การจำแนกหน่วยเสียงภาษาไทยโดยใช้การแทนเสียงแบบทั้งเซกเมนต์

นางสาวหนึ่งฤทัย เอกชัยวรสิน

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมคอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2549
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

THAI PHONEME CLASSIFICATION USING SEGMENTAL REPRESENTATION

Miss Nuengruethai Ekkachaiworrasin

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

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 2006
Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การจำแนกหน่วยเสียงภาษาไทยโดยใช้การแทนเสียงแบบทั้งเซกเมนต์
โดย	นางสาวหนึ่งฤทัย เอกชัยวรสิน
สาขาวิชา	วิศวกรรมคอมพิวเตอร์
อาจารย์ที่ปรึกษา	อาจารย์ ดร.อติวงศ์ สุชาโต
อาจารย์ที่ปรึกษาร่วม	อาจารย์ ดร. โปรดปราน บุณยพุกกณะ

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

..... คณบดีคณะวิศวกรรมศาสตร์
(ศาสตราจารย์ ดร.ดิเรก ลาวัณย์ศิริ)

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ
(รองศาสตราจารย์ ดร.บุญเสริม กิจศิริกุล)

..... อาจารย์ที่ปรึกษา
(อาจารย์ ดร.อติวงศ์ สุชาโต)

..... อาจารย์ที่ปรึกษาร่วม
(อาจารย์ ดร. โปรดปราน บุณยพุกกณะ)

..... กรรมการ
(อาจารย์ ดร.พิษณุ คนองชัยยศ)

..... กรรมการ
(ดร.ชัย วุฒิวิวัฒน์ชัย)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

หนึ่งฤทัย เอกชัยวรสิน : การจำแนกหน่วยเสียงภาษาไทยโดยใช้การแทนเสียงแบบทั้งเซกเมนต์. (THAI PHONEME CLASSIFICATION USING SEGMENTAL REPRESENTATION) อ. ที่ปรึกษา : อ.ดร.อติวงศ์ สุชาโต, อ.ที่ปรึกษาร่วม : อ.ดร. โปรดปราน บุณยพุกกณะ, 75 หน้า.

ในปัจจุบันนี้การหาลักษณะเฉพาะแบบอาศัยกรอบเวลานั้น ถือได้ว่าเป็นวิธีการที่มีความนิยมสูง แต่อย่างไรก็ตามหลักการทำงานดังกล่าว ยังคงมีสมมติฐานบางอย่างที่ยังไม่เหมาะสมกับธรรมชาติของเสียงพูด รวมไปถึงมีข้อจำกัดบางประการในการหาสมบัติของสัญญาณเสียงใดๆ ซึ่งปัญหาดังกล่าว สามารถแก้ไขด้วยหลักการทำงานแบบอาศัยเซกเมนต์ ที่มีการแบ่งสัญญาณเสียงออกเป็นเซกเมนต์ของหน่วยเสียงแทนการแบ่งตามกรอบเวลา ดังนั้นวิทยานิพนธ์นี้จึงได้นำเสนอหลักการหาลักษณะเฉพาะแบบอาศัยเซกเมนต์ ซึ่งได้จากการหาค่าสัมประสิทธิ์เมลฟรีเควนซีเคปตรอล 12 หลักรวมกับค่าพลังงาน 1 หลัก จากส่วนหน้า ส่วนกลางและส่วนหลังของเซกเมนต์ต่อกันรวม 39 หลัก ร่วมกับการใช้สมบัติความยาวของเซกเมนต์รวมเป็นลักษณะเฉพาะทั้งสิ้น 40 หลัก เมื่อทดสอบความสามารถในการแทนเสียงของลักษณะเฉพาะด้วยการจำแนกหน่วยเสียงภาษาไทย 52 หน่วยเสียง โดยใช้การวิเคราะห์ดิสคริมิแนนต์เชิงเส้น และให้ค่าโอกาสของการตอบหน่วยเสียงใดๆเท่ากัน ผลการจำแนกหน่วยเสียงที่ได้คือ 61.41 เปอร์เซนต์ และเมื่อให้ค่าโอกาสของการตอบหน่วยเสียงใดๆตามจำนวนหน่วยเสียงที่ใช้ฝึกฝนความถูกต้องที่ได้อยู่ที่ 66.14 เปอร์เซนต์ เมื่อเปรียบเทียบประสิทธิภาพความถูกต้องกับการจำแนกหน่วยเสียงแบบอาศัยกรอบเวลา ซึ่งมีความถูกต้องเฉลี่ย 56.95 เปอร์เซนต์ พบว่ามีความถูกต้องสูงกว่าถึง 9.19 เปอร์เซนต์ จากการวิเคราะห์ค่าความมีส่วนร่วมต่อการจำแนกหน่วยเสียงของลักษณะเฉพาะที่นำเสนอ พบว่าค่าความมีส่วนร่วมของส่วนกลางนั้น มีค่าเกือบครึ่งหนึ่งของการจำแนกหน่วยเสียงทั้งหมด ผลดังกล่าวสนับสนุนแนวความคิดในการหาลักษณะเฉพาะที่มีการแบ่งเซกเมนต์ออกเป็น 3 ส่วน ที่กล่าวว่าส่วนกลางของเซกเมนต์สามารถแทนหน่วยเสียงได้ดีที่สุด

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาควิชา.....วิศวกรรมคอมพิวเตอร์.....	ลายมือชื่อนิสิต.....หนึ่งฤทัย เอกชัยวรสิน.....
สาขาวิชา.....วิศวกรรมคอมพิวเตอร์.....	ลายมือชื่ออาจารย์ที่ปรึกษา.....อติวงศ์.....
ปีการศึกษา.....2549.....	ลายมือชื่ออาจารย์ที่ปรึกษาร่วม.....[illegible].....

497 06693 21 : MAJOR COMPUTER ENGINEERING

KEY WORD: PHONEME CLASSIFICATION/ SEGMENT-BASED SPEECH RECOGINTION/ THAI SPEECH RECOGNITION / FEATURE VECTOR REPRESENTATION

NUENGRUETHAI EKKACHAIWORRASIN : THAI PHONEME CLASSIFICATION USING SEGMENTAL REPRESENTATION. THESIS ADVISOR : ATIWONG SUCHATO, Ph.D., THESIS COADVISOR : PROADPRAN PUNYABUKKANA, Ph.D., 75 pp.

Nowadays, spectral feature extraction at a fixed frame rate is a highly popular technique for representing speech signals. However, some assumptions used by this technique are not suitable for natural speech. Also, the technique has various limitations in acquiring some types of acoustic properties. To avoid these limitations, a segmental representation method separates an acoustic speech signal into small segments according to the underlying phonemes before performing the feature extraction. In this thesis, an approach for extracting feature vectors using segmental representation has been proposed. By means of this approach, a 40-dimensional feature vector, consisting of 12 Mel Frequency Cepstral Coefficients and an energy feature of three regions: the frontal region of the segment, the middle region of the segment, and the rear region of the segment, together with the segment duration, is used to represent a speech segment. In the experiments, where 52 Thai phonemes are classified using Linear Discriminant Analysis, the classification accuracy is 66.14% when prior probabilities are used while it is 61.41% without prior probabilities. The best accuracy obtained using our segment-based approach is 9.19% higher than the one using a baseline frame-based approach, which is 56.95%. In addition, it is found as a result of our contribution analysis that features extracted from the middle region contribute the most to the classification.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Department .....Computer Engineering...	Student's Signature Nuengruethai Ekkachaiworrasin.
Field of study.....Computer Engineering...	Advisor's Signature.....
Academic year 2006	Co-advisor's Signature.....

กิตติกรรมประกาศ

การทำงานวิทยานิพนธ์นี้เกิดขึ้นจากแรงบันดาลใจจาก อ.ดร. อติวงศ์ สุชาโต และ อ.ดร. โปรดปราน บุณยพุกกณะ ที่ชี้เห็นถึงความสำคัญ ความน่าสนใจในการศึกษางานวิจัยที่เกี่ยวข้องกับการรู้จำเสียงโดยเฉพาะการรู้จำเสียงภาษาไทย ที่ถือว่ามีการพัฒนาขึ้นมานานแล้ว แต่กลับยังไม่เป็นที่แพร่หลาย รวมทั้งขอบพระคุณคณะกรรมการทุกท่าน ได้แก่ รศ.ดร. บุญเสริม กิจศริกุล อ.ดร. พิษณุ คนองชัยยศ และ ดร.ชัย วุฒิวิวัฒน์ชัย ที่สละเวลาอันมีค่ามาชี้แนะข้อบกพร่อง เพื่อเติมเต็มให้จุดประสงค์ของวิทยานิพนธ์นี้สมบูรณ์ยิ่งขึ้น

ขอขอบคุณภาควิชาวิศวกรรมคอมพิวเตอร์รวมถึงทีมงานผู้ดูแลระบบที่อำนวยความสะดวกในด้านการใช้งานเครื่องคอมพิวเตอร์คลัสเตอร์ ซึ่งสนับสนุนโดยศูนย์ไทยกริดแห่งชาติ(Thai National Grid Center : TNGC)

ขอขอบคุณพี่ๆเพื่อนๆสมาชิกห้องปฏิบัติการ SLS (Spoken Language system) ทุกคน รวมไปถึงพี่ๆเพื่อนๆระดับบัณฑิตศึกษาทุกท่านสำหรับคำแนะนำ บรรยากาศ และทัศนคติที่ดีในการทำงาน

ขอขอบคุณจุฬาลงกรณ์มหาวิทยาลัยและอาจารย์ทุกท่าน ผู้ประสิทธิประสาทความรู้กระบวนการความคิดอย่างวิศวกรรม ขอขอบคุณโรงเรียนซางตาครู้สคอนแวนท์ โรงเรียนมหิดลวิทยานุสรณ์ คุณครูและอาจารย์ทุกท่านที่ได้อบรมสั่งสอน หล่อหลอมความคิดอย่างวิทยาศาสตร์ซึ่งเป็นพื้นฐานต่อมุมมองความคิดทั้งปวงของผู้เขียน

สุดท้ายนี้ขอกราบขอบพระคุณคุณพ่อและคุณแม่ที่คอยเป็นกำลังใจ ช่วยเหลือ สนับสนุนและเห็นชอบต่อการกระทำในทุกๆด้านของผู้เขียน จนทำให้ผู้เขียนสามารถประสบความสำเร็จจนถึงทุกวันนี้

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ง
บทคัดย่อภาษาอังกฤษ ..... จ
กิตติกรรมประกาศ ..... ฉ
สารบัญ ..... ช
สารบัญภาพ ..... ฌ
สารบัญตาราง ..... ฎ
บทที่ 1 บทนำ 1
ความเป็นมาและความสำคัญของปัญหา 1
วัตถุประสงค์ของการวิจัย 1
ขอบเขตของการวิจัย 2
ขั้นตอนการวิจัย 2
ประโยชน์ที่คาดว่าจะได้รับ 3
ผลงานที่ตีพิมพ์จากวิทยานิพนธ์ 3
โครงร่างของวิทยานิพนธ์ 3
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง 5
ทฤษฎีที่เกี่ยวข้อง 5
1. หลักการรู้จำเสียง 5
2. การแทนสัญญาณเสียงด้วยสมบัติในโดเมนความถี่ 10
3. การวิเคราะห์ดิสคริมิแนนต์เชิงเส้น 11
4. การทดสอบนัยสำคัญ (Significance testing) 14
งานวิจัยที่เกี่ยวข้อง 15
1. ลักษณะเฉพาะที่มีการทำงานแบบอาศัยกรอบเวลา 15
2. ลักษณะเฉพาะที่มีการทำงานแบบอาศัยเซกเมนต์ 18
บทที่ 3 หน่วยเสียงภาษาไทย 20
นิยามหน่วยเสียงในภาษาไทย 20
1. เสียงสระ (Vowels) 20
2. เสียงพยัญชนะ (Consonants) 21
3. เสียงกึ่งสระ (Semi-vowels) 24
กลุ่มหน่วยเสียงภาษาไทยสำหรับการจำแนกหน่วยเสียง 26

หน้า

สัทอักษรสากล 27
บทที่ 4 การหาลักษณะเฉพาะ 30
การศึกษาพฤติกรรมของหน่วยเสียง 30
ลักษณะเฉพาะ 35
บทที่ 5 การทดลองจำแนกหน่วยเสียง 38
ฐานข้อมูลเสียง 38
การทดลองจำแนกหน่วยเสียงด้วยหลักการการวิเคราะห์ดิสคริมิแนนต์เชิงเส้นโดยใช้ลักษณะเฉพาะแบบต่างๆที่มีวิธีการทำงานแบบอาศัยเซกเมนต์ 39
1. ขั้นตอนการทดลอง 39
2. ผลการทดลอง 40
การวิเคราะห์ค่าความมีส่วนร่วมของลักษณะเฉพาะ 58
การทดลองจำแนกหน่วยเสียงด้วยหลักการทำงานแบบอาศัยกรอบเวลาเพื่อเปรียบเทียบประสิทธิภาพ 61
1. วิธีการทดลอง 61
2. ผลการทดลอง 61
การเปรียบเทียบประสิทธิภาพของการจำแนกหน่วยเสียงระหว่างวิธีการทำงานแบบเซกเมนต์และวิธีการทำงานแบบอาศัยกรอบเวลา 66
บทที่ 6 สรุปผลการวิจัย 68
ผลการวิจัย 68
คุณประโยชน์ของงานวิจัยนี้ต่อวงวิชาการ 69
ข้อเสนอแนะ 71
รายการอ้างอิง 73
ประวัติผู้เขียนวิทยานิพนธ์ 75

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญภาพ

หน้า

รูปที่ 2.1 ขั้นตอนการทำงานของหลักการรู้จำเสียง 5
รูปที่ 2.2 ลักษณะการตัดแบ่งสัญญาณเสียงตามการทำงานแบบอาศัยกรอบเวลา 6
รูปที่ 2.3 ลำดับของสัญญาณเสียงที่ทำการแบ่งเสียงเป็นหน่วยย่อยๆ 6
รูปที่ 2.4 แบบจำลองเสียงตามหลักการทำงานฮิดเดนมาร์คอฟ 7
รูปที่ 2.5 ลำดับของสัญญาณเสียงที่ทำการแบ่งเสียงเป็นหน่วยย่อยๆ 9
รูปที่ 2.6 แบบจำลองภาษาสำหรับการรู้จำเสียง 10
รูปที่ 2.7 ขั้นตอนการหาค่าเอ็มเอฟซีซีของสัญญาณเสียง 11
รูปที่ 2.8 ตัวอย่างการจำแนกข้อมูล 2 ประเภทเปรียบเทียบระหว่างการจำแนกข้อมูลบนปริภูมิดั้งเดิมกับการจำแนกข้อมูลบนปริภูมิใหม่ที่ผ่านกระบวนการการวิเคราะห์ดิสคริมิแนนต์เชิงเส้น . 12
รูปที่ 2.9 บริเวณของการกระจายตัวของค่าความน่าจะเป็นที่รวมเป็นค่า $\boldsymbol{P}$ 15
รูปที่ 3.1 สเปกโตรแกรมเสียงสระอา (/aa/) 20
รูปที่ 3.2 สเปกโตรแกรมเสียงสระอัว (/uua/) 21
รูปที่ 3.3 สเปกโตรแกรมเสียงเสียดแทรก ส,ศ,ษ,ซ (/s/) 22
รูปที่ 3.4 สเปกโตรแกรมเสียงพยัญชนะกัก “ข”, “ค”, “ฆ” (/kh/) 23
รูปที่ 3.5 สเปกโตรแกรมเสียงพยัญชนะกัก “ก” (/k/) 23
รูปที่ 3.6 สเปกโตรแกรมเสียงพยัญชนะกักแบ่งเป็นส่วนแถบพลังงานความถี่ต่ำ (/vb/) และ“ด” (/d/) 23
รูปที่ 3.7 สเปกโตรแกรมเสียงพยัญชนะกึ่งเสียดแทรก “จ” (/c/) 24
รูปที่ 3.8 สเปกโตรแกรมเสียงพยัญชนะนาสิก “ม” (/m/) 24
รูปที่ 3.9 สเปกโตรแกรมเสียงกึ่งสระ “ร” (/r/) 25
รูปที่ 3.10 สเปกโตรแกรมเสียงกึ่งสระ “ล” (/l/) 25
รูปที่ 4.1 สเปกโตรแกรมแสดงเสียงสระอาจากคำว่า “กาน” (/kaan^/) 32
รูปที่ 4.2 สเปกโตรแกรมแสดงเสียงสระอาจากคำว่า “ยาน” (/jaan^/) 32
รูปที่ 4.3 สเปกโตรแกรมแสดงเสียงสระอาจากคำว่า “จาน” (/caan^/) 33
รูปที่ 4.4 สเปกโตรแกรมแสดงเสียงสระอาจากคำว่า “กา” (/kaa/) ที่ตามด้วยหน่วยเสียงในแถบพลังงานความถี่ต่ำ (/vb/) 33
รูปที่ 4.5 สเปกโตรแกรมแสดงเสียงสระอาจากคำว่า “กา” (/kaa/) 33
รูปที่ 4.6 สเปกโตรแกรมแสดงเสียงสระอาจากคำว่า “จอน” (/c@@n^/) 34
รูปที่ 4.7 สเปกโตรแกรมแสดงเสียงสระอาจากคำว่า “ลาย” (/laaj^/) 34

หน้า

รูปที่ 4.8 การแบ่งสัดส่วนเซกเมนต์ในการหาลักษณะเฉพาะในแบบ ก4..... 37
รูปที่ 5.1 ขั้นตอนการทดลองด้วยหลักการการวิเคราะห์ดิสคริมิแนนต์เชิงเส้น โดยใช้ลักษณะเฉพาะที่มีการทำงานแบบอาศัยเซกเมนต์..... 40
รูปที่ 5.2 กราฟแสดงการกระจายตัวของสระเสียงสั้นและยาวตามความยาวของสระ..... 48
รูปที่ 5.3 กราฟแสดงความถี่ของค่าความน่าจะเป็นของหน่วยเสียงที่จำแนกถูกและผิด..... 49
รูปที่ 5.4 กราฟแสดงความสัมพันธ์ระหว่างค่าความน่าจะเป็นที่ตอบถูกเมื่อทราบค่าความน่าจะเป็นที่จะตอบหน่วยเสียงใดๆเมื่อรู้สัญญาณเสียงกับค่าความน่าจะเป็นที่จะตอบหน่วยเสียงใดๆ เมื่อรู้สัญญาณเสียง..... 49
รูปที่ 5.5 กราฟแสดงสัดส่วนค่าความมีส่วนร่วมรวมของส่วนต่างๆในเซกเมนต์..... 59
รูปที่ 6.1 ตัวอย่างเซกเมนต์หน่วยเสียง “ส”(/s/)..... 72

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญตาราง

หน้า

ตารางที่3.1 สระตามลักษณะตำแหน่งความสูงของลิ้น 20
ตารางที่3.2 พยัญชนะในแต่ละประเภทแบ่งตามพฤติกรรมการเปล่งเสียง 21
ตารางที่3.3 พยัญชนะกึ่งสระตามพฤติกรรมการเปล่งเสียง 25
ตารางที่3.4 สัทอักษรสากลที่ใช้แทนเสียงสระในภาษาไทย 27
ตารางที่3.5 สัทอักษรสากลที่ใช้แทนเสียงพยัญชนะต้นในภาษาไทย 28
ตารางที่3.6 สัทอักษรสากลที่ใช้แทนเสียงพยัญชนะตัวสะกดภาษาไทย 29
ตารางที่4.1 ตารางแสดงสัดส่วนในการแบ่งเซกเมนต์ของหน่วยเสียง 1 เซกเมนต์ 36
ตารางที่5.1 จำนวนหน่วยเสียงที่ใช้ในชุดฝึกฝนและชุดทดสอบ 39
ตารางที่5.2 ผลการจำแนกหน่วยเสียงด้วยลักษณะเฉพาะแบบต่างๆ 41
ตารางที่5.3 ผลการทดสอบนัยสำคัญระหว่างลักษณะเฉพาะแบบ ก3 กับลักษณะเฉพาะแบบอื่นๆ 42
ตารางที่5.4 ผลการจำแนกหน่วยเสียงที่ใช้ลักษณะเฉพาะแบบ ก3 เพื่อวัดประสิทธิภาพความถูกต้อง 45
ตารางที่5.5 ผลการจำแนกหน่วยเสียงแบบใช้ค่าน้ำหนักเท่ากันโดยแจกแจงตามความสับสนระหว่างกลุ่ม 50
ตารางที่5.6 ผลการจำแนกหน่วยเสียงแบบใช้ค่าน้ำหนักขึ้นกับจำนวนข้อมูลโดยแจกแจงตามความสับสนระหว่างกลุ่ม 54
ตารางที่5.7 ค่าความมีส่วนร่วมรวมในแต่ละส่วนของเซกเมนต์ 58
ตารางที่5.8 ค่าความมีส่วนร่วมในแต่ละมิติของลักษณะเฉพาะที่มีการทำงานแบบอาศัยเซกเมนต์สัดส่วน 20: 60: 20 โดยด้านซ้ายเรียงลำดับตามมิติ ด้านขวาเรียงลำดับตามขนาดของค่าความมีส่วนร่วม 60
ตารางที่5.9 ผลการจำแนกหน่วยเสียงด้วยแบบจำลองฮิดเดนมาร์คอฟโดยแจกแจงตามความสับสนระหว่างกลุ่ม 62

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 1
บทนำ

ความเป็นมาและความสำคัญของปัญหา

แบบจำลองฮิดเดนมาร์คอฟ (Hidden Markov models: HMMs) [1] นั้นถือได้ว่าเป็นเทคนิคในการรู้จำเสียงอัตโนมัติที่ได้รับความนิยมมากในปัจจุบัน ซึ่งแบบจำลองดังกล่าวมีลักษณะในการทำงานโดยได้นำหลักการในการทำงานแบบอาศัยกรอบเวลา (Frame-based approach) ซึ่งมีการกำหนดขนาดเฟรมคงที่ (Fixed-duration frame) โดยทั่วไปแล้วมักกำหนดให้เฟรมมีขนาดเล็กๆ เช่น 10 มิลลิวินาที ก่อนนำไปผ่านกระบวนการดึงลักษณะเฉพาะทางสเปกตรัม (Spectral feature vector) ของเสียงออกมา โดยหลักการรู้จำเสียงของแบบจำลองนี้จะรับลักษณะเฉพาะ (Feature vector) ดังกล่าวเข้ามาค้นหาลำดับคำภายในเครือข่ายของคำที่เป็นไปได้ทั้งหมดของแบบจำลอง

ถึงแม้ว่าหลักการการทำงานดังกล่าวของแบบจำลองจะดูเหมาะสมกับการรู้จำเสียง แต่ยังคงมีสมมติฐานหนึ่งที่เกี่ยวข้องกับพฤติกรรมของแบบจำลองที่กล่าวว่าเฟรมแต่ละเฟรมที่ติดกันนั้นมีความเป็นอิสระไม่ขึ้นต่อกัน ส่งผลให้สมมติฐานดังกล่าวนั้นไม่สอดคล้องกับลักษณะธรรมชาติของสัญญาณเสียงในภาษาต่างๆ เนื่องจากเฟรมที่ได้จากการแบ่งสัญญาณเสียงภาษาพูดดังกล่าวนั้นมีขนาดเล็กๆ ทำให้เฟรมแต่ละเฟรมที่อยู่ติดกันนั้นมักจะมาจากหน่วยเสียงเดียวกัน นอกจากนี้ลักษณะการทำงานดังกล่าวยังยากที่จะเก็บลักษณะต่างๆที่มีความสำคัญได้โดยตรงอีกด้วย เช่น ระยะเวลาของหน่วยเสียง ซึ่งถือว่าเป็นสมบัติหนึ่งที่บ่งบอกความแตกต่างระหว่างหน่วยเสียงในบางภาษา เช่น ภาษาไทย ที่มีทั้งสระเสียงสั้นและเสียงยาวซึ่งมีลักษณะเหมือนกันทุกประการจะแตกต่างกันเฉพาะระยะเวลาในการออกเสียงเท่านั้น

จากข้อจำกัดดังกล่าวข้างต้น การนำหลักการทำงานตัดแบ่งเสียงแบบแบ่งส่วนหรือการทำงานแบบอาศัยเซกเมนต์ (Segment-based approach) ซึ่งใช้วิธีการแบ่งเสียงที่เข้ามาออกเป็นหน่วยเสียงย่อยๆที่เล็กที่สุด (Phoneme) แล้วจึงนำไปทำการหาลักษณะเฉพาะของเสียงนั้น ถือได้ว่าเป็นอีกทางเลือกหนึ่งที่เหมาะสมในการรู้จำเสียง เนื่องจากความจริงที่ว่า พฤติกรรมของการเก็บลักษณะเฉพาะที่ได้จากเซกเมนต์ของหน่วยเสียงต่างๆนั้นสามารถเก็บสมบัติสำคัญต่างๆของหน่วยเสียงเพื่อใช้แทนสัญญาณเสียงใดๆนั้นสามารถทำได้โดยง่าย

ดังนั้นในงานวิจัยนี้ทำการนำเสนอลักษณะเฉพาะของหน่วยเสียงในภาษาไทยที่สนับสนุนหลักการทำงานแบบอาศัยเซกเมนต์ เพื่อนำไปใช้ในการจำแนกหน่วยเสียงในภาษาไทย

วัตถุประสงค์ของการวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อนำเสนอลักษณะเฉพาะของหน่วยเสียงแบบอาศัยเซกเมนต์เพื่อนำไปใช้ในระบบรู้จำเสียงพูดแบบอาศัยเซกเมนต์ โดยจะนำเสนอเกี่ยวกับวิธีการเก็บ

ลักษณะเฉพาะและขั้นตอนการนำลักษณะเฉพาะไปใช้ในการจำแนกเสียงพูด โดยอาศัยวิธีการจำแนกหน่วยเสียงด้วยการวิเคราะห์ดิสคริมิแนนต์เชิงเส้น เพื่อให้สามารถจำแนกหน่วยเสียงได้ถูกต้องมากกว่าการจำแนกหน่วยเสียงแบบอาศัยกรอบเวลา

ขอบเขตของการวิจัย

1. งานวิจัยนี้นำเสนอการหาลักษณะเฉพาะโดยเลือกใช้ค่าสัมประสิทธิ์เมลฟรีเควันซีเคปตรอลมาเป็นสมบัติหลักในการจับสมบัติของสัญญาณเสียง เพื่อนำไปใช้แทนหน่วยเสียงในภาษาไทยทั้งสิ้น 52 หน่วยเสียงตามที่ระบุไว้ในบทที่ 3หน่วยเสียงภาษาไทย
2. งานวิจัยนี้จะทำการทดสอบประสิทธิภาพความถูกต้องของลักษณะเฉพาะโดยใช้กระบวนการจำแนกหน่วยเสียงด้วยวิธีการการวิเคราะห์ดิสคริมิแนนต์เชิงเส้น (Linear Discriminant Analysis : LDA) เพื่อทำการจำแนกหน่วยเสียงที่ได้จากการแทนด้วยลักษณะเฉพาะที่ใช้หลักการแบบอาศัยเซกเมนต์ตามที่ได้นำเสนอในงานวิจัยนี้ ซึ่งงานวิจัยนี้จะไม่ได้มีการมุ่งเน้นการเปรียบเทียบผลการจำแนกหน่วยเสียงระหว่างวิธีการต่างๆ
3. งานวิจัยนี้ทำการวัดประสิทธิภาพความถูกต้องโดยเปรียบเทียบกับผลการจำแนกหน่วยเสียงภาษาไทยทั้งสิ้น 52 หน่วยเสียงตามหลักการแบบอาศัยเซกเมนต์เปรียบเทียบกับผลการจำแนกหน่วยเสียงที่ใช้ลักษณะเฉพาะแบบแบ่งตามกรอบเวลาโดยใช้หลักการทำงานที่ได้ระบุไว้ในงานวิจัยของเอกฤทธิ์ มณีน้อย[2]
4. งานวิจัยนี้ทำการทดลองโดยเลือกใช้ข้อมูลเสียงจากฐานข้อมูลเสียงโลตัส ชุดหน่วยเสียงสมดุล (Phonetically Balanced Set) ที่บันทึกด้วยไมโครโฟนแบบโคลสทอร์ค (Close-talk)คุณภาพสูงในสภาพแวดล้อมแบบห้องเงียบ โดยข้อมูลเสียงชุดนี้เป็นข้อมูลเสียงแบบต่อเนื่องที่ได้กำกับหน่วยเสียงแบ่งเป็นหน่วยเสียงย่อยๆแล้ว
5. งานวิจัยนี้ไม่คำนึงถึงโทนเสียงในภาษาไทย
6. งานวิจัยนี้ไม่ครอบคลุมถึงการสร้างหรือประกอบงานรวมกับการแบ่งเสียงเป็นหน่วยย่อยๆเป็นโปรแกรมสำเร็จรูปขึ้นเพื่อใช้งาน

ขั้นตอนการวิจัย

1. ศึกษาทฤษฎี สมบัติและข้อแตกต่างของหน่วยเสียงต่างๆในภาษาไทย
2. ศึกษาวิธีการแทนหน่วยเสียงด้วยลักษณะเฉพาะแบบสัมประสิทธิ์เมลฟรีเควันซีเคปตรอลและสมบัติทางเซกเมนต์อื่นๆของหน่วยเสียง เพื่อใช้ในการแทนหน่วยเสียงต่างๆในภาษาไทย

3. ทำการทดลองจำแนกหน่วยเสียงในภาษาไทย เพื่อทดสอบประสิทธิภาพความถูกต้องในการใช้ลักษณะเฉพาะดังกล่าวกับจำแนกหน่วยเสียง
4. ทำการปรับเปลี่ยนและพัฒนาลักษณะเฉพาะให้เหมาะสมกับการจำแนกหน่วยเสียงในภาษาไทย
5. วิเคราะห์ผลการทดลองที่ได้
6. ประเมินผลและวัดประสิทธิภาพการจำแนกหน่วยเสียงจากการใช้ลักษณะเฉพาะที่ได้เทียบกับผลการทดลอง
7. สรุปเป็นระเบียบวิธีในการแบ่งแยกประเภทหน่วยเสียงในภาษาไทยและลักษณะเฉพาะของหน่วยเสียงในภาษาไทยที่เหมาะสม
8. เรียบเรียงและจัดทำรายงานวิทยานิพนธ์

ประโยชน์ที่คาดว่าจะได้รับ

งานวิจัยนี้นำเสนอวิธีการหาลักษณะเฉพาะเพื่อใช้แทนหน่วยเสียงตามพฤติกรรมการเปล่งเสียงสำหรับแบบจำลองทางเสียงแบบอาศัยเซกเมนต์ เพื่อเพิ่มประสิทธิภาพความถูกต้องในการจำแนกหน่วยเสียงซึ่งถือเป็นส่วนหนึ่งสำหรับการสร้างระบบรู้จำเสียงภาษาไทย

ผลงานที่ตีพิมพ์จากวิทยานิพนธ์

หลักการส่วนหนึ่งของวิทยานิพนธ์นี้ได้รับการตอบรับให้ตีพิมพ์เป็นบทความทางวิชาการในหัวข้อเรื่อง “Phoneme Classification Study for Thai Segment-Based Acoustic Models” โดย หนึ่งฤทัย เอกชัยวรสิน โปรดปราน บุณยพุกกณะ และ อติวงศ์ สุชาโต ในงานประชุมวิชาการ “IEEE International Symposium on Communications and Information Technologies 2006 (ISCIT 2006)” ณ. กรุงเทพมหานคร ในระหว่างวันที่ 18-20 ตุลาคม 2549

โครงร่างของวิทยานิพนธ์

รายละเอียดในวิทยานิพนธ์ฉบับนี้จะประกอบด้วยกันทั้งสิ้น 6 บท ซึ่งจะกล่าวถึงเนื้อหาต่างๆที่เกี่ยวข้องกับงานวิจัยเรียงตามลำดับ ดังนี้

บทที่ 1 **บทนำ** บทนี้จะกล่าวเริ่มต้นด้วยความเป็นมาและความสำคัญของปัญหาที่เป็นสาเหตุให้มีการทำงานวิจัยชิ้นนี้พร้อมทั้งระบุวัตถุประสงค์ของการวิจัย ขอบเขตของการวิจัย แสดงขั้นตอนการวิจัย รวมถึงการอธิบายถึงประโยชน์ที่คาดว่าจะได้รับ นำเสนอรายละเอียดของผลงานที่ได้รับการตีพิมพ์ และสุดท้ายคือ ส่วนสรุปโดยย่อที่จะกล่าวถึงรายละเอียดของเนื้อหาโดยรวมทั้งหมดเพื่อให้เห็นลำดับเนื้อหาในส่วนต่างๆของวิทยานิพนธ์ฉบับนี้

บทที่ 2 **ทฤษฎีและงานวิจัยที่เกี่ยวข้อง** บทนี้จะแบ่งเป็น 2 ส่วนหลักๆ โดยส่วนแรกจะกล่าวถึง ทฤษฎีพื้นฐานและหลักการต่างๆที่มีความสำคัญต่อการเข้าใจถึงเนื้อหาการทำงานภายในงานวิจัยนี้ และอีกส่วนจะกล่าวถึงงานวิจัยต่างๆที่เกี่ยวข้องกับงานวิจัยนี้

บทที่ 3 **หน่วยเสียงภาษาไทย** บทนี้จะกล่าวถึงลักษณะพฤติกรรมของหน่วยเสียงแต่ละชนิด พร้อมทั้งกล่าวถึงหน่วยเสียงต่างๆที่นำมาใช้ในงานวิจัยนี้

บทที่ 4 **การหาลักษณะเฉพาะ** กล่าวถึงเนื้อหาสำคัญของงานวิจัย คือวิธีการหาลักษณะเฉพาะ โดยจะกล่าวถึงสาเหตุที่มาในการหาลักษณะเฉพาะดังกล่าวด้วย

บทที่ 5 **การทดลองจำแนกหน่วยเสียง** บทนี้จะนำลักษณะเฉพาะที่หาได้จากบทที่แล้วมาทดลองจำแนกหน่วยเสียง เพื่อทดสอบประสิทธิภาพโดยทำการจำแนกหน่วยเสียงภาษาไทย พร้อมทั้งวิเคราะห์ค่าความมีส่วนร่วมของลักษณะเฉพาะที่ใช้ และได้มีการจำแนกหน่วยเสียงด้วยฮิดเดนมาร์คอฟ ซึ่งมีหลักการทำงานแบบอาศัยกรอบเวลา เพื่อเปรียบเทียบประสิทธิภาพระหว่างผลที่ได้จากการทำงานแบบเซกเมนต์กับการทำงานแบบอาศัยกรอบเวลา

บทที่ 6 **สรุปผลการวิจัย** บทสุดท้ายนี้จะสรุปขั้นตอนการทำงานและผลการทำงานที่ได้ทั้งหมด และกล่าวถึงคุณประโยชน์ของงานวิจัยต่อวงวิชาการ พร้อมทั้งข้อเสนอแนะเพิ่มเติมที่เกี่ยวข้องกับการพัฒนางานวิจัยนี้ในภายภาคหน้า

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

บทนี้จะกล่าวถึงทฤษฎีพื้นฐานรวมทั้งหลักการต่างๆที่จำเป็นในการทำความเข้าใจต่อเนื้อหาที่จะนำเสนอในบทถัด ๆ ไป เช่น หลักการรู้จำเสียง การแทนสัญญาณเสียงด้วยสมบัติในโดเมนความถี่ การวิเคราะห์ดิสคริมิแนนต์เชิงเส้นและการทดสอบนัยสำคัญ (Significance testing) นอกจากนี้ ในบทนี้ยังกล่าวถึงงานวิจัยต่างๆที่มีส่วนเกี่ยวข้องกับงานวิจัยนี้ด้วย

ทฤษฎีที่เกี่ยวข้อง

1. **หลักการรู้จำเสียง**

ในกระบวนการรู้จำเสียงซึ่งไม่ว่าจะเป็นเสียงในภาษาใดๆนั้น จะมีหลักการการทำงานโดยรวมที่เหมือนกัน ซึ่งประกอบด้วยภาพรวม ดังนี้

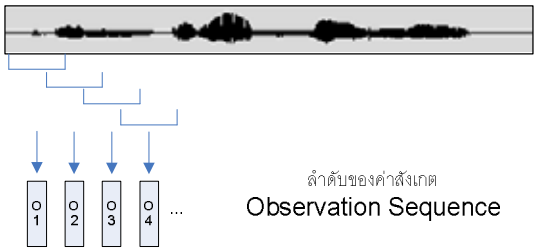
รูปที่ 2.1 ขั้นตอนการทำงานของหลักการรู้จำเสียง

