมขึ้น เช่นเดียวกันกับอักขระ “T” ที่อ่านออกเสียงได้หลายเสียงเช่น “t” และ “th” โดยความผิดพลาดทั้ง 3 อย่างที่กล่าวถึงนั้นรวมแล้วเป็นสัดส่วนถึงประมาณ 80% ของความผิดพลาดทั้งหมดแล้วโดยที่ยังไม่รวมความผิดพลาดในทำนายพยางค์เกินหรือขาด ที่มีสัดส่วนรวมกันอีกประมาณ 6.4 %

สำหรับความผิดพลาดที่มีโอกาสแก้ไขได้ซึ่งมีสัดส่วนประมาณ 12.62% นั้นสามารถแบ่งได้เป็น 6 ลักษณะ ซึ่งได้มีกล่าวไว้ในบทที่ 3 ซึ่งความผิดพลาดเหล่านี้จะถูกทำการพยายามแก้ไขจากกระบวนการประมวลผลหลังจากการทำนายหน่วยเสียง ที่ทำการแก้ไขความผิดพลาดเหล่านี้ได้บางส่วน ซึ่งเมื่อพิจารณาจากผลลัพธ์ที่แสดงในตารางที่ 4.1 และ 4.2 พบว่ากระบวนการประมวลผลหลังจากการทำนายหน่วยเสียงสามารถแก้ไขความผิดได้และเพิ่มความแม่นยำในระดับคำเฉลี่ยได้จาก 62.87% เป็น 64.23% ซึ่งผู้วิจัยได้วิเคราะห์ต่อไปว่ากระบวนการประมวลผลหลังจากการทำนายหน่วยเสียงสามารถแก้ไขความผิดที่มีโอกาสแก้ไขได้มีประสิทธิภาพมากน้อยอย่างไร โดยวิคราะห์จากผลลัพธ์จากโฟลด์ 1 พบว่ากระบวนการประมวลผลหลังจากการทำนายหน่วยเสียงสามารถแก้ไขความผิดที่มีโอกาสแก้ไขได้ประมาณ 35% หรือประมาณ 4.4% ของความผิดพลาดที่เกิดขึ้นทั้งหมด

4.3.2. การวิเคราะห์ผลลัพธ์จากโมเดลการทำนายเสียงวรรณยุกต์

ผลลัพธ์จากโมเดลการทำนายเสียงวรรณยุกต์ได้ผลลัพธ์ความแม่นยำในระดับพยางค์เท่ากับ 92.73% และความแม่นยำในระดับคำเท่ากับ 82.09% เมื่อเทียบกับระบบอ้างอิง [6] ซึ่งได้ความแม่นยำในระดับพยางค์และระดับคำเท่ากับ 92.13% และ 80.72% ตามลำดับ ซึ่งความแม่นยำที่มากขึ้น เป็นผลมาจากค่าลักษณะสำคัญ คือ มาตราพยัญชนะสะกด และ อักขระพยางค์ภาษาอังกฤษ ผู้วิจัยได้ทำการศึกษาจากข้อผิดพลาดที่เกิดขึ้นจากโมเดลของระบบอ้างอิงและโมเดลใหม่ เพื่อดูว่าค่าลักษณะสำคัญมาตราสะกดและพยางค์ภาษาอังกฤษสามารถช่วยให้โมเดลทำนายเสียงวรรณยุกต์ได้แม่นยำขึ้นตามที่คาดหวังไว้หรือไม่ พบว่ามาตราสะกดสามารถช่วยทำนายเสียงวรรณยุกต์ให้แม่นยำขึ้นในกรณีที่เกิดเสียงพยัญชนะสะกดที่ลงท้ายด้วยเสียงเอส เช่น ตัวอย่างจากข้อมูลที่ใช้ในการทดสอบ คำภาษาอังกฤษ “SERIES” มีคำอ่านคือ “s-ii-z^-0|r-ii-zs^-2” ที่พยางค์ที่สองจะปรากฏเสียงเอส “zs^” ที่ท้ายพยางค์ด้วย ซึ่งในข้อมูลทั้งหมด 8,283 คำ พบเสียง “r-ii-zs^-2” เกิดขึ้นเพียง 10 พยางค์ แต่เสียง “r-ii-z^-2” เกิดขึ้นถึง 101 พยางค์ ซึ่งทั้งสองพยางค์ออกเสียงวรรณยุกต์เดียวกัน ดังนั้นการจัดกลุ่มพยัญชนะสะกดจึงมีส่วนช่วยให้ทำนายเสียงวรรณยุกต์ของพยางค์ที่มีเสียงเอสที่ท้ายพยางค์ได้ดีขึ้นเพราะเป็นกรณีที่เกิดขึ้นน้อยและอาจจะไม่พอในการเรียนรู้ให้มีความแม่นยำ ในขณะที่โมเดลเก่าทำนายเสียงวรรณยุกต์ผิด

นอกจากนั้นค่าลักษณะสำคัญพยางค์ภาษาอังกฤษก็มีส่วนช่วยให้การทำนายแม่นยำขึ้นมาก จากข้อมูลที่ใช้ในการทดสอบพบว่า คำภาษาอังกฤษ “AIRMEN” ซึ่งมีคำอ่านคือ “z-xx-z^-0|m-e-n^-0” โมเดลจากระบบอ้างอิงไม่สามารถทำนายเสียงวรรณยุกต์ของพยางค์ “m-e-n^” ได้