ซึ่งในแต่ละขั้นตอน สามารถอธิบายความหมายและความสำคัญ ได้ดังนี้

1.2 **การแทนสัญญาณเสียง**

ในขั้นตอนนี้เป็นการแปลงสัญญาณเสียงที่เข้ามาให้อยู่ในลักษณะที่พร้อมใช้ในการรู้จำเสียง เพราะในการรู้จำเสียงนั้นไม่สามารถเอาสัญญาณเสียงเข้ามาใช้งานได้โดยตรงแต่จะเป็นการดึงเอาลักษณะเฉพาะของสัญญาณเสียงออกมา ซึ่งสามารถกระทำได้หลายวิธี เช่น การแทนสัญญาณเสียงด้วยสมบัติในโดเมนความถี่ ซึ่งจะกล่าวในหัวข้อต่อไป เป็นต้น โดยขั้นตอนในการเตรียมสัญญาณเสียงก่อนเข้าสู่กระบวนการแทนสัญญาณเสียงดังกล่าวนั้น มีหลักการทำงานอยู่ 2 ประเภท คือ

1.2.1 **การทำงานแบบอาศัยกรอบเวลา** กล่าวคือ เป็นการรับสัญญาณเสียงต่อเนื่องเข้ามา แล้วทำการแบ่งสัญญาณดังกล่าวออกเป็นเฟรมๆขนาดเท่าๆกันก่อนนำสัญญาณเสียงในแต่ละเฟรมไปหาลักษณะเฉพาะต่อไป ซึ่งลำดับของค่าลักษณะเฉพาะที่ได้จากเฟรมแต่ละเฟรมที่ต่อเนื่องกันนั้นจะเรียกว่า ลำดับของค่าสังเกต (Observation Sequence) ดังตัวอย่างในรูป

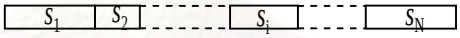


รูปที่ 2.2 ลักษณะการตัดแบ่งสัญญาณเสียงตามการทำงานแบบอาศัยกรอบเวลา

ดังนั้นเมื่อมีสัญญาณเสียงเข้ามา จะได้ลำดับของค่าสังเกตคือ $O = o_1, o_2, \ldots, o_T$ โดยที่ o_t คือ เวกเตอร์ลักษณะเฉพาะที่เวลา t ไปใช้ในกระบวนการรู้จำต่อไป

นอกจากนี้ยังสังเกตได้ว่าการทำงานแบบอาศัยกรอบเวลานี้ จะตัดแบ่งสัญญาณเสียงออกเป็นเฟรม โดยไม่มีการคำนึงถึงลักษณะโครงสร้างของสัญญาณเสียงในการแบ่งสัญญาณเสียงออกเป็นเฟรมๆแต่อย่างใด

1.2.2 **การทำงานแบบอาศัยเซกเมนต์** เป็นการทำงานที่มีการแบ่งสัญญาณเสียงโดยการคำนึงถึงโครงสร้างของสัญญาณเสียง ก่อนนำไปหาลักษณะเฉพาะ โดยจะพยายามตัดแบ่งสัญญาณเสียงดังกล่าวก็เป็นส่วนย่อยๆที่เรียกว่า หน่วยเสียง (Phoneme) ด้วยกระบวนการตัดแบ่งหน่วยเสียง (Phoneme segmentation) โดยลักษณะการแบ่งสัญญาณเสียงออกเป็นหน่วยเสียงย่อยๆ สามารถอธิบายได้ดังรูปที่ 2.3



รูปที่ 2.3 ลำดับของสัญญาณเสียงที่ทำการแบ่งเสียงเป็นหน่วยย่อยๆ

ซึ่งจะได้หน่วยเสียงลำดับของหน่วยเสียง $S = \{s_1, s_2, \ldots, s_N\}$ แล้วจึงนำแต่ละ s_i ไปหาลักษณะเฉพาะต่อไป

1.3 การรู้จำเสียง

หลังจากที่ได้ลักษณะเฉพาะซึ่งเป็นการแทนสัญญาณเสียงด้วยค่าสำคัญออกมาแล้วนั้น ก็จะเข้าสู่กระบวนการการรู้จำเสียง ซึ่งสามารถอธิบายได้ ดังนี้

เมื่อสัญญาณเสียงที่เป็นคำหรือประโยคใดๆเข้ามา จะสามารถคำนวณหาคำตอบว่าเป็นคำใดๆได้จาก

$$\arg\max_i \{P(w_i \mid O)\} \qquad (2.1)$$

โดยที่ w_i คือ คำหรือประโยคที่เป็นไปได้ลำดับที่ i ที่เป็นไปได้ ซึ่งค่าความน่าจะเป็นดังกล่าวสามารถคำนวณได้จากกฎของเบส์(Bayes' Rule) ที่ว่า

$$P(w_i \mid O) = \frac{P(O \mid w_i)P(w_i)}{P(O)} \qquad (2.2)$$

ซึ่งในการหาค่าความน่าจะเป็นดังกล่าว ประกอบด้วยเทคนิค 2 ส่วนที่เข้ามาช่วยในการรู้จำเสียง ได้แก่ แบบจำลองเสียงและแบบจำลองภาษา ซึ่งแต่ละส่วนมีหน้าที่ในการรู้จำเสียง ดังนี้

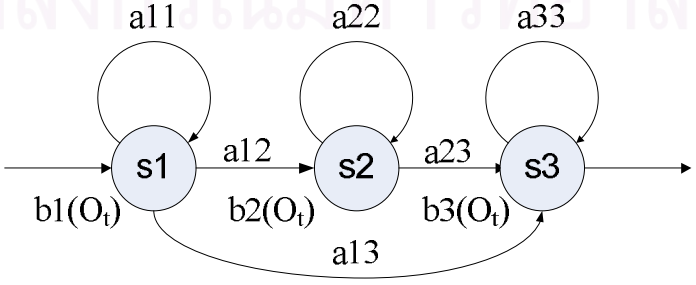
1.3.1 **แบบจำลองเสียง** เป็นแบบจำลองที่อธิบายถึงลักษณะของหน่วยเสียงแต่ละหน่วย โดยแบบจำลองเสียง 1 แบบ จะอธิบายลักษณะ พฤติกรรม ความเป็นหน่วยเสียงได้ 1 หน่วยเสียง ดังนั้นจะมีจำนวนแบบจำลองเสียงเท่ากับจำนวนหน่วยเสียงทั้งหมด

ซึ่งเมื่อมีการป้อนลักษณะเฉพาะของสัญญาณเสียงใดๆเข้ามาแล้ว แบบจำลองเสียงแต่ละแบบจำลองจะคืนค่า $P(O \mid u_i)$ กล่าวคือ เป็นค่าไลค์ลีฮูด (Likelihood) ของ u_i ที่สัมพันธ์กับ O เมื่อ O คือ ค่าสังเกต (Observation) หรือลักษณะเฉพาะ และ u_i คือ หน่วยเสียงใดๆ ซึ่งในที่นี้พฤติกรรม ลักษณะของหน่วยเสียงนั้นๆถูกแทนด้วยฟังก์ชันการกระจายตัวของความน่าจะเป็น (Probability density function) ที่เกิดจากการฝึกฝนข้อมูลจากหน่วยเสียงนั้นๆ ซึ่งในที่นี้เรียกว่าแบบจำลองเสียง ซึ่งค่า $P(O \mid u_i)$ นี้จะนำไปใช้หาค่า $P(O \mid w_i)$ ตามหลักการการรู้จำเสียง กล่าวคือ $P(O \mid w_i)$ หาได้จากการนำ $P(O \mid u_i)$ ที่สามารถต่อเข้ากันได้ในระดับคำมาคูณกันทุกตัว

ซึ่งรายละเอียดการหาค่า $P(O \mid u_i)$ ในส่วนนี้จะขึ้นกับวิธีการในการเตรียมสัญญาณเสียงก่อนเข้าสู่กระบวนการแทนสัญญาณเสียง โดยแต่ละแบบจำลองเสียงของแต่ละประเภท มีลักษณะดังนี้

1. **แบบจำลองเสียงแบบอาศัยกรอบเวลา** (Frame-based Acoustic Models)

สำหรับการหาค่า $P(a \mid u_i)$ ของแบบจำลองเสียงแบบอาศัยกรอบเวลานี้จะขออ้างอิงจากหลักการทำงานของแบบจำลองฮิดเดนมาร์คอฟ[4] ซึ่งเป็นเทคนิคในการรู้จำเสียงที่มีการทำงานแบบอาศัยกรอบเวลาที่เป็นที่นิยมมาก โดยแบบจำลองฮิดเดนมาร์คอฟจะอธิบายแบบจำลองเสียงต่างๆ ด้วยเครื่องจักรสถานะจำกัด (Finite state machine) ซึ่งในที่นี้จะขออธิบายโดยใช้ 3 สถานะ โดยมีการทำงานจากซ้ายไปขวา (Left-to-Right HMM)



รูปที่ 2.4 แบบจำลองเสียงตามหลักการทำงานฮิดเดนมาร์คอฟ

เมื่อทำการหาลักษณะเฉพาะของสัญญาณเสียงที่ถูกแบ่งเป็นเฟรมๆแล้ว ในขณะนี้จะได้ลำดับของค่าสังเกต (Observation sequence) ซึ่งเป็นลำดับของลักษณะเฉพาะในทุกๆหน่วยเวลาใดๆ นั่นเอง โดยลำดับของค่าสังเกตจะถูกป้อนเข้าทางด้านซ้ายของเครื่องสถานะจำกัด ซึ่งจะมีการเปลี่ยนสถานะในทุกๆหน่วยเวลา ซึ่งในขณะที่เวลา t ใดๆที่มีการเปลี่ยนเข้าสู่สถานะ j เวกเตอร์ของค่าสังเกตในช่วงเวลา t นั้นหรือ $\boldsymbol{O}_t$ จะได้ค่าความน่าจะเป็น $\boldsymbol{b}_j(\boldsymbol{O}_t)$ รวมไปถึงจะมีค่าความน่าจะเป็นจากการเปลี่ยนจากสถานะ i ไปเป็นสถานะ j ด้วยค่า a_{ij} ด้วยเช่นกัน

ดังนั้น ค่าความน่าจะเป็นร่วมที่ $\boldsymbol{O}$ จะเป็นหน่วยเสียงตามแบบจำลอง โดยผ่านสถานะต่างๆเป็นลำดับๆหรือลำดับของสถานะใดๆ $\boldsymbol{X}$ (State sequence) สามารถหาได้จาก

$$\boldsymbol{P}(\boldsymbol{O},\boldsymbol{X}\mid \mathrm{u}) = \boldsymbol{a}_{12}\boldsymbol{b}_2(\boldsymbol{O}_1)\boldsymbol{a}_{22}\boldsymbol{b}_2(\boldsymbol{O}_2)\boldsymbol{a}_{23}\boldsymbol{b}_3(\boldsymbol{O}_3)\ldots \tag{2.3}$$

แต่เนื่องจากในความเป็นจริงที่ว่า เราไม่สามารถทราบลำดับของสถานะ $\boldsymbol{X}$ ดังกล่าวได้ ดังนั้นจึงต้องทำการคำนวณทุกลำดับ $\boldsymbol{X}$ ที่เป็นไปได้ จะได้ว่า

$$\boldsymbol{P}(\boldsymbol{O}\mid \mathrm{u}) = \sum_{\boldsymbol{X}} \boldsymbol{a}_{\boldsymbol{x}(0)\boldsymbol{x}(1)} \prod_{\boldsymbol{t}=1}^{\boldsymbol{T}} \boldsymbol{b}_{\boldsymbol{x}(\boldsymbol{t})}(\boldsymbol{O}_{\boldsymbol{t}})\boldsymbol{a}_{\boldsymbol{x}(\boldsymbol{t})\boldsymbol{x}(\boldsymbol{t}+1)} \tag{2.4}$$

โดยที่ $\boldsymbol{x}(0)$ คือสถานะตั้งต้นก่อนเข้าแบบจำลอง และ $\boldsymbol{x}(\boldsymbol{T}+1)$ เป็นสถานะออกจากแบบจำลอง

ทำการหาค่าความน่าจะเป็น $\boldsymbol{P}(\boldsymbol{O}\mid \boldsymbol{u}_i)$ ของแต่ละแบบจำลองเสียง เพื่อนำไปใช้ในการหาค่าความน่าจะเป็นในระดับคำต่อไป

แต่หากเป็นกรณีที่ต้องการจำแนกหน่วยเสียง ซึ่งถือว่าเป็นการรู้จำโดยใช้เฉพาะแบบจำลองเสียงเพียงอย่างเดียว จะสามารถจำแนกหน่วยเสียงต่างๆได้ โดยคำนวณจาก

$$\arg\max_i\{\boldsymbol{P}(\boldsymbol{u}_i\mid \boldsymbol{O})\} \tag{2.5}$$

ซึ่ง $\boldsymbol{P}(\boldsymbol{u}_i\mid \boldsymbol{O})$ สามารถคำนวณได้จากกฎของเบย์(Bayes'Rule) ที่ว่า

$$\boldsymbol{P}(\boldsymbol{u}_i\mid \boldsymbol{O}) = \frac{\boldsymbol{P}(\boldsymbol{O}\mid \boldsymbol{u}_i)\boldsymbol{P}(\boldsymbol{u}_i)}{\boldsymbol{P}(\boldsymbol{O})} \tag{2.6}$$

หากกำหนดให้หน่วยเสียงทุกเสียงมีความน่าจะเป็นในการเกิดเท่าๆกัน จะได้ว่า

$$\arg\max_i\{\boldsymbol{P}(\boldsymbol{u}_i\mid \boldsymbol{O})\} = \arg\max_i\{\boldsymbol{P}(\boldsymbol{O}\mid \boldsymbol{u}_i)\} \tag{2.7}$$

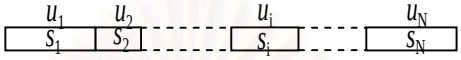
2. **แบบจำลองเสียงแบบอาศัยเซกเมนต์** (Segment-based Acoustic Models)

สำหรับการรู้จำเสียงแบบอาศัยเซกเมนต์นั้น ลักษณะแบบจำลองเสียงจะขึ้นกับวิธีที่ใช้ในการจำแนกหน่วยเสียงใดๆ โดยวิธีการจำแนกหน่วยเสียงดังกล่าวจะต้องสามารถคำนวณค่าความน่าจะเป็น $\boldsymbol{P}(\boldsymbol{a}\mid \boldsymbol{u}_i)$ เพื่อนำไปใช้ในขั้นตอนการรู้จำเสียงในระดับต่อๆไปได้ โดยที่ $\boldsymbol{a}$ คือลักษณะเฉพาะของสัญญาณเสียงที่หาได้จากการตัดแบ่งสัญญาณเสียงออกเป็นหน่วยเสียงใดๆ ซึ่งเปรียบได้กับ $\boldsymbol{O}$ ในการทำงานแบบอาศัยกรอบเวลา ซึ่งในที่นี้จะขอยกตัวอย่างวิธีการจำแนกหน่วยเสียงที่ใช้ในงานวิจัย

นี้ คือ วิธีการการวิเคราะห์ดิสคริมิแนนต์เชิงเส้น ซึ่งได้อธิบายรายละเอียดไว้ในหัวข้อที่ 3 เรื่องการวิเคราะห์ดิสคริมิแนนต์เชิงเส้น

สำหรับการรู้จำเสียงด้วยแบบจำลองเสียงแบบอาศัยเซกเมนต์เพียงอย่างเดียว สามารถอธิบายได้ ดังนี้

เมื่อสัญญาณเสียงผ่านกระบวนการตัดแบ่งเป็นเสียงเป็นหน่วยย่อยๆ [3] แล้ว จะได้ลำดับของหน่วยเสียงของสัญญาณเสียงนั้นๆ โดยแต่ละส่วนนั้นจะเป็นหน่วยเสียง 1 หน่วย ดังรูปที่ 2.5



รูปที่ 2.5 ลำดับของสัญญาณเสียงที่ทำการแบ่งเสียงเป็นหน่วยย่อยๆ

สมมติให้สัญญาณเสียงที่ถูกแบ่งแล้วเป็น S= $\{s_1,s_2,\ldots,s_N\}$, ดังรูปที่ 2.5 ให้ u_i เป็นหน่วยเสียงที่สัมพันธ์กับ s_i ดังนั้นจึงกล่าวได้ว่า

$$\boldsymbol{P(U \mid A) = P(u_1 u_2 \ldots u_N \mid a_1 a_2 \ldots a_N)} \tag{2.8}$$

เมื่อ $\boldsymbol{a_i}$ เป็นสัญญาณเสียงในหน่วยเสียงลำดับที่ i สมมติให้ในแต่ละ u_i มีความอิสระต่อกันและ u_i ไม่ขึ้นกับ $\boldsymbol{a_j}$ เมื่อ $i \neq j$ ดังนั้นจะได้ว่า

$$\boldsymbol{P(U \mid A) = \prod_{i=1}^{N} P(u_i \mid a)} \tag{2.9}$$

ดังนั้นแบบจำลองเสียงในโครงสร้างแบบอาศัยเซกเมนต์(Segment-based framework) นี้ต้องคำนวณหาความน่าจะเป็นที่หน่วยเสียงในลำดับที่ i จะเป็น $\boldsymbol{u_i}$ เมื่อรู้ $\boldsymbol{a_i}$ ซึ่ง สามารถหาค่า $\boldsymbol{P(u_i \mid a)}$ สำหรับ $\boldsymbol{a_i}$ ใดๆได้จาก

$$\arg\max_i\{\boldsymbol{P(u_i \mid a)}\} \tag{2.10}$$

ซึ่ง $\boldsymbol{P(u_i \mid a)}$ สามารถคำนวณได้จากกฎของเบย์(Bayes'Rule) ที่ว่า

$$\boldsymbol{P(u_i \mid a) = \frac{P(a \mid u_i)P(u_i)}{P(a)}} \tag{2.11}$$

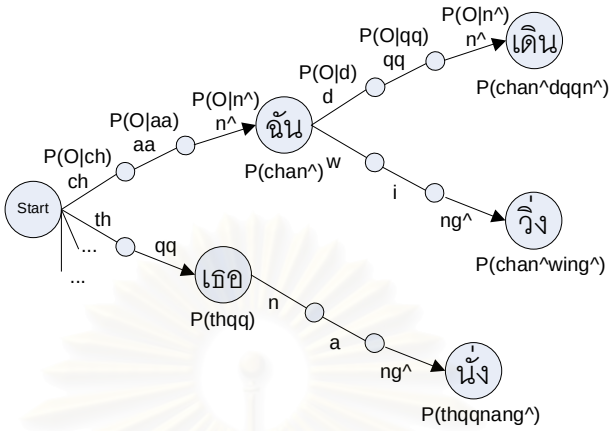
ดังนั้นในกรณีที่การรู้จำเสียงใช้เพียงแบบจำลองเสียงแบบอาศัยเซกเมนต์เพียงแบบจำลองเดียว และมีการกำหนดให้หน่วยเสียงทุกเสียงมีความน่าจะเป็นในการเกิดเท่าๆกัน จะได้ว่า

$$\arg\max_i\{\boldsymbol{P(u_i \mid a)}\} = \arg\max_i\{\boldsymbol{P(a \mid u_i)}\} \tag{2.12}$$

1.4 แบบจำลองภาษา

สำหรับแบบจำลองภาษานี้ เป็นแบบจำลองที่อาศัยโครงสร้างของภาษาและคำเข้ามาช่วยในการรู้จำด้วย กล่าวคือ เป็นแบบจำลองที่ประกอบจากคำต่างๆขึ้นเป็นประโยค ดังนั้นความเป็นไปได้ที่จะเกิดคำนั้นๆในแต่ละตำแหน่งบนประโยคจึงมีค่าไม่เท่ากัน ส่งผลให้หลักการรู้จำเสียงจึงมีค่า $\boldsymbol{P(w_i)}$ เข้ามาเกี่ยวข้องด้วย ซึ่งค่า $\boldsymbol{P(w_i)}$ ดังกล่าวนี้จะได้จากการกำหนดค่าความเป็นไปได้ของคำ

แต่ละคำตามตำแหน่งที่เป็นไปได้ของคำบนประโยคที่กำหนดไว้ในโครงสร้างของแบบจำลองภาษา โดยตัวอย่างของแบบจำลองภาษามีลักษณะ ดังนี้



รูปที่ 2.6 แบบจำลองภาษาสำหรับการรู้จำเสียง

จากรูปจะมี $P(w_i)$ ได้แก่ $P(chan\hat{})$, $P(chan\hat{}\,dqqn\hat{})$, $P(chan\hat{}\,wing\hat{})$, $P(thqq)$, $P(thqqnang\hat{})$ เป็นต้น โดยมีรูปประโยคที่เป็นไปได้คือ ฉันเดิน, ฉันวิ่ง, เธอนั่ง เป็นต้น ซึ่งหากการรู้จำเสียงอยู่ในสถานะที่รู้จำคำว่า “ฉัน” แล้ว ตำแหน่งคำต่อไปของประโยค จะมีคำที่เป็นไปได้อยู่ 2 คำเท่านั้นคือ คำว่า “เดิน” และ “วิ่ง” ซึ่งแบบจำลองภาษาถือได้ว่าเป็นส่วนช่วยเพิ่มประสิทธิภาพให้มีความแม่นยำในการรู้จำเพิ่มมากขึ้น

กล่าวโดยสรุปแล้วเราสามารถหาค่า $P(O \mid w_i)$ ได้จากแบบจำลองเสียง และ $P(w_i)$ จากแบบจำลองภาษา ซึ่งก็จะสามารถทำให้หาค่า $P(w_i \mid O)$ สำหรับการรู้จำเสียงของทั้งระบบได้

2. **การแทนสัญญาณเสียงด้วยสมบัติในโดเมนความถี่**

เนื่องจากตำแหน่งและพฤติกรรมของอวัยวะที่ใช้ในการเปล่งเสียงเป็นตัวที่กำหนดลักษณะของสัญญาณเสียงต่างๆจนประกอบเป็นหน่วยเสียงได้นั้น เป็นพฤติกรรมที่ทำให้เกิดสัญญาณที่อยู่ภายใต้โดเมนความถี่ ซึ่งแต่ละสัญญาณเสียงที่เกิดขึ้นจากการเปล่งเสียงนั้นมีความถี่สั่นพ้องที่แตกต่างกัน ส่งผลให้ความสำคัญในช่วงความถี่ต่างๆมีความแตกต่างกันในหน่วยเสียงต่างชนิดกัน

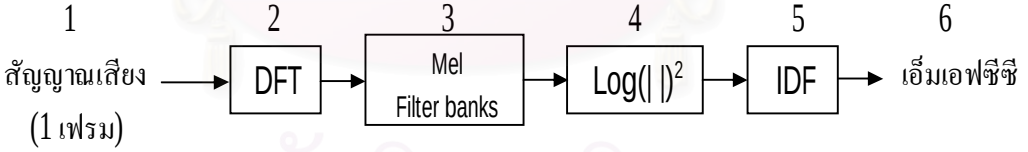
ดังนั้นในด้านการรู้จำสัญญาณเสียง จึงถือว่าการวิเคราะห์สัญญาณเสียงทางโดเมนความถี่นั้นเป็นวิธีการหนึ่งที่มีความสามารถในการใช้สมบัติดังกล่าวแทนสัญญาณเสียงเพื่อนำมาใช้ในการรู้จำหน่วยเสียง จำแนกหน่วยเสียงต่างๆได้ โดยการวิเคราะห์สัญญาณเสียงทางโดเมนความถี่ดังกล่าว ถือเป็นการวิเคราะห์เชิงสเปกตรัม (Spectral analysis) โดยสามารถคำนวณหาได้จากผลการแปลงฟูริเยร์ (Fourier transform) ซึ่งเป็นหลักการการแยกองค์ประกอบทางความถี่ แต่เนื่องจากหลักการดังกล่าวเป็นหลักการที่ทำงานบนสเปกตรัมต่อเนื่องที่มีการกำหนดค่าอินทริกรัล (integral) บนช่วงระยะอนันต์ ซึ่งการประมวลผลบนเครื่องคอมพิวเตอร์ไม่สามารถแก้ปัญหาดังกล่าวได้ ดังนั้นจึงได้มีการนำหลักการแปลงฟูริเยร์แบบไม่ต่อเนื่อง (Discrete Fourier transform: DFT) ซึ่ง

เป็นหลักการที่ทำการประมาณค่าจากผลการแปลงฟูริเยร์ให้สามารถดำเนินการด้วยกระบวนการแบบไม่ต่อเนื่องแทน

2.1 **สัมประสิทธิ์เมลฟรีเคว็นซีเคพสตรอล** (Mel Frequency Cepstral Coefficients: MFCCs) [5-6]

หนึ่งในหลักการที่ดึงลักษณะเฉพาะทางสเปกตรัมของสัญญาณเสียงออกมา เพื่อนำไปใช้แทนสัญญาณเสียงต่างๆด้วยหลักการวิเคราะห์ทางโดเมนความถี่ ซึ่งได้ใช้วิธีการแปลงฟูริเยร์เป็นส่วนหนึ่งในกระบวนการนั้น คือ การหาค่าสัมประสิทธิ์เมลฟรีเคว็นซีเคพตรอลหรือเอ็มเอฟซีซี ซึ่งเป็นการคำนวณค่าสัมประสิทธิ์ของสัญญาณเสียงใดๆ ที่สามารถคำนวณได้จากการขั้นตอนต่างๆ ดังนี้

1. ตัดแบ่งสัญญาณเสียงที่เข้ามาออกเป็นเฟรม
2. คำนวณหาค่าผลการแปลงฟูเรียร์แบบไม่ต่อเนื่อง ซึ่งเป็นการแปลงสัญญาณให้อยู่บนโดเมนความถี่
3. นำผลที่ได้ข้างต้นไปผ่านกระบวนการฟิลเตอร์แบงค์ (Filter bank processing) ซึ่งเป็นอาเรย์ของแบนพาสฟิลเตอร์(band-pass filters) ที่แบ่งสัญญาณขาเข้าเป็นหลายๆส่วน โดยแต่ละส่วนจะมี 1 ช่วงสัญญาณความถี่ของสัญญาณดั้งเดิม
4. หาค่าลอการิทึมของค่าสัมบูรณ์ยกกำลังสองของค่าที่ได้จากกระบวนการฟิลเตอร์แบงค์
5. นำไปหาค่าย้อนกลับของผลการแปลงฟูเรียร์แบบไม่ต่อเนื่อง (Inverse DFT)
6. ได้ค่าสัมประสิทธิ์เมลฟรีเคว็นซีเคปสตรอลซึ่งเป็นลักษณะเฉพาะของสัญญาณเสียง โดยขั้นตอนต่างๆข้างต้น สามารถอธิบายได้ ดังรูปที่ 2.7



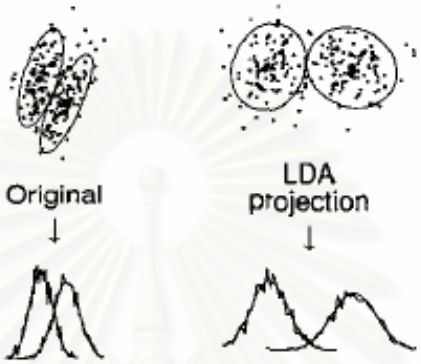
รูปที่ 2.7 ขั้นตอนการหาค่าเอ็มเอฟซีซีของสัญญาณเสียง

กล่าวโดยสรุปแล้วเอ็มเอฟซีซีจะทำการแปลงสัญญาณเสียงจากทางเวลาไปเป็นสัญญาณเสียงทางความถี่ เพื่อเก็บลักษณะต่างๆของสัญญาณเสียงในแต่ละช่วงสัญญาณความถี่ แล้วจึงแปลงกลับไปเป็นสัญญาณทางเวลาดังเดิม

3. **การวิเคราะห์ดิสคริมิแนนต์เชิงเส้น**

การวิเคราะห์ดิสคริมิแนนต์เชิงเส้นหรือแอลดีเอ [7] เป็นกระบวนการการแปลงข้อมูลดั้งเดิมให้ไปอยู่ในปริภูมิใหม่เพื่อให้ข้อมูลแต่ละประเภทมีการแบ่งแยกกัน (Separation) มากที่สุด ซึ่งการวิเคราะห์ดิสคริมิแนนต์เชิงเส้นจะใช้หลักการในการเพิ่มอัตราส่วนของความแปรปรวน

ระหว่างกลุ่มกับความแปรปรวนภายในกลุ่มให้มีค่ามากๆ เพื่อให้มั่นใจได้ว่าจะสามารถแบ่งแยกข้อมูลในกลุ่มต่างๆได้อย่างชัดเจนขึ้น ดังรูปที่ 2.8 ซึ่งจะเห็นว่าชุดข้อมูลเก่ามีความซ้อนทับกันระหว่างกลุ่ม 2 กลุ่มค่อนข้างมาก จนเมื่อได้นำข้อมูลดังกล่าวมาผ่านกระบวนการการวิเคราะห์ดิสคริมิแนนต์เชิงเส้น ซึ่งจะแปลงข้อมูลดังกล่าวไปอยู่ในปริภูมิใหม่นั้น ส่งผลให้ข้อมูลทั้ง 2 กลุ่มมีการกระจายตัวแยกออกจากกันมากขึ้น ทำให้สามารถแบ่งแยกชุดข้อมูลทั้ง 2 กลุ่มออกจากกันได้ง่ายขึ้น



รูปที่ 2.8 ตัวอย่างการจำแนกข้อมูล 2 ประเภทเปรียบเทียบระหว่างการจำแนกข้อมูลบนปริภูมิดั้งเดิมกับการจำแนกข้อมูลบนปริภูมิใหม่ที่ผ่านกระบวนการการวิเคราะห์ดิสคริมิแนนต์เชิงเส้น [8]

โดยวิธีการดังกล่าวใช้หลักการที่เรียกว่าฟังก์ชันดิสคริมิแนนต์บัญญัติ (Canonical discriminant functions) หรือ ตัวแปรบัญญัติ (Canonical variables) ($\boldsymbol{L_n}$) ซึ่งหลักการนี้จะทำการแปลงข้อมูลในปริภูมิดั้งเดิมไปอยู่ในรูปข้อมูลที่อยู่บนปริภูมิใหม่ โดยฟังก์ชันดิสคริมิแนนต์บัญญัติลำดับที่ n (F_n) สามารถเขียนได้ตามสมการ ดังนี้

$$\boldsymbol{L_n} = \lambda_n \boldsymbol{F_n} = \sum_{i=1}^{m} \boldsymbol{a_{ni}} (\boldsymbol{x_i} - \mu_i) \qquad (2.13)$$

โดยที่ n คือ มิติของข้อมูลบนปริภูมิใหม่โดย $n \leqslant m$เมื่อ m มิติของข้อมูลบนปริภูมิดั้งเดิม

$\boldsymbol{x_i}$ คือ ค่าข้อมูลบนปริภูมิดั้งเดิมมิติที่ i

$\boldsymbol{a_{ni}}$ คือ น้ำหนักความสำคัญของ $(x_i - \mu_i)$ ซึ่งเป็นค่าที่ได้จากการแก้ปัญหาการหาค่าเหมาะที่สุด (Optimization) หรือที่เรียกว่าค่าสัมประสิทธิ์ดิสคริมิแนนต์ (Discriminant coefficients) โดยเวกเตอร์ $\boldsymbol{a_n}$ สามารถคำนวณได้จากค่าที่มากสุดของ

$$\frac{\lambda_n \mathrm{H} \boldsymbol{a_n}}{\lambda_n \mathrm{E} \boldsymbol{a_n}} \qquad (2.14)$$

โดยที่

$$\mathrm{E} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (\mathrm{y}_{ij} - \mathrm{y}_{i.})(\mathrm{y}_{ij} - \mathrm{y}_{i.})' \qquad (2.15)$$

และ

$$H = \sum_{i=1}^{k} \boldsymbol{n}_i (y_{i.} - \bar{y}_{..})(y_{ij} - \bar{y}_{..})' \quad (2.16)$$

k คือ จำนวนกลุ่มที่ต้องการจำแนก

$\boldsymbol{n}_i$ คือ จำนวนข้อมูลในกลุ่มที่ i

y_{ij} คือ เวกเตอร์ $[\boldsymbol{y}_1, \boldsymbol{y}_2, ..., \boldsymbol{y}_N]$ ของข้อมูลชุดที่ j ในกลุ่มที่ i

$y_{i.}$ คือ เวกเตอร์เฉลี่ยของกลุ่มที่ i

$\bar{y}_{..}$ คือ เวกเตอร์เฉลี่ยของข้อมูลทั้งหมดใน k กลุ่ม

λ_n คือ น้ำหนักความสำคัญของแต่ละ F_n

ซึ่ง $\lambda_n H \boldsymbol{a}_n$ นั้นเป็นการคำนวณความแปรปรวนระหว่างกลุ่ม จึงจำเป็นต้องมีค่ามากๆ ส่วน $\lambda_n E \boldsymbol{a}_n$ เป็นการคำนวณความแปรปรวนภายในกลุ่ม จึงควรต้องมีค่าน้อยๆ ซึ่งสามารถแก้ปัญหาดังกล่าวได้ด้วยสมการไอเก็น (Eigen-equation)

$$| H - \lambda_n E | = 0 \quad (2.17)$$

ซึ่งจะให้ค่าลักษณะเฉพาะ (Eigenvalues) λ_n ที่สัมพันธ์กับเวกเตอร์ลักษณะเฉพาะ (Eigenvectors) $\boldsymbol{a}_n$ ที่ทำให้ค่าดังกล่าวเป็นค่ามากสุดเนื่องจากเวกเตอร์ลักษณะเฉพาะ $\boldsymbol{a}_n$ แต่ละตัวจะตั้งฉากซึ่งกันและกัน ดังนั้นค่าความมีส่วนร่วมของการจำแนกกลุ่มจะไม่ซ้ำซ้อนกัน ดังนั้น $\boldsymbol{L}_n$ จึงเป็นการแบ่งแยกค่าเฉลี่ยระหว่างกลุ่มที่ขึ้นกับข้อมูลที่ได้รับมากที่สุด

ค่าความมีส่วนร่วม (Contribution) ซึ่งแสดงถึงความสำคัญของข้อมูลในมิติที่ i ที่มีต่อ F_n ใดๆในการแบ่งแยกข้อมูลออกตามประเภท สามารถคำนวณได้จาก

$$contribution_{ni} = \frac{a_{ni}}{\sum_{i=1}^{m} a_{ni}} \quad (2.18)$$

ซึ่งข้อมูลที่แปลงไปอยู่บนปริภูมิใหม่หรือ F_n นี้ก็ยังคงมีการให้ค่าความสามารถในการแบ่งแยกกลุ่มข้อมูลบนปริภูมิใหม่ (P_j) ของแต่ละ F_n ในแต่ละหลักด้วยเช่นเดียวกัน ซึ่งสามารถคำนวณได้จาก

$$P_i = \frac{\lambda_i}{\sum_{i=1}^{n} \lambda_i} \quad (2.19)$$

โดยที่ λ_i คือ น้ำหนักความสำคัญของแต่ละ F_n

n คือ มิติของข้อมูลบนปริภูมิใหม่

ดังนั้นจึงสามารถกล่าวได้ว่าค่าความสำคัญในแต่ละมิติของข้อมูลบนปริภูมิดั้งเดิมเมื่อผ่านกระบวนการการวิเคราะห์ดิสคริมิแนนต์เชิงเส้นเพื่อแปลงไปอยู่ในรูปแบบข้อมูลบนปริภูมิใหม่แล้วจะได้ค่าความสำคัญในแต่ละมิติของข้อมูลบนปริภูมิปัจจุบันว่า

$$contribution_i = \sum_{j=1}^{n} P_j \times contribution_{ji} \quad (2.20)$$

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

และความน่าจะเป็นของความถูกต้องของคำตอบที่ได้จากการจำแนกข้อมูลได้ ดังนี้

$$P(\omega_i \mid X) = \frac{e^{\frac{d_i}{2}}}{\sum_{k=1}^{K} e^{\frac{d_k}{2}}} \qquad (2.21)$$

โดยที่ $P(\omega_i \mid X)$ คือ ความน่าจะเป็นที่ข้อมูล X เป็นประเภท ω_i เมื่อทราบข้อมูลชุดนั้น โดย i=1, 2,..., K

K คือ จำนวนประเภทของข้อมูล

d_i คือระยะห่างในปริภูมิที่ถูกแผ่ออกโดยฟังก์ชันดิสคริมิแนนต์ ภายในกลุ่ม ω_i

4. **การทดสอบนัยสำคัญ** (Significance testing)[9]

การทดสอบนัยสำคัญเป็นเทคนิคในการหาค่านัยสำคัญระหว่างขั้นตอนวิธี (Algorithm) ในการจำแนกกลุ่ม 2 วิธีใดๆ ซึ่งจะเป็นการเปรียบเทียบจากผลความแตกต่างที่เกิดขึ้นระหว่างการใช้ 2 วิธีดังกล่าวจำแนกกลุ่มใดๆ เนื่องจากการสังเกตเฉพาะขนาดของความแตกต่างนั้น ไม่สามารถบอกค่านัยสำคัญได้ เนื่องจากความแตกต่างกันในปริมาณมากอาจไม่มีนัยสำคัญหากเลือกใช้ข้อมูลทดสอบขนาดเล็ก ในขณะเดียวกัน ความแตกต่างเพียงเล็กน้อยอาจมีนัยสำคัญได้หากใช้ข้อมูลทดสอบขนาดใหญ่