ถูกต้อง โดยโมเดลทำนายเป็นเสียงวรรณยุกต์ตรี ในขณะที่โมเดลใหม่สามารถทำนายได้ถูกต้องคือเสียงวรรณยุกต์สามัญ เนื่องจากพยางค์ที่มีหน่วยเสียง “m-e-n^” ส่วนมากจะเป็นเสียงวรรณยุกต์ตรีที่มาจากพยางค์ “MENT” ซึ่งพบได้บ่อยในคำภาษาอังกฤษ ดังนั้นการมีเพียงค่าลักษณะสำคัญที่มาจากเฉพาะหน่วยเสียงจะทำให้ไม่สามารถแยกความต่างของพยางค์ได้ ดังนั้นจึงทำให้โมเดลไม่สามารถทำนายเสียงวรรณยุกต์ได้ถูก แต่พยางค์ภาษาอังกฤษสามารถช่วยให้โมเดลสามารถแยกความต่างของพยางค์ในกรณีนี้ได้ เช่นเดียวกันกับ “APPROXIMATE” ที่มีคำอ่านคือ “z-x-p^-3|phr-@-k^-3|s-i-z^-1|m-ee-t^-1” การทำนายเสียงวรรณยุกต์ให้กับพยางค์ที่สาม “s-i-z^” ค่อนข้างมีความกำกวม เพราะจากข้อมูลพยางค์นี้ถูกกำกับด้วยเสียงวรรณยุกต์เอกและตรีในจำนวนที่เท่า ๆ กัน แต่พยางค์ภาษาอังกฤษ “XI” เกือบทั้งหมดถูกกำกับด้วยเสียงวรรณยุกต์เอก ดังนั้นจึงทำให้โมเดลทำนายเสียงวรรณยุกต์ได้ถูกต้องในขณะที่โมเดลจากระบบอ้างอิงทำนายเสียงตรี ในขณะที่ “APPOINTEE” และ “COMMITEE” ที่พยางค์ “TEE” ซึ่งเป็นเสียงสามัญ แต่หากไม่มีพยางค์ภาษาอังกฤษ โมเดลจะทำนายเป็นเสียงโท เนื่องจากมีหน่วยเสียงเดียวกันกับ “TY” ซึ่งพบบ่อยมากในคำภาษาอังกฤษ จากตัวอย่างทั้งหมดที่กล่าวมาก็จะแสดงให้เห็นถึงความสามารถของค่าลักษณะทั้งสองที่มีส่วนช่วยให้โมเดลการทำนายเสียงวรรณยุกต์มีความแม่นยำมากขึ้น

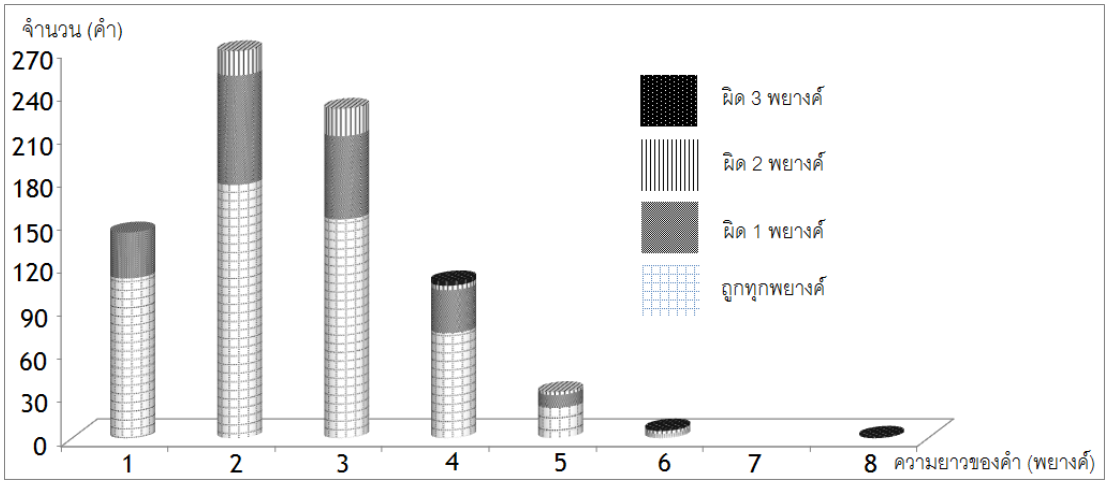
4.3.3. การวิเคราะห์ผลลัพธ์รวม

เนื่องจากคำอ่านที่สมบูรณ์เกิดจากผลลัพธ์จากทั้งสองโมเดล โดยการทำนายหน่วยเสียงจากโมเดลการทำนายหน่วยเสียงก่อน จากนั้นจึงนำผลลัพธ์ไปทำนายเสียงวรรณยุกต์จากโมเดลการทำนายเสียงวรรณยุกต์ ซึ่งจะเห็นได้ว่าหากหน่วยเสียงที่ถูกทำนายออกมานั้นผิด จะส่งผลต่อผลลัพธ์รวมทันที ไม่ว่าจะทำนายเสียงวรรณยุกต์ถูกหรือผิด ผลลัพธ์รวมก็จะผิดอย่างแน่นอน ผลลัพธ์จากโมเดลการทำนายเสียงวรรณยุกต์ใช้ข้อมูลที่นำมาทดสอบเป็นข้อมูลที่มีหน่วยเสียงถูกต้องทั้งหมด แต่หากข้อมูลหน่วยเสียงเป็นข้อมูลที่มาจากโมเดลทำนายหน่วยเสียงซึ่งจะมีความผิดพลาดปะปนอยู่ด้วย ก็ส่งผลให้โมเดลการทำนายเสียงวรรณยุกต์มีความแม่นยำน้อยลงอย่างแน่นอน ดังนั้นผู้วิจัยได้ทดลองนำผลลัพธ์หน่วยเสียงที่ได้จากโมเดลการทำนายหน่วยเสียงโดยทดลองจากข้อมูลโฟลด์ 1 เพื่อนำไปทดสอบกับโมเดลทำนายเสียงวรรณยุกต์ ผลลัพธ์ที่ได้คือความแม่นยำในระดับพยางค์ลดลงเหลือ 86.09% จาก 92.23% และความแม่นยำในระดับคำลดลงเหลือ 70.65% จาก 81.04% ซึ่งจะเห็นว่าการทำนายเสียงวรรณยุกต์จากหน่วยเสียงที่มีความผิดพลาดปะปนมานั้นจะทำให้การทำนายแม่นยำน้อยลงมากถึงประมาณ 10% ในระดับคำ สิ่งเหล่านี้ทำให้ผลลัพธ์รวมต่ำลงมากเมื่อเทียบกับความแม่นยำในแต่ละโมเดลแยกกัน และ

นอกจากนั้นจะเห็นว่าโมเดลการทำนายหน่วยเสียงมีความแม่นยำน้อยกว่าโมเดลการทำนายเสียงวรรณยุกต์ในระดับคำถึงเกือบ 20 % เนื่องจากผลลัพธ์รวมเกิดจากการรวมผลลัพธ์จากทั้งสองโมเดลจึงทำให้ความแม่นยำรวมเฉลี่ยในระดับคำลดลงไปอยู่ที่ประมาณ 54%

ประเด็นที่มีความน่าสนใจอีกประเด็นก็คือ เมื่อพิจารณาความแม่นยำเฉลี่ยในระดับหน่วยเสียงและพยางค์ของโมเดลการทำนายหน่วยเสียงซึ่งสูงถึง 90.92% และ 81.30% แต่ความแม่นยำเฉลี่ยในระดับคำกลับต่ำกว่ามาก อยู่ที่ 64.23% ดังนั้นผู้วิจัยจึงได้ทำการวิเคราะห์ต่อไปถึงจำนวนพยางค์ที่ผิดในแต่ละคำว่ามีลักษณะการผิดพลาดอย่างไร โดยวิเคราะห์จากการทดลองโฟลด์ 1 ซึ่งมีข้อมูลทดลองทั้งสิ้น 828 คำ มีคำที่ทำนายหน่วยเสียงผิดทั้งสิ้น 301 คำ โดยเป็นการเกิดความผิดพลาดแบบแทนที่ทั้งหมด 262 คำ และเป็นความผิดพลาดแบบขาดรวมกับความผิดพลาดแบบเกิน 39 คำ จากรูปที่ 4.1 แสดงจำนวนพยางค์ที่ผิดพลาดเฉพาะที่เป็นความผิดพลาดแบบแทนที่ โดยแกนนอนแสดงจำนวนพยางค์ในคำ และแกนตั้งแสดงจำนวนคำ สัดส่วนในรูปทรงกระบอกคือจำนวนพยางค์ที่ทำนายผิดในคำ โดยจะเห็นว่าคำที่มี 1 พยางค์ จะมีจำนวนพยางค์ผิดพลาดมากสุดได้แค่ 1 ในขณะที่คำที่มี 2 พยางค์ก็จะมีจำนวนพยางค์ที่ผิดพลาดได้มากที่สุดคือ 2 เป็นเช่นนี้ไปเรื่อย ๆ ซึ่งจากข้อมูล 828 มีจำนวนพยางค์เฉลี่ยที่ 2.5 พยางค์ และมีจำนวนพยางค์ที่มากที่สุดเท่ากับ 8 พยางค์ แต่จำนวนพยางค์ที่ผิดพลาดมากที่สุดต่อ 1 คำ เท่ากับ 3 พยางค์เท่านั้น โดยจะเห็นว่าการผิดถึง 3 พยางค์ใน 1 คำนั้นเกิดขึ้นน้อยมาก โดยส่วนใหญ่จะเป็นการผิดพลาดเพียงแค่พยางค์เดียวต่อ 1 คำ ซึ่งเป็นสัดส่วนถึง 80% ของความผิดพลาดทั้งหมด ด้วยเหตุนี้จึงเป็นเหตุผลที่ทำให้ความแม่นยำในระดับคำลดลงมาก ส่วนความผิดพลาดที่ผิดทั้งคำนั้นเกิดเฉพาะกับคำที่มีจำนวนพยางค์ 1 หรือ 2 คำเท่านั้น ซึ่งมีสัดส่วนประมาณ 13% ของความผิดพลาดทั้งหมดเท่านั้น และในส่วนของความผิดแบบขาดและเกินทั้ง 39 คำนั้น สามารถแจกแจงตามจำนวนพยางค์ที่ขาดหรือเกินได้ดังรายละเอียดในตารางที่ 4.7 และ 4.8 ซึ่งค่าเฉลี่ยพยางค์ที่เกินหรือขาดอยู่ที่ 1.2 พยางค์ ซึ่งหมายความว่าการขาดหรือเกินของพยางค์นั้นเกิดเพียงประมาณ 1 พยางค์ต่อคำเท่านั้น ซึ่งถือว่าน้อยมากและไม่ทำให้คำอ่านเกิดความผิดพลาดมาก

จากข้อมูลเหล่านี้จึงแสดงให้เห็นว่าถึงแม้ความแม่นยำของการทำนายคำอ่านจะยังไม่สูงมาก แต่ความผิดพลาดที่เกิดขึ้นในแต่ละคำก็เกิดขึ้นเพียงเล็กน้อยเท่านั้น และผลลัพธ์จากการทดสอบความพึงพอใจจากผู้ทดสอบก็พบว่าสูงถึงประมาณ 91% เป็นผลลัพธ์ที่ยอมรับได้ ซึ่งแสดงให้เห็นว่าโมเดลการทำนายคำอ่านนี้สามารถใช้งานจริงได้อย่างมีประสิทธิภาพ



รูปที่ 4.1 สัดส่วนของคำที่ผิดพลาดแบบแทนที่ แบ่งตามจำนวนพยางค์ที่ผิดและความยาวของคำ

ตารางที่ 4.7 ความผิดพลาดแบบเกินแจกแจงตามจำนวนพยางค์ที่เกิน

จำนวนพยางค์ในคำ	เกิน 1 พยางค์ (คำ)	เกิน 2 พยางค์ (คำ)	เกิน 3 พยางค์ (คำ)	เกิน 4 พยางค์ (คำ)
1	8	1	1	1
2	12	-	-	-
3	4	1	-	-
4	1	-	-	-
5	1	-	-	-

ตารางที่ 4.8 ความผิดพลาดแบบขาดแจกแจงตามจำนวนพยางค์ที่ขาด

จำนวนพยางค์ในคำ	ขาด 1 พยางค์(คำ)	ขาด 2 พยางค์(คำ)
2	2	-
3	3	1
4	2	-
8	-	1

4.3.4. การวิเคราะห์ผลลัพธ์เทียบกับระบบอ้างอิง

ผลลัพธ์ของโมเดลการทำนายคำอ่านมีความแม่นยำเพิ่มขึ้นมากเมื่อเปรียบเทียบกับระบบอ้างอิง [6] ความแม่นยำที่เพิ่มขึ้นมานั้นเกิดขึ้นมาจากสิ่งต่าง ๆ เหล่านี้