สำหรับขั้นตอนการหานัยสำคัญของ 2 วิธีการใดๆนั้น จำเป็นที่จะต้องใช้ข้อมูลฝึกฝนและข้อมูลทดสอบชุดเดียวกันทั้ง 2 วิธี เพื่อให้การทดสอบนัยสำคัญที่ได้เป็นการทดสอบเปรียบเทียบวิธีการทั้ง 2 วิธีจริงๆ ไม่ใช่เป็นผลที่ได้จากการวัดความแตกต่างของชุดข้อมูล นอกจากนี้ การวัดค่าความถูกต้องของวิธีการใดๆ จำเป็นว่าจะต้องสามารถระบุผลการจำแนกของข้อมูบแต่ละตัวได้ เนื่องจากในการทดสอบนัยสำคัญนี้ จำเป็นต้องนับจำนวนของข้อมูลที่สามารถจำแนกได้ถูกต้องในวิธีการใดๆวิธีการหนึ่งเท่านั้น

โดยการทดสอบนัยสำคัญมีสมมติฐานตั้งต้นว่า ความน่าจะเป็นของผลที่สามารถจำแนกได้ถูกต้อง (p_i) ของทั้ง 2 วิธีมีค่าเท่ากัน

ตัวอย่างวิธีทดสอบนัยสำคัญ ได้แก่ วิธีทดสอบนัยสำคัญของแมคนีมา (McNemar) ซึ่งมีวิธีการคำนวณ โดย

นับจำนวนข้อมูลทั้งหมด ซึ่งสามารถแบ่งประเภทของข้อมูลได้ 4 ลักษณะ ดังนี้

- จำนวนข้อมูลที่สามารถจำแนกได้ถูกต้องทั้ง 2 วิธี ($\boldsymbol{n}_{00}$)
- จำนวนข้อมูลที่วิธีที่ 1 สามารถจำแนกถูกเท่านั้น ($\boldsymbol{n}_{01}$)
- จำนวนข้อมูลที่วิธีที่ 2 สามารถจำแนกถูกเท่านั้น ($\boldsymbol{n}_{10}$)
- จำนวนข้อมูลที่จำแนกผิดทั้ง 2 วิธี ($\boldsymbol{n}_{11}$)

เพื่อที่จะเปรียบเทียบขั้นตอนวิธีทั้ง 2 วิธี จึงตั้งสมมติฐานตั้งต้นว่า

$$p_1 = p_2 \text{ หรือ } \boldsymbol{n}_{01} = \boldsymbol{n}_{10} \text{ หรือ } \boldsymbol{q} = \frac{\boldsymbol{n}_{01}}{\boldsymbol{n}_{01} + \boldsymbol{n}_{10}} = \frac{1}{2} \qquad (2.22)$$

ดังนั้นเมื่อรู้สมมติฐานข้างต้นแล้ว ค่าความน่าจะเป็นของการที่ข้อมูล k ตัวมีการจำแนกออกจากกันอย่างอสมมาตรในขนาด $\boldsymbol{n} = \boldsymbol{n}_{01} + \boldsymbol{n}_{10}$ มีค่าฟังก์ชันมวลความน่าจะเป็นทวินาม (Binomial probability mass function: Binomial PMF) แสดงเป็นสมการ ได้ว่า

$$P(k) = \binom{n}{k}\left(\frac{1}{2}\right)^n \qquad (2.23)$$

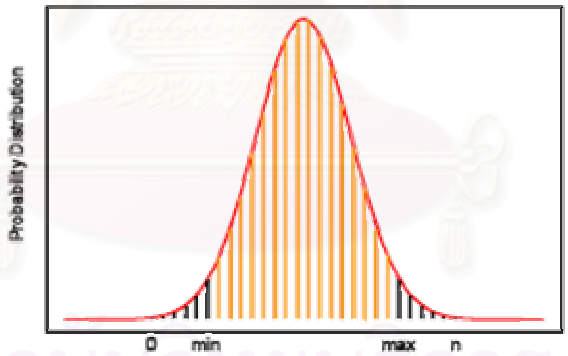
โดยการทดสอบนัยสำคัญของแมคนีมา จะวัดค่าความน่าจะเป็น $\boldsymbol{P}$ ของกรณีที่มีค่าเกินกว่าการกระจายอสมมาตร ซึ่ง $\boldsymbol{P} < \alpha$ โดยที่ α คือเกณฑ์ที่กำหนดขึ้น เพื่อวัดว่าวิธีการการทั้ง 2 วิธีมีนัยสำคัญต่อกันหรือไม่ ซึ่งโดยทั่วไปมักกำหนดให้ค่า $\alpha = 0.005$ หมายความว่า หากในกรณีที่ $\boldsymbol{P} = 0.001$ จะแสดงว่า ทั้ง 2 วิธีนี้ ความแตกต่างมีนัยสำคัญต่อกัน เป็นต้น

ซึ่งค่าความน่าจะเป็น $\boldsymbol{P}$ สามารถคำนวณได้จาก

$$\boldsymbol{P} = \sum_{k=0}^{l} \boldsymbol{P}(\boldsymbol{k}) + \sum_{k=m}^{n} \boldsymbol{P}(\boldsymbol{k}) \qquad (2.24)$$

โดยที่ $\boldsymbol{l} = \min(\boldsymbol{n}_{01}, \boldsymbol{n}_{10})$ และ $\boldsymbol{m} = \max(\boldsymbol{n}_{01}, \boldsymbol{n}_{10})$

ซึ่งถือว่าเป็นการรวมค่าที่บริเวณปลายของค่าฟังก์ชันกลุ่มความน่าจะเป็น แสดงดังรูป



รูปที่ 2.9 บริเวณของการกระจายตัวของค่าความน่าจะเป็นที่รวมเป็นค่า $\boldsymbol{P}$ [9]

งานวิจัยที่เกี่ยวข้อง

สำหรับงานวิจัยที่เกี่ยวข้องนี้ จะกล่าวถึงงานวิจัยที่เกี่ยวข้องกับการจำแนกหน่วยเสียง ที่ใช้ลักษณะเฉพาะที่มีการทำงานแบบอาศัยเซกเมนต์และใช้ลักษณะเฉพาะที่มีการทำงานแบบอาศัยกรอบเวลาเป็นหลัก

1. **ลักษณะเฉพาะที่มีการทำงานแบบอาศัยกรอบเวลา**

งานวิจัยของเม็ง (Meng) และซู (Zue) [10] ในปี 1991 ต้องการศึกษาประโยชน์ในการดึงเอาสมบัติทางเสียง (Acoustic attributes) บนการใช้ข้อมูลทางสเปกตรัมที่มีหลักการทำงานในการ

จัดการข้อมูลแบบอาศัยกรอบเวลา เพื่อทำการจัดแบ่งประเภทของหน่วยเสียง รวมทั้งศึกษาประโยชน์ในการใช้สมบัติลักษณะเฉพาะ (Distinctive features) เพื่อทำการจัดแบ่งประเภทหน่วยเสียงเช่นกัน โดยได้ทำการทดสอบโดยการทดลองจำแนกสระ 13 ตัวในภาษาอังกฤษด้วยใช้วิธีข่ายงานประสาทเทียม (Artificial neural network) ในการจำแนกสระดังกล่าว ในงานวิจัยนี้ได้ทำการทดลองโดยใช้ทีมิท(TIMIT) เป็นฐานข้อมูลเสียงในการทดลอง โดยผลการทดลองที่ได้นั้นสรุปว่าผลจากการใช้สมบัติทางเสียงได้ผลใกล้เคียงกับการใช้ข้อมูลสเปกตรัมโดยตรง แต่จะมีข้อดีกว่าตรงที่ประหยัดพลังงานในการคำนวณ ส่วนการทดลองศึกษาประโยชน์ในการใช้สมบัติลักษณะเฉพาะได้ผลใกล้เคียงกับการจำแนกสระโดยตรงเช่นกัน แต่เป็นเครื่องมือที่มีความยืดหยุ่นสูงกว่า

สำหรับงานวิจัยนี้ถือได้ว่าเป็นงานวิจัยที่มีความต้องการพัฒนาลักษณะเฉพาะที่ได้จากการแทนข้อมูลแบบสเปกตรัมเช่นกัน แต่เป็นการหาลักษณะเฉพาะจากการหาค่าอัตราค่าเฉลี่ย (Mean Rate) มาเป็นการแทนข้อมูลแบบสเปกตรัมดังกล่าว แต่ งานวิจัยนี้ทำอยู่บนโดเมนที่แคบกว่า นั่นคือการจำแนกหน่วยเสียงเฉพาะเสียงสระเท่านั้น ซึ่งพฤติกรรมของสระเป็นลักษณะที่สามารถเห็นค่าความถี่สั่นพ้องได้อย่างชัดเจนในทุกๆสระ

งานวิจัยต่อมาในปี 2003 ซึ่งเป็นงานวิจัยส่วนหนึ่งในวิทยานิพนธ์ของเอกฤทธิ์ มณีน้อย [2]ที่ได้ทำการทดลองการรู้จำหน่วยเสียงภาษาไทยโดยใช้หลักการแบบจำลองทางเสียง(Acoustic Models) อย่างเดียวเท่านั้น กล่าวคือ ยังไม่คำนึงถึงการใช้หลักการแบบจำลองทางภาษา (Language models) เข้ามาช่วยในการรู้จำหน่วยเสียง โดยการทดลองดังกล่าวจะรับสัญญาณเสียงต่อเนื่องเข้ามาแล้วจึงจำแนกออกเป็นหน่วยเสียง ด้วยแบบจำลองฮิดเดนมาร์คอฟ ซึ่งถือว่ามีลักษณะการทำงานแบบอาศัยกรอบเวลาและใช้สัมประสิทธิ์เมลฟรีเคว็นซีเคปสตรอลขนาด 12 หลักและพลังงาน 1 หลักรวมกับสัมประสิทธิ์เดลต้าเป็นลักษณะเฉพาะซึ่งได้ทำการทดลองกับผู้พูดหลายระบบ ได้แก่ ระบบผู้พูดชายอย่างเดียวแบบขึ้นกับผู้พูด ระบบผู้พูดชายอย่างเดียวแบบไม่ขึ้นกับผู้พูด ระบบผู้พูดหญิงอย่างเดียวแบบขึ้นกับผู้พูดและระบบผู้พูดหญิงอย่างเดียวแบบไม่ขึ้นกับผู้พูด

จากลักษณะการทดลองดังกล่าวข้างต้นที่ได้ใช้เฉพาะหลักการแบบจำลองทางเสียงอย่างเดียว โดยยังไม่คำนึงถึงแบบจำลองทางภาษา จึงถือได้ว่าการทดลองดังกล่าวเป็นการจำแนกหน่วยเสียงในภาษาไทยเช่นเดียวกับวิทยานิพนธ์นี้ แต่งานวิจัยดังกล่าวได้ใช้แบบจำลองฮิดเดนมาร์คอฟซึ่งถือว่าเป็นหลักการทำงานแบบอาศัยกรอบเวลา ดังนั้นในวิทยานิพนธ์นี้จะใช้งานวิจัยดังกล่าวเป็นมาตรฐานในการเปรียบเทียบประสิทธิภาพความถูกต้องระหว่างการจำแนกหน่วยเสียงภาษาไทยโดยใช้หลักการทำงานแบบอาศัยกรอบเวลากับการจำแนกหน่วยเสียงภาษาไทยโดยใช้หลักการทำงานแบบอาศัยเซกเมนต์ โดยจะปรับเปลี่ยนลักษณะบางส่วน เพื่อให้การทดลองดังกล่าวมีความครอบคลุมเพื่อสามารถนำมาใช้ในการเปรียบเทียบได้อย่างดีที่สุด

ต่อมาในปี 2004 ได้มีงานวิจัยบนภาษาไทยอีกเช่นกัน โดยศวิต กาสุริยะ (Kasuriya) และคณะ[11]ได้ทำการเสนอแบบจำลองเสียง (Acoustic models) ที่เหมาะสมกับการรู้จำเสียงพูดในภาษาไทยในหลายๆแบบจำลอง ได้แก่ แบบจำลองเสียงเดี่ยว (Single phone models) แบบจำลองเสียงแบบพื้นฐาน (Base phone models) ซึ่งทั้ง 2 แบบจำลองเป็นแบบจำลองที่ไม่คำนึงถึงโทนเสียง และแบบจำลองหน่วยเสียงที่จำแนกตามโทนเสียง (Phone with tone modeling) ซึ่งลักษณะแบบจำลองทั้งหมดที่กล่าวถึงนี้เป็นการเสนอแบบจำลองที่รวมกลุ่มภาษาไทยตามพฤติกรรมที่เหมือนกัน ดังนี้ แบบจำลองเสียงเดี่ยว จะเป็นแบบจำลองที่คำนึงถึงหน่วยเสียงขึ้นพื้นฐานที่สุด กล่าวคือ สระเสียงสั้นเสียงยาวนั้นถือว่าเป็นสระประเภทเดียวกัน เพราะมีลักษณะออกเสียงที่เหมือนกัน สระเสียงยาวก็เกิดจากการนำสระเสียงสั้นมาต่อกันพยัญชนะควบกล้ำก็เกิดจากการนำพยัญชนะ 2 ตัวมาต่อกันเช่นเดียวกัน รวมไปถึงพยัญชนะตัวสะกดก็ถือว่าเป็นลักษณะเช่นเดียวกับพยัญชนะต้น จึงกล่าวได้ว่าแบบจำลองเสียงเดี่ยวนี้ ถือว่าเป็นแบบจำลองที่มีขนาดเล็กที่สุด ซึ่งมีขนาดหน่วยเสียงทั้งสิ้น 31 หน่วยเสียง

แบบจำลองต่อมาคือ แบบจำลองพื้นฐาน ซึ่งเป็นแบบจำลองที่เพิ่มเติมจากแบบจำลองเสียงเดี่ยว กล่าวคือ มีการจำแนกสระเป็นประเภทสระเสียงสั้น สระเสียงยาว และ สระผสม แบ่งพยัญชนะออกเป็น พยัญชนะต้น พยัญชนะควบกล้ำ และพยัญชนะตัวสะกด จึงทำให้แบบจำลองดังกล่าวมีหน่วยเสียงทั้งสิ้น 75 หน่วยเสียง

แบบจำลองชุดต่อมาคือแบบจำลองที่กล่าวถึงโทนเสียงภาษาไทยที่มีทั้งหมด 5 โทนด้วย ซึ่งจะเป็นการนำ เสียงสระในแบบจำลองเสียงเดี่ยวและแบบจำลองพื้นฐานมาจำแนกออกเป็นหน่วยเสียงประเภทต่างๆตามโทนเสียงเสียงละ 5 ประเภท

นอกจากนี้งานวิจัยนี้ได้ทำการทดลองเพื่อทำการวัดประสิทธิภาพของการรู้จำคำในภาษาไทยกับแบบจำลองเสียงต่างๆ โดยมีการเปลี่ยนแปลงจำนวนส่วนประกอบ(Mixture) เป็นค่าต่างๆ ซึ่งในที่นี้ได้ใช้เอชทีเค (HTK) เป็นเครื่องมือในการสร้างแบบจำลองฮิดเดนมาร์คอฟโดยใช้สัมประสิทธิ์เมลฟรีเคว็นซีเคปสตรอล12 หลัก + ลอการิทึมของพลังงาน + สัมประสิทธิ์เดลต้า 13 หลัก + สัมประสิทธิ์ความเร่ง 13 หลัก รวมทั้งสิ้น 39 หลักเป็นลักษณะเฉพาะโดยผลการทดลองในการรู้จำคำในภาษาไทยสำหรับแบบจำลองที่ไม่มีโทนที่ดีที่สุดคือ 87.03 เปอร์เซนต์ ซึ่งได้จากแบบจำลองเสียงเดี่ยว และผลการรู้จำคำในภาษาไทยสำหรับแบบจำลองที่มีโทนที่ดีที่สุดที่ได้ คือ 82.03 เปอร์เซนต์ จากแบบจำลองแบบพื้นฐาน

ถึงแม้ว่างานวิจัยนี้เป็นการทดลองกับเสียงภาษาไทยและได้เลือกใช้เอ็มเอฟซีซีเป็นลักษณะเฉพาะเช่นเดียวกันก็ตาม แต่งานวิจัยนี้กลับได้ทำการทดลองในลักษณะรู้จำในระดับคำในภาษาไทย ซึ่งไม่ใช่การทดลองจำแนกหน่วยเสียงภาษาไทยโดยตรงตามแบบจำลองที่ได้เสนอไว้ รวมทั้งงานวิจัยดังกล่าวได้ใช้แบบจำลองฮิดเดนมาร์คอฟเป็นเทคนิคในการรู้จำจึงถือว่าเป็นการทำงานแบบอาศัยกรอบเวลา

2. ลักษณะเฉพาะที่มีการทำงานแบบอาศัยเซกเมนต์

ในปี 1993 ลูง (Leung) เช็คจิเยอร์ (Chigier) และกลาส (Glass)[12] ได้ทำการทดลองรู้จำเสียงโดยมีเป้าหมายในการทดสอบ 2 เรื่อง คือ เรื่อง การแทนสัญญาณเสียง (Representation) และการจำแนกเสียง(Signal classification)โดยในการแทนสัญญาณเสียงนั้น ได้ทดลองใช้ลักษณะเฉพาะซึ่งสามารถแบ่งเป็นกลุ่มหลักๆ 2 กลุ่ม คือ กลุ่มการแทนสัญญาณเสียงซึ่งมีพื้นฐานแบบสเปกตรัม (The spectrally-based representations) และกลุ่มการแทนสัญญาณเสียงซึ่งมีพื้นฐานแบบเคพทรัม (The cepstrally-based representations) โดยแต่ละกลุ่มได้มีการใช้ลักษณะเฉพาะอีกกลุ่มละ 5 วิธี รวมลักษณะเฉพาะทั้งสิ้น 10 รูปแบบ ส่วนวิธีการจำแนกเสียง ในงานวิจัยนี้ได้ใช้วิธีการจำแนกเสียงทั้งหมด 3 วิธี ทำการทดลองโดยใช้ทีมิท(TIMIT) และเอ็นทีมิท(NTIMIT) เป็นฐานข้อมูล ซึ่งข้อมูลดังกล่าวเป็นชื่อคน 25 ชื่อ ซึ่งรับเสียงดังกล่าวจากทางโทรศัพท์ โดยผู้พูดมีทั้งผู้ชาย ผู้หญิงและเด็ก

ในงานวิจัยนี้ถือว่าเป็นการเปรียบเทียบทั้งลักษณะเฉพาะแบบต่างๆ และใช้วิธีการจำแนกหน่วยเสียงแบบต่างๆ ซึ่งในงานวิจัยนี้ได้เริ่มต้นด้วยการทำงานแบบอาศัยเซกเมนต์ตัดแบ่งสัญญาณเสียงออกมาเป็นเซกเมนต์แล้วแบ่งเซกเมนต์ดังกล่าวออกเป็น 3 ส่วนก่อนจึงใช้วิธีแบบอาศัยกรอบเวลาโดยการแบ่งเซกเมนต์ออกเป็นเฟรมเฟรมละ 28.5 มิลลิวินาทีเพื่อนำมาทำการหาลักษณะเฉพาะและสมบัติของเซกเมนต์อื่นในแต่ละส่วนต่อไป ดังนั้นจากวิธีการดังกล่าวจึงถือได้ว่างานวิจัยนี้ถึงแม้ว่าจะมีการทำงานแบบอาศัยเซกเมนต์ก็ตาม แต่ก็ยังคงมีการอ้างอิงหรือขึ้นกับการทำงานแบบอาศัยกรอบเวลาอยู่

อีกงานวิจัยหนึ่งเป็นงานวิจัยของซาโฮเรียน (Zahorian) ซิลสบี (Silsbee) และหวัง (Wang) [12] ที่ได้มีการนำวิธีการแบบอาศัยเซกเมนต์มาใช้ โดยในงานวิจัยนี้ ได้ทำการแบ่งหมวดหมู่ (classification) หน่วยเสียงภาษาอังกฤษทั้งสิ้น 39 หมวดหมู่ โดยใช้ทีมิทและเอ็นทีมิทเป็นฐานข้อมูลในการทำงาน ในงานวิจัยนี้ได้ใช้สมบัติทางสเปกตรัมและทางเวลา (Spectral/temporal feature) ซึ่งเป็นค่าที่ได้จากการและทางทำการแปลงโคไซน์แบบไม่ต่อเนื่อง (discrete cosine transform: DCT) มาใช้เป็นลักษณะเฉพาะในการวิเคราะห์การแบ่งหมวดหมู่ ส่วนวิธีในการแบ่งหมวดหมู่นั้นงานวิจัยนี้ได้ทำการทดลองโดยการใช้ทั้งวิธีเครือข่ายทางประสาทที่แบ่งแบบไบนารีแพร์(Binary-pair partitioned neural network) และวิธีเครือข่ายทางประสาทแบบซิงเกิลลาจ(Single large neural network) ซึ่งผลการทดลองที่ได้ของทั้ง 2 วิธีดังกล่าวมีความใกล้เคียงกัน แต่วิธีเครือข่ายทางประสาทที่แบ่งแบบไบนารีแพร์ใช้เวลาในการฝึกฝนน้อยกว่า โดยผลการทดลองที่ดีที่สุดที่ได้คือ 77 เปอร์เซนต์ สำหรับฐานข้อมูลทีมิทและ 67.4 เปอร์เซนต์ สำหรับฐานข้อมูลเอ็นทีมิท

สำหรับงานวิจัยนี้ได้ทำการทดลองทั้งการศึกษาลักษณะเฉพาะและวิธีการจำแนกเสียง รวมทั้งได้ใช้สมบัติของเซกเมนต์เช่นเดียวกัน โดยงานวิจัยนี้ได้ใช้เทคนิคในการตัดเฉพาะส่วนกลางของหน่วยเสียงขนาด 300 มิลลิวินาที แล้วจึงทำการแบ่งส่วนดังกล่าวออกเป็นเฟรมเฟรมละ 10 มิลลิวินาทีก่อนนำไปหาค่าทางสเปกตรัมและทางเวลา ซึ่งถือว่าก็ยังคงมีการอ้างอิงการทำงานในลักษณะแบบอาศัยกรอบเวลาอยู่ นอกจากนี้ในงานวิจัยนี้ได้ทำการปรับเลือกใช้จำนวนค่าลักษณะเฉพาะทางสเปกตรัมและทางเวลาไม่เท่ากัน เพื่อทดสอบว่าค่าใดมีความสำคัญในการจำแนกหน่วยเสียงมากกว่ากัน ซึ่งในการทดลองดังกล่าวได้ผลว่าการปรับเลือกใช้จำนวนค่าทางสเปกตรัมเวลานั้นไม่ได้ส่งผลต่อประสิทธิภาพในการจำแนกหน่วยเสียงเท่าไรนัก

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 3

หน่วยเสียงภาษาไทย

บทนี้จะอธิบายถึงลักษณะหน่วยเสียงต่างๆในภาษาไทยและกล่าวถึงกลุ่มหน่วยเสียงที่จะนำมาจำแนกหน่วยเสียงภาษาไทยเพื่อวัดประสิทธิภาพของลักษณะเฉพาะในงานวิจัยนี้

นิยามหน่วยเสียงในภาษาไทย

เสียงในภาษาไทยนั้นสามารถแบ่งเป็น 3 กลุ่มใหญ่ๆ ตามพฤติกรรมการเปล่งเสียง คือ เสียงสระ เสียงพยัญชนะ และ เสียงกึ่งสระ โดยแต่ละกลุ่มมีลักษณะต่างๆ [14] ดังนี้

1. **เสียงสระ** (Vowels)

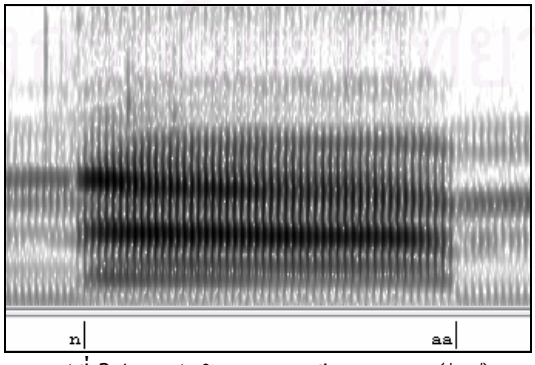
เสียงสระคือ เสียงที่เกิดจากการกระตุ้นอวัยวะที่ใช้ในการเปล่งเสียง (Vocal tract) ด้วยอากาศที่มีลักษณะเป็นคาบ ซึ่งมีสาเหตุมาจากการสั่นของเส้นเสียง (Vocal cords) โดยภายในกลุ่มเสียงสระนั้นหน่วยเสียงแต่ละเสียงจะแตกต่างกันตรงตำแหน่งความสูง-ต่ำและตำแหน่งหน้า-หลังของลิ้นในการเปล่งเสียง รวมไปถึงความยาวของสระเสียงสั้นและเสียงยาวด้วย

ซึ่งในภาษาไทยมีสระทั้งหมด 24 ตัว แบ่งเป็นสระเสียงเดี่ยว 18 ตัว และสระผสม 6 ตัว โดยแสดงสระต่างๆแบ่งตามตำแหน่งของลิ้นในการออกเสียงได้ ดังตารางที่3.1

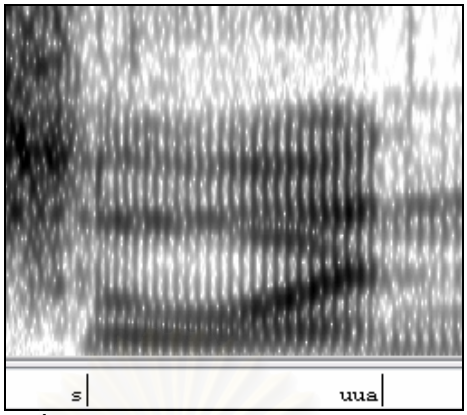
ตารางที่3.1 สระตามลักษณะตำแหน่งความสูงของลิ้น

ตำแหน่งความสูงของลิ้น	ตำแหน่ง		
	หน้า(เสียงสั้น, ยาว)	กลาง(เสียงสั้น, ยาว)	หลัง(เสียงสั้น, ยาว)
สูง	i, ii	v, vv	u, uu
กลาง	e, ee	q, qq	o, oo
ต่ำ	x, xx	a, aa	@, @@
สระผสม	ia, iia	va, vva	ua, uua

ตัวอย่างสัญญาณเสียงสระสามารถแสดงได้ด้วยสเปกโตรแกรม (Spectrogram) ดังรูป



รูปที่ 3.1 สเปกโตรแกรมเสียงสระอา (/aa/)



รูปที่ 3.2 สเปกโตรแกรมเสียงสระอัว (/uua/)

2. **เสียงพยัญชนะ** (Consonants)

เสียงพยัญชนะเกิดจากกลไกการปิดหรือการหดตัวเป็นช่องแคบๆภายในช่องปาก ซึ่งอวัยวะที่ใช้ในการพูดจะเคลื่อนไหวอย่างรวดเร็วในการเปล่งเสียงพยัญชนะ โดยเสียงพยัญชนะสามารถแบ่งกลุ่มย่อยตามพฤติกรรมการเปล่งเสียงได้ 4 กลุ่มและแต่ละกลุ่มนั้นยังสามารถแบ่งแยกย่อยตามพฤติกรรมการสั่นของเส้นเสียง (สั่นและไม่สั่น) ได้ ดังตารางที่3.2

ตารางที่3.2 พยัญชนะในแต่ละประเภทแบ่งตามพฤติกรรมการเปล่งเสียง

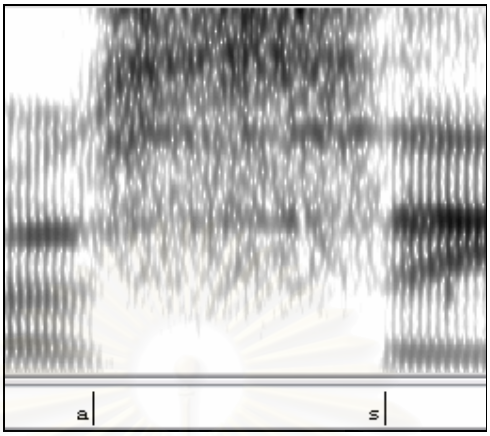
พฤติกรรมการเปล่งเสียง		**พยัญชนะต้น**	**ตัวสะกด**
พยัญชนะเสียงเสียดแทรก	สั่น	-	-
	ไม่สั่น	f, s	f^, s^
พยัญชนะกัก	สั่น	b, d	-
	ไม่สั่น	p, t, k, ph, th, kh	p^, t^, k^
พยัญชนะเสียงกึ่งเสียดแทรก	สั่น	ch	ch^
	ไม่สั่น	c	-
พยัญชนะนาสิก		n, m, ng	n^, m^, ng^

โดยลักษณะของพยัญชนะแต่ประเภทมีลักษณะพฤติกรรมการเปล่งเสียงที่แตกต่างกัน ดังนี้

2.1 **พยัญชนะเสียงเสียดแทรก** (Fricative consonants)

พยัญชนะเสียงเสียดแทรกจะมีลักษณะของสัญญาณเป็นเสียงรบกวน (Noise) ที่เกิดจากกระแสลมถูกขับผ่านช่องแคบ โดยความแตกต่างของพยัญชนะเสียงเสียดแทรกแต่ละตัวจะขึ้นกับตำแหน่งของช่องแคบที่ใช้ในการสร้างเสียงเสียดแทรกนั้นๆซึ่งจะส่งผลให้พลังงานของสัญญาณรบกวนจะหนาแน่นในช่วงความถี่ใดๆต่างกัน ซึ่งการเปล่งเสียงเสียดแทรกในภาษาไทยทุกตัวจะไม่พบหน่วยเสียงที่มีการสั่นของเส้นเสียง

ตัวอย่างสัญญาณเสียงเสียดแทรกสามารถแสดงได้ด้วยสเปกโตรแกรม ดังรูป



รูปที่ 3.3 สเปกโตรแกรมเสียงเสียดแทรก ส,ศ,ษ,ซ (/s/)

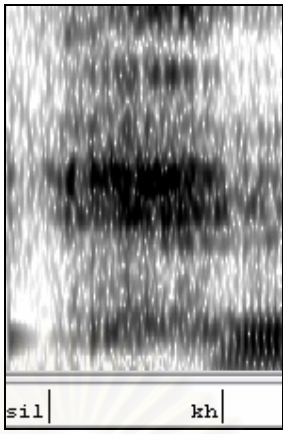
2.2 **พยัญชนะกัก** (Stop consonants)

พยัญชนะกักแบบไม่สั่น ในช่วงแรกนั้นจะมีลักษณะพลังงานตลอดทั้งช่วงความถี่หายไป ซึ่งเกิดจากการสร้างช่องปิดในกรณีที่เป็นพยัญชนะกักแบบไม่สั่น ส่วนในขณะที่ช่องปิดเปิดออก ถ้ามีการปล่อยลมออกมาจะพบลักษณะสัญญาณเสียงเป็นเสียงรบกวนกระจัดกระจายในบริเวณช่วงความถี่สูง พยัญชนะกักในกรณีนี้ได้แก่ หน่วยเสียง /ph/, /th/, /kh/ หากเป็นกรณีที่จุดกักลมดังกล่าวไม่เปิดออก แล้วลมที่ถูกกักไว้ถูกกลืนลงไปแทน ทำให้มีลักษณะเหมือนเสียงเงียบ และเมื่อช่องปิดดังกล่าวถูกเปิดออกมาอย่างรวดเร็วนั้น อาจทำให้เกิดลักษณะสัญญาณเป็นเสียงรบกวนสั้นๆ ซึ่งพยัญชนะกักในกรณีนี้ได้แก่ หน่วยเสียง /p/, /t/, /k/

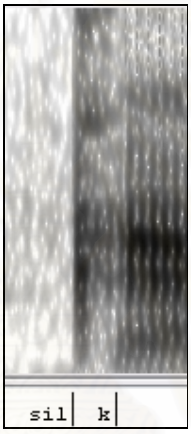
สำหรับพยัญชนะกักแบบสั่นนั้น ในช่วงแรกจะปรากฏพลังงานในช่วงความถี่ต่ำอันเกิดจากการสั่นของเส้นเสียงในขณะที่เกิดช่องปิด ซึ่งพลังงานที่ถูกส่งผ่านออกมาในอากาศนี้จะผ่านออกมาจากการแผ่รังสี (Radiation) จากกระพุ้งแก้ม ไม่ใช่ลมจากช่องปากโดยตรง ก่อนที่จะมีการระเบิดลมออกมา ซึ่งพยัญชนะกักในกรณีนี้ได้แก่ หน่วยเสียง /b/, /d/

โดยในงานวิจัยนี้จะทำการแบ่งหน่วยเสียงดังกล่าวออกเป็น 2 ส่วนคือ ส่วนที่เกิดจากการกักลมซึ่งในกรณีที่เป็นพยัญชนะกักแบบไม่สั่นจะถือเป็นหน่วยเสียงเงียบ /sil/ ส่วนพยัญชนะกักแบบสั่นที่พบแถบพลังงานบริเวณความถี่ต่ำจะแทนด้วยหน่วยเสียง /vb/ และส่วนที่เกิดจากการระเบิดลมออกมาจะถือเป็นพยัญชนะกัก ซึ่งแทนด้วยหน่วยเสียงที่เป็นพยัญชนะกักตามเดิม ดังตัวอย่างในรูปที่ 3.6 แสดงการแบ่งส่วนหน่วยเสียง “ด” ออกเป็น 2 ส่วนคือ ส่วนแถบพลังงานความถี่ต่ำแทนสัญลักษณ์ด้วย /vb/ และส่วนที่มีการระเบิดลมออกมาแทนสัญลักษณ์ด้วย /d/

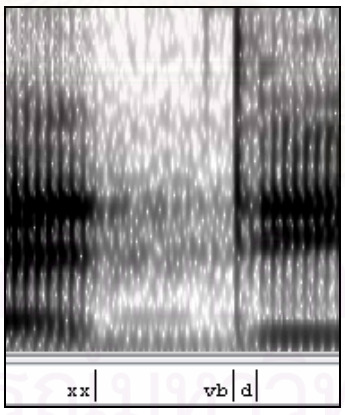
สำหรับตัวอย่างสเปกโตรแกรมของพยัญชนะกัก“ข”, “ค”, “ฆ” (/kh/) และ“ก” (/k/) แสดงในรูปที่ 3.4 และรูปที่ 3.5 ตามลำดับ



รูปที่ 3.4 สเปกโตรแกรมเสียงพยัญชนะกัก “ข”, “ค”, “ฆ” (/kh/)



รูปที่ 3.5 สเปกโตรแกรมเสียงพยัญชนะกัก “ก” (/k/)

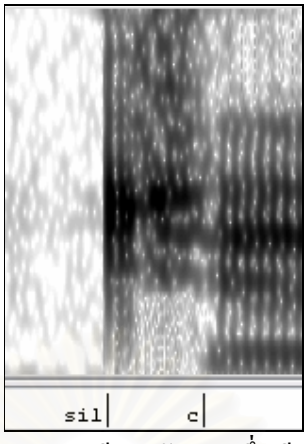


รูปที่ 3.6 สเปกโตรแกรมเสียงพยัญชนะกักแบ่งเป็นส่วนแถบพลังงานความถี่ต่ำ (/vb/) และ“ด”(/d/)

2.3 **พยัญชนะกึ่งเสียดแทรก** (Affricates)

พยัญชนะเสียงกึ่งเสียดแทรกจะเกิดจากการสร้างช่องปิด เพื่อเตรียมเปล่งเสียงพยัญชนะกัก แล้วตามด้วยเสียงเสียดแทรก หลังจากที่ช่องปิดดังกล่าวถูกปล่อยออก

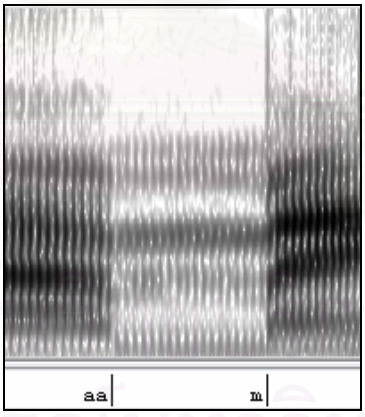
ตัวอย่างสัญญาณของพยัญชนะกักสามารถแสดงได้ด้วยสเปกโตรแกรม ดังรูป



รูปที่ 3.7 สเปกโตรแกรมเสียงพยัญชนะกึ่งเสียดแทรก “จ” (/c/)

2.4 **พยัญชนะนาสิก** (Nasal consonants)

พลังงานเสียงพยัญชนะนาสิกในช่วงความถี่กลางถึงสูงจะลดต่ำลงจากระดับพลังงานของสระที่อยู่ใกล้เคียงเล็กน้อย เนื่องจากการสูญเสียพลังงานส่วนหนึ่งในโพรงจมูก นอกจากนี้เสียงนาสิกที่ตามหลังสระจะมีลูกคลื่นในทางเวลามีขนาดต่ำลงกว่าลูกคลื่นของสระที่นำหน้าอยู่และมีการหายไปของส่วนประกอบความถี่สูง (ลูกคลื่นมีรูปร่างขรุขระน้อยลง) ซึ่งถ้าพิจารณา สเปกโตรแกรมของสระในส่วนที่ติดกับพยัญชนะนาสิกซึ่งถูกทำให้เสียงส่วนหนึ่งผ่านโพรงจมูกนั้น จะพบว่าแถบพลังงานความถี่ของฟอร์แมนท์ที่หนึ่ง (F1) ของสระนั้นจะกว้างขึ้นกว่าสระในส่วนที่ไม่ผ่านโพรงจมูก



รูปที่ 3.8 สเปกโตรแกรมเสียงพยัญชนะนาสิก “ม” (/m/)

3. **เสียงกึ่งสระ** (Semi-vowels)

เสียงกึ่งสระจะเป็นเสียงที่อยู่กำกึ่งระหว่าง 2 กลุ่มข้างต้น โดยการเปล่งเสียงกึ่งสระนั้นจะเกี่ยวข้องกับรูปแบบการหดคลายตัวของอวัยวะที่ใช้ในการเปล่งเสียง แต่จะมีปริมาณในการหดคลายตัวและความเร็วในการเคลื่อนที่ของอวัยวะในการเปล่งเสียงต่ำกว่าการเปล่งเสียงพยัญชนะ และมีการเปลี่ยนแปลงของลักษณะโครงสร้างกลุ่มความถี่สั่นพ้อง (Formant frequency) ที่รวดเร็ว

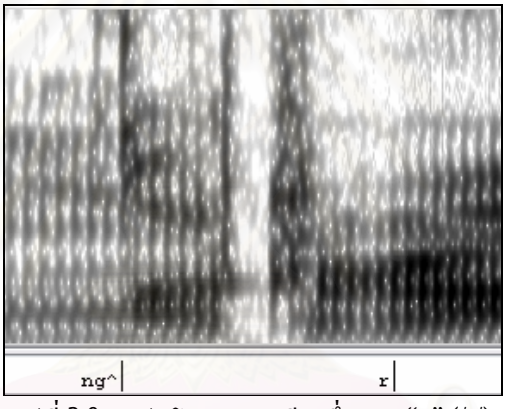
มากกว่าการเปลี่ยนแปลงที่เกิดในสระ เนื่องจากการสร้างช่องแคบที่แคบกว่า แต่มักจะไม่มีความไม่ต่อเนื่องของกลุ่มความถี่สั่นพ้องต่างๆ นอกจากนี้ยังอาจมีการเปลี่ยนแปลงทางขนาดของสัญญาณที่กะทันหันเหมือนกับการเปลี่ยนแปลงที่เกิดในพยัญชนะได้ด้วย

โดยเสียงกึ่งสระสามารถแบ่งแยกตามพฤติกรรมการเปล่งเสียงได้ ดังตารางที่3.3

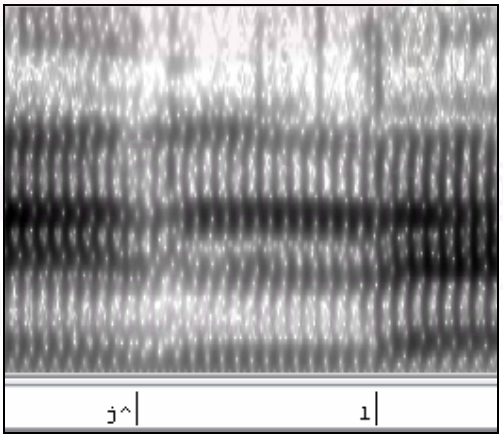
ตารางที่3.3 พยัญชนะกึ่งสระตามพฤติกรรมการเปล่งเสียง

พฤติกรรมการเปล่งเสียง	**พยัญชนะต้นเสียงกึ่งสระ**	**ตัวสะกดเสียงกึ่งสระ**
Lateral	l	l^
Trill	r	-
Aspirant	h	-
Glide	w, j	w^, j^

ตัวอย่างสัญญาณของเสียงกึ่งสระสามารถแสดงได้ด้วยสเปกโตรแกรม ดังรูป



รูปที่ 3.9 สเปกโตรแกรมเสียงกึ่งสระ “ร” (/r/)



รูปที่ 3.10 สเปกโตรแกรมเสียงกึ่งสระ “ล” (/l/)

กลุ่มหน่วยเสียงภาษาไทยสำหรับการจำแนกหน่วยเสียง

สำหรับในงานวิจัยนี้จะทำการทดสอบประสิทธิภาพของลักษณะเฉพาะในการแทนหน่วยเสียงต่างๆในภาษาไทย โดยจะทำการจำแนกหน่วยเสียง โดยจะใช้หน่วยเสียงภาษาไทยทั้งสิ้น 52 หน่วยเสียง ได้แก่ เสียงสระ 21 หน่วยเสียง เสียงพยัญชนะ 22 หน่วยเสียง เสียงกึ่งสระ 8 หน่วยเสียง และเสียงเงียบ 1 หน่วยเสียง โดยได้มีการตัดหน่วยเสียงบางหน่วยจากที่แสดงไว้ในส่วนนิยามหน่วยเสียงในภาษาไทย ได้แก่ สระผสมเสียงสั้น /ua/, /va/, /ia/ เนื่องจากไม่มีหน่วยเสียงดังกล่าวในฐานข้อมูลโลตัสชุดหน่วยเสียงสมดุล และเนื่องจากพฤติกรรมทางสัญญาณเสียงพยัญชนะกักประเภทตัวสะกด /p^/, /t^/, /k^/ มีพฤติกรรมเช่นเดียวกับเสียงเงียบ ดังนั้นพยัญชนะดังกล่าวจะถูกรวมเข้าไปกับหน่วยเสียงข้างเคียงตามความเหมาะสมกับพฤติกรรมของหน่วยเสียงนั้นๆ นอกจากหน่วยเสียงภาษาไทยที่ได้ระบุไว้ในส่วนนิยามหน่วยเสียงในภาษาไทยแล้ว ในภาษาไทยยังมีพยัญชนะประเภทควบกล้ำด้วย ซึ่งในงานวิจัยนี้ได้ทำการแยกพยัญชนะควบกล้ำดังกล่าวออกเป็น 2 หน่วยเสียงคือ หน่วยเสียงพยัญชนะและหน่วยเสียงกึ่งสระ ก่อนนำมาทำการจำแนกเสียง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สัทอักษรสากล

สัทอักษรสากลของหน่วยเสียงภาษาไทยที่ใช้ในงานวิจัยนี้ สามารถแสดงได้ดังตารางต่อไปนี้

ตารางที่3.4 สัทอักษรสากลที่ใช้แทนเสียงสระในภาษาไทย

สระในภาษาไทย	สัทอักษรสากล	ตัวอักษร ASCII
อิ	i	/i/
อี	iː	/ii/
เอะ	e	/e/
เอ	eː	/ee/
แอะ	ɛ หรือ æ	/x/
แอ	ɛː หรือ æː	/xx/
อึ	ɯ หรือ ɨ	/v/
อือ	ɯː หรือ ɨː	/vv/
เออะ	ɤ หรือ ə	/q/
เออ	ɤː หรือ əː	/qq/
อะ	a	/a/
อา	aː	/aa/
อุ	u	/u/
อู	uː	/uu/
โอะ	o	/o/
โอ	oː	/oo/
เอาะ	ɔ	/@/
ออ	ɔː	/@@/
เอีย	iːa	/iia/
เอือ	ɯːa หรือ ɨːa	/vva/
อัว	uːa	/uua/

ตารางที่3.5 สัทอักษรสากลที่ใช้แทนเสียงพยัญชนะต้นในภาษาไทย

พยัญชนะต้นในภาษาไทย	ลักษณะ	สัทอักษรสากล	ตัวอักษร ASCII
ป	Voiceless unaspirated bilabial stop release	p	/p/
ต ฏ	Voiceless unaspirated apico alveolar stop release	t	/t/
ก	Voiceless unaspirated dorso velar stop release	k	/k/
พ ภ ผ	Voiceless aspirated bilabial stop release	p^h	/ph/
ท ธ ฒ ฑ ถ ฐ	Voiceless aspirated apico alveolar stop release	t^h	/th/
ข ค ฆ	Voiceless aspirated dorso velar stop release	k^h	/kh/
-	Voiced stop closure	-	/vb/
-	Silence, Voiceless stop closure	-	/sil/
บ	voiced bilabial stop release	b	/b/
ด ฎ ฑ	Voiced apico alveolar stop release	d	/d/
ม หม	Voiced bilabial nasal	m	/m/
น ณ หน	Voiced apico alveolar nasal	n	/n/
ง หง	Voiced dorso velar nasal	ŋ	/ng/
จ จร	Voiceless unaspirated laminal palatal affricate	c	/c/
ช ฌ ฉ	Voiceless apico alveolar affricate	c^h	/ch/
ฝ ฟ	Voiceless labio dental fricative	f	/f/
ส ศ ษ ซ ทร สร ศร	Voiceless apico alveolar fricative	s	/s/
ห ฮ	Voiceless glottal fricative	h	/h/
ร หร ฤ	Voiced apico alveolar trill	r	/r/
ล ฬ หล ฦ	Voiced apico alveolar lateral	l	/l/
ว หว	Voiced labio velar approximant	w	/w/
ย ญ หย หญ อย	Voiced fronto palatal approximant	j	/j/

ตารางที่3.6 สัทอักษรสากลที่ใช้แทนเสียงพยัญชนะตัวสะกดภาษาไทย

พยัญชนะตัวสะกดภาษาไทย	**สัทอักษรสากล**	**ตัวอักษร** ASCII
แม่กน (น ณ ญ ร ล ฬ)	n*	/n^/
แม่กง (ง)	ŋ*	/ng^/
แม่กม (ม)	m*	/m^/
แม่เกอย (ย)	j*	/j^/
แม่เกอว (ว)	w*	/w^/
ตัวสะกดเสียงทับศัพท์ ฟ	f*	/f^/
ตัวสะกดเสียงทับศัพท์ ส	s*	/s^/
ตัวสะกดเสียงทับศัพท์ ช	ch*	/ch^/
ตัวสะกดเสียงทับศัพท์ ล	l*	/l^/

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 4

การหาลักษณะเฉพาะ

ในบทที่แล้ว ได้กล่าวถึงลักษณะของหน่วยเสียงในภาษาไทยที่เกิดขึ้นตามพฤติกรรมการเปล่งเสียง พร้อมทั้งแสดงตัวอย่างลักษณะของหน่วยเสียงภาษาไทยด้วยสเปกโตรแกรมที่เกิดขึ้นในลักษณะอุดมคติ ดังนั้นในบทนี้จะกล่าวถึงการหารูปแบบของลักษณะเฉพาะที่เหมาะสมกับการนำไปใช้แทนหน่วยเสียงทุกๆหน่วยเสียง ซึ่งจะเริ่มต้นด้วยการศึกษาพฤติกรรมโดยรวมของหน่วยเสียงทุกๆเสียงก่อน แล้วจึงจะทำการสรุปเป็นลักษณะเฉพาะที่นำไปใช้แทนหน่วยเสียง เพื่อให้สามารถนำไปใช้ในขั้นตอนการจำแนกหน่วยเสียงต่อไป

การศึกษาพฤติกรรมของหน่วยเสียง

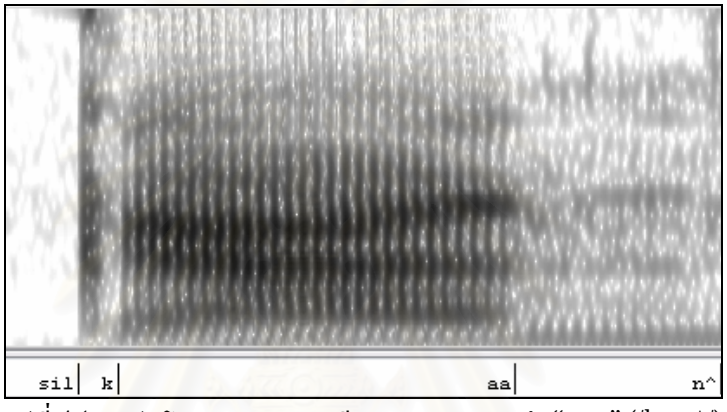
สำหรับการศึกษาพฤติกรรมของหน่วยเสียงในส่วนนี้นั้น จะเป็นการสังเกตพฤติกรรมของหน่วยเสียงในแต่ละเซกเมนต์ที่เกิดขึ้นจริงนำการไปใช้รู้จำเสียง ซึ่งในการรู้จำเสียงพูดใดๆก็ตามนั้น หน่วยที่เล็กที่สุดสำหรับเสียงพูดมักเป็นหน่วยคำ ซึ่งเกิดจากการประกอบกันของหน่วยเสียงย่อยๆ ได้แก่ พยัญชนะกับสระ(CV) หรือ พยัญชนะ สระและตัวสะกด (CVC) เรียงตามลำดับ ซึ่งเป็นลักษณะการออกเสียงหน่วยเสียงต่างๆติดต่อกันทันที ทำให้อวัยวะในการออกเสียงจะต้องมีการเคลื่อนที่จากตำแหน่งที่ใช้เปล่งเสียงของหน่วยเสียงตัวแรกไปยังหน่วยเสียงตัวที่สองทันทีเช่นกัน ส่งผลให้สมบัติของหน่วยเสียงที่บริเวณขอบเขตระหว่างรอยต่อของหน่วยเสียงที่กำลังจะออกเสียงนั้นมีความขึ้นกันกับหน่วยเสียงก่อนหน้านี้ ทำให้เซกเมนต์ดังกล่าวเกิดความไม่สม่ำเสมอกันภายในเซกเมนต์ ซึ่งมีผลให้ส่วนที่สามารถแสดงสมบัติของหน่วยเสียงดังกล่าวได้ดีที่สุดคือบริเวณส่วนกลางของเซกเมนต์ที่เกิดจากอวัยวะในการออกเสียงได้เคลื่อนที่มายังบริเวณที่ต้องการออกเสียงโดยสมบูรณ์แล้ว ก่อนที่จะมีการเคลื่อนที่เพื่อไปออกเสียงในหน่วยเสียงถัดไปหรือบริเวณส่วนปลายของเซกเมนต์

ซึ่งสาเหตุดังกล่าวนอกจากจะทำให้เซกเมนต์ต่างๆเกิดความไม่สม่ำเสมอกันในแต่ละเซกเมนต์แล้ว ยังส่งผลให้เซกเมนต์บริเวณรอยต่อช่วงที่มีการเปลี่ยนหน่วยเสียงทั้งทางต้นเสียงและท้ายเสียงในหน่วยเสียงเดียวกันใดๆมีความแตกต่างกันด้วย เนื่องจากหน่วยเสียงก่อนหน้าและหน่วยเสียงถัดไปสำหรับการเปล่งเสียงในหน่วยเสียงใดๆ เป็นคนละหน่วยเสียงกัน ทำให้ตำแหน่งของอวัยวะที่เคลื่อนเข้ามาเพื่อเปล่งเสียงหน่วยเสียงนั้น และตำแหน่งของอวัยวะที่เคลื่อนที่ออกเพื่อเปล่งเสียงหน่วยเสียงถัดไปเป็นคนละตำแหน่งกัน ดังตัวอย่างในรูปที่ 4.1 ซึ่งเป็นสเปกโตรแกรมแสดงเสียงสระอา จากคำว่า “กาน” (/kaan^/) นั้นจะเห็นว่า เซกเมนต์ของสัญญาณเสียงสระอามีกลุ่มค่าความถี่สั่นพ้อง (Formant) ไม่คงที่สม่ำเสมอกัน โดยในช่วงต้นของสัญญาณเสียงสระอา พบว่าค่าความถี่สั่นพ้อง (Formant frequency) ที่ได้นั้น เกิดจากในขณะที่เปล่งเสียงดังกล่าว ตำแหน่งของ

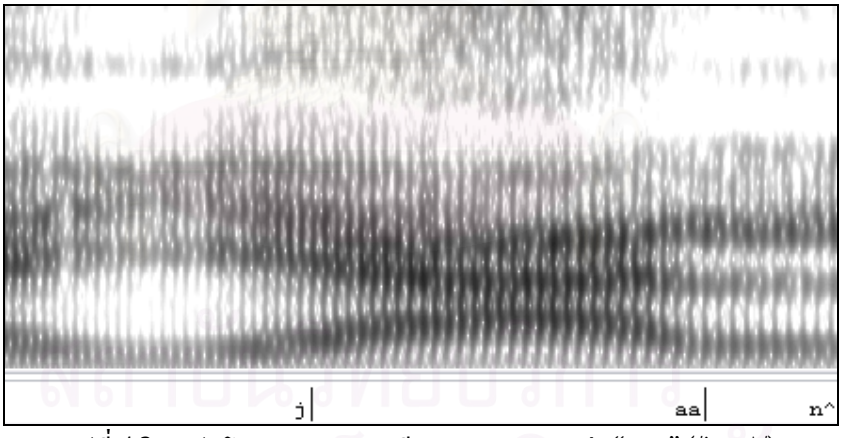
ลิ้น ซึ่งเป็นอวัยวะสำคัญในการเปล่งเสียงสระ กำลังเคลื่อนที่ออกจากตำแหน่งของการเปล่งเสียง“ก” (/k/) ซึ่งเป็นหน่วยเสียงพยัญชนะกักที่มีฐานเสียงอยู่ที่เพดานแข็งไปยังตำแหน่งการเปล่งเสียงสระอา จึงทำให้ค่าความถี่สั่นพ้องที่ได้มีการเลื่อนจากกลุ่มค่าความถี่สั่นพ้องที่เกิดจากอวัยวะการเปล่งเสียงที่อยู่ตำแหน่งการเปล่งเสียง “ก” (/k/) ไปยังกลุ่มค่าความถี่สั่นพ้องของสระอา จนกระทั่งในช่วงกลางของเซกเมนต์ ซึ่งลิ้นได้เคลื่อนที่มายังตำแหน่งของการเปล่งเสียงสระอาแล้ว จึงทำให้ค่าความถี่สั่นพ้องที่เกิดขึ้นในบริเวณดังกล่าวมีค่าคงที่สม่ำเสมอกัน และหลังจากนั้นกลุ่มค่าความถี่สั่นพ้อง จึงเริ่มมีการเปลี่ยนแปลงอีกครั้งหนึ่งในบริเวณท้ายเซกเมนต์ เนื่องจากลิ้นได้เริ่มมีการเคลื่อนที่ไปยังตำแหน่งของการเปล่งเสียงตัวสะกด “น” (/n^/) จึงทำให้สัญญาณเสียงในบริเวณนี้แสดงการเลื่อนที่ของกลุ่มค่าความถี่สั่นพ้องจากกลุ่มค่าความถี่สั่นพ้องของสระอาไปยังกลุ่มค่าความถี่สั่นพ้องของหน่วยเสียงตัวสะกด “น” (/n^/)

และเมื่อทดลองศึกษาเปรียบเทียบสเปกโตรแกรมของหน่วยเสียงสระอากับเซกเมนต์อื่นๆ ซึ่งเกิดจากคำที่แตกต่างกันเฉพาะพยัญชนะต้นเท่านั้น เช่น คำว่า“ยาน” (/jaan^/) และ “จาน” (/caan^/) ดังตัวอย่างในรูปที่ 4.2 และ รูปที่ 4.3 ตามลำดับ พบว่า แนวโน้มของลักษณะสัญญาณของหน่วยเสียงสระอาที่ได้นั้น ยังคงมีแนวโน้มเดิม กล่าวคือ กลุ่มค่าความถี่สั่นพ้องที่เกิดขึ้นภายในเซกเมนต์ ยังมีลักษณะการเลื่อนจากกลุ่มค่าความถี่สั่นพ้องของหน่วยเสียงก่อนหน้ามายังกลุ่มความถี่สั่นพ้องหน่วยเสียงสระอาในบริเวณตอนต้นของเซกเมนต์และมีค่าคงที่ในบริเวณส่วนกลางของเซกเมนต์ ซึ่งเกิดจากอวัยวะในการเปล่งเสียงได้เคลื่อนที่มายังตำแหน่งที่ใช้ในการเปล่งเสียงของเซกเมนต์นั้นๆอย่างสมบูรณ์แล้ว แต่ในส่วนหน้าของเซกเมนต์ทั้งสองมีจุดเริ่มต้นของกลุ่มค่าความถี่สั่นพ้องที่แตกต่างกันตามลักษณะของหน่วยเสียงก่อนหน้า แล้วจึงมีการเลื่อนที่ของกลุ่มความถี่สั่นพ้องอีกในช่วงท้ายของเซกเมนต์ ซึ่งเกิดจากการเคลื่อนที่ของลิ้นไปยังตำแหน่งในการเปล่งเสียงของหน่วยเสียงตัวต่อไป ซึ่งเมื่อเปรียบเทียบหน่วยเสียงสระอาจากเซกเมนต์อื่นๆ ซึ่งเกิดจากคำที่มีตัวสะกดต่างกัน เช่น เปรียบเทียบกับคำว่า “กา” (/kaa/) ที่ตามด้วยหน่วยเสียงในแถบพลังงานความถี่ต่ำ (/vb/) และคำว่า “กา” (/kaa/) ที่ตามด้วยเสียงเงียบ (/sil/) ดังรูปที่ 4.4 และรูปที่ 4.5 ตามลำดับ พบว่ากลุ่มค่าความถี่สั่นพ้องมีแนวโน้มการเปลี่ยนแปลงไปยังกลุ่มค่าความถี่สั่นพ้องของหน่วยเสียงในเซกเมนต์ถัดไป และเมื่อเปรียบเทียบระหว่างหน่วยเสียงที่มีสระที่แตกต่างกัน เช่น ระหว่างคำว่า “จาน” (/caan^/) และคำว่า “จอน” (/c@@n^/) ดังตัวอย่างสเปกโตรแกรมในรูปที่ 4.3 และรูปที่ 4.6 ตามลำดับนั้นพบว่า พฤติกรรมของสัญญาณเสียงที่เกิดขึ้นนั้น ยังคงเป็นไปในลักษณะเดียวกัน กล่าวคือ มีการเลื่อนที่ของกลุ่มค่าความถี่สั่นพ้องในช่วงหน้าและหลังภายในเซกเมนต์และส่วนกลางของเซกเมนต์เป็นส่วนที่มีสัญญาณคงที่ตามกลุ่มค่าความถี่สั่นพ้องของหน่วยเสียงสระนั้นๆ นอกจากนี้ยังสามารถเปรียบเทียบเซกเมนต์เสียงสระอา จากคำที่มีพยัญชนะและตัวสะกดคนละตัวกัน เช่น เปรียบเทียบระหว่างคำว่า “กาน” (/kaan^/) กับคำว่า “ลาย” (/laaj^/) ก็พบว่า มีลักษณะพฤติกรรมการเลื่อนที่ของกลุ่มค่าความถี่สั่นพ้องเช่นเดียวกัน โดยจะเห็นว่าใน

ส่วนกลางของเซกเมนต์จะมีกลุ่มค่าความถี่สั่นพ้องของสระอาเช่นเดียวกัน ทำให้สามารถกล่าวได้ว่าลักษณะการเลื่อนที่ของกลุ่มค่าความถี่สั่นพ้องภายในเซกเมนต์ สามารถเกิดขึ้นได้กับการออกเสียงหน่วยเสียงใดๆที่จำเป็นต้องมีการเคลื่อนที่ของอวัยวะที่ใช้ในการเปล่งเสียงจากตำแหน่งของหน่วยเสียงก่อนหน้ามายังหน่วยเสียงปัจจุบัน และมีการเคลื่อนที่ออกของอวัยวะไปยังตำแหน่งการเปล่งเสียงหน่วยเสียงถัดไป จึงทำให้วิธีการหนึ่งที่สามารถช่วยในการหาลักษณะเฉพาะที่มีการทำงานแบบอาศัยเซกเมนต์ คือ การแบ่งสัญญาณเสียงของเซกเมนต์ดังกล่าวออกเป็น 3 ส่วน เพื่อหาลักษณะเฉพาะในแต่ละส่วน ที่มีพฤติกรรมเดียวกันภายในเซกเมนต์

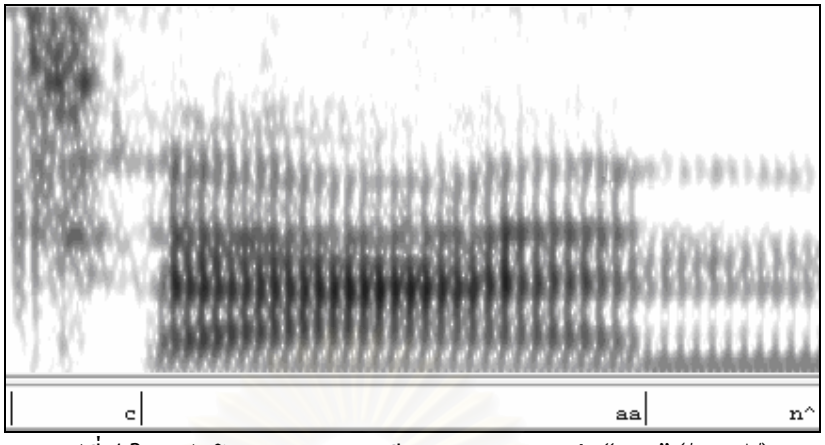


รูปที่ 4.1 สเปกโตรแกรมแสดงเสียงสระอาจากคำว่า “กาน” (/kaan^/)

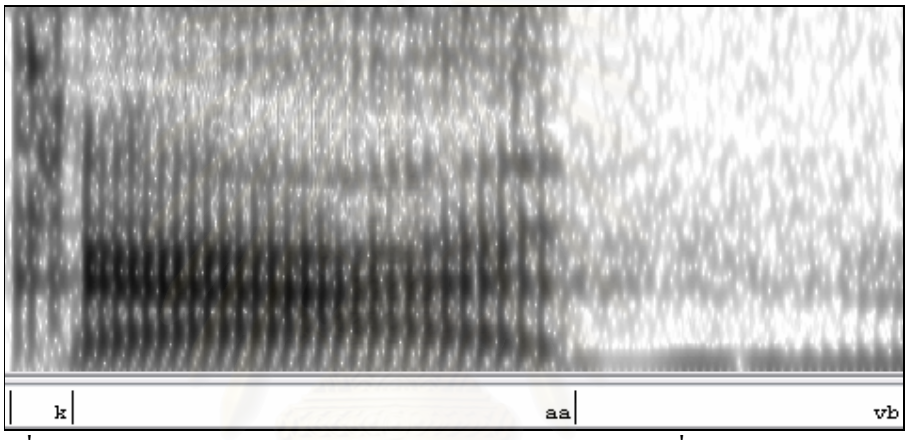


รูปที่ 4.2 สเปกโตรแกรมแสดงเสียงสระอาจากคำว่า “ยาน” (/jaan^/)

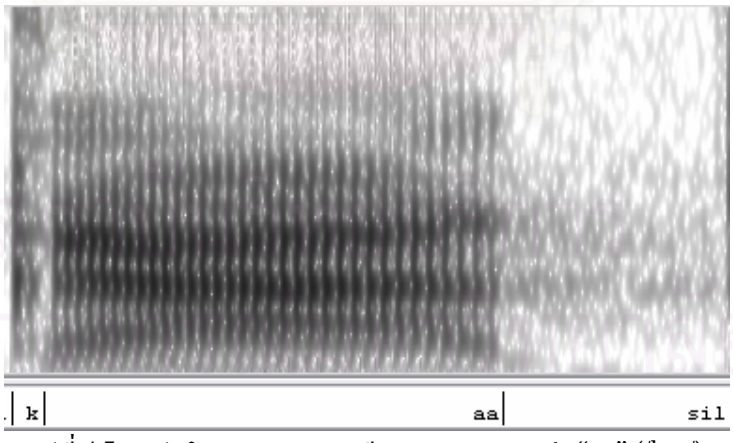
จุฬาลงกรณ์มหาวิทยาลัย



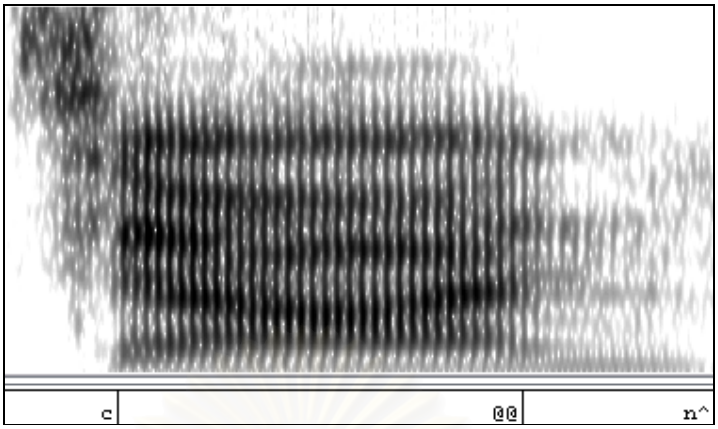
รูปที่ 4.3 สเปกโตรแกรมแสดงเสียงสระอาจากคำว่า “จาน” (/caan^/)



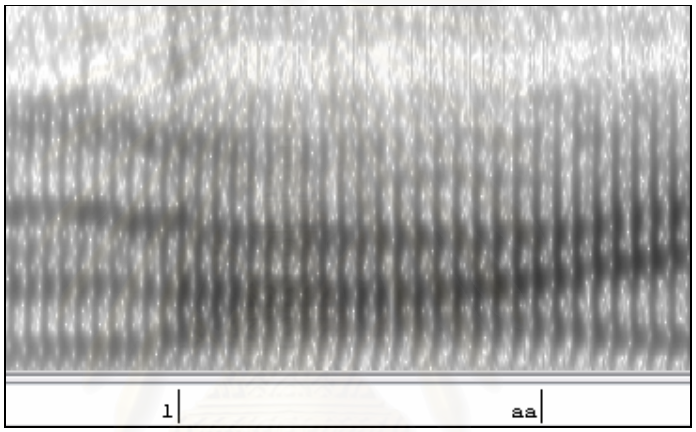
รูปที่ 4.4 สเปกโตรแกรมแสดงเสียงสระอาจากคำว่า “กา” (/kaa/) ที่ตามด้วยหน่วยเสียงในแถบพลังงานความถี่ต่ำ (/vb/)



รูปที่ 4.5 สเปกโตรแกรมแสดงเสียงสระอาจากคำว่า “กา” (/kaa/)



รูปที่ 4.6 สเปกโตรแกรมแสดงเสียงสระอาจากคำว่า “จอน” (/c@@n^/)



รูปที่ 4.7 สเปกโตรแกรมแสดงเสียงสระอาจากคำว่า “ลาย” (/laaj^/)

ตัวอย่างข้างต้น เป็นเพียงการยกตัวอย่างหน่วยเสียงประเภทสระเท่านั้น พฤติกรรมของสัญญาณเสียงที่กล่าวถึงจึงมีลักษณะเป็นกลุ่มค่าความถี่สั่นพ้อง ดังนั้น ในกรณีที่หน่วยเสียงดังกล่าวเป็นหน่วยเสียงประเภทอื่นๆแล้ว ลักษณะพฤติกรรมที่เกิดขึ้นภายในเซกเมนต์ ก็จะมีลักษณะตามหน่วยเสียงประเภทนั้นๆ

แต่อย่างไรก็ตาม พฤติกรรมความไม่สม่ำเสมอกันในเซกเมนต์ดังกล่าวนั้น เป็นเพียงการกล่าวถึงเซกเมนต์ของหน่วยเสียงโดยรวมส่วนใหญ่เท่านั้น ไม่ได้เจาะจงเฉพาะเซกเมนต์ของหน่วยเสียงใดๆ ดังนั้นในบางกรณีที่ผู้พูดมีการออกเสียงชัดเจนก็สามารถทำให้เซกเมนต์ของหน่วยเสียงดังกล่าวมีความสม่ำเสมอกันตลอดทั้งเซกเมนต์ได้

ดังนั้น สำหรับวิธีการแบ่งส่วนเซกเมนต์ใดๆออกเป็น 3 ส่วนนั้น จึงเป็นการศึกษาลักษณะของเซกเมนต์ที่เกิดขึ้นโดยทั่วๆไป เพื่อช่วยในการหาลักษณะเฉพาะแบบอาศัยเซกเมนต์โดยรวม เนื่องจากในการรู้จำเสียงหรือการจำแนกหน่วยเสียงในความเป็นจริงนั้น ไม่สามารถทราบเบื้องต้นได้เลยว่า เซกเมนต์ดังกล่าวจะมีพฤติกรรมสัญญาณเสียงในลักษณะใด ซึ่งถึงแม้ว่าสัญญาณเสียงที่เข้ามาจะมีความสม่ำเสมอกันตลอดทั้งเซกเมนต์นั้น ก็ยังคงสามารถใช้หลักการแบ่งเซกเมนต์

ออกเป็น 3 ส่วนได้โดยไม่ขัดแย้งต่อหลักการ เนื่องจากในแต่ละส่วนของเซกเมนต์ก็ยังคงเป็นส่วนที่มีความสม่ำเสมอกันของสัญญาณเสียงอยู่

นอกจากหลักการและเหตุผลดังกล่าวที่สนับสนุนต่อแนวความคิดในการแบ่งเซกเมนต์ออกเป็น 3 ส่วนแล้ว พบว่างานวิจัยที่ทำการจำแนกหน่วยเสียงซึ่งมีหลักการทำงานแบบอาศัยเซกเมนต์อื่นๆนั้น ได้มีการหาลักษณะเฉพาะจากการแบ่งเซกเมนต์ดังกล่าวออกเป็น 3 ส่วนด้วยเช่นเดียวกัน ได้แก่ งานวิจัยของฮาลเบอแสตด (Halberstadt) และกลาส (Glass) [15]ในปี 1998 งานวิจัยของสตรอม (Strom) เฮ็ดเธอริงตัน (Hetherington) ฮาเซ็น (Hazen) แซนเนส (Sandness) และกลาส (Glass)[16]ในปี 1999 และงานวิจัยของกลาส (Glass)[3] ในปี 2002 ได้หาลักษณะเฉพาะจากการแบ่งเซกเมนต์ออกเป็น 3 ส่วน โดยการกำหนดให้ส่วนหน้าและหลังมีขนาดเท่ากันคือ ขนาด 30 เปอร์เซนต์ ทำให้ส่วนกลางมีขนาดเป็น 40 เปอร์เซนต์ของเซกเมนต์ งานวิจัยของลูง (Leung) เช็ค จิเยอร์ (Chigier) และกลาส(Glass) [11] ในปี 1993 และงานวิจัยของชาง (Chang) และ กลาส (Glass) [17] ในปี 1997 ได้หาลักษณะเฉพาะจากการแบ่งเซกเมนต์ออกเป็น 3 ส่วนเช่นเดียวกัน โดยกำหนดให้ทั้ง 3 ส่วนมีขนาดเท่าๆกัน รวมทั้งงานวิจัยของซาโฮเรียน (Zahorian) ซิลสบี (Silsbee) และหวัง (Wang) [12]ในปี 1997 ก็ได้เลือกใช้สัญญาณเสียงเฉพาะส่วนกลางของเซกเมนต์ในการหาลักษณะเฉพาะเพื่อแทนหน่วยเสียงในเซกเมนต์นั้น

ลักษณะเฉพาะ

จากการศึกษาพฤติกรรมโดยรวมของสัญญาณเสียงข้างต้น ดังนั้นในงานวิจัยนี้จึงได้เลือกใช้วิธีการหาลักษณะเฉพาะของสัญญาณเสียง โดยการแบ่งเซกเมนต์ของหน่วยเสียงใดๆออกเป็น 3 ส่วนย่อย แล้วจึงนำทั้งส่วนย่อยแต่ละส่วนไปหาค่าสัมประสิทธิ์เมลฟรีเคว็นซีเคปสตรอลซึ่งเป็นการดึงสมบัติทางสเปกตรัมของสัญญาณเสียงใดๆออกมา โดยเลือกใช้ค่าสัมประสิทธิ์ขนาด 12 หลัก + ลอการิทึมของพลังงาน 1 หลัก ของทั้งส่วนในแต่ละส่วน รวม 39 หลัก และเพิ่มความยาวของสัญญาณเสียงแต่ละหน่วยเสียง (Duration) ซึ่งถือว่าเป็นสมบัติที่ได้จากทำงานแบบอาศัยเซกเมนต์อีก 1 หลัก รวมทั้งสิ้น 40 หลัก

สำหรับการแบ่งเซกเมนต์ใดๆออกเป็น 3 ส่วนนั้น งานวิจัยนี้ได้ทำการแบ่งส่วนออกเป็นหลายสัดส่วนด้วยกัน โดยพิจารณาการแบ่งสัดส่วนออกเป็นในการแบ่งออกเป็น 3 ลักษณะ ได้แก่