1. ค่าลักษณะสำคัญของของโมเดลการทำนายหน่วยเสียง

ค่าลักษณะสำคัญที่ใช้ในระบบอ้างอิงใช้เฉพาะบริบทของอักขระในการทำนายหน่วยเสียงเท่านั้น ซึ่งในอีกโมเดลที่ทำการพัฒนาขึ้นมาของระบบอ้างอิงได้นำเอาบริบทหน่วยเสียงภาษาอังกฤษของคำมาเป็นค่าลักษณะสำคัญร่วมกันกับบริบทของอักขระและทำให้ผลลัพธ์มีความแม่นยำเพิ่มขึ้น แต่ก็จะเห็นพบว่าการใช้ค่าลักษณะสำคัญเฉพาะกับตัวอักขระ อาจยังไม่เพียงพอ ดังนั้นในโมเดลของผู้วิจัยจึงใช้ค่าลักษณะสำคัญ ที่นำเอาข้อมูลในส่วนของหน่วยเสียงมาใช้ในการเรียนรู้ร่วมกันกับอักขระด้วย ค่าลักษณะสำคัญต่าง ๆ ที่ใช้เป็นดังนี้ คือ บริบทของอักขระ, เอ็นแกรม, ลิเนียร์เซน, มาคอฟออร์เดอร์ และ จอยนต์เอ็นแกรม โดยค่าลักษณะสำคัญ ลิเนียร์เซน, มาคอฟออร์เดอร์ และ จอยนต์เอ็นแกรม เป็นค่าลักษณะสำคัญที่ใช้ข้อมูลของหน่วยเสียงร่วมกันกับข้อมูลของอักขระเพื่อใช้ในการเรียนรู้ด้วย ซึ่งส่งผลให้โมเดลทำนายหน่วยเสียงแม่นยำขึ้นอย่างเห็นได้ชัด

2. อัลกอริทึมในการเรียนรู้ของโมเดลทำนายหน่วยเสียง

อัลกอริทึมที่ใช้ในการเรียนรู้ของระบบอ้างอิงใช้อัลกอริทึมต้นไม้ ซึ่งเป็นอัลกอริทึมแรก ๆ ในงานวิจัยการทำนายหน่วยเสียงทั้งในงานภาษาอังกฤษ และงานภาษาไทย เนื่องจากอัลกอริทึมต้นไม้ประมวลผลรวดเร็วและให้ผลลัพธ์ที่ดีพอสมควร ดังนั้นจึงถูกเลือกใช้อย่างแพร่หลาย แต่ในปัจจุบันโมเดลอื่น ๆ ได้ถูกพัฒนาขึ้นมา อย่างเช่น โมเดลแบบดิสคริมิเนทีฟ เช่น คอนดิชันนอลแรนดอมฟิลด์ และ ออนไลน์ดิสคริมิเนทีฟเทรนนิงเฟรมเวิร์ค ซึ่งโมเดลเหล่านี้มีความเหมาะสมกับงานประเภทนี้มากกว่าโมเดลแบบต้นไม้ เพราะโมเดลต้นไม้ออกแบบในลักษณะของต้นไม้ตัดสินใจซึ่งจะเหมาะกับงานที่ผลลัพธ์มีเพียง 2 คลาส มากกว่างานที่ผลลัพธ์มีหลาย ๆ คลาส ออนไลน์ดิสคริมิเนทีฟเทรนนิงเฟรมเวิร์ค มีการออกแบบสำหรับการใช้ประโยชน์จากข้อมูลหน่วยเสียงในการเรียนรู้ร่วมกันกับข้อมูลตัวอักขระด้วย และแสดงให้เห็นถึงประสิทธิภาพที่ดีกว่าอัลกอริทึมต้นไม้อย่างมาก ในออนไลน์ดิสคริมิเนทีฟเทรนนิงเฟรมเวิร์คออกแบบให้รองรับกับการปรับแนวแบบกลุ่มต่อกลุ่ม ซึ่งการปรับแนวแบบนี้แสดงความสัมพันธ์ระหว่างตัวอักขระและหน่วยเสียงได้ดีกว่าการปรับแนวแบบหนึ่งต่อหนึ่งซึ่งเป็นการปรับแนวที่ใช้กันอย่างแพร่หลายในงานวิจัยก่อน ๆ และใน

ระบบอ้างอิงก็ใช้การปรับแนวแบบหนึ่งต่อหนึ่ง เหตุที่การปรับแนวแบบกลุ่มต่อกลุ่มให้ผลลัพธ์ที่ดีกว่าก็เพราะว่า ในบางคำมีตัวอักขระมากกว่า 1 ตัวอักขระที่สัมพันธ์หน่วยเสียงเพียงหน่วยเสียงเดียว อย่างเช่น พยัญชนะประสม เป็นต้น ซึ่งในการปรับแนวแบบหนึ่งต่อหนึ่งจำเป็นต้องใช้หน่วยเสียงว่างเพื่อจับคู่ตัวอักขระตัวใดตัวหนึ่งที่ไม่มีคู่ สิ่งเหล่านี้ทำให้การเรียนรู้จากบริบทเกิดความสับสนกำกวมได้

3. ค่าลักษณะสำคัญกลุ่มของพยัญชนะสะกด และ พยางค์ภาษาอังกฤษ

สำหรับโมเดลการทำนายเสียงวรรณยุกต์ทั้งในระบบอ้างอิงและระบบที่ผู้วิจัยพัฒนาขึ้นมาใช้อัลกอริทึมเดียวกันในการเรียนรู้ ดังนั้นความแตกต่างที่ทำให้ผลลัพธ์มีความแม่นยำเพิ่มขึ้นอย่างเห็นได้ชัดก็คือ ค่าลักษณะสำคัญที่เพิ่มเติมจากระบบอ้างอิง ค่าลักษณะสำคัญแรกก็คือ กลุ่มของพยัญชนะสะกด เนื่องจากคำภาษาอังกฤษมีเสียงพยัญชนะสะกดมาก กว่าเสียงพยัญชนะสะกดในภาษาไทยมาก อย่างเช่น เสียงเอส, เสียงแอล และ เสียงเอฟ ที่เพิ่มเติมเข้ามาเป็นต้น และการทำนายเสียงวรรณยุกต์ของคำอ่านในภาษาไทยก็คำนึงถึงเสียงพยัญชนะสะกดไทย ดังนั้นหากโมเดลพิจารณาจากเฉพาะเสียงพยัญชนะสะกดที่รวมเอาเสียงพยัญชนะสะกดภาษาอังกฤษเข้ามาด้วย อาจทำให้มีข้อมูลในส่วนเหล่านั้นในการเรียนรู้น้อย และทำให้การทำนายเสียงวรรณยุกต์ของคำเหล่านั้นไม่แม่นยำ ซึ่งจากการพิจารณาจากผลลัพธ์ที่เกิดขึ้นจากโฟล์ด 1 ผู้ทดลองได้ทำการทดลองโดยเพิ่มเฉพาะค่าลักษณะสำคัญของกลุ่มพยัญชนะสะกดเพียงค่าลักษณะสำคัญเดียวแล้วทำการดูผลลัพธ์ที่เกิดขึ้น พบว่า หน่วยเสียงที่ลงท้ายด้วยเสียงเอส อย่างเช่น s^, zs^ และ ngs^ หน่วยเสียงเหล่านี้เกิดขึ้นในข้อมูลน้อย และในโมเดลของระบบอ้างอิงก็ทำนายเสียงวรรณยุกต์ของหน่วยเสียงเหล่านี้ผิดเป็นส่วนใหญ่ แต่ในโมเดลที่เพิ่มค่าลักษณะสำคัญของกลุ่มพยัญชนะสะกดเข้าไปสามารถทำนายเสียงวรรณยุกต์ของหน่วยเสียงเหล่านี้ได้แม่นยำขึ้น และนอกจากนั้นเมื่อผู้วิจัยได้ทำการทดลองโดยเพิ่มค่าลักษณะสำคัญของพยางค์ภาษาอังกฤษในการเรียนรู้ ความแม่นยำในการทำนายเสียงวรรณยุกต์ดีขึ้นอย่างเห็นได้ชัด เนื่องจากการอ่านออกเสียงวรรณยุกต์ของคำภาษาอังกฤษของหน่วยเสียงไทย ตัวอักขระของคำหรือพยางค์ภาษาอังกฤษมีผลต่อการอ่านอย่างมาก ดังได้มีการวิเคราะห์การเกิดเสียงวรรณยุกต์ที่มีความสัมพันธ์กับพยางค์ภาษาอังกฤษในบทที่ 3 เมื่อรวมค่าลักษณะสำคัญของระบบอ้างอิงที่เป็นค่าลักษณะที่เกี่ยวกับหลักภาษาไทยกับค่าลักษณะสำคัญกลุ่มของพยัญชนะสะกดและพยางค์ภาษาอังกฤษ ทำให้ผลลัพธ์มีความแม่นยำเพิ่มขึ้นเมื่อเปรียบเทียบกับระบบอ้างอิง