ก. กำหนดให้ส่วนหน้าและหลังเท่ากันทั้งหมด	7	สัดส่วน คือ	ก1 - ก7
ข. กำหนดให้ส่วนหน้าเล็กกว่าส่วนหลังทั้งหมด	15	สัดส่วน คือ	ข1 - ข15
ค. กำหนดให้ส่วนหน้าใหญ่กว่าส่วนหลังทั้งหมด	15	สัดส่วน คือ	ค1 - ค15

รวมลักษณะเฉพาะทั้งสิ้น 37 แบบโดยสัดส่วนในการแบ่งของแต่ละแบบแสดงในตารางที่ 4.1 ดังนี้

ตารางที่4.1 ตารางแสดงสัดส่วนในการแบ่งเซกเมนต์ของหน่วยเสียง 1 เซกเมนต์

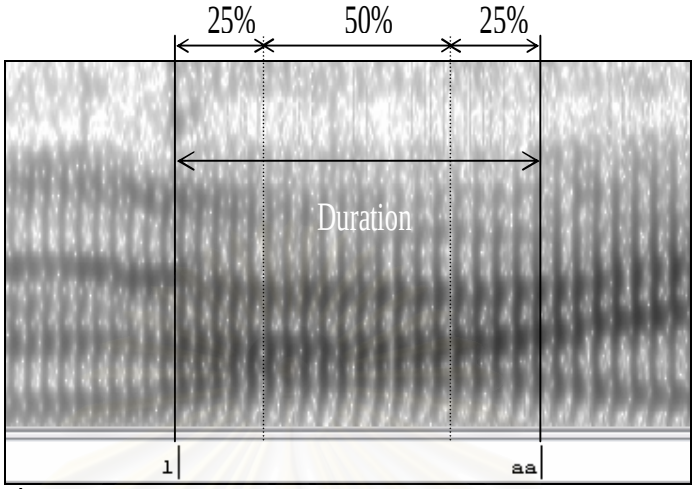
ลักษณะเฉพาะแบบ	สัดส่วนการแบ่งเซกเมนต์ หน้า:กลาง:หลัง(เปอร์เซนต์)
ก1	10: 80: 10
ก2	15: 70: 15
ก3	20: 60: 20
ก4	25: 50: 25
ก5	30: 40: 30
ก6	เท่ากัน
ก7	35: 30: 35
ข1	10: 70: 20
ข2	10: 60: 30
ข3	10: 50: 40
ข4	10: 40: 50
ข5	10: 30: 60
ข6	15: 60: 25
ข7	15: 50: 35
ข8	15: 40: 45
ข9	15: 30: 55
ข10	20: 50: 30
ข11	20: 40: 40
ข12	20: 30: 50

ลักษณะเฉพาะแบบ	สัดส่วนการแบ่งเซกเมนต์ หน้า:กลาง:หลัง(เปอร์เซนต์)
ข13	25: 40: 35
ข14	25: 30: 45
ข15	30: 30: 40
ค1	20: 70: 10
ค2	25: 60: 15
ค3	30: 60: 10
ค4	30: 50: 20
ค5	35: 50: 15
ค6	35: 40: 25
ค7	40: 50: 10
ค8	40: 40: 20
ค9	40: 30: 30
ค10	45: 40: 15
ค11	45: 30: 25
ค12	50: 40: 10
ค13	50: 30: 20
ค14	55: 30: 15
ค15	60: 30: 10

ซึ่งสัดส่วนเปอร์เซนต์ดังกล่าวจะเทียบกับสัดส่วนของเซกเมนต์หน่วยเสียง1 เซกเมนต์ โดยหลักการกำหนดสัดส่วนดังกล่าว มีวิธีการกำหนดดังนี้

1. กำหนดให้ทั้งส่วนหน้าและหลังมีขนาดไม่ต่ำกว่า 10 เปอร์เซนต์ของเซกเมนต์
2. กำหนดให้ส่วนกลางมีขนาดไม่ต่ำกว่า 30 เปอร์เซนต์ของเซกเมนต์
3. ในการขยับสัดส่วนให้ส่วนหน้าและส่วนหลังไม่เท่ากันนั้น จะอ้างอิงจากการแบ่งสัดส่วนในกลุ่ม ก กล่าวคือ กำหนดให้ส่วนกลางมีขนาด 70, 60, 50, 40, 30 เปอร์เซนต์ โดยจะขยับส่วนข้างของแต่ละกลุ่มครั้งละ 5 เปอร์เซนต์

โดยตัวอย่างการแบ่งสัดส่วนเซกเมนต์แสดงดังรูปที่ 4.8 ซึ่งเป็นสปกแกรมหน่วยเสียงสระอาที่มีการแบ่งสัดส่วนของลักษณะเฉพาะตามแบบ ก4



รูปที่ 4.8 การแบ่งสัดส่วนเซกเมนต์ในการหาลักษณะเฉพาะในแบบ ก4

ซึ่งการเลือกใช้สัดส่วนการแบ่งเซกเมนต์ออกเป็น 3 ส่วนด้วยสัดส่วนต่างๆ เป็นเพียงการทดลองสังเกตถึงค่าความแตกต่างที่เกิดขึ้นต่อการแบ่งเซกเมนต์ เพื่อเลือกสัดส่วนการแบ่งเซกเมนต์วิธีหนึ่งสำหรับใช้ในขั้นตอนการหาค่าลักษณะเฉพาะของหน่วยเสียงในแต่ละเซกเมนต์ เพื่อทำการจำแนกหน่วยเสียงเท่านั้น ไม่ได้มีจุดประสงค์ในการนำเสนอสัดส่วนที่เหมาะสมที่สุดสำหรับการแบ่งเซกเมนต์ในการหาลักษณะเฉพาะ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 5
การทดลองจำแนกหน่วยเสียง

ในบทนี้จะทำการทดลองจำแนกหน่วยเสียงภาษาไทย ซึ่งแบ่งออกเป็น 5 ส่วนหลักๆ ด้วยกัน โดยส่วนแรกจะกล่าวถึงฐานข้อมูลเสียงที่นำมาใช้ในการทดลอง ส่วนที่สองคือ ส่วนการทดลองจำแนกหน่วยเสียงด้วยหลักการการวิเคราะห์ดิสคริมิแนนต์เชิงเส้นโดยใช้ลักษณะเฉพาะแบบต่างๆที่มีวิธีการทำงานแบบอาศัยเซกเมนต์ ส่วนที่สาม คือ การวิเคราะห์ค่าความมีส่วนร่วมของลักษณะเฉพาะที่เลือกใช้ ส่วนที่สี่ คือ การจำแนกหน่วยเสียงด้วยวิธีการแบบอาศัยกรอบเวลา และส่วนสุดท้าย คือ การวิเคราะห์เปรียบเทียบผลการจำแนกระหว่างวิธีการทั้ง 2 วิธีนี้

ฐานข้อมูลเสียง

งานวิจัยนี้เลือกใช้ฐานข้อมูลเสียงโลตัส (Large vOcabulary Thai continUous Speech recognition Corpus: LOTUS Corpus) [18] พัฒนาโดยเนคเทค ซึ่งถือว่าเป็นฐานข้อมูลเสียงภาษาไทยขนาดใหญ่ เพื่อใช้สำหรับการพัฒนาระบบเสียงพูดต่อเนื่อง (Large Vocabulary Continuous Speech Recognition: LVCSR) โดยเลือกใช้ข้อมูลในชุดหน่วยเสียงสมดุล ซึ่งได้มีการกำกับหน่วย (Label) เสียงอัตโนมัติ (Automatic Phoneme Labeler) บันทึกด้วยไมโครโฟนแบบโคลสทอล์ค (Close-talk) คุณภาพสูง ซึ่งบันทึกเสียงสภาพแวดล้อมแบบห้องเงียบ โดยมีผู้พูดเพศชาย 20 คนและเพศหญิง 20 คนโดยแต่ละคนจะบันทึกเสียงเป็นประโยคต่อเนื่องคนละ 35 ประโยค รวมจำนวนหน่วยเสียงทั้งหมด 70,670 หน่วยเสียง โดยแบ่งเป็นชุดข้อมูลฝึกฝนและชุดข้อมูลทดสอบ ซึ่งชุดข้อมูลฝึกฝนจะทำหน้าที่ในการฝึกฝน เพื่อสร้างแบบจำลองเสียงวิธีการต่างๆตามการทดลองต่างๆ รวมไปถึงเป็นชุดข้อมูลที่นำไปใช้ทดสอบ เพื่อใช้ในการปรับปรุงวิธีการทำงานด้วย ส่วนชุดทดสอบจะนำมาใช้ทดสอบการจำแนกหน่วยเสียง เพื่อใช้วัดประสิทธิภาพการทำงานที่ได้จริงๆ โดยแต่ละชุดจะประกอบด้วยหน่วยเสียงของผู้พูดเพศชาย 10 คน และเพศหญิง 10 คน โดยชุดฝึกฝนมีจำนวนหน่วยเสียงทั้งหมด 35,331 หน่วยเสียง ชุดทดสอบมีจำนวน 35,339 หน่วยเสียง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่5.1 จำนวนหน่วยเสียงที่ใช้ในชุดฝึกฝนและชุดทดสอบ

กลุ่มหน่วยเสียง	จำนวนหน่วยเสียง	
	ชุดฝึกฝน	ชุดทดสอบ
a	2933	2862
i	491	506
e	318	342
o	497	494
u	263	259
v	255	228
q	13	16
@	208	206
x	240	216
aa	2351	2319
ii	680	693
ee	203	181
oo	242	208

กลุ่มหน่วยเสียง	จำนวนหน่วยเสียง	
	ชุดฝึกฝน	ชุดทดสอบ
uu	331	297
vv	187	198
qq	180	178
@@	741	749
xx	369	346
iia	149	180
vva	253	258
uua	364	305
f	104	104
s	1030	1016
p	511	519
t	635	592
k	919	951

กลุ่มหน่วยเสียง	จำนวนหน่วยเสียง	
	ชุดฝึกฝน	ชุดทดสอบ
b	73	76
d	204	187
ph	649	620
th	883	862
kh	973	977
c	489	481
ch	322	332
m	607	604
n	826	764
ng	93	107
vb	786	785
m^	689	683
n^	1658	1621

กลุ่มหน่วยเสียง	จำนวนหน่วยเสียง	
	ชุดฝึกฝน	ชุดทดสอบ
ng^	1268	1257
j^	1169	1134
w^	314	295
l^	4	6
f^	18	24
s^	24	22
ch^	2	5
l	831	807
r	1138	1185
h	257	253
w	418	436
j	324	262
sil	6845	7331
รวม	**35331**	**35339**

การทดลองจำแนกหน่วยเสียงด้วยหลักการการวิเคราะห์ดิสคริมิแนนต์เชิงเส้นโดยใช้ลักษณะเฉพาะแบบต่างๆที่มีวิธีการทำงานแบบอาศัยเซกเมนต์

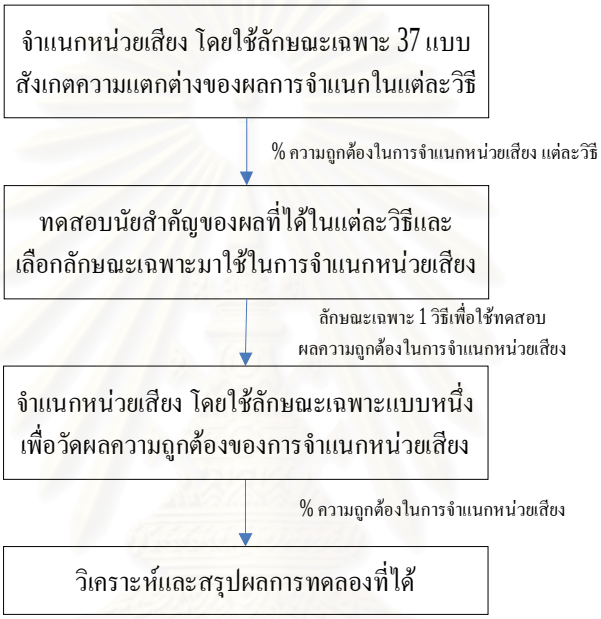
1. **ขั้นตอนการทดลอง**

สำหรับขั้นตอนการทดลองนั้น จะเริ่มจากการนำลักษณะเฉพาะแต่ละวิธีที่เสนอไว้ในบทที่ 4 การหาลักษณะเฉพาะ ในหัวข้อ ลักษณะเฉพาะ มาทดสอบความสามารถในการแทนสัญญาณเสียงต่างๆ เพื่อพิจารณาว่าลักษณะเฉพาะที่มีการแบ่งสัดส่วนที่แตกต่างกันในแต่ละแบบนั้น มีความสำคัญหรือมีความแตกต่างกันหรือไม่ ด้วยวิธีการจำแนกหน่วยเสียงโดยใช้หลักการการวิเคราะห์ดิสคริมิแนนต์เชิงเส้น ซึ่งในขั้นตอนนี้จะใช้เฉพาะชุดข้อมูลฝึกฝนมาแบ่งข้อมูลออกเป็น 2 กลุ่ม คือ กลุ่มข้อมูลที่ใช้ในการฝึกฝนจำนวน 24,735 หน่วยเสียง ประกอบด้วยหน่วยเสียงประเภทต่างๆ ประเภทละประมาณ 70 เปอร์เซนต์ของจำนวนหน่วยเสียงในแต่ละประเภทของชุดข้อมูลฝึกฝน และกลุ่มข้อมูลที่ใช้ในการพัฒนาจำนวน 10,596 หน่วยเสียง เนื่องจากเป็นการทดสอบความสามารถในการแทนสัญญาณเสียงต่างๆ ในการใช้ลักษณะเฉพาะที่มีการแบ่งสัดส่วนที่แตกต่างกัน จึงถือว่ายังไม่ใช่การทดสอบประสิทธิภาพความถูกต้องในการจำแนกหน่วยเสียง หลังจากนั้นจะนำผลการจำแนกหน่วยเสียงจากการใช้ลักษณะเฉพาะแบบต่างๆไปผ่านกระบวนการทดสอบนัยสำคัญที่ได้นำเสนอไว้ในบทที่ 2 ทฤษฎีที่เกี่ยวข้อง ในหัวข้อ การทดสอบนัยสำคัญ (Significance testing) เพื่อวัดค่าความแตกต่างที่เกิดขึ้นระหว่าง 2 วิธีการใดๆนั้นว่ามีนัยสำคัญหรือไม่ ซึ่งการทดลองทั้งหมดข้างต้นนั้น เป็นเพียงการทดลอง เพื่อทดสอบว่าสัดส่วนที่แตกต่างกันจะให้ผลการจำแนกหน่วยเสียงที่แตกต่างกันหรือไม่ ไม่ได้มีจุดประสงค์ในการหาสัดส่วนที่เหมาะสมที่สุดในการแบ่งเซกเมนต์ออกเป็น 3 ส่วน

หลังจากทำการทดลองข้างต้นแล้ว การทดลองต่อไปจะเป็นการทดสอบประสิทธิภาพความถูกต้องในการจำแนกหน่วยเสียงของลักษณะเฉพาะจริงๆ ซึ่งจะทำการจำแนกหน่วยเสียง โดยเลือกใช้ชุดข้อมูลฝึกฝนทั้งชุดในการฝึกฝน และใช้ชุดข้อมูลทดสอบทั้งชุดในการวัด % *ความถูกต้อง* ของการจำแนกหน่วยเสียงที่ได้จากการใช้ลักษณะเฉพาะตามงานวิจัยนี้

โดยลักษณะวิธีการวัด % *ความถูกต้อง*ในแต่ละครั้งนั้น จะวัดจากการนับจำนวนหน่วยเสียงทั้งหมดที่สามารถจำแนกได้ถูกต้องหารด้วยจำนวนหน่วยเสียงทั้งหมดที่นำไปจำแนก

ขั้นตอนการทดลองดังกล่าว สามารถเขียนอธิบายด้วยแผนภาพ ดังรูป



รูปที่ 5.1 ขั้นตอนการทดลองด้วยหลักการการวิเคราะห์ดิสคริมิแนนต์เชิงเส้น โดยใช้ลักษณะเฉพาะที่มีการทำงานแบบอาศัยเซกเมนต์

2. **ผลการทดลอง**

2.1 **ผลการจำแนกหน่วยเสียงภาษาไทยด้วยลักษณะเฉพาะสัดส่วนต่างๆบนข้อมูลฝึกฝน**

การทดลองแรกเป็นการทดลองจำแนกหน่วยเสียง โดยใช้ลักษณะเฉพาะสัดส่วนต่างๆ ตามที่ได้เสนอไว้ในบทที่ 4 การหาลักษณะเฉพาะ มาแทนสัญญาณเสียง โดยทำการจำแนกหน่วยเสียงภาษาไทยทั้งหมด 52 หน่วยเสียง ซึ่งผลการทดลองในการจำแนกหน่วยเสียง โดยใช้ลักษณะเฉพาะที่มีอัตราส่วนแบบต่างๆทั้ง 6 แบบนั้น ได้ % *ความถูกต้อง* ในการจำแนกหน่วยเสียง ดังตารางที่5.2 นี้

ตารางที่5.2 ผลการจำแนกหน่วยเสียงด้วยลักษณะเฉพาะแบบต่างๆ

ลักษณะเฉพาะแบบ	% *ความถูกต้อง* ในการจำแนกหน่วยเสียง	ลักษณะเฉพาะแบบ	% *ความถูกต้อง* ในการจำแนกหน่วยเสียง
ก1	58.91	ข13	58.64
ก2	59.89	ข14	57.45
ก3	60.00	ข15	58.23
ก4	59.23	ค1	58.82
ก5	58.68	ค2	58.48
ก6	58.19	ค3	58.30
ก7	58.54	ค4	59.60
ข1	59.11	ค5	58.64
ข2	59.38	ค6	58.59
ข3	59.01	ค7	58.32
ข4	58.14	ค8	58.56
ข5	55.96	ค9	58.25
ข6	59.21	ค10	58.99
ข7	58.74	ค11	58.80
ข8	58.14	ค12	58.26
ข9	57.65	ค13	58.75
ข10	59.54	ค14	58.70
ข11	58.73	ค15	58.45
ข12	57.85		

จากผลการทดลองข้างต้นแสดงให้เห็นว่าผล % *ความถูกต้อง* ที่ได้จากการจำแนกหน่วยเสียงโดยรวมมีค่า% *ความถูกต้อง* อยู่ในช่วง 58-60 เปอร์เซนต์ โดยสัดส่วนของลักษณะเฉพาะที่มีค่า% *ความถูกต้อง* สูงที่สุดคือ ลักษณะเฉพาะที่มีการแบ่งส่วนหน้าและหลังเท่ากันแบบ ก3 ซึ่งมีการแบ่งสัดส่วนด้วยขนาด 20:60:20 เปอร์เซนต์ ส่วนสัดส่วนที่ให้ค่า% *ความถูกต้อง* ต่ำสุด คือ ลักษณะเฉพาะแบบ ข5 ที่มีการแบ่งสัดส่วน 10: 30: 60 เปอร์เซนต์ โดยลักษณะเฉพาะในกลุ่ม ก ซึ่งเป็นกลุ่มที่มีการแบ่งส่วนหน้าและหลังเท่ากันนั้น พบว่า สัดส่วนที่ค่า % *ความถูกต้อง*ในการจำแนกหน่วยเสียงเกิน 59.00% อยู่ถึง 3 แบบจาก 7 แบบ ได้แก่ ลักษณะเฉพาะแบบ ก2 ก3 และ ก4 ในขณะที่ลักษณะในกลุ่ม ข ซึ่งมีขนาดส่วนหน้าเล็กกว่าส่วนหลัง พบว่า สัดส่วนที่ค่า % *ความถูกต้อง*ใน

การจำแนกหน่วยเสียงเกิน 59.00% อยู่ 5 แบบ ได้แก่ ลักษณะเฉพาะแบบ ข1 ข2 ข3 ข6 และ ข10 ส่วนลักษณะเฉพาะในกลุ่ม ค ที่มีการกำหนดให้ส่วนหน้าใหญ่กว่าส่วนหลัง พบว่า สัดส่วนที่ค่า % *ความถูกต้อง*ในการจำแนกหน่วยเสียงเกิน 59.00% มีอยู่เพียงแบบเดียวเท่านั้น คือ ลักษณะเฉพาะแบบ ค4

2.2 การทดสอบนัยสำคัญของลักษณะเฉพาะที่มีสัดส่วนต่างกัน

ในขั้นตอนนี้ จะนำผลการทดลองในการจำแนกหน่วยเสียงโดยใช้ลักษณะเฉพาะสัดส่วนต่างๆข้างต้นนั้น มาวิเคราะห์ค่าความแตกต่างระหว่าง 2 วิธีใดๆว่ามีนัยสำคัญต่อกันหรือไม่ โดยจะทำการทดสอบความแตกต่างที่เกิดขึ้นระหว่างวิธีการใช้ลักษณะเฉพาะแบบ ก3 ซึ่งให้ผลการจำแนกสูงสุด กับผลการใช้ลักษณะเฉพาะแบบอื่นๆ ว่ามีนัยสำคัญหรือไม่ ด้วยวิธีการทดสอบนัยสำคัญ ที่มีวิธีการทำงานที่ได้กล่าวไว้ในบทที่ 2ทฤษฎีและงานวิจัยที่เกี่ยวข้องโดยกำหนดค่า α = 0.005 ซึ่งผลการทดสอบนัยสำคัญทั้งหมดแสดงในตารางที่5.3 โดยกำหนดให้

- $\boldsymbol{n}_{00}$ = จำนวนข้อมูลที่สามารถจำแนกได้ถูกต้องทั้ง 2 วิธี
- $\boldsymbol{n}_{01}$ = จำนวนข้อมูลที่วิธีที่ 1 สามารถจำแนกถูกเท่านั้น
- $\boldsymbol{n}_{10}$ = จำนวนข้อมูลที่วิธีที่ 2 สามารถจำแนกถูกเท่านั้น
- $\boldsymbol{n}_{11}$ = จำนวนข้อมูลที่จำแนกผิดทั้ง 2 วิธี

ตารางที่5.3 ผลการทดสอบนัยสำคัญระหว่างลักษณะเฉพาะแบบ ก3 กับลักษณะเฉพาะแบบอื่นๆ

วิธีที่ 2	$\boldsymbol{n}_{00}$	$\boldsymbol{n}_{01}$	$\boldsymbol{n}_{10}$	$\boldsymbol{n}_{11}$	ค่า P ที่ได้
ก1	5729	629	513	3725	0.00066005
ก2*	5825	533	521	3717	0.73476
ก4*	5809	549	467	3771	0.011011
ก5	5656	702	562	3676	9.1014×10^{-5}
ก6	5569	789	597	3641	2.7727×10^{-7}
ก7	5522	836	653	3585	2.3333×10^{-7}
ข1	5780	578	483	3755	0.0038833
ข2*	5646	712	646	3592	0.077717
ข3*	5531	827	722	3516	0.0082100
ข4	5411	947	749	3489	1.6782×10^{-5}
ข5	5188	1170	742	3496	1.0592×10^{-22}
ข6*	5775	583	499	3739	0.011592
ข7	5639	719	585	3653	0.00022775

ตารางที่ 5.3 ผลการทดสอบนัยสำคัญระหว่างลักษณะเฉพาะแบบ ก3 กับลักษณะเฉพาะแบบอื่นๆ

วิธีที่ 2	n_{00}	n_{01}	n_{10}	n_{11}	ค่า P ที่ได้
ข8	5483	875	678	3560	6.3596×10^{-7}
ข9	5375	983	734	3504	2.0320×10^{-9}
ข10*	5893	465	416	3822	0.10579
ข11	5649	709	549	3689	7.1699×10^{-6}
ข12	5431	927	657	3581	2.8450×10^{-9}
ข13	5615	743	598	3640	8.2899×10^{-5}
ข14	5431	927	657	3581	1.2446×10^{-11}
ข15	5511	847	619	3619	1.4025×10^{-6}
ค1	5707	651	473	3765	1.2262×10^{-7}
ค2	5724	634	473	3765	1.4575×10^{-6}
ค3	5549	809	628	3610	1.99103×10^{-6}
ค4*	5918	440	397	3841	0.14653
ค5	5640	718	573	3665	6.0295×10^{-5}
ค6	5637	721	571	3667	3.3300×10^{-5}
ค7	5491	867	689	3549	7.06410×10^{-6}
ค8	5696	662	509	3729	8.6746×10^{-6}
ค9	5522	836	650	3588	15468×10^{-6}
ค10*	5574	784	677	3561	0.0055327
ค11	5540	818	690	3548	0.0010672
ค12	5423	935	750	3488	7.2326×10^{-6}
ค13	5594	764	631	3607	0.00040531
ค14	5475	883	745	3493	0.00068066
ค15	5397	961	796	3442	9.0331×10^{-6}

* วิธีที่ผลที่ได้ไม่แตกต่างอย่างมีนัยสำคัญกับผลที่ได้จากวิธี ก3

จากผลการทดสอบนัยสำคัญกับระหว่างลักษณะแบบ ก3 กับลักษณะเฉพาะแบบอื่นๆอีก 36 แบบ พบว่ามีลักษณะเฉพาะทั้งหมด 8 แบบ ได้แก่ ลักษณะเฉพาะแบบ ก2 ก3 ก4 ข2 ข3 ข6 ข10 ค4 และ ค10 ที่คำนวณค่า P ได้สูงกว่าค่า α ที่ได้กำหนดไว้ แสดงว่าความแตกต่างที่เกิดขึ้นระหว่างการใช้ลักษณะเฉพาะแบบ ก3 กับการใช้ลักษณะเฉพาะแบบต่างๆทั้ง 8 แบบดังกล่าวไม่มีนัยสำคัญ ซึ่งสังเกตได้ว่า ลักษณะเฉพาะทั้ง 8 แบบนั้น มีค่า % *ความถูกต้อง*ในการจำแนกหน่วยเสียงสูงใกล้เคียง

กับลักษณะเฉพาะแบบ ก3 โดยมี % *ความถูกต้อง*ในการจำแนกหน่วยเสียงอยู่ที่ประมาณ 59 เปอร์เซนต์

ดังนั้น สามารถกล่าวได้ว่า สัดส่วนที่สามารถนำไปใช้ในการแบ่งส่วนเซกเมนต์ที่สามารถให้ผล % *ความถูกต้อง*ในการจำแนกหน่วยเสียงที่ความแตกต่างไม่มีนัยสำคัญกับการจำแนกหน่วยเสียงที่ใช้ลักษณะเฉพาะแบบ ก3 ที่มีสัดส่วน 20: 60:20 ได้แก่ ลักษณะเฉพาะที่มีการแบ่งสัดส่วน 15: 70: 15, 25: 50: 25, 10: 60: 30, 10: 50: 40, 15: 60: 25, 20: 50: 30, 30: 50: 20 และ 45: 40: 15 ดังนั้น การเลือกใช้ลักษณะเฉพาะแบบใดๆภายใน 9 แบบนี้ในการแทนสัญญาณเสียงแต่ละหน่วยเสียงย่อมไม่มีความแตกต่างกัน ดังนั้นในงานวิจัยนี้จะเลือกใช้ลักษณะเฉพาะแบบ ก3 ซึ่งแบ่งสัญญาณเสียงออกเป็น 3 ส่วนด้วยสัดส่วน 20: 60: 20 เปอร์เซนต์ของเซกเมนต์หน่วยเสียงในการแทนสัญญาณเสียงแต่ละหน่วยเสียง เพื่อใช้ในการทดลองจำแนกหน่วยเสียงเพื่อวัดประสิทธิภาพความถูกต้องในการจำแนกหน่วยเสียงที่มีการทำงานแบบอาศัยเซกเมนต์ ซึ่งการเลือกใช้สัดส่วนดังกล่าว ไม่ได้เป็นการนำเสนอว่าเป็นสัดส่วนที่ดีที่สุดในการแบ่งเซกเมนต์ แต่ถือว่าเป็นเพียงวิธีการหนึ่งที่สามารถนำมาใช้ในการแบ่งส่วนภายในเซกเมนต์ของหน่วยเสียง เพื่อนำไปหาค่าลักษณะเฉพาะที่มีการทำงานแบบอาศัยเซกเมนต์ที่สามารถให้ผลการจำแนกหน่วยเสียงที่ดีวิธีหนึ่ง

2.3 **การทดลองจำแนกหน่วยเสียงด้วยชุดข้อมูลทดสอบเพื่อวัดประสิทธิภาพความถูกต้องในการจำแนกหน่วยเสียงของลักษณะเฉพาะที่นำเสนอ**

สำหรับการทดลองนี้จะทำการจำแนกหน่วยเสียงภาษาไทย 52 หน่วยเสียงโดยใช้ชุดข้อมูลฝึกฝนทั้งชุดในการฝึกฝน และใช้ข้อมูลทดสอบในการทดสอบจำแนกหน่วยเสียงจริงๆ เพื่อวัดประสิทธิภาพความถูกต้องที่ได้จากการใช้ลักษณะเฉพาะในงานวิจัยนี้ โดยการใช้ลักษณะเฉพาะแบบ ก3 ที่มีสัดส่วน 20: 60: 20 มาแทนสัญญาณเสียงในชุดข้อมูลฝึกฝนและชุดข้อมูลทดสอบ เพื่อวัดประสิทธิภาพความถูกต้องในการจำแนกหน่วยเสียงที่มีการทำงานแบบอาศัยเซกเมนต์

การทดลองนี้ได้แบ่งวิธีการจำแนกหน่วยเสียงออกเป็น 2 วิธี ได้แก่

- การจำแนกหน่วยเสียงโดยการกำหนดให้ค่าน้ำหนักหรือความน่าจะเป็น ในการที่หน่วยเสียงหนึ่งๆจะมีโอกาสตอบว่าเป็นหน่วยเสียงประเภทใดมีค่าเท่าๆกันหรือ

$$\boldsymbol{prior}_i = \frac{1}{\boldsymbol{n}} \qquad (5.1)$$

โดยที่ $\boldsymbol{n}$ คือ จำนวนหน่วยเสียงที่ต้องการจำแนก ในที่นี้ $\boldsymbol{n} = 52$

- การจำแนกหน่วยเสียงโดยการกำหนดให้ค่าน้ำหนักหรือความน่าจะเป็น ของโอกาสที่จะตอบว่าเป็นหน่วยเสียงประเภทใดไม่เท่ากัน ซึ่งจะกำหนดค่าความน่าจะเป็นดังกล่าวขึ้นกับจำนวนหน่วยเสียงที่ใช้ในการฝึกฝนของแต่ละกลุ่มหน่วยเสียงเทียบกับจำนวนหน่วยเสียงทั้งหมดทุกกลุ่มที่ใช้ในชุดฝึกฝน เขียนเป็นสมการได้ ดังนี้

$$\boldsymbol{prior_i} = \frac{\boldsymbol{w_i}}{\sum_{i=1}^{n} \boldsymbol{w_i}} \qquad (5.2)$$

โดยที่ $\boldsymbol{w_i}$ คือ จำนวนข้อมูลในชุดข้อมูลฝึกฝน ของหน่วยเสียงตัวที่ i

ซึ่งผล% *ความถูกต้อง*เฉลี่ยของการจำแนกหน่วยเสียงทั้ง 2 วิธี ได้ผล ดังนี้

ตารางที่5.4 ผลการจำแนกหน่วยเสียงที่ใช้ลักษณะเฉพาะแบบ ก3 เพื่อวัดประสิทธิภาพความถูกต้อง

ประเภทของค่าน้ำหนัก	*% ความถูกต้อง* **ในการจำแนกหน่วยเสียง**
เท่ากัน	61.14
ขึ้นกับจำนวนข้อมูล	66.41

จากผลการจำแนกหน่วยเสียงข้างต้น พบว่า *% ความถูกต้อง* ในการจำแนกหน่วยเสียงที่ได้สูงสุดคือ การใช้ลักษณะเฉพาะแบบ ก3 สัดส่วน 20: 60: 20 ทั้งการกำหนดให้ค่าน้ำหนักในการตอบหน่วยเสียงใดๆเท่าๆกัน และการตอบหน่วยเสียงใดๆขึ้นกับจำนวนข้อมูลที่ใช้ในการฝึกฝนของหน่วยเสียงนั้นๆ โดยมีค่า *% ความถูกต้อง* อยู่ที่ 61.14 เปอร์เซนต์และ 66.41 เปอร์เซนต์ ตามลำดับ โดยผลการจำแนกหน่วยเสียงโดยแจกแจงผลตามความสับสนระหว่างหน่วยเสียงต่างๆของลักษณะเฉพาะแบบ ก3 แสดงในตารางที่5.5 และตารางที่5.6 ซึ่งเป็นผลการจำแนกหน่วยเสียงโดยใช้ค่าน้ำหนักเท่ากันและใช้ค่าน้ำหนักตามจำนวนข้อมูลของแต่ละกลุ่มตามลำดับ ซึ่งแสดงข้อมูลของผลการจำแนกของหน่วยเสียงแต่ละหน่วยเสียงในแนวนอน กล่าวคือในแต่ละแถวจะแสดงผลการจำแนกหน่วยเสียงแต่ละหน่วยว่ามีความสับสนกับหน่วยเสียงใดๆเท่าไร มีจำนวนทั้งหมด 52 แถว ตามจำนวนกลุ่มหน่วยเสียงที่ต้องการจำแนก

โดยผลการทดลองจำแนกหน่วยเสียงที่ให้น้ำหนักของแต่ละกลุ่มเท่ากัน พบว่า การจำแนกหน่วยเสียงประเภทสระมักมีความสับสนอยู่ระหว่างภายในกลุ่มสระและกึ่งสระสูงกว่าหน่วยเสียงประเภทพยัญชนะ ซึ่งเป็นไปตามลักษณะของหน่วยเสียงที่มีความใกล้เคียงกันมากกว่าประเภทพยัญชนะ โดยเฉพาะความสับสนระหว่างสระเสียงสั้นและสระเสียงยาวที่มีสูงกว่ากลุ่มอื่นๆ ซึ่งจากผลการทดลองสามารถสังเกตได้ว่า สระเสียงยาวของทุกๆเสียงนั้นสามารถจำแนกได้ถูกต้องกว่าสระเสียงสั้นในประเภทเดียวกัน โดยผลการจำแนกที่ดีที่สุดในกลุ่มสระเสียงยาวคือ เสียงสระอี(/ii/) จำแนกได้ถูกต้อง 73.74 เปอร์เซนต์ และรองลงมาคือสระออ (/vv/) ได้ 71.72 เปอร์เซนต์ สำหรับสระเสียงสั้น สระแอะ (/x/) สามารถจำแนกได้ถูกต้องสูงสุด 51.39 เปอร์เซนต์ ส่วนสระอิ (/i/) ซึ่งเป็นสระคู่เสียงกับสระอี(/ii/) นั้น สามารถจำแนกได้ถูกต้องเป็นอันดับรองลงมาคือ 48.42 เปอร์เซนต์ ส่วนหน่วยเสียงที่สามารถจำแนกได้ถูกต้องน้อยที่สุดคือ สระเออะ (/q/)ได้เพียง 25.00 เปอร์เซนต์ เท่านั้น ซึ่งอาจมีสาเหตุจากชุดข้อมูลที่ใช้ฝึกฝนน้อยมาก เมื่อเปรียบเทียบกับหน่วยเสียง

อื่นๆในกลุ่มเดียวกัน จึงอาจทำให้ข้อมูลต่างๆมีไม่เพียงพอต่อการจำแนกหน่วยเสียงประเภทนี้ออกจากหน่วยเสียงในกลุ่มเดียวกันได้อย่างชัดเจน ส่วนหน่วยเสียงในกลุ่มสระผสมนั้นได้ *% ความถูกต้อง*เกิน 70 เปอร์เซนต์ ถึง 2 ตัวจากทั้งหมด 3 ตัว โดยสระเอือ (/vva/) เป็นหน่วยเสียงที่มีความถูกต้องสูงที่สุดที่ได้คือ 73.64 เปอร์เซนต์

ส่วนผลการจำแนกในกลุ่มพยัญชนะพบว่า พยัญชนะแต่ละประเภทมักมีความสับสนกันระหว่างพยัญชนะและตัวสะกดประเภทเดียวกันรวมไปถึงพยัญชนะที่มีพฤติกรรมการออกเสียงที่คล้ายคลึงกันด้วย โดยหน่วยเสียงในกลุ่มพยัญชนะเสียงเสียดแทรกที่สามารถจำแนกผลได้ดีที่สุดคือพยัญชนะตัวสะกด “ส” (/s^/) สามารถจำแนกได้ถูกต้องถึง 86.36 เปอร์เซนต์ ส่วนในกลุ่มพยัญชนะกักนั้น ถือว่ามีความสับสนภายในกลุ่มค่อนข้างมาก โดยหน่วยเสียงที่มีผลการจำแนกที่ดีที่สุดคือหน่วยเสียง “ด” (/d/) ซึ่งมีค่าเพียง 56.15 เปอร์เซนต์ สำหรับกลุ่มพยัญชนะเสียงกึ่งเสียดแทรกนอกจากที่จะมีความสับสนภายในกลุ่มแล้ว ยังมีบางส่วนที่มีความสับสนกับกลุ่มพยัญชนะเสียงเสียดแทรก ซึ่งถือว่ามีพฤติกรรมของลักษณะการเปล่งเสียงที่ค่อนข้างใกล้เคียงกันด้วย โดยพยัญชนะเสียง “ช”, “ฌ” (/ch/) สามารถจำแนกได้ถูกต้อง 66.27 เปอร์เซนต์ หน่วยเสียงกลุ่มสุดท้ายคือ กลุ่มพยัญชนะนาสิก ซึ่งในที่นี้จะขอวิเคราะห์รวมกับหน่วยเสียงที่เกิดจากการสั่นของเส้นเสียง (/vb/) ก่อนที่มีการเปล่งเสียง “ด” (/b/), “ต”(/d/) เข้าไปด้วย ซึ่งหน่วยเสียงกลุ่มนี้ ก็ยังคงความสับสนกันอยู่ภายในกลุ่มเช่นกัน โดยหน่วยเสียงที่จำแนกได้ดีที่สุดคือหน่วยเสียง (/vb/) ที่สูงถึง 80.64 เปอร์เซนต์ แต่หน่วยเสียงในกลุ่มที่มีผลการจำแนกได้ต่ำสุด คือ หน่วยเสียง “ง” (/ng/) นั้นพบว่าการที่หน่วยเสียงดังกล่าวมีผลที่ต่ำมากนั้น ซึ่งอาจเกิดจากจำนวนหน่วยเสียงที่นำมาใช้ในการฝึกฝนนั้นน้อยมาก เมื่อเทียบกับจำนวนหน่วยเสียงอื่นๆในกลุ่มพยัญชนะนาสิก

กลุ่มหน่วยเสียงต่อมาคือ กลุ่มหน่วยเสียงกึ่งสระซึ่งมีพฤติกรรมการเปล่งเสียงที่คล้ายคลึงกับกลุ่มสระ จึงทำให้มีหน่วยเสียงบางส่วนสับสนกับกลุ่มสระ โดยมักจะเป็นสระประเภทเสียงสั้นมากกว่าสระเสียงยาว รวมไปถึงมีบางส่วนที่สับสนกับพยัญชนะเสียงนาสิกซึ่งเป็นหน่วยเสียงที่มีพฤติกรรมของลักษณะหน่วยเสียงที่คล้ายคลึงกันด้วยเช่นกัน โดยหน่วยเสียงที่สามารถจำแนกได้ถูกต้องสูงสุดคือ หน่วยเสียง “ย”, “ญ” (/j/) และหน่วยเสียงตัวสะกด “ย” (/j^/) ได้ 69.47 เปอร์เซนต์ และ 66.67 เปอร์เซนต์ ตามลำดับ

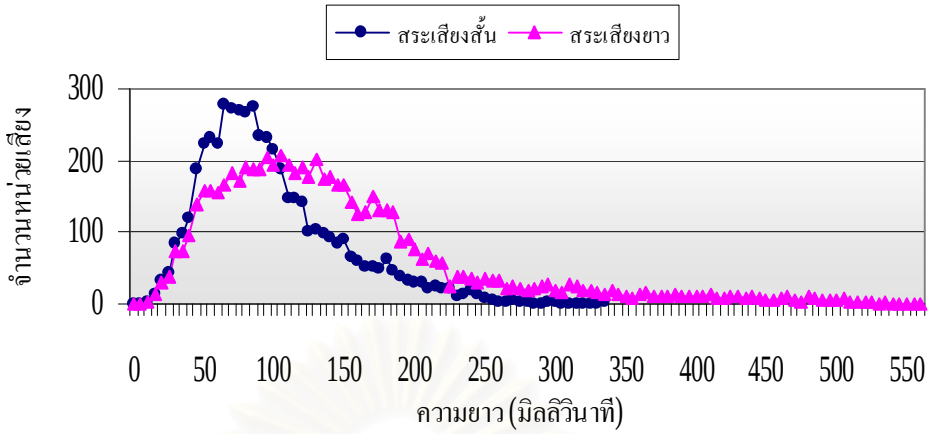
สุดท้ายคือ หน่วยเสียงเงียบ (/sil/) ซึ่งถือว่าหน่วยเสียงนี้เป็นหน่วยเสียงที่มีเอกลักษณ์มากที่สุด และเป็นหน่วยเสียงที่สามารถจำแนกได้ค่าความถูกต้องสูงที่สุดคือ 91.27 เปอร์เซนต์ โดยส่วนที่มีความสับสนเพียงเล็กน้อยนั้นจะสับสนกับหน่วยเสียงพยัญชนะเท่านั้น

และเมื่อทดลองจำแนกหน่วยเสียงโดยให้ค่านำหนักตามจำนวนข้อมูลที่นำไปใช้ฝึกฝน พบว่า ผลการจำแนกมีความถูกต้องเท่ากับ 66.41 เปอร์เซนต์ เพิ่มขึ้นจากเดิมเกือบ 5 เปอร์เซนต์ ซึ่งลักษณะผลการจำแนกที่ได้นั้น ยังมีลักษณะเดิม กล่าวคือ ยังคงสับสนอยู่ภายในกลุ่มที่มีลักษณะที่คล้ายคลึงกัน แต่หน่วยเสียงที่มีความสับสนนั้น มีแนวโน้มที่จะตอบเป็นหน่วยเสียงที่มีจำนวน

ข้อมูลฝึกฝนมากมากขึ้น เนื่องจากเป็นหน่วยเสียงที่มีค่าน้ำหนักมาก ซึ่งลักษณะดังกล่าวส่งผลให้เกิดทั้งข้อดีและข้อเสีย กล่าวคือ หน่วยเสียงที่มีเข้ามาจำแนกนั้น หากเป็นหน่วยเสียงที่มีค่าน้ำหนักมาก จะส่งผลให้หน่วยเสียงที่เคยสับสนกับกลุ่มอื่นนั้น มีผลการจำแนกถูกต้องมากขึ้นเนื่องจากได้ให้ค่าความน่าจะเป็นของโอกาสที่จะตอบเป็นหน่วยเสียงนั้นมีมากขึ้น โดยตัวอย่างที่เห็นได้ชัดคือ สระอะ (/a/) ซึ่งเป็นกลุ่มที่มีจำนวนข้อมูลในฝึกฝนจำนวนมากกว่าหน่วยเสียงสระอื่นๆ ทำให้หน่วยเสียงสระอะ (/a/) ที่เคยมีความสับสนเป็นเสียงสระตัวอื่นสามารถจำแนกได้ถูกต้องมากขึ้นจากเดิมถึง 20.69 เปอร์เซนต์ แต่จำนวนหน่วยเสียงที่มีการจำแนกผิดเป็นสระเสียงยาวคือ สระอา (/aa/) ยังคงมีจำนวนสูงอยู่ เนื่องจาก หน่วยเสียงสระอา (/aa/) ที่นำมาใช้ในการฝึกฝนนั้นมีมากจึงทำให้ค่าน้ำหนักของหน่วยเสียงชนิดนี้สูงเช่นกัน นอกจากหน่วยเสียงสระอะ (/a/) แล้ว หน่วยเสียงสระออ (/@@/) ก็สามารถจำแนกได้ถูกต้องเพิ่มขึ้นจากเดิมจำนวนมากเช่นกัน โดยเพิ่มขึ้นเกือบ 20 เปอร์เซนต์ ซึ่งหน่วยเสียงนี้ถือว่ามีค่าน้ำหนักเป็นอันดับรองๆลงมา แต่ในขณะเดียวกันพฤติกรรมดังกล่าวก็ส่งผลให้หน่วยเสียงบางหน่วยเสียงกลับมีผลในการจำแนกลดลง เนื่องจากหน่วยเสียงดังกล่าวมีค่าน้ำหนักน้อยกว่า ซึ่งทำให้เวลาจำแนกหน่วยเสียงนั้น ผลการจำแนกหน่วยเสียงก็มีโอกาสเบี่ยงเบนไปยังกลุ่มที่มีน้ำหนักมากด้วยเช่นกัน ตัวอย่างเช่น หน่วยเสียง สระเออะ(/q/), สระออ (/@/), สระแอะ(/x/) ที่มีจำนวนหน่วยเสียงที่ใช้ในการฝึกฝนน้อยกว่าหน่วยเสียงอื่นๆภายในกลุ่มหน่วยเสียงสระอย่างชัดเจน จึงทำให้ผลการจำแนกที่ได้มี % *ความถูกต้อง* ลดลง

จากผลการวิเคราะห์ผลการจำแนกข้างต้นนั้น จะเห็นว่าความสับสนระหว่างสระเสียงสั้นและสระเสียงยาวเป็นประเด็นหนึ่งที่มีผลสำคัญต่อค่าความถูกต้องในการจำแนกหน่วยเสียง หากรวมผลการจำแนกที่มีความผิดพลาดกันระหว่างกลุ่มสระเสียงสั้นและสระเสียงยาวประเภทเดียวกันทั้ง 9 ประเภท รวม 18 หน่วยเสียง จะปรากฏว่าผลการจำแนกที่ได้สูงขึ้นเป็น 71.47 เปอร์เซนต์ เพิ่มขึ้นจากเดิมถึง 5.06 เปอร์เซนต์ และเมื่อลองนำความยาวของสระเสียงสั้นและสระเสียงยาวมาลงจุดเป็นกราฟแสดงการกระจายตัวของสระเสียงสั้นและสระเสียงยาวตามความยาวแล้วได้ผล ดังรูป

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

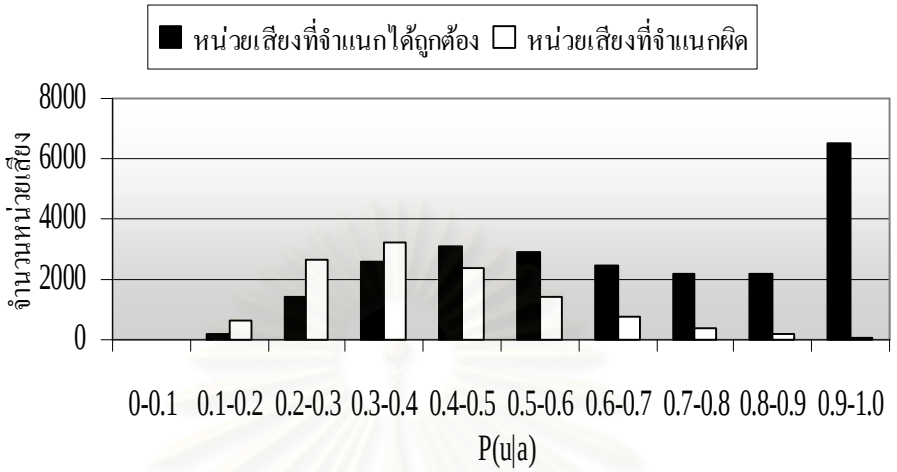


รูปที่ 5.2 กราฟแสดงการกระจายตัวของสระเสียงสั้นและยาวตามความยาวของสระ

จากกราฟจะเห็นว่าค่าการกระจายของทั้งสองกลุ่มใกล้เคียงกันมากจนซ้อนทับกัน โดยกลุ่มสระเสียงสั้นมีค่าฐานนิยมอยู่ที่ความยาว 65 มิลลิวินาทีในขณะที่กลุ่มสระเสียงยาวอยู่ที่ 105 มิลลิวินาที ดังนั้นจึงกล่าวได้ว่าการจำแนกหน่วยเสียงระหว่างสระเสียงสั้นและสระเสียงยาวทำได้ลำบาก เนื่องจากสระทั้ง 2 ชนิดมีพฤติกรรมที่เหมือนกัน โดยจะต่างกันเฉพาะความยาวเท่านั้น ซึ่งในความเป็นจริงแล้ว ความยาวหน่วยเสียงของสระเสียงยาวและสระเสียงสั้นมีค่าการกระจายที่ใกล้เคียงกันมาก ดังนั้นสมบัติความยาวจึงยังไม่มีความชัดเจนมากพอที่จะนำมาใช้จำแนกหน่วยเสียง และนอกจากนี้ความสมบัติความยาวของสระก็เป็นสิ่งที่ยังไม่ชัดเจนตายตัวอีกเช่นกัน ดังจะเห็นจากการกระจายตัวของสระเสียงยาวที่เริ่มตั้งแต่ความยาวไม่มากจนกระทั่งความยาวมากๆ ซึ่งสาเหตุดังกล่าว ส่วนหนึ่งอาจเกิดจากความเร็วของผู้พูดแต่ละคนแตกต่างกัน ซึ่งอาจทำให้สระเสียงยาวที่เกิดจากผู้พูดคนหนึ่งหรือคำๆหนึ่งมีความยาวใกล้เคียงกับสระเสียงสั้นที่เกิดจากผู้พูดอีกคนหนึ่งหรือคำอีกคำหนึ่งได้เช่นกัน ดังนั้นการที่จะจำแนกสระเสียงสั้นและสระเสียงยาวนั้น จึงถือว่าเป็นเรื่องที่ทำได้ลำบาก

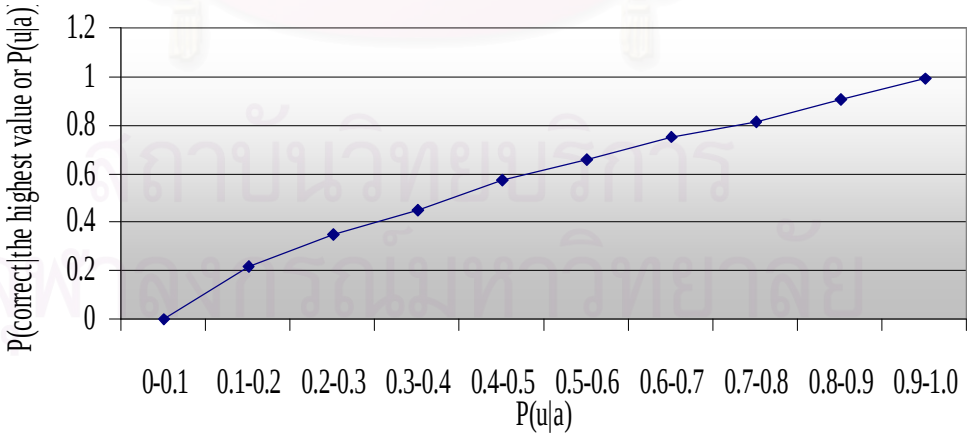
ขั้นตอนต่อมาจะทำการวิเคราะห์ว่า หลักการการวิเคราะห์ดิสคริมิแนนต์เชิงเส้นมีความมั่นใจในคำตอบที่ได้มากน้อยเท่าใด โดยจะนำค่าความน่าจะเป็นที่ได้จากการจำแนกหน่วยเสียงที่ได้จากการคำนวณตามสมการที่ 2.21 มาแบ่งออกเป็น 2 กลุ่ม คือ กลุ่มที่สามารถจำแนกได้ผลถูกต้องและกลุ่มที่จำแนกผิด เพื่อดูว่าค่าน่าจะเป็นของคำตอบที่ถูกต้องมีความมั่นใจเท่าไหร่ ในขณะที่ตอบผิด วิธีการนี้มีความมั่นใจเท่าไร โดยจะนำเสนอด้วยกราฟแสดงความถี่ของค่าความน่าจะเป็นของหน่วยเสียงที่จำแนกถูกและผิด ดังรูปที่ 5.3 ซึ่งปรากฏว่ากลุ่มที่สามารถจำแนกหน่วยเสียงได้ถูกต้องมีค่าเฉลี่ยความน่าจะเป็นอยู่ที่ 0.67 และมีค่าเบี่ยงเบนมาตรฐานอยู่ที่ 0.25 ส่วนกลุ่มที่จำแนกผิดมีค่าเฉลี่ยความน่าจะเป็นอยู่ที่ 0.41 และมีค่าเบี่ยงเบนมาตรฐานอยู่ที่ 0.16 ผลที่ได้ดังกล่าว แสดงให้เห็นว่าเมื่อมีการจำแนกหน่วยเสียงที่ถูกต้องแล้ว วิธีการวิเคราะห์ดิสคริมิแนนต์เชิงเส้นมีความมั่นใจในคำตอบมากกว่ากรณีที่เลือกคำตอบที่ไม่ถูกต้อง และนอกจากนี้ยังจะเห็นว่า คำตอบที่หลักการนี้

สามารถจำแนกได้ถูกต้อง และมีความมั่นใจมากถึง 0.9-1.0 นั้นมีค่าสูงกว่าในค่าความน่าจะเป็นในช่วงอื่นๆ ในปริมาณมากกว่าอย่างเห็นได้ชัด



รูปที่ 5.3 กราฟแสดงความถี่ของค่าความน่าจะเป็นของหน่วยเสียงที่จำแนกถูกและผิด

และเมื่อนำจำนวนหน่วยเสียงที่จำแนกได้ถูกต้องในช่วงความน่าจะเป็นต่างๆมาหารด้วยจำนวนหน่วยเสียงทั้งหมดที่มีค่าความมั่นใจในช่วงนั้นๆ จะได้กราฟแสดงความสัมพันธ์ระหว่างค่าความน่าจะเป็นที่ตอบถูกเมื่อทราบค่าความน่าจะเป็นที่จะตอบหน่วยเสียงใดๆเมื่อรู้สัญญาณเสียงกับค่าความน่าจะเป็นที่จะตอบหน่วยเสียงใดๆเมื่อรู้สัญญาณเสียง ดังกราฟในรูปที่ 5.4 ซึ่งกราฟมีแนวโน้มเพิ่มขึ้นเรื่อยๆ ทำให้สามารถอธิบายได้ว่า การจำแนกหน่วยเสียงด้วยวิธีนี้มีแนวโน้มที่จะจำแนกถูกต้องมากขึ้นเรื่อยๆ เมื่อมีค่าความน่าจะเป็นที่จะตอบหน่วยเสียงนั้นๆ มีค่ามากขึ้นเรื่อยๆ



รูปที่ 5.4 กราฟแสดงความสัมพันธ์ระหว่างค่าความน่าจะเป็นที่ตอบถูกเมื่อทราบค่าความน่าจะเป็นที่จะตอบหน่วยเสียงใดๆเมื่อรู้สัญญาณเสียงกับค่าความน่าจะเป็นที่จะตอบหน่วยเสียงใดๆเมื่อรู้สัญญาณเสียง

ตารางที่5.5 ผลการจำแนกหน่วยเสียงแบบใช้ค่าน้ำหนักเท่ากันโดยแจกแจงตามความสับสนระหว่างกลุ่ม

	a	**i**	**e**	**o**	**u**	**v**	**q**	**@**	**x**	**aa**	**ii**	**ee**	**oo**	**uu**
a	889	16	66	36	9	41	94	208	181	723	5	0	10	0
i	0	245	1	0	3	34	0	0	0	0	121	7	0	0
e	0	5	144	0	0	55	9	0	1	0	4	50	0	0
o	4	0	1	226	7	3	11	34	0	0	0	0	82	17
u	0	11	0	4	103	11	0	0	0	0	0	0	9	60
v	0	26	5	0	6	102	2	0	0	0	0	6	0	0
q	1	0	1	0	0	1	4	0	1	0	0	0	1	0
@	8	0	0	21	0	0	1	98	0	15	0	0	1	0
x	20	0	2	0	0	0	5	0	111	11	0	0	0	0
aa	220	1	10	6	1	0	31	81	81	1612	0	0	1	0
ii	0	117	0	0	1	9	0	0	0	0	511	0	0	0
ee	0	9	20	0	0	13	0	0	0	0	10	101	0	0
oo	0	0	1	47	3	0	6	8	0	0	0	0	79	24
uu	0	0	1	4	52	0	2	0	0	0	2	0	19	168
vv	0	14	0	0	4	15	0	0	0	0	2	1	0	0
qq	3	0	3	0	2	3	30	0	1	0	0	2	4	0
@@	14	0	0	80	0	0	2	183	0	43	0	0	43	1
xx	22	0	8	0	0	0	5	0	55	26	0	0	0	0
iia	0	1	2	0	0	9	0	0	0	0	11	18	0	0
vva	3	0	1	0	3	11	12	0	0	0	0	1	1	0
uua	0	0	0	16	9	0	8	0	0	1	0	0	26	37
f	0	0	0	0	0	0	0	0	0	0	0	0	0	0
s	0	0	0	0	0	0	0	0	0	0	0	0	0	0
p	0	0	0	0	0	0	0	0	0	2	0	0	0	0
t	0	0	0	0	0	0	1	0	0	0	0	0	0	0
k	0	0	0	0	0	0	0	1	0	0	0	0	0	0
b	0	0	0	0	0	0	0	0	0	0	0	0	0	0
d	0	0	2	0	0	2	1	0	0	0	0	0	0	0
ph	1	0	0	0	1	0	0	1	0	0	0	0	0	0
th	1	0	0	0	0	0	2	0	0	0	0	0	0	0
kh	0	0	0	2	1	0	0	0	0	0	0	0	0	0
c	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ch	0	0	0	0	0	0	0	0	0	0	0	0	0	0
m	0	0	0	0	4	0	0	0	0	0	0	0	0	1
n	0	0	0	0	3	5	0	0	0	0	0	0	0	0
ng	0	1	0	0	2	0	0	0	0	0	0	0	0	0
vb	0	0	0	0	2	0	0	0	0	0	0	0	0	0
m^	0	0	0	0	4	0	0	0	0	0	0	0	0	0
n^	1	7	0	0	11	12	0	0	1	0	1	0	0	3
ng^	4	4	0	0	40	2	0	1	0	0	2	0	0	7
j^	10	116	48	0	8	53	4	0	19	0	38	8	1	0
w^	9	1	0	17	18	5	3	8	0	0	0	0	8	23
l^	0	0	0	0	0	0	0	0	0	0	0	0	0	0
f^	1	0	0	0	0	0	0	0	0	0	0	0	0	0
s^	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ch^	0	0	0	0	0	0	0	0	0	0	0	0	0	0
l	18	13	4	5	19	33	13	3	1	0	4	2	6	0
r	56	2	8	12	19	15	12	2	1	2	1	0	12	0
h	1	0	0	0	0	0	0	0	0	0	2	0	0	0
w	0	0	0	6	4	1	0	5	0	1	0	0	4	9
j	0	8	0	0	0	4	0	0	0	0	13	1	0	0
sil	0	0	0	0	0	0	0	0	0	0	0	0	0	0

ตารางที่ 5.5 ผลการจำแนกหน่วยเสียงแบบใช้ค่าน้ำหนักเท่ากันโดยแจกแจงตามความสับสนระหว่างกลุ่ม

	vv	**qq**	**@@**	**xx**	**iia**	**vva**	**uua**	**f**	**s**	**p**	**t**	**k**	**b**	**d**
a	1	36	33	106	1	37	2	0	0	4	0	1	2	9
i	5	0	0	0	2	0	0	0	0	0	0	0	1	2
e	1	3	0	6	10	12	0	0	0	0	0	0	0	2
o	0	4	4	0	0	1	38	0	0	0	0	0	0	0
u	12	0	0	0	0	1	4	0	0	0	0	0	0	0
v	37	2	0	0	1	5	0	0	0	0	0	0	0	0
q	0	5	0	0	0	0	1	0	0	0	0	0	0	0
@	0	0	48	1	0	0	2	0	0	0	0	0	0	0
x	0	9	0	50	1	3	0	0	0	0	0	0	0	0
aa	0	21	81	129	3	6	2	0	0	0	0	0	0	0
ii	9	0	0	0	11	0	0	0	0	0	0	0	2	0
ee	3	1	0	0	7	2	0	0	0	0	0	0	0	0
oo	1	7	3	0	0	0	20	0	0	0	0	0	0	0
uu	7	2	0	0	0	0	11	0	0	0	0	0	0	0
vv	142	1	0	0	0	0	0	0	0	0	0	0	0	0
qq	9	106	0	1	0	8	3	0	0	0	0	0	0	0
@@	0	1	353	0	0	1	15	0	0	0	0	0	0	0
xx	0	21	1	193	2	10	0	0	0	0	0	0	0	1
iia	2	0	0	0	127	9	0	0	0	0	0	0	0	0
vva	13	4	0	2	6	190	4	0	0	0	0	0	0	0
uua	3	2	0	0	0	1	192	0	0	0	0	0	0	0
f	0	0	0	0	0	0	0	69	8	1	3	1	0	1
s	0	0	0	0	0	0	0	35	679	0	1	0	0	0
p	0	0	0	0	0	0	0	7	0	253	41	26	58	35
t	0	0	0	0	0	0	0	29	0	70	214	96	21	64
k	0	0	0	0	0	0	0	41	7	26	60	467	3	34
b	0	0	0	0	0	0	0	0	0	9	1	1	28	9
d	0	0	0	0	0	0	0	4	0	2	11	8	15	105
ph	0	0	0	0	0	0	0	18	0	65	5	18	5	23
th	0	0	0	0	0	0	0	29	9	10	46	26	0	17
kh	0	0	2	0	0	0	0	41	8	23	20	84	8	8
c	0	0	0	0	0	0	0	18	65	5	18	5	0	9
ch	0	0	0	0	0	0	0	5	54	1	4	0	0	0
m	6	0	0	0	0	0	0	1	0	9	0	0	5	1
n	14	0	0	0	0	0	0	0	0	15	0	0	7	5
ng	2	0	0	1	0	0	0	0	0	0	0	0	0	0
vb	0	0	0	0	0	0	0	1	0	22	1	1	14	9
m^	15	0	0	0	0	0	0	0	0	0	0	0	2	1
n^	24	0	0	0	0	1	0	0	0	0	0	0	0	0
ng^	4	0	1	0	0	0	0	0	0	0	0	0	0	0
j^	6	5	0	11	3	0	0	0	0	0	0	0	0	0
w^	3	1	12	0	0	0	1	0	0	0	0	0	0	0
l^	0	0	1	0	0	0	0	0	0	0	0	0	0	0
f^	0	0	0	0	0	0	0	2	0	0	0	0	0	0
s^	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ch^	0	0	0	0	0	0	0	0	0	0	0	0	0	0
l	23	0	0	1	3	24	3	5	0	13	1	0	15	7
r	7	5	1	2	1	8	8	11	0	24	7	6	32	41
h	0	0	1	0	0	0	0	17	0	11	0	3	3	1
w	0	0	3	0	0	0	9	0	0	6	0	0	12	2
j	1	0	0	0	5	1	0	1	0	0	0	0	0	5
sil	0	0	0	0	0	0	0	44	0	127	29	21	24	10

ตารางที่ 5.5 ผลการจำแนกหน่วยเสียงแบบใช้ค่าน้ำหนักเท่ากันโดยแจกแจงตามความสับสนระหว่างกลุ่ม

	ph	**th**	**kh**	**c**	**ch**	**m**	**n**	**ng**	**vb**	**m^**	**n^**	**ng^**	**j^**	**w^**
a	10	1	3	0	0	2	4	8	1	1	8	20	56	71
i	0	1	0	0	0	4	17	1	2	2	20	0	11	0
e	0	0	0	0	0	0	0	0	0	0	3	0	30	0
o	0	0	0	0	0	0	0	3	0	0	0	9	1	29
u	0	0	0	0	0	2	1	1	0	6	7	10	0	3
v	0	0	0	0	0	1	1	2	0	3	8	0	6	1
q	0	0	0	0	0	0	0	0	0	0	0	0	0	0
@	1	0	0	0	0	0	0	0	0	0	0	0	0	6
x	0	0	0	0	0	0	0	0	0	0	0	0	3	0
aa	4	1	1	0	0	0	0	0	0	0	1	2	1	7
ii	0	0	0	0	0	1	1	0	0	0	3	0	1	0
ee	0	0	0	0	0	0	0	0	0	0	0	0	12	0
oo	0	0	0	0	0	0	0	0	0	0	0	0	1	2
uu	0	0	0	0	0	2	2	1	0	3	2	8	4	2
vv	0	0	0	0	0	5	0	0	0	4	2	0	2	0
qq	0	0	0	0	0	0	0	0	0	0	0	2	0	0
@@	1	0	0	0	0	0	0	0	0	0	0	0	1	7
xx	0	1	0	0	0	0	0	0	0	0	0	0	1	0
iia	0	0	0	0	0	0	0	0	0	0	0	0	0	0
vva	0	0	0	0	0	1	0	1	0	0	0	0	1	0
uua	0	0	0	0	0	0	0	1	0	0	0	0	0	0
f	1	1	2	4	0	0	0	0	0	0	0	0	0	0
s	0	0	1	117	51	0	0	0	0	0	0	0	0	1
p	36	3	4	0	0	0	1	0	10	0	0	1	0	1
t	10	28	12	13	2	0	0	0	0	0	0	1	1	1
k	31	36	131	55	1	0	1	0	0	0	0	0	1	1
b	0	0	0	0	0	2	1	0	17	0	0	0	0	0
d	4	2	0	2	0	1	1	1	15	0	0	0	0	0
ph	324	41	26	2	1	1	5	2	10	0	2	4	1	3
th	96	439	45	57	22	1	1	0	0	0	0	1	3	0
kh	47	84	542	32	18	0	0	0	0	0	0	0	0	8
c	0	19	0	277	41	0	0	0	0	0	0	0	0	0
ch	0	6	0	29	220	0	0	0	0	0	0	0	0	0
m	0	0	0	0	0	331	56	11	93	47	4	15	0	0
n	3	0	0	0	0	112	353	37	121	14	44	10	0	0
ng	0	0	0	0	0	15	29	17	10	2	3	9	0	0
vb	3	0	2	0	0	23	14	7	633	5	29	4	0	0
m^	4	0	0	0	0	43	10	5	46	446	56	47	0	0
n^	3	0	0	0	0	40	114	13	110	212	965	70	1	6
ng^	9	0	1	0	0	57	18	25	38	126	212	643	2	11
j^	0	1	0	0	0	0	0	0	0	3	22	9	756	2
w^	0	0	0	0	0	9	2	1	0	5	2	12	4	136
l^	0	0	0	0	0	0	0	0	1	0	0	0	0	2
f^	1	1	1	0	0	0	0	0	0	0	1	0	0	0
s^	0	0	0	0	1	0	0	0	0	0	0	0	0	0
ch^	0	0	0	0	0	0	0	0	0	0	0	0	0	0
l	1	0	0	0	0	32	41	13	8	4	0	0	2	2
r	12	3	2	0	0	14	14	7	7	1	0	3	5	6
h	27	6	8	0	0	3	0	2	6	1	0	3	0	0
w	6	0	5	0	0	33	31	13	29	0	0	1	0	1
j	0	1	0	0	0	1	16	2	11	0	1	0	3	0
sil	69	20	13	7	1	1	0	0	56	1	3	1	0	1

ตารางที่ 5.5 ผลการจำแนกหน่วยเสียงแบบใช้ค่าน้ำหนักเท่ากันโดยแจกแจงตามความสับสนระหว่างกลุ่ม

	l^	f^	s^	ch^	l	r	h	w	j	sil	%
a	1	3	0	0	30	56	60	10	7	0	31.06
i	0	2	0	0	6	0	0	0	18	1	48.42
e	0	0	0	0	1	0	0	0	6	0	42.11
o	9	0	0	0	5	1	1	4	0	0	45.75
u	2	0	0	0	8	1	0	3	0	0	39.77
v	0	0	0	0	12	1	0	0	1	0	44.74
q	0	0	0	0	0	1	0	0	0	0	25.00
@	0	0	0	0	0	0	2	2	0	0	47.57
x	0	0	0	0	0	1	0	0	0	0	51.39
aa	0	1	0	0	1	5	7	2	0	0	69.51
ii	0	0	0	0	1	0	0	0	26	0	73.74
ee	0	0	0	0	1	0	0	0	2	0	55.80
oo	0	0	0	0	0	2	0	4	0	0	37.98
uu	2	0	0	0	3	0	0	0	0	0	56.57
vv	0	0	0	0	2	0	0	0	4	0	71.72
qq	0	0	0	0	0	1	0	0	0	0	59.55
@@	4	0	0	0	0	0	0	0	0	0	47.13
xx	0	0	0	0	0	0	0	0	0	0	55.78
iia	0	0	0	0	0	0	0	0	1	0	70.56
vva	0	0	0	0	2	1	0	0	1	0	73.64
uua	1	0	0	0	2	0	0	6	0	0	62.95
f	0	7	3	0	0	0	2	0	0	1	66.35
s	0	3	128	0	0	0	0	0	0	0	66.83
p	0	7	0	0	2	14	4	4	0	10	48.75
t	0	17	0	0	1	6	2	1	0	2	36.15
k	1	24	14	0	0	5	6	0	0	5	49.11
b	0	0	0	0	3	1	0	3	0	1	36.84
d	0	1	0	0	1	4	2	1	2	0	56.15
ph	0	14	7	0	1	2	27	2	0	8	52.26
th	0	11	25	6	0	1	14	0	0	0	50.93
kh	4	16	6	1	0	1	6	5	0	10	55.48
c	0	1	22	1	0	0	0	0	0	0	57.59
ch	0	0	12	1	0	0	0	0	0	0	66.27
m	0	0	0	0	7	1	5	6	1	0	54.80
n	0	0	0	0	12	3	5	0	1	0	46.20
ng	0	0	0	0	8	1	6	0	1	0	15.89
vb	0	0	0	0	1	0	0	5	1	8	80.64
m^	0	1	0	0	2	0	1	0	0	0	65.30
n^	0	1	0	0	10	3	6	0	2	4	59.53
ng^	2	4	0	0	16	1	21	5	0	1	51.15
j^	0	2	0	0	2	5	0	0	2	0	66.67
w^	8	0	0	0	1	3	0	2	0	1	46.10
l^	2	0	0	0	0	0	0	0	0	0	33.33
f^	0	14	2	0	0	0	0	0	0	1	58.33
s^	0	2	19	0	0	0	0	0	0	0	86.36
ch^	0	0	2	3	0	0	0	0	0	0	60.00
l	0	0	0	0	299	68	12	39	70	0	37.05
r	1	0	0	0	148	524	48	75	20	0	44.22
h	0	3	0	0	2	5	147	0	1	0	58.10
w	0	0	0	0	11	2	11	228	2	1	52.29
j	0	0	0	0	5	0	0	1	182	0	69.47
sil	0	202	3	0	0	0	5	2	0	6691	91.27
										% ความถูกต้องเฉลี่ย	61.41

ตารางที่5.6 ผลการจำแนกหน่วยเสียงแบบใช้ค่าน้ำหนักขึ้นกับจำนวนข้อมูลโดยแจกแจงตามความสับสนระหว่างกลุ่ม

	a	**i**	**e**	**o**	**u**	**v**	**q**	**@**	**x**	**aa**	**ii**	**ee**	**oo**	**uu**
a	1481	12	35	40	2	20	1	39	25	749	6	0	6	0
i	0	226	1	0	2	26	0	0	0	0	137	2	0	0
e	15	6	140	0	0	49	1	0	0	0	6	20	0	0
o	23	0	1	271	5	0	1	12	0	0	0	0	48	17
u	1	12	0	13	77	8	0	0	0	0	0	0	9	64
v	6	33	7	0	4	96	0	0	0	0	2	2	0	0
q	7	0	0	0	0	0	0	0	0	0	0	0	1	0
@	24	0	0	22	0	0	0	24	0	28	0	0	1	0
x	77	0	0	0	0	0	0	0	62	30	0	0	0	0
aa	406	1	8	3	1	0	0	4	13	1727	0	0	2	0
ii	0	112	0	0	2	3	0	0	0	0	530	0	0	0
ee	1	12	19	0	0	13	0	0	0	0	15	74	0	0
oo	2	0	2	55	2	0	0	3	0	0	0	0	72	23
uu	0	0	0	6	45	0	0	0	0	0	2	0	16	168
vv	0	14	0	0	3	13	0	0	0	0	10	1	0	1
qq	31	0	4	2	2	4	1	0	0	1	0	2	3	0
@@	25	0	0	79	0	0	0	33	0	61	0	0	26	2
xx	85	0	9	0	0	0	0	0	18	71	0	0	0	0
iia	0	2	3	0	0	9	0	0	0	0	16	18	0	0
vva	14	0	2	0	3	11	0	0	0	0	0	0	1	0
uua	4	0	0	20	9	1	1	0	0	1	0	0	18	38
f	0	0	0	0	0	0	0	0	0	0	0	0	0	0
s	0	0	0	0	0	0	0	0	0	0	0	0	0	0
p	1	0	0	0	0	0	0	0	0	1	0	0	0	0
t	1	0	0	0	0	0	0	0	0	0	0	0	0	0
k	0	0	0	0	0	0	0	0	0	0	0	0	0	0
b	0	0	0	0	0	0	0	0	0	0	0	0	0	0
d	1	0	2	0	0	0	0	0	0	0	0	0	0	0
ph	5	0	0	0	1	0	0	0	0	0	0	0	0	0
th	3	0	0	0	0	0	0	0	0	0	1	0	0	0
kh	1	0	0	3	1	0	0	0	0	0	0	0	0	0
c	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ch	0	0	0	0	0	0	0	0	0	0	0	0	0	0
m	0	0	0	0	2	0	0	0	0	0	0	0	0	1
n	1	0	0	0	2	1	0	0	0	0	0	0	0	0
ng	1	1	0	0	1	0	0	0	0	0	0	0	0	0
vb	0	0	0	0	1	0	0	0	0	0	0	0	0	0
m^	0	0	0	0	2	0	0	0	0	0	0	0	0	0
n^	6	5	0	0	4	5	0	0	0	0	1	0	0	1
ng^	8	1	0	0	24	1	0	0	0	0	2	0	0	4
j^	41	87	27	0	7	33	0	0	5	0	43	1	1	0
w^	17	3	0	24	12	3	0	1	0	1	0	0	6	23
l^	0	0	0	0	0	0	0	0	0	0	0	0	0	1
f^	1	0	0	0	0	0	0	0	0	0	0	0	0	0
s^	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ch^	0	0	0	0	0	0	0	0	0	0	0	0	0	0
l	48	17	2	4	14	23	0	0	0	0	7	0	4	0
r	110	2	6	8	10	9	0	0	0	3	1	0	5	0
h	7	0	0	0	0	0	0	0	0	0	2	0	0	0
w	0	0	0	8	3	0	0	1	0	1	0	0	3	8
j	0	7	0	0	0	2	0	0	0	0	21	1	0	0
sil	0	0	0	0	0	0	0	0	0	0	0	0	0	0