นอกจากนั้นเมื่อผู้วิจัยทำการวิเคราะห์เพื่อดูว่าระบบที่พัฒนาขึ้นมาได้แก้ไขข้อผิดพลาดที่เกิดขึ้นในระบบเก่าได้อย่างไร เมื่อดูรวม ๆ ก็พบว่าความแม่นยำที่เพิ่มขึ้นมากถึงประมาณ 25% ในระดับคำ ก็แสดงให้เห็นว่าระบบใหม่ได้แก้ไขความผิดพลาดได้เยอะมากในภาพรวม โดยเมื่อพิจารณาถึงชนิดของความผิดพลาดที่เกิดขึ้นจากระบบอ้างอิง กล่าวไว้ว่าเกิดขึ้นมากใน 2 ลักษณะ คือ 1. ความกำกวมของเสียงพยัญชนะต้นและเสียงพยัญชนะสะกด และ 2. ความสั้นยาวของเสียงสระ ซึ่งเมื่อพิจารณาจากตารางที่ 3.2 ในบทที่ 3 ที่ได้ทำการวิเคราะห์แจกแจงความผิดพลาดทั้งหมดของโมเดลทำนายหน่วยเสียงของผู้วิจัยพบว่า ความผิดพลาดที่เกิดบ่อยสุด คือ การทำนายเสียงสระผิด เกิดขึ้นถึงเกือบ 50 % ของความผิดพลาดทั้งหมด ในจำนวนนี้ประมาณ 25 % เป็นการทำนายความสั้นยาวของเสียงสระผิด หรือประมาณ 12.5 % ของความผิดพลาดทั้งหมด ซึ่งจะเห็นได้ว่าความผิดพลาดที่ความสั้นยาวของเสียงสระในโมเดลเกิดขึ้นไม่มาก ในขณะที่ความผิดพลาดในเรื่องของความกำกวมของเสียงพยัญชนะต้นและเสียงพยัญชนะสะกดของโมเดลใหม่เกิดขึ้นประมาณ 33 % ของความผิดพลาดทั้งหมด ซึ่งก็เกิดขึ้นมากรองจากความผิดพลาดที่เสียงสระ ดังนั้นจึงจะเห็นได้ว่าโมเดลใหม่ได้ทำการแก้ไขความผิดพลาดที่เกิดขึ้นบ่อยในโมเดลของระบบอ้างอิงได้ดีขึ้น สำหรับความผิดพลาดในการทำนายเสียงสระผิดและความกำกวมของเสียงพยัญชนะต้นและเสียงพยัญชนะสะกดที่เกิดขึ้นในโมเดลใหม่เป็นความผิดพลาดที่แก้ไขได้ยาก เนื่องจากความกำกวมของอักขระภาษาอังกฤษตัวเดียวกันที่สามารถอ่านออกเสียงสระได้หลายเสียงและตัวอักขระที่อาจเป็นเสียงพยัญชนะต้นของพยางค์ที่พิจารณาหรือเสียงพยัญชนะสะกดของพยางค์ถัดไปซึ่งเกิดขึ้นแบบไม่แน่นอน ทำให้เกิดความกำกวม การแก้ไขจำเป็นต้องเพิ่มข้อมูลในการเรียนรู้ให้มากขึ้น