ตารางที่ 5.6 ผลการจำแนกหน่วยเสียงแบบใช้ค่าน้ำหนักขึ้นกับจำนวนข้อมูลโดยแจกแจงตามความสับสนระหว่างกลุ่ม

	vv	**qq**	**@@**	**xx**	**iia**	**vva**	**uua**	**f**	**s**	**p**	**t**	**k**	**b**	**d**
a	0	12	75	31	0	20	2	0	0	2	0	1	0	4
i	2	0	0	0	2	0	0	0	0	0	0	0	1	2
e	0	3	0	6	7	9	0	0	0	0	0	0	0	1
o	0	1	21	0	0	1	37	0	0	0	0	0	0	0
u	8	0	0	0	0	0	4	0	0	0	0	0	0	0
v	24	1	0	0	0	6	0	0	0	0	0	0	0	0
q	0	4	0	0	0	0	1	0	0	0	0	0	0	0
@	0	0	97	0	0	0	1	0	0	0	0	0	0	0
x	0	3	0	38	1	1	0	0	0	0	0	0	0	0
aa	0	3	96	25	0	8	2	0	0	0	0	0	0	0
ii	5	0	0	0	10	0	0	0	0	0	0	0	0	0
ee	2	0	0	1	6	2	0	0	0	0	0	0	0	0
oo	1	5	10	0	0	0	21	0	0	0	0	0	0	0
uu	5	0	0	0	0	0	12	0	0	0	0	0	0	0
vv	125	0	0	0	0	0	1	0	0	0	0	0	0	0
qq	11	103	0	0	0	6	2	0	0	0	0	0	0	0
@@	0	0	491	0	0	1	18	0	0	0	0	0	0	0
xx	0	10	0	140	2	7	0	0	0	0	0	0	0	1
iia	2	0	0	0	117	12	0	0	0	0	0	0	0	0
vva	10	4	0	2	4	184	5	0	0	0	0	0	0	0
uua	1	0	0	0	0	3	195	0	0	0	0	0	0	0
f	0	0	0	0	0	0	0	42	25	1	4	3	0	1
s	0	0	0	0	0	0	0	1	861	0	1	0	0	0
p	0	0	0	0	0	0	0	2	0	265	44	33	13	31
t	0	0	0	0	0	0	0	12	2	69	225	123	5	46
k	0	0	0	0	0	0	0	14	21	15	55	507	1	24
b	0	0	0	0	0	0	0	0	0	12	1	1	20	11
d	0	0	0	0	0	0	0	2	0	3	13	14	3	105
ph	0	0	0	0	0	0	0	5	2	56	4	26	1	13
th	0	0	0	0	0	0	0	12	22	7	48	27	0	16
kh	0	0	1	0	0	0	0	10	19	16	23	85	2	1
c	0	0	0	0	0	0	0	6	91	7	19	10	0	4
ch	0	0	0	0	0	0	0	3	79	1	4	0	0	0
m	1	0	0	0	0	0	0	1	0	8	0	0	2	0
n	5	0	0	0	0	0	0	0	0	13	1	0	1	2
ng	0	0	0	0	0	0	0	0	0	0	0	0	0	0
vb	0	0	0	0	0	0	0	0	1	18	2	1	0	10
m^	10	0	0	0	0	0	0	0	0	0	0	0	0	0
n^	7	0	0	0	0	0	0	0	0	0	0	0	0	0
ng^	2	0	1	0	0	0	0	0	0	0	0	0	0	0
j^	3	2	0	7	2	0	0	0	0	0	1	0	0	0
w^	1	0	21	0	0	0	1	0	0	0	0	0	0	0
l^	0	0	1	0	0	0	0	0	0	0	0	0	0	0
f^	0	0	0	0	0	0	0	2	3	0	0	0	0	0
s^	0	0	0	0	0	0	0	0	6	0	0	0	0	0
ch^	0	0	0	0	0	0	0	0	2	0	0	0	0	0
l	8	0	2	2	1	10	2	1	0	15	1	1	4	7
r	4	0	0	0	1	3	6	6	0	24	8	11	6	27
h	0	0	1	0	0	0	0	6	0	12	1	7	0	1
w	0	0	7	0	0	0	10	0	0	8	0	0	3	2
j	1	0	0	0	4	1	0	1	0	0	0	0	0	4
sil	0	0	0	0	0	0	0	11	4	35	9	18	1	9

ตารางที่ 5.6 ผลการจำแนกหน่วยเสียงแบบใช้ค่าน้ำหนักขึ้นกับจำนวนข้อมูลโดยแจกแจงตามความสับสนระหว่างกลุ่ม

	ph	**th**	**kh**	**c**	**ch**	**m**	**n**	**ng**	**vb**	**m^**	**n^**	**ng^**	**j^**	**w^**
a	5	4	4	0	0	1	7	1	1	1	13	29	81	33
i	0	1	0	0	0	3	21	0	1	1	33	1	21	0
e	0	0	0	0	0	0	0	0	0	0	3	0	64	0
o	0	0	0	0	0	0	0	1	0	0	0	14	1	28
u	0	0	0	0	0	6	2	0	0	7	13	15	0	3
v	0	0	0	0	0	1	2	0	0	3	10	1	9	1
q	0	0	0	0	0	0	0	0	0	0	0	0	0	0
@	1	0	0	0	0	0	0	0	0	0	0	2	0	2
x	0	0	0	0	0	0	0	0	0	0	0	0	3	0
aa	4	1	0	0	0	0	0	0	0	0	1	1	1	3
ii	0	0	0	0	0	1	4	0	0	0	3	0	3	0
ee	0	0	0	0	0	0	0	0	0	0	0	0	34	0
oo	0	0	0	0	0	0	0	0	0	0	0	0	1	2
uu	0	0	0	0	0	1	3	0	0	4	4	16	6	2
vv	0	0	0	0	0	9	1	0	0	6	4	0	2	0
qq	0	0	0	0	0	0	0	0	0	0	0	2	0	0
@@	1	0	0	0	0	0	0	0	0	0	0	2	2	6
xx	0	1	0	0	0	0	0	0	0	0	0	0	1	0
iia	0	0	0	0	0	0	0	0	0	0	0	0	0	0
vva	0	0	0	0	0	1	0	0	0	0	0	0	1	0
uua	0	0	0	0	0	0	0	1	0	0	0	0	0	0
f	2	6	3	6	0	0	0	0	0	0	0	0	0	0
s	0	1	1	99	34	0	0	0	0	0	0	0	0	1
p	37	4	5	0	1	0	1	0	13	0	1	5	0	0
t	12	39	16	8	1	0	0	0	0	0	1	1	2	0
k	28	53	144	54	1	0	1	0	0	0	0	0	1	0
b	0	0	0	0	0	2	1	0	19	0	0	0	0	0
d	4	1	0	3	0	0	2	0	16	0	0	1	0	0
ph	329	52	33	5	0	1	4	0	12	0	5	8	3	0
th	89	470	60	58	18	1	1	0	1	0	1	2	3	0
kh	44	98	579	31	14	1	0	0	0	0	0	1	0	5
c	0	26	0	278	35	0	0	0	0	0	0	0	0	0
ch	1	8	0	28	208	0	0	0	0	0	0	0	0	0
m	0	0	0	0	0	305	75	0	106	51	13	21	0	0
n	2	1	0	0	0	80	371	0	124	8	114	13	0	0
ng	0	0	0	0	0	13	37	4	11	1	3	17	0	0
vb	4	0	2	0	0	17	18	0	627	3	58	3	0	0
m^	4	0	0	0	0	37	10	0	38	384	137	55	0	0
n^	2	0	0	0	0	21	75	1	81	98	1217	62	3	2
ng^	8	0	2	0	0	46	14	2	26	68	304	698	3	7
j^	0	1	0	0	0	0	0	0	0	2	32	10	818	1
w^	1	0	0	0	0	8	1	0	0	5	7	22	3	125
l^	0	0	0	0	0	0	0	0	1	0	0	1	0	2
f^	0	3	1	0	0	0	0	0	0	0	1	0	0	0
s^	0	0	0	1	1	0	0	0	0	0	0	0	0	0
ch^	0	0	0	0	0	0	0	0	0	0	0	0	0	0
l	1	0	0	0	0	29	50	2	8	5	4	1	5	0
r	9	5	2	0	0	16	14	1	8	1	2	5	6	2
h	34	10	13	2	0	6	0	0	7	1	0	6	1	0
w	8	0	8	0	0	27	44	0	37	0	0	4	0	1
j	0	1	0	0	0	4	21	0	13	0	1	0	3	0
sil	28	16	6	9	2	1	1	0	21	0	4	1	0	0

ตารางที่ 5.6 ผลการจำแนกหน่วยเสียงแบบใช้ค่าน้ำหนักขึ้นกับจำนวนข้อมูลโดยแจกแจงตามความสับสนระหว่างกลุ่ม

	l^	f^	s^	ch^	l	r	h	w	j	sil	%
a	0	2	0	0	37	53	18	3	6	0	51.75
i	0	1	0	0	6	1	0	0	14	1	44.66
e	0	0	0	0	4	3	0	0	5	0	40.94
o	0	0	0	0	7	1	0	4	0	0	54.86
u	0	0	0	0	14	2	0	1	0	0	29.73
v	0	0	0	0	18	1	0	0	1	0	42.11
q	0	0	0	0	2	1	0	0	0	0	0.00
@	0	0	0	0	0	0	2	2	0	0	11.65
x	0	0	0	0	0	1	0	0	0	0	28.70
aa	0	1	0	0	1	5	1	1	0	0	74.47
ii	0	0	0	0	3	0	0	0	17	0	76.48
ee	0	0	0	0	1	0	0	0	1	0	40.88
oo	0	0	0	0	3	2	0	4	0	0	34.62
uu	0	0	0	0	6	0	0	1	0	0	56.57
vv	0	0	0	0	6	0	0	0	2	0	63.13
qq	0	0	0	0	2	2	0	0	0	0	57.87
@@	1	0	0	0	1	0	0	0	0	0	65.55
xx	0	0	0	0	0	1	0	0	0	0	40.46
iia	0	0	0	0	0	0	0	0	1	0	65.00
vva	0	0	0	0	13	3	0	0	0	0	71.32
uua	0	0	0	0	6	1	0	6	0	0	63.93
f	0	2	0	0	0	0	1	0	0	8	40.38
s	0	1	16	0	0	0	0	0	0	0	84.74
p	0	0	0	0	1	18	2	4	0	37	51.06
t	0	4	0	0	3	10	0	2	0	10	38.01
k	0	3	2	0	0	6	3	1	1	16	53.31
b	0	0	0	0	2	2	0	3	0	2	26.32
d	0	0	0	0	5	8	1	0	2	1	56.15
ph	0	2	3	0	1	4	18	2	1	24	53.06
th	0	2	9	1	0	2	8	0	0	0	54.52
kh	0	4	1	0	0	1	4	4	0	28	59.26
c	0	0	4	0	0	0	0	0	0	1	57.80
ch	0	0	0	0	0	0	0	0	0	0	62.65
m	0	0	0	0	10	3	2	3	0	0	50.50
n	0	0	0	0	17	6	2	0	0	0	48.56
ng	0	0	0	0	12	1	4	0	1	0	3.74
vb	0	0	0	0	2	0	0	4	0	14	79.87
m^	0	0	0	0	5	0	0	0	0	1	56.22
n^	0	0	0	0	13	3	2	0	2	10	75.08
ng^	0	0	0	0	17	1	9	6	0	3	55.53
j^	0	0	0	0	4	6	0	0	0	0	72.13
w^	0	0	0	0	5	3	0	1	0	1	42.37
l^	0	0	0	0	0	0	0	0	0	0	0.00
f^	0	7	0	0	0	0	0	0	0	6	29.17
s^	0	2	12	0	0	0	0	0	0	0	54.55
ch^	0	0	1	2	0	0	0	0	0	0	40.00
l	0	0	0	0	351	81	9	37	51	0	43.49
r	0	0	0	0	161	612	31	47	13	0	51.65
h	0	0	0	0	2	9	123	0	1	1	48.62
w	0	0	0	0	15	7	7	222	1	1	50.92
j	0	0	0	0	15	1	0	1	160	0	61.07
sil	0	9	0	0	0	1	1	1	0	7143	97.44
										% ความถูกต้องเฉลี่ย	66.14

การวิเคราะห์ค่าความมีส่วนร่วมของลักษณะเฉพาะ

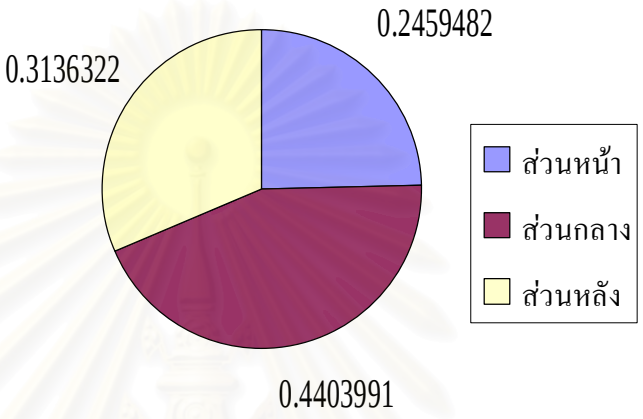
หลังจากที่ได้ทำการทดลองจำแนกหน่วยเสียง เพื่อวัดประสิทธิภาพความถูกต้องในการจำแนกหน่วยเสียง โดยใช้ลักษณะเฉพาะที่มีการทำงานแบบอาศัยเซกเมนต์แทนสัญญาณเสียงแล้ว ในหัวข้อนี้จะทำการคำนวณหาค่าความมีส่วนร่วมในแต่ละมิติของลักษณะเฉพาะสัดส่วน 20: 60: 20 ในชุดข้อมูลฝึกฝน ซึ่งเป็นชุดข้อมูลที่ใช้ในการแทนลักษณะของหน่วยเสียง เพื่อใช้สร้างแบบจำลองเสียงของแต่ละหน่วยเสียง โดยการคำนวณหาค่าความมีส่วนร่วมนี้ เป็นการวิเคราะห์หาความสำคัญในแต่ละมิติว่ามีความสำคัญต่อการแทนหน่วยเสียงในแบบจำลองเสียงต่างๆมากน้อยเพียงไร ซึ่งสำหรับการทดลองนี้ จะทำการหาค่าความมีส่วนร่วมโดยมุ่งเน้นดูค่าความมีส่วนร่วมโดยรวมของลักษณะเฉพาะที่มีการทำงานแบบอาศัยเซกเมนต์แต่ละส่วน ได้แก่ ส่วนหน้า ส่วนกลาง และส่วนหลัง เพื่อวิเคราะห์ถึงความสำคัญในแต่ละส่วนของเซกเมนต์ว่ามีความสำคัญต่อการสร้างแบบจำลองเสียงที่มีการทำงานแบบอาศัยเซกเมนต์อย่างไร โดยการคำนวณหาค่าความมีส่วนร่วมดังกล่าว สามารถหาได้จากวิธีการที่ได้นำเสนอไปในหัวข้อ การวิเคราะห์ดิสคริมิแนนต์เชิงเส้น ในส่วนการคำนวณหาค่าความมีส่วนร่วม ซึ่งผลการคำนวณหาค่าความมีส่วนร่วมของลักษณะเฉพาะที่มีการทำงานแบบอาศัยเซกเมนต์ที่มีอัตราส่วนในการแบ่ง 3 ส่วนด้วยสัดส่วน 20: 60: 20 แสดงในตารางที่5.8 โดยชุดทางด้านซ้ายมือเป็นค่าความมีส่วนร่วมที่เรียงลำดับตามมิติของลักษณะเฉพาะ ส่วนด้านขวามือเป็นค่าความมีส่วนร่วมที่ได้เรียงลำดับความสำคัญของแต่ละมิติแล้ว โดยใน 13 มิติแรกเป็นค่าสัมประสิทธิ์เมลฟรีเคว็นซีเคพตรอลที่ได้จากเซกเมนต์ในส่วนหน้า มิติที่ 14-26 เป็นค่าสัมประสิทธิ์เมลฟรีเคว็นซีเคพตรอลที่ได้จากเซกเมนต์ในส่วนกลาง มิติที่ 27-39 เป็นค่าสัมประสิทธิ์เมลฟรีเคว็นซีเคพตรอลที่ได้จากเซกเมนต์ในส่วนหลัง และมิติที่ 40 เป็นขนาดเซกเมนต์ของหน่วยเสียงนั้นๆ

จากผลการคำนวณค่าความมีส่วนร่วมตามตารางที่5.8 นั้น จะเห็นได้ชัดเจนว่าใน 10 อันดับแรกที่มีความสำคัญหรือมีค่าความมีส่วนร่วมสูงสุดนั้น มาจากเซกเมนต์ในส่วนกลางถึง 6 มิติ ซึ่งผลดังกล่าวเป็นสิ่งสำคัญที่ช่วยสนับสนุนผลการวิเคราะห์ลักษณะสัญญาณเสียงที่เสนอในขั้นตอนการหาลักษณะเฉพาะว่า ส่วนกลางของเซกเมนต์นั้นเป็นส่วนที่สามารถแทนสัญญาณเสียงหน่วยเสียงต่างๆได้อย่างชัดเจนที่สุด โดยผลรวมของค่าความมีส่วนร่วมของมิติต่างๆในเซกเมนต์ส่วนต่างๆ มีค่า ดังนี้

ตารางที่5.7 ค่าความมีส่วนร่วมรวมในแต่ละส่วนของเซกเมนต์

ตำแหน่งของส่วนในเซกเมนต์	ค่าความมีส่วนร่วมรวม
ส่วนหน้า	0.2459482
ส่วนกลาง	0.4403991
ส่วนหลัง	0.3136322

จากตารางแสดงค่าความมีส่วนร่วมรวมในแต่ละส่วนของเซกเมนต์ข้างต้นนั้น จะเห็นว่าส่วนกลางเป็นส่วนที่มีค่าความมีส่วนร่วมรวมสูงที่สุด ส่วนค่าความมีส่วนร่วมรวมของส่วนหน้าและส่วนหลังมีค่าที่ใกล้เคียงกัน ซึ่งเมื่อเปรียบเทียบกับค่าความมีส่วนร่วมในส่วนกลางของเซกเมนต์นั้น พบว่า ค่าความมีส่วนร่วมของทั้งส่วนหน้าและส่วนหลังมีขนาดแตกต่างกันกับส่วนกลางเกือบเท่าตัว โดยสัดส่วนของค่าความมีส่วนร่วมของแต่ละส่วนของเซกเมนต์ทั้ง 3 ส่วน ดังกล่าวสามารถแสดงได้ดังรูปที่ 5.5



รูปที่ 5.5 กราฟแสดงสัดส่วนค่าความมีส่วนร่วมรวมของส่วนต่างๆในเซกเมนต์

จากรูปดังกล่าวข้างต้นสามารถแสดงให้เห็นได้อย่างชัดเจนว่า ค่าความมีส่วนร่วมในส่วนกลางนั้นมีขนาดสัดส่วนเกือบครึ่งหนึ่งของพื้นที่ทั้งหมด ซึ่งจากผลการหาค่าความมีส่วนร่วมรวมดังกล่าวสามารถบ่งชี้และยืนยันได้ว่า ส่วนกลางของเซกเมนต์ เป็นส่วนที่มีความสำคัญต่อการมีส่วนร่วมในการจำแนกหน่วยเสียงต่างๆที่สุดจริงตามที่ได้เสนอไว้ในขั้นตอนการวิเคราะห์พฤติกรรมของเซกเมนต์ของหน่วยเสียงภาษาไทยเพื่อทำการออกแบบการหาลักษณะเฉพาะ ที่ได้ทำการแบ่งเซกเมนต์ของหน่วยเสียงต่างๆออกเป็น 3 ส่วน ซึ่งกล่าวว่า ส่วนกลางของเซกเมนต์จะเป็นส่วนที่มีความสามารถในบ่งบอกลักษณะพฤติกรรมของหน่วยเสียงนั้นๆ ได้อย่างชัดเจนที่สุด ส่วนค่าความมีส่วนร่วมรวมของส่วนหน้าและส่วนหลัง ซึ่งเป็นส่วนที่สัญญาณเสียงมีความขึ้นกันกับหน่วยเสียงข้างเคียงนั้น พบว่าค่าความมีส่วนร่วมที่ได้มีค่าใกล้เคียงกัน

ส่วนค่าความมีส่วนร่วมของขนาดของเซกเมนต์ ซึ่งเป็นลักษณะเฉพาะในมิติที่ 40 นั้น พบว่า มีค่าน้อยมาก เมื่อเทียบกับมิติอื่นๆ ซึ่งอาจมีสาเหตุมาจากขนาดของเซกเมนต์แต่ละหน่วยเสียงมีขนาดไม่สม่ำเสมอ กล่าวคือ มีค่าการกระจายของขนาดของเซกเมนต์ภายในกลุ่มนั้นๆมาก ซึ่งจะเห็นได้ชัดจากหน่วยเสียงสระเสียงสั้นและสระเสียงยาวดังกราฟแสดงจำนวนหน่วยเสียงสระเสียงสั้นและเสียงยาวในขนาดต่างๆในรูปที่ 5.2 ซึ่งรูปดังกล่าวยังแสดงให้เห็นว่า ขนาดของสระเสียงสั้นและสระเสียงยาวนั้นยังมีขนาดที่ใกล้เคียงกันมาก ส่งผลให้ความยาวหรือขนาดของเซกเมนต์ของหน่วยเสียงนั้นๆ ยังคงเป็นลักษณะเฉพาะที่ยังไม่ดีพอที่จะใช้จำแนกหน่วยเสียงได้

ถึงแม้ว่าความยาวหรือขนาดของเซกเมนต์ของหน่วยเสียงจะเป็นสมบัติเดียวที่บ่งบอกถึงความแตกต่างระหว่างสระเสียงสั้นและสระเสียงยาวตามหลักการความหมายของสระเสียงสั้นและสระเสียงยาว จึงส่งผลให้ค่าความมีส่วนร่วมที่คำนวณได้นั้นมีค่าน้อยมาก เมื่อเปรียบเทียบกับค่าความมีส่วนร่วมในมิติอื่นๆ

ตารางที่5.8 ค่าความมีส่วนร่วมในแต่ละมิติของลักษณะเฉพาะที่มีการทำงานแบบอาศัยเซกเมนต์สัดส่วน 20: 60: 20โดยด้านซ้ายเรียงลำดับตามมิติ ด้านขวาเรียงลำดับตามขนาดของค่าความมีส่วนร่วม

มิติ	ค่าความมีส่วนร่วม
1	0.014259
2	0.069775
3	0.028345
4	0.027809
5	0.019945
6	0.0071263
7	0.01695
8	0.015841
9	0.01223
10	0.0044255
11	0.0092418
12	0.011599
13	0.0084016
14	0.040785
15	0.14232
16	0.03802
17	0.050194
18	0.035083
19	0.014498
20	0.027486
21	0.015823
22	0.038306
23	0.0093156
24	0.01918
25	0.0093885
26	0.01746
27	0.026566
28	0.062729
29	0.022697
30	0.033649
31	0.029646
32	0.016844
33	0.023646
34	0.018225
35	0.011359
36	0.012639
37	0.0090902
38	0.014988
39	0.014094
40	1.51E-05

มิติ	ค่าความมีส่วนร่วม
15	0.14232
2	0.069775
28	0.062729
17	0.050194
14	0.040785
22	0.038306
16	0.03802
18	0.035083
30	0.033649
31	0.029646
3	0.028345
4	0.027809
20	0.027486
27	0.026566
33	0.023646
29	0.022697
5	0.019945
24	0.01918
34	0.018225
26	0.01746
7	0.01695
32	0.016844
8	0.015841
21	0.015823
38	0.014988
19	0.014498
1	0.014259
39	0.014094
36	0.012639
9	0.01223
12	0.011599
35	0.011359
25	0.0093885
23	0.0093156
11	0.0092418
37	0.0090902
13	0.0084016
6	0.0071263
10	0.0044255
40	1.51E-05

การทดลองจำแนกหน่วยเสียงด้วยหลักการทำงานแบบอาศัยกรอบเวลาเพื่อเปรียบเทียบประสิทธิภาพ

ในหัวข้อนี้จะทำการทดลองจำแนกหน่วยเสียงด้วยแบบจำลองฮิดเดนมาร์คอฟ ซึ่งมีกระบวนการการทำงานแบบอาศัยกรอบเวลาในการหาลักษณะเฉพาะที่นำมาใช้แทนสัญญาณเสียงตามงานวิจัยหนึ่งในวิทยานิพนธ์ของเอกฤทธิ์ มณีน้อย เพื่อเปรียบเทียบประสิทธิภาพความถูกต้องที่เกิดขึ้นกับผลการทดลองจำแนกหน่วยเสียงที่มีการทำงานแบบอาศัยเซกเมนต์

1. **วิธีการทดลอง**

สำหรับการจำแนกหน่วยเสียงตามวิธีการทำงานแบบอาศัยกรอบเวลาต่างๆ ด้วยแบบจำลองฮิดเดนมาร์คอฟนั้น มีการกำหนดค่าต่างๆ ดังนี้

- แบ่งสัญญาณเสียงที่เข้ามาเป็นเฟรมๆเฟรมละ 10 มิลลิวินาที
- หาค่าลักษณะเฉพาะจากค่าสัมประสิทธิ์เมลฟรีเคว็นซีเคปตรอล 13 หลัก ค่าเดลต้า 13 หลัก และค่าสัมประสิทธิ์ความเร่ง 13 หลัก รวมทั้งสิ้น 39 หลัก ทุกๆ 5 มิลลิวินาที
- ใช้แบบจำลองฮิดเด็นมาร์คอฟแบบ 3 สถานะจากซ้ายไปขวา (3 state left-to-right HMM)
- ใช้ค่าการกระจายเกาส์เซียนแบบเดี่ยว (Single Gaussian output distribution)
- เนื่องจากในการทดลองนี้เป็นการทดลองเพื่อจำแนกหน่วยเสียงเท่านั้น ดังนั้นจึงใช้เฉพาะแบบจำลองเสียงเท่านั้น จะไม่มีการคำนึงถึงแบบจำลองภาษา รวมทั้งสัญญาณเสียงที่นำเข้านั้น เป็นหน่วยเสียงที่ได้จากการตัดแบ่งสัญญาณเสียงต่อเนื่องออกเป็นหน่วยเสียงเรียบร้อยแล้ว

2. **ผลการทดลอง**

เมื่อทำการจำแนกหน่วยเสียงภาษาไทยจำนวน 52 หน่วยเสียงโดยกำหนดค่าต่างๆตามหัวข้อข้างต้น ปรากฏว่าได้ *% ความถูกต้อง*อยู่ที่ 56.95 เปอร์เซนต์ แสดงในตารางที่5.9 โดยแต่ละแถวจะแสดงผลการทดลองจำแนกหน่วยเสียงแต่ละหน่วยเสียงว่ามีความสับสนกับหน่วยเสียงใดๆ เท่าไร ซึ่งมีจำนวนแถวทั้งหมด 52 แถวตามจำนวนกลุ่มหน่วยเสียงที่ต้องการจำแนก ซึ่งแบ่งเป็น 4 ตารางย่อย โดยตำแหน่งของตารางดังกล่าว วางเรียงลำดับแต่ละตารางจากซ้ายไปขวา

ซึ่งเมื่อเปรียบเทียบกับหลักการทำงานแบบอาศัยเซกเมนต์แล้วพบว่า การจำแนกหน่วยเสียงโดยใช้ลักษณะเฉพาะที่ใช้หลักการแบบอาศัยเซกเมนต์ที่ให้ค่าน้ำหนักโอกาสในการตอบหน่วยเสียงใดๆเท่ากันให้ผลการจำแนกที่ดีกว่าอยู่ 4.46 เปอร์เซนต์ ส่วนผลการจำแนกหน่วยเสียงโดยให้ค่าน้ำหนักโอกาสในการตอบหน่วยเสียงใดๆขึ้นกับจำนวนหน่วยเสียงที่ใช้ฝึกฝนนั้น ให้ผลความถูกต้องสูงกว่าถึง 9.19 เปอร์เซนต์

ตารางที่5.9 ผลการจำแนกหน่วยเสียงด้วยแบบจำลองฮิดเดนมาร์คอฟโดยแจกแจงตามความสับสนระหว่างกลุ่ม

	a	**i**	**e**	**o**	**u**	**v**	**q**	**@**	**x**	**aa**	**ii**	**ee**	**oo**	**uu**
a	791	10	69	33	14	27	9	158	136	734	8	10	21	0
i	0	231	7	0	4	18	0	0	0	0	125	7	0	0
e	4	7	159	0	1	26	1	0	0	0	2	51	0	0
o	5	0	0	164	35	1	0	6	0	0	0	0	45	37
u	0	7	0	4	98	4	0	0	0	0	1	0	6	62
v	0	12	8	0	6	107	0	0	0	0	0	8	0	0
q	0	0	0	1	0	0	1	0	0	0	0	0	1	0
@	3	0	0	19	1	0	0	83	1	13	0	0	3	0
x	34	0	0	0	0	0	0	0	82	15	0	1	0	0
aa	368	0	7	13	0	0	2	115	86	1349	1	1	1	0
ii	0	138	1	0	4	10	0	0	0	0	463	1	0	0
ee	1	12	20	0	0	18	0	0	0	0	7	71	0	0
oo	0	0	1	34	23	1	1	2	0	0	1	0	55	18
uu	0	0	0	5	70	2	0	1	0	0	0	0	14	160
vv	0	8	0	0	6	23	0	0	0	0	2	0	0	1
qq	4	0	4	4	2	4	8	0	0	0	0	0	1	0
@@	8	0	0	106	2	0	0	176	2	27	0	1	13	1
xx	46	0	11	0	0	0	1	0	67	37	0	0	0	0
iia	0	13	10	1	0	9	0	0	0	0	9	9	0	0
vva	6	0	2	0	5	23	1	0	1	0	1	2	6	0
uua	3	0	0	25	24	0	1	2	0	0	0	0	11	21
f	0	0	0	0	0	0	0	0	0	0	0	0	0	0
s	0	0	0	1	0	0	0	0	0	0	0	0	0	0
p	2	0	0	0	0	0	0	0	0	0	0	0	0	0
t	0	0	0	0	0	0	0	0	0	0	0	0	0	0
k	0	0	0	0	0	0	0	0	0	0	0	0	0	0
b	0	0	0	0	0	0	0	0	0	0	0	0	0	0
d	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ph	0	0	0	0	0	0	0	0	0	0	0	0	0	0
th	0	0	0	0	0	0	0	0	0	0	0	0	0	0
kh	0	0	0	0	0	0	0	0	0	0	0	0	0	0
c	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ch	0	0	0	0	0	0	0	0	0	0	0	0	0	0
m	0	0	0	0	6	6	0	0	0	0	1	0	0	6
n	0	1	0	0	3	9	0	0	0	0	0	0	0	2
ng	1	0	0	0	3	3	0	0	0	0	0	0	0	0
vb	0	0	0	0	0	0	0	0	0	0	0	0	0	0
m^	0	0	0	0	0	4	0	0	0	0	3	0	0	0
n^	2	0	0	1	8	28	0	0	0	2	3	0	0	1
ng^	5	0	0	2	38	10	0	2	1	0	1	0	0	6
j^	3	97	75	0	3	42	0	0	19	1	73	32	1	0
w^	2	0	0	12	11	3	0	4	0	2	0	1	4	28
l^	0	0	0	0	1	0	0	0	0	0	0	0	0	0
f^	0	0	0	0	0	0	0	0	0	0	0	0	0	0
s^	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ch^	0	0	0	0	0	0	0	0	0	0	0	0	0	0
l	9	6	3	4	18	31	1	2	0	1	23	4	8	3
r	27	2	2	5	20	3	0	6	8	0	1	1	7	1
h	0	0	0	0	0	0	0	0	0	1	2	0	0	0
w	1	0	0	5	6	1	0	5	0	0	0	0	2	12
j	0	9	2	0	0	2	0	0	0	0	37	2	0	0
sil	0	0	0	0	0	0	0	0	0	0	0	1	0	0

ตารางที่ 5.9 ผลการจำแนกหน่วยเสียงด้วยแบบจำลองฮิดเดนมาร์คอฟโดยแจกแจงตามความสับสนระหว่างกลุ่ม

	vv	**qq**	**@@**	**xx**	**iia**	**vva**	**uua**	**f**	**s**	**p**	**t**	**k**	**b**	**d**
a	1	28	121	154	25	72	12	0	0	8	1	6	0	3
i	9	0	0	0	2	2	0	0	0	2	0	2	0	10
e	1	3	0	8	15	19	0	0	0	0	0	0	0	0
o	0	3	6	0	0	5	71	0	0	0	0	0	0	0
u	9	0	0	0	2	1	8	0	0	0	0	0	0	1
v	46	2	0	0	1	13	0	0	0	0	0	1	0	0
q	0	3	0	0	0	9	1	0	0	0	0	0	0	0
@	0	0	66	1	0	0	1	0	0	0	0	0	0	0
x	0	4	1	70	1	2	0	0	0	0	0	0	0	0
aa	0	24	93	177	11	18	7	0	0	1	0	0	0	0
ii	1	0	0	0	9	0	0	0	0	0	0	0	0	1
ee	0	2	0	2	9	5	0	0	0	0	0	0	0	0
oo	0	4	4	0	0	4	27	0	0	0	0	0	0	0
uu	2	1	0	0	0	1	7	0	0	0	0	0	0	0
vv	106	0	0	0	1	14	0	0	0	0	0	0	0	0
qq	4	85	0	1	0	40	1	0	0	0	0	0	0	0
@@	0	2	353	1	0	1	15	0	0	0	0	0	0	0
xx	0	2	0	168	4	3	0	0	0	0	0	0	0	1
iia	1	0	0	0	104	9	0	0	0	0	0	0	0	0
vva	19	15	0	0	9	141	2	0	0	0	0	0	0	0
uua	0	1	2	0	0	3	180	0	0	0	0	0	0	0
f	0	0	0	0	0	0	0	86	8	0	0	0	0	0
s	0	0	0	0	0	0	0	34	755	0	0	0	0	0
p	0	0	0	0	0	0	0	0	0	5	1	0	0	1
t	0	0	0	0	0	0	0	1	0	0	5	0	0	2
k	0	0	0	0	0	0	0	8	1	5	8	23	0	4
b	0	0	0	0	0	0	0	0	0	0	0	0	3	1
d	0	0	0	0	0	0	0	0	0	1	0	0	0	5
ph	0	0	0	0	0	0	0	5	0	20	1	2	1	0
th	0	0	0	0	0	0	0	24	2	4	22	2	0	5
kh	0	0	0	0	0	0	0	17	12	6	5	15	0	0
c	0	0	0	0	0	0	0	0	4	1	4	1	0	2
ch	0	0	0	0	0	0	0	4	15	0	1	0	0	0
m	9	0	0	0	0	2	0	0	0	3	0	0	5	0
n	18	0	0	0	0	1	0	0	0	7	0	0	6	3
ng	4	0	0	0	0	0	1	0	0	0	0	0	4	0
vb	0	0	0	0	0	0	0	0	0	8	1	0	18	8
m^	18	0	0	0	0	1	0	0	0	0	0	0	6	0
n^	36	0	0	0	0	3	0	0	0	3	2	0	5	3
ng^	25	2	1	0	0	2	1	0	0	6	1	0	6	1
j^	11	6	0	22	17	22	0	0	0	1	0	1	0	7
w^	6	3	7	0	0	2	6	0	0	0	0	0	0	3
l^	0	0	0	0	0	0	0	0	0	0	0	0	1	0
f^	0	0	0	0	0	0	0	13	0	0	0	0	0	0
s^	0	0	0	0	0	0	0	0	11	0	0	0	0	0
ch^	0	0	0	0	0	0	0	0	1	0	0	0	0	0
l	52	6	4	0	2	26	4	0	0	14	4	1	2	15
r	6	0	4	1	1	16	5	5	0	71	26	19	2	48
h	0	0	0	0	0	1	0	1	0	2	0	1	2	0
w	1	0	6	0	0	0	4	0	0	5	1	0	28	2
j	2	0	0	0	3	2	0	0	0	0	0	0	0	11
sil	1	0	0	0	0	0	0	32	5	386	203	44	8	121