บทที่ 5
บทสรุปผลการวิจัย และข้อเสนอแนะ

5.1 สรุปผลการวิจัย

วิทยานิพนธ์เล่มนี้ได้เสนอวิธีการหาคำอ่านในหน่วยเสียงไทยจากคำภาษาอังกฤษ ซึ่งระบบที่นำเสนอประกอบไปด้วยกระบวนการหลัก 4 ส่วน คือ 1. กระบวนการปรับแนวระหว่างอักขระภาษาอังกฤษและหน่วยเสียงไทย 2. โมเดลการทำนายหน่วยเสียง 3. กระบวนการประมวลผลหลังจากการทำนายหน่วยเสียง และ 4. โมเดลการทำนายเสียงวรรณยุกต์ กระบวนการปรับแนวเป็นกระบวนการที่จะเชื่อมโยงความสัมพันธ์ระหว่างระหว่างอักขระอังกฤษและหน่วยเสียงไทยก่อนจะนำข้อมูลที่ได้ทำการปรับแนวไปใช้ในการสร้างโมเดลการทำนายหน่วยเสียง ซึ่งในกระบวนการนี้จะใช้การปรับแนวแบบกลุ่มต่อกลุ่ม กระบวนการทำนายคำอ่านจะเริ่มจากการทำนายหน่วยเสียงก่อน จากนั้นทำการแก้ไขหน่วยเสียงที่ผิดพลาดตามหลักการพยางค์ในภาษาไทยด้วยกระบวนการประมวลผลหลังจากการทำนายหน่วยเสียง และสุดท้ายคือการทำนายเสียงวรรณยุกต์ให้กับทุกพยางค์ ซึ่งโมเดลการทำนายหน่วยเสียงถูกสร้างจากอัลกอริทึมออนไลน์ดิสคริมิเนทีฟเทรนนิงเฟรมเวิร์ค ด้วยค่าลักษณะสำคัญ 5 อย่าง คือ บริบท, เอ็นแกรม, มาคอฟออร์เดอร์, ลิเนียร์เชน และจอยนต์เอ็นแกรม และโมเดลการทำนายเสียงวรรณยุกต์ถูกสร้างด้วยอัลกอริทึมต้นไม้ ด้วยค่าลักษณะสำคัญ 13 อย่าง คือ เสียงพยัญชนะต้นของพยางค์ที่กำลังพิจารณา, เสียงสระของพยางค์ที่กำลังพิจารณา, เสียงพยัญชนะสะกดของพยางค์ที่กำลังพิจารณา , กลุ่มของพยัญชนะต้นของพยางค์ที่กำลังพิจารณา, ความยาวของเสียงสระของพยางค์ที่กำลังพิจารณา, มาตราสะกดของพยางค์ที่กำลังพิจารณา, จำนวนพยางค์ของคำที่กำลังพิจารณา, ตำแหน่งของพยางค์ที่กำลังพิจารณา, คำเป็นคำตายของพยางค์ที่กำลังพิจารณา, เสียงวรรณยุกต์ของพยางค์ก่อนหน้า, อักขระพยางค์ภาษาอังกฤษของพยางค์ที่กำลังพิจารณา, อักขระพยางค์ภาษาอังกฤษของพยางค์ก่อนหน้า และอักขระพยางค์ภาษาอังกฤษของพยางค์ถัดไป ผลลัพธ์ของทั้งสองโมเดลมีความแม่นยำที่เพิ่มขึ้นมากเมื่อเทียบกับระบบอ้างอิง และทำให้ผลลัพธ์รวมมีความแม่นยำเพิ่มขึ้นมากเช่นกัน โดยมีความแม่นยำเฉลี่ยในระดับพยางค์เท่ากับ 76.03%, ความแม่นยำเฉลี่ยในระดับคำเท่ากับ 53.93% และผลลัพธ์ความพึงพอใจที่ 90.95% ซึ่งสามารถทำนายได้ผลลัพธ์ที่แม่นยำเพิ่มขึ้นอย่างเห็นได้ชัดเมื่อเปรียบเทียบกับระบบอ้างอิง

ความแม่นยำที่เพิ่มขึ้นเป็นผลมาจากค่าลักษณะสำคัญที่เสนอคือ กลุ่มของพยัญชนะสะกดในภาษาไทยและพยางค์ภาษาอังกฤษ ที่มีส่วนช่วยให้โมเดลการทำนายเสียงวรรณยุกต์ทำนายเสียงวรรณยุกต์ได้แม่นยำมากขึ้นเมื่อเทียบกับระบบอ้างอิง และนอกจากนั้นการประยุกต์ใช้

โมเดลออนไลน์ดิสคริมิเนทีฟเทรนนิงเฟรมเวิร์ค ซึ่งเหมาะสมกับงานประเภทการเปลี่ยนแปรสายอักขระ เช่น งานการแปลงอักขระเป็นคำอ่าน และให้ผลลัพธ์ที่ดีในงานวิจัยภาษาอังกฤษ ก็พัฒนาให้โมเดลการทำนายเสียงวรรณยุกต์มีความแม่นยำเพิ่มขึ้นอย่างเห็นได้ชัด และการสังเกตข้อผิดพลาดที่เกิดขึ้นจากโมเดลทำนายหน่วยเสียงก็สามารถนำไปสร้างเป็นกฎแบบง่าย เพื่อใช้แก้ไขข้อผิดพลาดเหล่านี้ได้บางส่วน ทั้งหมดนี้ล้วนแล้วแต่ทำให้โมเดลการหาคำอ่านของคำภาษาอังกฤษในหน่วยเสียงภาษาไทยพัฒนาขึ้น ซึ่งก็เป็นไปตามที่ผู้วิจัยได้คาดหวังไว้ และแสดงให้เห็นว่าระบบที่เสนอนั้นสามารถนำไปใช้ในการหาคำอ่านในหน่วยเสียงภาษาไทยจากคำภาษาอังกฤษในระบบสังเคราะห์เสียงภาษาไทยได้อย่างมีประสิทธิภาพ

5.2 **ข้อเสนอแนะ**

ผลลัพธ์ของโมเดลการทำนายหน่วยเสียง ให้ผลลัพธ์ที่ไม่สูงเมื่อเทียบกับโมเดลทำนายเสียงวรรณยุกต์ จึงทำให้เกิดความผิดพลาดรวมมากกว่าที่คาดการณ์ไว้ ดังนั้นการพัฒนาต่อไปควรจะให้ความสำคัญกับโมเดลการทำนายหน่วยเสียง เพื่อปรับให้มีความแม่นยำเพิ่มขึ้น และจะเป็นส่วนสำคัญที่ทำให้ความแม่นยำรวมดีขึ้น นอกจากนั้นความผิดพลาดที่เกิดขึ้นเยอะที่สุดก็คือการทำนายเสียงสระผิด ซึ่งเกิดขึ้นถึงครึ่งหนึ่งของความผิดพลาดทั้งหมด ดังนั้นการพัฒนาต่อจึงควรจะเน้นที่การแก้ปัญหานี้เป็นหลัก

ระบบที่นำเสนออาศัยผลลัพธ์จากสองโมเดลที่แยกจากกัน จึงทำให้การรวมผลลัพธ์เกิดความผิดพลาดขึ้นมาก ดังนั้นการสร้างโมเดลในการหาคำอ่านที่รวมเป็นโมเดลเดียว อาจจะลดความซับซ้อนและทำให้การทำนายคำอ่านมีความแม่นยำมากขึ้นจากการอาศัยข้อมูลค่าลักษณะสำคัญที่มาจากทั้งของโมเดลพร้อม ๆ กัน

รายการอ้างอิง

[1] Chalamandaris, A., Karabetsos, S., Tsiakoulis, P., and Raptis, S. A Unit Selection Text-to-Speech Synthesis System Optimized for Use with Screen Readers. Proceedings of IEEE Transactions on Consumer Electronics, pp. 1890-1897, 2010.

[2] Chomtip Pornpanomchai, Nichakant Soontharanont, Charnchai Langla and Narunat Wongsawat. A Dictionary-Based Approach for Thai Text to Speech (TTTS). Proceedings of 3rd International Conference on Measuring Technology and Mechatronics Automation (ICMTMA), pp. 40-43, 2011.

[3] Chai Wutiwiwatchai, Ausdang Thangthai, Ananlada Chotimongkol, Chatchawarn Hansakunbuntheung and Nattanun Thatphithakkul. Accent Level Adjustment in Bilingual Thai-English Text-to-Speech Synthesis. Proceedings of Automatic Speech Recognition and Understanding (ASRU), pp. 295-299, 2011.

[4] วิโรจน์ อรุณมานะกุล. อักขรวิธีไทยและการถอดอักษรระหว่างภาษาไทยและภาษาอังกฤษ. กรุงเทพฯ: คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2551.

[5] Wikipedia, International Phonetic Alphabet [online]. Available from: http://en.wikipedia.org/wiki/IPA [8 October 2012].

[6] Ausdang Thangthai, Chai Wutiwiwatchai, Anocha Ragchatjaroen and Sittipong Saychum. A learning method for Thai phonetization of English words. Proceedings of INTERSPEECH, pp. 1777-1780, 2007.

[7] The Carnegie Mellon University, The CMU Pronouncing Dictionary [online]. Available from: http://www.speech.cs.cmu.edu/cgi-bin/cmudict [15 October 2012].

[8] Wirote Aroonmanakun, Nuttakorn Thapthong, Pakaket Wattuya, Benjawan Kasisopa and Sudaporn Luksaneeyanawin. Generating Thai Transcriptions of English Words. Proceedings of SEALS 14, May, 2004.

[9] Bosch, A.V.D., and Daelemans, W. Data-Oriented Methods for Grapheme-to-Phoneme Conversion. Proceedings of European Chapter of the ACL, Utrecht, April, 1993.

[10] Chu, M., Peng, H., Zhou, Y., Niu Z., and Chang, E. Microsoft Mulan- A Bilingual

TTS System. Proceedings of Acoustics Speech and Signal Processing 2003 (ICASSP '03), pp 264-267, Hong Kong, China, 2003.

[11] Pradit Mittrapiyanuruk, Chatchawarn Hansakunbuntheung, Virongrong Tesprasit and Virach Sornlertlamvanich. Issues in Thai Text-to-Speech Synthesis: The NECTEC Approach. Proceedings of NECTEC Annual Conference 2000, Bangkok Thailand, 2000.

[12] Black, A., Lenzo, K., and Pagel, V. Issues in Building General Letter to Sound Rules. Proceedings of 3rd ESCA Speech Synthesis Workshop, pp. 77-80, Jenolan Caves, Australia, 1998.

[13] Ananlada Chotimongkol and Black, A.W. Statistically Trained Orthographic to Sound Models for Thai. Proceedings of ICSLP 2000, Beijing, China, 2000.

[14] Taylor, P. Hidden Markov Models for Grapheme to Phoneme Conversion. In Proceedings of Interspeech, Lisbon, Portugal, 2005.

[15] Wang, D., and King, S. Letter-to-Sound Pronunciation Prediction Using Conditional Random Fields. Proceedings of IEEE Signal Processing Letters, vol. 18, no. 2, pp. 122–125, 2011.

[16] Sittichai Jiampojamarn, Cherry, C., and Kondrak, G. Joint Processing and Discriminative Training for Letter-to-Phoneme Conversion. Proceedings of ACL, pp. 905–913, 2008.

[17] Sittichai Jiampojamarn, Cherry, C., and Kondrak, G. Integrating Joint n-gram Features into a Discriminative Training Framework. Proceedings of NAACL, 2010.

[18] Song, Y., and Kit, C. Does Joint Decoding Really Outperform Cascade Processing in English-to-Chinese Transliteration Generation? The Role of Syllabification. Proceedings of Machine Learning and Cybernetics (ICMLC), pp 3323-3328, Qingdao, China, 2010.

[19] Bisani, M., and Ney, H. Investigations on Joint-Multigram Models for Grapheme-to-Phoneme Conversion. Proceedings of 7th International Conference on Spoken Language Processing, pp. 105-108, 2002.

[20] Sittichai Jiampojamarn, Kondrak, G., and Sherif, T. Applying Many-to-Many Alignments and Hidden Markov Models to Letter-to-Phoneme Conversion. Proceedings of HLT-NAACL, pp. 372–379, 2007.

[21] Damper, R.I., Marchand, Y., Marsters, J.DS., and Bazin, A.I. Aligning Text and Phonemes for Speech Technology Applications Using an EM-Like Algorithm. Proceedings of International Journal of Speech Technology, 2005.

[22] Zens, R., and Ney, H. Improvements in Phrase-Based Statistical Machine Translation. Proceedings of HLT NAACL 2004: Main Proceedings, pp. 257–264, Boston, Massachusetts, USA, 2004.

[23] Collins, M. Discriminative Training Methods for Hidden Markov Models: Theory and Experiments with Perceptron Algorithms. In EMNLP '02: Proceedings of the ACL-02 conference on Empirical methods in natural language processing, pp. 1-8, Morristown, NJ, USA, 2002.

[24] Crammer, K., and Singer, Y. Ultraconser-Vative Online Algorithms for Multiclass Problems. The Journal of Machine Learning Research, 2003.

[25] Taylor, P., Caley, R., Black, A.W., and King, S. Edinburgh Speech Tools Library System Documentation. University of Edinburgh, 1997.

ภาคผนวก

คำภาษาอังกฤษและคำอ่านจากระบบการทำนายคำอ่านที่ใช้ทดสอบความพึงพอใจ

คำที่	คำภาษาอังกฤษ	ผลลัพธ์คำอ่านจากระบบการทำนายคำอ่าน
1	ADOBE	z-a-z^-1|d-oo-z^-0|b-ii-z^-2|
2	ADVANCE	z-x-t^-3|w-aa-n^-3|
3	APPLE	z-x-p^-3|ph-q-l^-2|
4	APPLICATION	z-x-p^-3|phl-ii-z^-0|kh-ee-z^-0|ch-a-n^-2|
5	ASEAN	z-aa-z^-0|s-iia-n^-2|
6	BANGKOK	b-x-ng^-0|kh-@-k^-3|
7	BLACK	bl-x-k^-3|
8	BODY	b-@@-z^-0|d-ii-z^-2|
9	CERTIFICATE	s-qq-z^-0|th-i-z^-3|f-i-z^-3|kh-ee-t^-1|
10	CHINESE	ch-i-n^-0|z-ii-s^-3|
11	CHRISTMAS	khr-i-s^-3|m-a-s^-2|
12	CLICK	khl-i-k^-3|
13	COMMUNICATION	kh-@-m^-0|m-i-w^-0|n-i-z^-3|kh-ee-z^-0|ch-a-n^-2|
14	COMMUNITY	kh-@-m^-0|m-i-w^-0|n-i-z^-3|t-ii-z^-2|
15	COMPUTER	kh-@-m^-0|ph-i-w^-0|t-qq-z^-2|
16	DAY	d-ee-z^-0|
17	DECORATION	d-e-k^-3|kh-qq-z^-0|r-ee-z^-0|ch-a-n^-2|
18	ECONOMIC	z-ii-z^-0|kh-oo-z^-0|n-oo-z^-0|m-i-k^-1|
19	EXPO	z-e-k^-3|p-oo-z^-2|
20	FACEBOOK	f-ee-s^-3|b-u-k^-3|
21	FOOD	f-uu-t^-3|
22	GARDEN	k-aa-z^-0|d-e-n^-2|
23	HAPPY	h-x-p^-3|ph-ii-z^-2|
24	HOME	h-oo-m^-0|
25	HONDA	h-@-n^-0|d-aa-z^-2|
26	HYBRID	h-a-j^-0|br-i-t^-1|
27	IDEA	z-a-j^-0|d-iia-z^-0|

28	INFO	z-i-n^-0|f-oo-z^-0|
29	INFORMATION	z-i-n^-0|f-@@-z^-0|m-ee-z^-0|ch-a-n^-2|
30	IPHONE	z-i-z^-1|f-oo-n^-0|
31	JAVA	c-aa-z^-0|w-aa-z^-2|
32	JUNIOR	c-uu-z^-0|n-iia-z^-1|
33	JUST	c-a-s^-3|
34	KEYBOARD	kh-ii-z^-0|b-@@-t^-1|
35	KNOW	n-oo-w^-0|
36	LEARNING	l-qq-n^-0|n-i-ng^-2|
37	LIMITED	l-i-z^-3|m-i-z^-3|th-e-t^-2|
38	LIVE	l-ii-f^-3|
39	MACBOOK	m-x-k^-3|b-u-k^-3|
40	MAIL	m-e-l^-3|
41	MARKET	m-aa-z^-0|k-e-t^-2|
42	MOBILE	m-oo-z^-0|b-aa-j^-0|
43	MOVIE	m-uu-z^-0|w-ii-z^-2|
44	MUSIC	m-i-w^-0|s-i-k^-1|
45	NASA	n-aa-z^-0|s-aa-z^-2|
46	NATION	n-ee-z^-0|ch-a-n^-2|
47	NETWORK	n-e-t^-3|w-qq-k^-3|
48	NEW	n-i-w^-0|
49	OLYMPIC	z-oo-z^-0|l-i-m^-0|p-i-k^-1|
50	ONLINE	z-@@-n^-0|l-aa-j^-0|
51	OUTLET	z-a-w^-3|l-e-t^-1|
52	OUTLOOK	z-a-w^-3|l-u-k^-3|
53	PANASONIC	ph-aa-z^-0|n-aa-z^-0|s-oo-z^-0|n-i-k^-1|
54	PARAGON	ph-xx-z^-0|r-aa-z^-0|k-@-n^-2|
55	PHOTOSHOP	f-oo-z^-0|t-oo-z^-0|ch-@-p^-3|
56	PROFESSIONAL	phr-oo-z^-0|f-e-s^-3|ch-a-n^-0|n-@-l^-2|

57	QUEEN	kw-ii-n^-0|
58	RATE	r-ee-t^-1|
59	RISK	r-i-s^-3|
60	SERVICE	s-qq-z^-0|w-i-s^-3|
61	SHOW	ch-oo-w^-0|
62	SIAM	s-a-z^-1|j-aa-m^-0|
63	SMART	s-a-z^-1|m-aa-t^-3|
64	SOCIAL	s-oo-z^-0|ch-iia-l^-2|
65	SOLUTION	s-oo-z^-0|l-uu-z^-0|ch-a-n^-2|
66	STAR	s-a-z^-1|t-aa-z^-0|
67	TAX	th-x-k^-3|
68	TECHNOLOGY	th-e-k^-3|n-oo-z^-0|l-oo-z^-0|j-ii-z^-2|
69	THAILAND	th-a-j^-3|l-xx-n^-0|
70	THE	d-q-z^-1|
71	TOYOTA	t-oo-z^-0|j-oo-z^-0|t-aa-z^-2|
72	TRANSCRIPT	thr-aa-n^-0|s-a-z^-1|khr-i-p^-3|
73	TREND	thr-e-n^-0|
74	TRUE	thr-uu-z^-0|
75	ULTRA	z-a-l^-0|tr-aa-z^-2|
76	UPDATE	z-a-p^-3|d-ee-t^-1|
77	VIRUS	w-a-j^-0|r-a-s^-3|
78	VOICE	w-@@-js^-0|
79	VOTE	w-oo-t^-1|
80	WORLD	w-qq-l^-0|
81	XENON	s-ii-z^-0|n-@-n^-2|
82	YAHOO	j-a-z^-3|h-uu-z^-0|
83	YEAR	j-iia-z^-0|
84	ZONE	s-oo-n^-0|

ประวัติผู้เขียนวิทยานิพนธ์

นายกฤตนันท์ พิทักษ์ภวัตกุล เกิดวันศุกร์ที่ 13 กรกฎาคม พ.ศ.2533 ที่กรุงเทพมหานคร สำเร็จการศึกษาระดับมัธยมศึกษาจาก โรงเรียนสารวิทยา กรุงเทพมหานคร สำเร็จการศึกษาระดับปริญญาบัณฑิต เกียรตินิยมอันดับหนึ่ง สาขาวิชาวิศวกรรมคอมพิวเตอร์ จากคณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย กรุงเทพมหานคร ในปีการศึกษา 2554