ตารางที่ 5.9 ผลการจำแนกหน่วยเสียงด้วยแบบจำลองฮิดเดนมาร์คอฟโดยแจกแจงตามความสับสนระหว่างกลุ่ม

	ph	**th**	**kh**	**c**	**ch**	**m**	**n**	**ng**	**vb**	**m^**	**n^**	**ng^**	**j^**	**w^**
a	16	2	7	0	0	0	3	1	0	0	1	3	39	29
i	0	0	0	0	0	0	3	0	2	0	1	0	24	0
e	0	0	0	0	0	0	0	0	0	0	0	0	32	0
o	0	0	0	0	0	0	0	1	0	0	0	1	1	90
u	0	0	1	0	0	3	1	2	0	1	3	12	1	10
v	0	0	0	0	0	0	1	0	0	1	0	1	5	1
q	0	0	0	0	0	0	0	0	0	0	0	0	0	0
@	0	0	1	0	0	0	0	0	0	0	0	0	0	9
x	0	0	0	0	0	0	0	0	0	0	0	1	1	0
aa	1	2	0	0	0	0	0	0	0	0	0	0	2	4
ii	0	0	0	0	0	0	2	0	0	0	0	0	15	0
ee	0	0	0	0	0	0	0	0	0	0	0	0	27	0
oo	0	0	0	0	0	0	0	1	0	0	0	0	0	25
uu	0	0	0	0	0	1	0	5	0	2	2	13	0	3
vv	0	0	0	0	0	4	2	0	1	8	5	4	3	0
qq	0	0	0	0	0	0	0	0	0	0	0	5	0	0
@@	0	0	0	0	0	0	0	0	0	0	0	0	0	37
xx	0	0	0	0	0	0	0	1	0	0	0	0	1	0
iia	0	0	0	0	0	0	0	0	0	0	0	0	4	0
vva	0	0	0	0	0	0	0	0	0	0	0	2	0	0
uua	0	0	0	0	0	1	0	4	0	0	0	3	0	5
f	1	1	0	1	2	0	0	0	0	0	0	0	0	0
s	0	0	0	18	101	0	0	0	0	0	0	0	0	0
p	2	0	0	0	0	0	0	0	0	0	0	0	0	0
t	2	4	0	2	0	0	0	0	0	0	0	0	0	0
k	11	8	45	18	3	0	0	0	0	0	0	0	0	0
b	0	0	0	0	0	1	0	0	4	1	0	0	0	0
d	0	0	0	0	0	0	0	0	4	0	0	0	0	0
ph	331	63	64	1	15	1	0	0	0	0	0	0	0	0
th	183	301	31	34	75	0	0	0	0	0	0	0	0	0
kh	74	107	624	9	40	0	0	0	1	0	0	0	0	0
c	0	9	1	199	14	0	0	0	0	0	0	0	0	0
ch	0	4	1	39	268	0	0	0	0	0	0	0	0	0
m	2	0	0	0	0	262	40	11	74	101	18	19	1	0
n	2	0	0	0	0	73	289	19	92	69	97	6	1	0
ng	0	0	0	0	0	13	15	13	7	3	11	20	1	0
vb	2	0	2	1	0	9	7	0	496	14	23	0	0	0
m^	3	0	0	0	0	31	8	1	84	387	84	44	1	0
n^	2	0	0	0	0	11	107	8	273	197	838	40	6	0
ng^	7	0	1	0	0	30	17	29	54	91	203	640	8	28
j^	1	3	0	0	0	0	0	1	0	0	9	5	632	2
w^	1	0	0	0	0	6	0	0	0	10	0	14	3	145
l^	0	0	0	0	0	0	0	0	0	0	0	0	0	4
f^	1	1	1	0	0	0	0	0	0	0	0	0	0	0
s^	0	1	0	0	2	0	0	0	0	0	0	0	0	0
ch^	0	0	0	0	4	0	0	0	0	0	0	0	0	0
l	7	4	3	0	0	17	19	17	7	4	5	3	5	1
r	31	12	7	7	1	4	5	9	1	3	2	0	2	3
h	45	32	18	0	1	1	1	0	3	0	0	1	0	1
w	5	0	8	0	0	15	3	13	30	2	5	7	0	0
j	0	1	0	2	1	0	1	0	7	0	0	0	1	0
sil	188	44	67	39	1	0	0	0	99	3	5	3	1	0

ตารางที่ 5.9 ผลการจำแนกหน่วยเสียงด้วยแบบจำลองฮิดเดนมาร์คอฟโดยแจกแจงตามความสับสนระหว่างกลุ่ม

	l^	f^	s^	ch^	l	r	h	w	j	sil	%
a	0	0	0	0	34	69	83	6	15	0	28.67
i	0	0	0	0	2	3	0	1	34	2	47.05
e	0	0	0	0	4	2	1	0	5	0	46.63
o	3	0	0	0	3	4	0	15	0	0	33.06
u	2	0	0	0	3	5	0	10	0	0	38.13
v	0	0	0	0	11	0	0	0	2	0	47.35
q	0	0	0	0	0	0	0	0	0	0	6.25
@	1	0	0	0	1	0	0	1	0	0	40.69
x	0	0	0	0	0	2	2	0	0	0	37.96
aa	0	0	0	0	4	8	25	0	0	0	58.15
ii	0	0	0	0	0	0	0	0	47	0	66.91
ee	0	0	0	0	3	0	0	0	4	0	39.23
oo	0	0	0	0	4	1	0	2	0	0	26.44
uu	3	0	0	0	3	0	0	3	0	0	53.69
vv	0	0	0	0	6	0	0	0	4	0	53.54
qq	0	0	0	0	14	0	1	0	0	0	47.75
@@	1	0	0	0	0	2	1	0	0	0	47.13
xx	0	0	0	0	0	1	3	0	0	0	48.55
iia	0	0	0	0	1	0	0	0	10	0	57.78
vva	0	0	0	0	17	4	0	1	1	0	54.65
uua	2	0	0	0	6	0	0	11	0	0	59.02
f	0	3	1	1	0	0	0	0	0	0	82.69
s	0	1	69	37	0	0	0	0	0	0	74.31
p	0	0	0	0	1	2	0	0	0	0	35.71
t	0	0	0	0	0	2	0	0	0	1	26.32
k	0	1	0	0	0	1	0	0	0	0	16.91
b	0	0	0	0	0	0	0	0	0	0	30.00
d	0	0	0	0	0	0	0	0	0	0	50.00
ph	0	2	0	0	0	0	10	2	0	4	63.41
th	0	3	3	1	0	0	17	0	0	0	42.57
kh	0	2	0	0	0	0	4	0	0	0	68.12
c	0	0	2	2	0	0	0	0	0	0	83.26
ch	0	0	0	0	0	0	0	0	0	0	80.72
m	0	0	0	0	16	0	1	6	1	0	44.41
n	0	0	0	0	30	3	2	0	3	0	39.27
ng	0	0	0	0	3	3	0	0	0	0	12.38
vb	0	0	0	0	1	0	0	1	1	2	83.50
m^	0	0	0	0	4	0	0	0	2	0	56.83
n^	0	0	0	0	16	7	0	2	7	1	51.99
ng^	2	0	0	0	7	1	10	9	3	0	51.20
j^	0	0	0	0	7	6	4	0	9	0	56.83
w^	6	0	0	0	3	2	3	2	0	1	50.00
l^	0	0	0	0	0	0	0	0	0	0	0.00
f^	0	8	0	0	0	0	0	0	0	0	33.33
s^	0	0	2	6	0	0	0	0	0	0	9.09
ch^	0	0	0	0	0	0	0	0	0	0	0.00
l	0	0	0	0	275	61	9	20	25	0	37.93
r	0	0	0	0	89	482	10	26	5	0	48.88
h	0	0	0	0	1	1	135	0	0	1	53.78
w	3	0	0	0	7	4	10	165	0	0	46.35
j	0	0	0	0	2	0	1	1	166	0	65.61
sil	1	140	7	4	1	49	18	1	0	5401	78.58
										% ความถูกต้องเฉลี่ย	56.95

การเปรียบเทียบประสิทธิภาพของการจำแนกหน่วยเสียงระหว่างวิธีการทำงานแบบเซกเมนต์และวิธีการทำงานแบบอาศัยกรอบเวลา

จากผลการทดลองทั้งหมดข้างต้นนั้น ถือว่าผลความถูกต้องโดยรวมของการจำแนกหน่วยเสียงโดยใช้ลักษณะเฉพาะที่มีการทำงานแบบอาศัยเซกเมนต์มีประสิทธิภาพในการจำแนกสูงกว่าการจำแนกหน่วยเสียงที่ใช้ลักษณะเฉพาะแบบอาศัยกรอบเวลา โดยการจำแนกหน่วยเสียงที่ใช้ลักษณะเฉพาะหลักการแบบอาศัยเซกเมนต์และให้ค่าน้ำหนักโอกาสในการตอบหน่วยเสียงใดๆ เท่ากันให้ผลการจำแนกได้ดีกว่าอยู่ 4.46 เปอร์เซนต์ ส่วนผลการจำแนกหน่วยเสียงโดยให้ค่าน้ำหนักโอกาสในการตอบหน่วยเสียงใดๆขึ้นกับจำนวนหน่วยเสียงที่ใช้ฝึกฝนนั้น ให้ผลความถูกต้องสูงกว่าถึง 9.19 เปอร์เซนต์

โดยจากการสังเกตผลการจำแนกหน่วยเสียงด้วยลักษณะเฉพาะที่มีการทำงานแบบอาศัยเซกเมนต์ ซึ่งจำแนกด้วยหลักการการวิเคราะห์ดิสคริมิแนนต์เชิงเส้นแบบใช้ค่าน้ำหนักที่มีโอกาสจำแนกเป็นหน่วยเสียงใดๆเท่าๆกันกับการจำแนกหน่วยเสียงด้วยแบบจำลองฮิดเดนมาร์คอฟ ซึ่งใช้ลักษณะเฉพาะที่มีการทำงานแบบอาศัยกรอบเวลา พบว่า ผลการจำแนกหน่วยเสียงในส่วนสระโดยรวมแล้ว ถือว่าวิธีการทำงานแบบอาศัยเซกเมนต์สามารถจำแนกผลได้ดีกว่าถึง 7 หน่วยเสียงจากจำนวนหน่วยเสียงสระเสียงสั้น 9 หน่วยเสียง ซึ่ง 3 ใน 7 หน่วยเสียงนั้น เป็นหน่วยเสียงที่ใช้ชุดข้อมูลฝึกฝนน้อยที่สุด 3 อันดับแรก เมื่อเทียบกับจำนวนข้อมูลหน่วยเสียงอื่นๆในกลุ่มสระ ได้แก่ สระเออะ(/q/), สระแอะ (/x/), สระเอาะ (/@/) รวมไปถึงสระที่มีจำนวนข้อมูลในการฝึกฝนมากที่สุดเป็น 3 อันดับแรกก็สามารถจำแนกได้ดีกว่าการทำงานแบบอาศัยกรอบเวลาด้วย ส่วนสระเสียงยาวนั้นปรากฏว่า การใช้ลักษณะเฉพาะที่มีการทำงานแบบอาศัยเซกเมนต์สามารถทำได้ดีกว่าทุกตัว ยกเว้นสระออ (/@@/) ที่สามารถจำแนกได้ผลเท่ากัน ส่วนสระผสมทั้ง 3 ตัวนั้น ลักษณะเฉพาะแบบอาศัยเซกเมนต์สามารถทำได้ดีกว่าทุกๆตัว

หน่วยเสียงต่อมาคือ หน่วยเสียงในกลุ่มพยัญชนะ โดยเริ่มจากพยัญชนะเสียดแทรกปรากฏว่า หน่วยเสียงเสียดแทรกประเภทพยัญชนะตัวสะกดนั้น ลักษณะเฉพาะแบบอาศัยเซกเมนต์ให้ผลการจำแนกดีกว่าทั้ง 2 ตัว แต่ในทางกลับกันสำหรับพยัญชนะต้น ลักษณะเฉพาะแบบอาศัยกรอบเวลากลับจำแนกได้ดีกว่า เช่นเดียวกับกลุ่มพยัญชนะเสียงเสียดแทรก โดยเฉพาะตัวสะกด “ช” (/ch^/) ที่มีจำนวนข้อมูลในการฝึกฝนน้อยมากๆพบว่าคือ ลักษณะเฉพาะแบบอาศัยกรอบเวลาไม่สามารถจำแนกได้เลย ส่วนกลุ่มพยัญชนะกักนั้นพบว่า ลักษณะเฉพาะแบบอาศัยเซกเมนต์ก็สามารถทำได้ดีกว่าเกือบทุกตัว ยกเว้นหน่วยเสียง “พ”, “ภ”, “ผ” (/ph/), “ค”, “ข”, “ฆ” (/kh/) 2 หน่วยเสียงที่ลักษณะเฉพาะแบบอาศัยกรอบเวลาทำได้ดีกว่าโดยพยัญชนะกักตัวที่ลักษณะเฉพาะแบบอาศัยเซกเมนต์ทำได้ดีกว่าสูงมากๆคือ หน่วยเสียง “ก” (/k/) ที่สูงกว่าถึง 32.20 เปอร์เซนต์ ซึ่งเป็นหน่วยเสียงที่ลักษณะเฉพาะแบบอาศัยกรอบเวลาสามารถทำได้ต่ำที่สุดเพียง 16.91 เปอร์เซนต์ เท่านั้น

ในขณะที่งานวิจัยนี้ทำได้ 49.11 เปอร์เซนต์ นอกจากนี้ ลักษณะเฉพาะแบบอาศัยเซกเมนต์ของพยัญชนะกักตัวอื่นๆที่สามารถจำแนกได้ดีกว่า ก็ให้ผล *% ความถูกต้อง* ที่ดีกว่าแตกต่างกันอย่างชัดเจนเช่นเดียวกัน สำหรับกลุ่มพยัญชนะกลุ่มสุดท้ายคือ กลุ่มพยัญชนะนาสิก ซึ่งลักษณะเฉพาะแบบอาศัยเซกเมนต์ยังคงทำได้ดีกว่าในเกือบทุกๆหน่วยเสียงอีกเช่นกัน โดยในหลายหน่วยเสียงจะเห็นว่าลักษณะเฉพาะแบบอาศัยเซกเมนต์สามารถทำได้ดีกว่าอย่างชัดเจน ได้แก่ หน่วยเสียงพยัญชนะ “ม” (/m/) ที่ต่างกัน 10.39 เปอร์เซนต์

สำหรับหน่วยเสียงในกลุ่มกึ่งสระนั้น ลักษณะเฉพาะแบบอาศัยเซกเมนต์ให้ผลการจำแนกหน่วยเสียงได้ดีกว่า 5 ใน 8 หน่วยเสียง นอกจากนี้ยังสามารถสังเกตได้ว่า ในหน่วยเสียงตัวสะกด “ล” (/l^/) ซึ่งถือได้ว่ามีจำนวนหน่วยเสียงที่น้อยมากๆนั้น ลักษณะเฉพาะแบบอาศัยกรอบเวลาไม่สามารถจำแนกได้เลย

หน่วยเสียงกลุ่มสุดท้าย คือ เสียงเงียบ (/sil/) ซึ่งเป็นอีกหน่วยเสียงหนึ่งที่ลักษณะเฉพาะแบบอาศัยเซกเมนต์สามารถทำได้ดีกว่าสูงถึง 12.69 เปอร์เซนต์ โดยหน่วยเสียงนี้อาจวิเคราะห์ได้ว่ามีความเป็นไปได้ว่าการทำงานแบบอาศัยเซกเมนต์จะมีความเหมาะสมมากกว่า เนื่องจากลักษณะเสียงเงียบนั้น ควรเป็นหน่วยเสียงที่ไม่ปรากฏสัญญาณใดๆเลย แต่ในความเป็นจริงนั้น สัญญาณที่เป็นเสียงเงียบดังกล่าวอาจเกิดสัญญาณรบกวนได้ ซึ่งถ้าเป็นการทำงานแบบอาศัยกรอบเวลาที่มีการตัดเฟรมขนาดเล็กๆแล้วก็อาจทำให้เฟรมๆนั้นมีลักษณะที่ใกล้เคียงกับหน่วยเสียงอื่นมากกว่าทำให้เกิดความสับสนได้ แต่ลักษณะเฉพาะแบบอาศัยเซกเมนต์จะเป็นการเก็บลักษณะของสัญญาณเสียงที่มองหน่วยเสียงดังกล่าวในภาพรวมของเซกเมนต์ ดังนั้น ถึงแม้ว่าจะมีสัญญาณรบกวนเข้ามาปะปนในบางบริเวณของเซกเมนต์เสียงเงียบ แต่ลักษณะเฉพาะที่ได้ก็ยังคงสามารถแทนความเป็นเสียงเงียบได้ดีอยู่

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 6
สรุปผลการวิจัย

ผลการวิจัย

งานวิจัยนี้เป็นการหาลักษณะเฉพาะที่รองรับหลักการทำงานแบบอาศัยเซกเมนต์ ซึ่งหลักการดังกล่าวจะทำการแบ่งสัญญาณเสียงเป็นหน่วยย่อยๆที่เรียกว่า หน่วยเสียง ก่อนนำไปเข้าสู่กระบวนการการหาลักษณะเฉพาะ โดยงานวิจัยนี้ได้ทำการศึกษาหาลักษณะเฉพาะที่เหมาะสมกับหน่วยเสียงภาษาไทยที่ถูกตัดแบ่งออกเป็นเซกเมนต์ดังกล่าว และได้ทำการทดลองจำแนกหน่วยเสียงภาษาไทยทั้งหมด 52 หน่วยเสียงด้วยหลักการการวิเคราะห์ดิสครีมิแนนต์เชิงเส้นโดยใช้ลักษณะเฉพาะตามที่ได้ออกแบบไว้ รวมทั้งได้ทำการทดลองจำแนกหน่วยเสียงด้วยแบบจำลองฮิดเดนมาร์คอฟซึ่งมีการใช้ลักษณะเฉพาะที่มีการทำงานแบบอาศัยกรอบเวลา เพื่อเปรียบเทียบประสิทธิภาพที่ได้ระหว่างการใช้ลักษณะเฉพาะที่มีการทำงานแบบอาศัยเซกเมนต์และลักษณะเฉพาะที่มีการทำงานแบบอาศัยกรอบเวลา โดยรายละเอียดขั้นตอนการทำงานในส่วนต่างๆโดยสรุปมีดังนี้

สำหรับการทำงานวิจัยนี้ในขั้นแรกเริ่มต้นด้วยการศึกษาลักษณะของสัญญาณเสียงที่ถูกแบ่งออกเป็นหน่วยเสียง ซึ่งพบว่า ในความเป็นจริงแล้วหน่วยเสียงต่างๆในสัญญาณเสียงพูดต่อเนื่องนั้นพบว่า ไม่ได้มีความสม่ำเสมอกันทั้งเซกเมนต์ รวมไปถึงในหน่วยเสียงเดียวกันแต่ละเซกเมนต์ก็อาจมีรูปร่างต่างกันได้ โดยเฉพาะบริเวณขอบของเซกเมนต์ เนื่องจากในการพูดทุกๆหน่วยเสียงจำเป็นต้องมีการพูดหรือออกเสียงกันอย่างต่อเนื่อง จึงจะสามารถออกมาเป็นคำที่ใช้สื่อสารกันได้ จึงส่งผลให้บริเวณขอบของหน่วยเสียงแต่ละหน่วยมีการแปรผันกับหน่วยเสียงที่อยู่ติดกัน เนื่องจากการเลื่อนที่ของอวัยวะที่ใช้ในการออกเสียงจากหน่วยเสียงหนึ่งไปยังอีกหน่วยเสียงหนึ่ง ดังนั้นงานวิจัยนี้จึงมีแนวคิดในการแบ่งเซกเมนต์ของหน่วยเสียงดังกล่าวออกเป็น 3 ส่วนก่อนนำไปหาค่าสัมประสิทธิ์เมลฟรีเคว็นซีเคพสตรอลของแต่ละส่วนรวม 39 หลักและความยาวของเซกเมนต์อีก 1 หลัก รวมเป็นลักษณะเฉพาะขนาด 40 หลักในการแทนสัญญาณเสียง 1 เซกเมนต์ ซึ่งได้ทำการทดลองแบ่งสัดส่วนการแบ่งเซกเมนต์ออกเป็นสัดส่วนต่างๆ 37 แบบ เพื่อนำไปหาลักษณะเฉพาะของหน่วยเสียงต่างๆ

ขั้นต่อมาจึงได้นำลักษณะเฉพาะสัดส่วนต่างๆไปทำการทดลองจำแนกหน่วยเสียงภาษาไทยทั้งหมด 52 หน่วยเสียง เพื่อสังเกตผลการจำแนกหน่วยเสียงที่ใช้ลักษณะเฉพาะในแต่ละแบบแทนหน่วยเสียง เท่านั้น ยังไม่ได้เป็นการวัดประสิทธิภาพความถูกต้องในการจำแนกหน่วยเสียงจริงๆ ซึ่งปรากฏว่า สัดส่วนลักษณะเฉพาะที่ให้ผลการจำแนกสูงที่สุดในจำนวน 37 สัดส่วน คือ ลักษณะเฉพาะที่มีการแบ่ง 3 ส่วนในสัดส่วน 20: 60: 20 ซึ่งสามารถจำแนกหน่วยเสียงได้ 60.00 เปอร์เซนต์

ส่วนลักษณะเฉพาะแบบ ข5 ที่มีการแบ่งสัดส่วน 10: 30: 60 พบว่าได้ % *ความถูกต้อง* ในการจำแนกหน่วยเสียงต่ำสุดอยู่ที่ 55.96 เปอร์เซนต์

หลังจากนั้นจึงได้ทำการทดสอบนัยสำคัญ เพื่อวัดความแตกต่างของแต่ละวิธีว่า ถึงแม้ว่าผลการจำแนกที่ได้มีค่าใกล้เคียงกัน แต่ความแตกต่างที่เกิดขึ้นดังกล่าวนั้น จะมีนัยสำคัญต่อกันหรือไม่ โดยได้ทดสอบนัยสำคัญเฉพาะระหว่างลักษณะเฉพาะที่ให้ผลการจำแนกดีที่สุดกับลักษณะเฉพาะแบบอื่นๆซึ่งจากการทดสอบปรากฏว่า ความแตกต่างที่เกิดขึ้นไม่มีนัยสำคัญต่อการใช้ลักษณะเฉพาะที่มีสัดส่วน 20: 60: 20 คือ การใช้ลักษณะเฉพาะ15: 70: 15, 25: 50: 25, 10: 60: 30, 10: 50: 40, 15: 60: 25, 20: 50: 30, 30: 50: 20 และ 45: 40: 15 รวม 8 แบบ ดังนั้นจึงกล่าวได้ว่า การเลือกใช้ลักษณะเฉพาะ 9 แบบ ไม่มีความแตกต่างกัน

หลังจากการทดสอบนัยสำคัญดังกล่าวแล้ว งานวิจัยนี้ได้เลือกลักษณะเฉพาะแบบ ก3 ซึ่งมีสัดส่วน 20: 60: 20 มาใช้ในการแทนสัญญาณเสียง เพื่อทำการจำแนกหน่วยเสียงหน่วยเสียงภาษาไทย โดยใช้ชุดข้อมูลทดสอบมาทดสอบจริง เพื่อวัดประสิทธิภาพความถูกต้อง ซึ่งปรากฏว่าการใช้ลักษณะเฉพาะแบบ ก3 สามารถจำแนกได้สูงสุด 61.14 เปอร์เซนต์ เมื่อให้ค่าของโอกาสที่จะตอบว่าเป็นหน่วยเสียงใดๆเท่าๆกัน และสามารถจำแนกได้สูงสุดถึง 66.41 เปอร์เซนต์ เมื่อให้ค่าของโอกาสที่จะตอบว่าเป็นหน่วยเสียงใดๆขึ้นกับจำนวนข้อมูลที่ใช้ในการฝึกฝน นอกจากนี้ เมื่อทดลองรวมจำนวนหน่วยเสียงที่จำแนกสับสนกันระหว่างสระเสียงสั้นและสระเสียงยาวจากผลการทดลองที่กำหนดให้ค่าของโอกาสตอบเป็นหน่วยเสียงใดๆไม่เท่ากันเข้าด้วยกันนั้น พบว่าผลที่ได้มีความถูกต้องเพิ่มขึ้นถึง 71.47 เปอร์เซนต์

สำหรับการทดลองสุดท้าย เป็นการทดลองจำแนกหน่วยเสียงที่ใช้ลักษณะเฉพาะที่มีหลักการทำงานแบบอาศัยกรอบเวลา เพื่อวัดเปรียบเทียบประสิทธิภาพที่เกิดขึ้นระหว่างการทำงานแบบอาศัยเซกเมนต์และการทำงานแบบอาศัยกรอบเวลา โดยผลการจำแนกดังกล่าวได้ผลอยู่ที่ 56.95 เปอร์เซนต์ ซึ่งถือว่าผลที่ได้น้อยกว่าผลการจำแนกจากงานวิจัยนี้อยู่ถึง 9.19 เปอร์เซนต์

คุณประโยชน์ของงานวิจัยนี้ต่อวงวิชาการ

กล่าวโดยสรุปแล้ว ถือว่างานวิจัยนี้เป็นงานวิจัยที่นำเสนอลักษณะเฉพาะที่มีการทำงานแบบอาศัยเซกเมนต์วิธีหนึ่ง ซึ่งมีความพยายามที่จะช่วยพัฒนาปรับปรุงให้ระบบรู้จำมีประสิทธิภาพดีขึ้น เนื่องจากลักษณะการทำงานที่เป็นที่นิยมใช้ในปัจจุบันนั้นเป็นลักษณะการทำงานแบบอาศัยกรอบเวลา ซึ่งยังคงมีพฤติกรรมหลายอย่างที่ขัดกับลักษณะธรรมชาติของเสียงพูด รวมไปถึงการทำงานแบบอาศัยกรอบเวลานั้นยังยากที่จะดึงลักษณะสำคัญบางประการออกจากสัญญาณเสียงด้วย ดังนั้นสิ่งเหล่านี้อาจเป็นสาเหตุหนึ่งที่ส่งผลให้ลักษณะเฉพาะดังกล่าวขาดสมบัติสำคัญในการแทนสัญญาณเสียงของหน่วยเสียงต่างๆได้อย่างมีประสิทธิภาพ ทำให้ส่งผลกระทบโดยตรงต่อประสิทธิภาพความสามารถในการทำงานของระบบรู้จำ

ดังนั้นในงานวิจัยนี้จึงได้นำหลักการการทำงานแบบอาศัยเซกเมนต์เข้ามาช่วยในการหาลักษณะเฉพาะแทน ซึ่งเป็นวิธีการที่ทำงานอยู่บนสัญญาณเสียง 1 หน่วยเสียง ทำให้สามารถดึงสมบัติความเป็นเซกเมนต์เข้ามาช่วยในการแทนสัญญาณเสียงเพื่อเพิ่มศักยภาพในการสร้างแบบจำลองเสียงให้สามารถแทนหน่วยเสียงต่างๆที่สามารถจำแนกหน่วยเสียงต่างๆออกจากกันได้อย่างชัดเจนยิ่งขึ้น ซึ่งจะส่งผลให้ประสิทธิภาพการทำงานของระบบรู้จำโดยรวมดียิ่งขึ้นด้วย

ซึ่งคุณประโยชน์ต่อวงวิชาการของงานวิจัยชิ้นนี้ สามารถกล่าวโดยสรุปเป็นหัวข้อต่างๆได้ดังนี้

- **ออกแบบการหาลักษณะเฉพาะที่เหมาะสมกับการทำงานแบบอาศัยเซกเมนต์**

ซึ่งในงานวิจัยนี้ได้ออกแบบวิธีการหาลักษณะเฉพาะโดยการแบ่งเซกเมนต์ของหน่วยเสียงใดๆออกเป็น 3 ส่วน โดยมีการกำหนดสัดส่วนในการแบ่งส่วนในอัตราส่วนต่างๆทั้งหมด 37 แบบ ซึ่งสามารถแบ่งเป็น 3 กลุ่มได้แก่ สัดส่วนที่ส่วนหน้าและส่วนหลังมีขนาดเท่ากัน สัดส่วนที่ส่วนหน้ามีขนาดเล็กกว่าส่วนหลัง และสัดส่วนที่ส่วนหน้ามีขนาดใหญ่กว่าส่วนหลัง เพื่อเลือกสัดส่วนหนึ่งขึ้นมาใช้ในการแบ่งสัญญาณเสียงในขั้นตอนการหาลักษณะเฉพาะ โดยได้นำลักษณะเฉพาะทั้ง 37 แบบมาทำการทดลองจำแนกหน่วยเสียง รวมทั้งในงานวิจัยนี้ได้เลือกลักษณะเฉพาะขึ้นมา 1 แบบคือ ลักษณะเฉพาะที่มีการแบ่งสัดส่วนแบบ 20: 60: 20 มาทดลองจำแนกหน่วยเสียงเพื่อสังเกตผล % *ความถูกต้อง* ในการจำแนกหน่วยเสียงที่มีการทำงานแบบอาศัยเซกเมนต์ ทั้งนี้ในการแบ่งสัดส่วนด้วยวิธีดังกล่าวไม่ได้ถือว่าเป็นสัดส่วนที่ดีที่สุดสำหรับการแบ่งเซกเมนต์ แต่ถือว่าเป็นเพียงวิธีหนึ่งที่สามารถแบ่งส่วนเซกเมนต์ เพื่อใช้ในการหาลักษณะเฉพาะที่มีการทำงานแบบอาศัยเซกเมนต์ที่ดีวิธีหนึ่งเท่านั้น

- **ลักษณะเฉพาะที่ได้มีจำนวนมิติคงที่และน้อยกว่าลักษณะเฉพาะที่มีการทำงานแบบอาศัยกรอบเวลา**

เนื่องจากลักษณะเฉพาะที่มีการทำงานแบบอาศัยกรอบเวลานั้นมีจำนวนมากและขนาดไม่คงที่ กล่าวคือ ขนาดของลักษณะเฉพาะที่ใช้นั้นจะขึ้นกับจำนวนเฟรมของสัญญาณเสียงที่ต้องการรู้จำ ซึ่งโดยปกติแล้ว ในการแบ่งเฟรมดังกล่าวขึ้น จะทำการแบ่งเป็นเฟรมที่มีขนาดเล็กๆ ซึ่งโดยทั่วไปนั้นอยู่ที่ขนาด 10 มิลลิวินาที ซึ่งถือว่าเป็นขนาดที่เล็กมาก ทำให้สามารถแบ่งสัญญาณเสียงออกเป็นเฟรมๆได้จำนวนมาก ซึ่งส่งผลทำให้ขนาดของลักษณะเฉพาะที่จะใช้ในการรู้จำเสียงทั้งหมดมีขนาดมากตามไปด้วย ซึ่งต่างกับการหาลักษณะเฉพาะที่มีการทำงานแบบอาศัยเซกเมนต์ ที่สามารถกำหนดขนาดของลักษณะเฉพาะให้มีขนาดคงที่ได้ รวมไปถึงกำหนดขนาดของลักษณะเฉพาะที่จะใช้ทั้งหมดได้ เนื่องจากวิธีการแบบอาศัยเซกเมนต์นี้เป็นจะทราบขนาดของหน่วยเสียงแต่ละหน่วยซึ่งเป็นสมบัติของการทำงานแบบอาศัยเซกเมนต์ ทำให้สามารถกำหนดวิธีการหาลักษณะเฉพาะของแต่ละเซกเมนต์ได้ โดยในงานวิจัยนี้ ได้ใช้ลักษณะเฉพาะที่มีขนาดคงที่คือ 40 หลัก ซึ่งคำนวณได้จากค่าสัมประสิทธิ์เมล

ฟรีเควิ้นซีเคปตรอล 12 หลักรวมกับค่าพลังงานในแต่ละส่วนของทั้ง 3 ส่วน รวมทั้งสิ้น 39 หลัก และรวมกับขนาดของเซกเมนต์อีก 1 หลัก

- **ลดเวลาการทำงานในส่วนขั้นตอนการหาลักษณะเฉพาะ**

เนื่องจากในงานวิจัยนี้ใช้วิธีการหาค่าสัมประสิทธิ์เมลฟรีเควิ้นซีเคปตรอล โดยกำหนดขนาดเฟรมเป็นขนาดของสัดส่วนที่กำหนดของเซกเมนต์ทั้งส่วน ส่งผลให้การหาค่าลักษณะเฉพาะของ 1 เซกเมนต์ใดๆมีจำนวนคงที่ กล่าวคือใช้เวลาในการหาเพียง 3 ครั้งต่อ 1 เซกเมนต์เท่านั้น เปรียบเทียบกับการหาลักษณะเฉพาะที่มีการทำงานแบบอาศัยกรอบเวลานั้นจะมีจำนวนที่ไม่แน่นอนโดยจะขึ้นกับจำนวนเฟรมที่สามารถแบ่งได้

- **เพิ่มประสิทธิภาพความถูกต้องในการจำแนกหน่วยเสียง ซึ่งเป็นขั้นตอนหนึ่งในกระบวนการรู้จำเสียง**

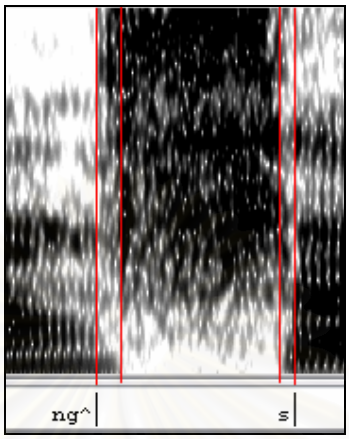
โดยในงานวิจัยนี้ได้ทำการทดลองเปรียบเทียบผลการจำแนกหน่วยเสียงโดยการใช้ลักษณะเฉพาะที่มีการทำงานแบบอาศัยเซกเมนต์ตามที่ได้นำเสนอในงานวิจัยนี้ เปรียบเทียบกับผลการทดลองจำแนกหน่วยเสียงใช้ลักษณะเฉพาะที่มีการทำงานแบบอาศัยกรอบเวลาในการแทนสัญญาณเสียงนั้น ปรากฏว่าผลการทดลองจำแนกหน่วยเสียงตามการทดลองในงานวิจัยนี้มีประสิทธิภาพความถูกต้องสูงกว่าอยู่ถึง 9.19 เปอร์เซนต์

ข้อเสนอแนะ

ในงานวิจัยนี้ยังสามารถพัฒนาส่วนอื่นๆ เพื่อแก้ปัญหาที่ยังเป็นอุปสรรคต่อผลการจำแนกหน่วยเสียงได้อีก เช่น ปัญหาเรื่องสระเสียงสั้นและเสียงยาว ซึ่งถึงแม้ว่างานวิจัยนี้สามารถทำได้ดีกว่าหลักการทำงานแบบอาศัยกรอบเวลาแต่ก็ยังคงมีความสับสนกันอยู่ ซึ่งอาจแก้ปัญหาโดยการคิดหาลักษณะเฉพาะอื่นๆเพิ่มเติม หรืออาจใช้หลักการแบบจำลองภาษาเข้ามาช่วยตัดสินใจในการตอบผลลัพธ์ในการจำแนกหน่วยเสียงได้

นอกจากนี้ยังอาจสามารถพัฒนากระบวนการการแบ่งส่วนลักษณะเฉพาะเพิ่มเติมได้ โดยไม่ต้องกำหนดสัดส่วนตายตัวตามงานวิจัยนี้ เช่น การแบ่งสัดส่วนของเซกเมนต์โดยอาศัยพฤติกรรมของสัญญาณเสียงแบบต่างๆเป็นเกณฑ์ในแบ่งสัดส่วน ตัวอย่างเช่น การหาค่าอัตสหสัมพันธ์ (Autocorrelation) ซึ่งเป็นเครื่องมือที่ใช้วัดความเป็นรายคาบของสัญญาณ (Periodicity) เนื่องจากในกรอบสัญญาณที่เป็นเสียงก้องนั้นจะมีค่าอัตสหสัมพันธ์ในช่วงความถี่ต่ำมากกว่ากรอบสัญญาณเสียงแบบไม่ก้อง ดังนั้นค่าอัตสหสัมพันธ์ของกรอบสัญญาณเสียงนี้จึงสามารถช่วยในการแบ่งส่วนภายในหน่วยเสียงพยัญชนะแบบไม่ก้อง ที่อาจมีส่วนของสัญญาณเสียงก้องปนเข้ามาได้ ตัวอย่างดังรูปที่ 6.1 ซึ่งเป็นเซกเมนต์ของหน่วยเสียง “ส” (/s/) ที่พบว่ามีสัญญาณเสียงก้องเข้ามาปะปนอยู่ภายในเซกเมนต์ด้วย ดังนั้นในกรณีนี้จึงสามารถใช้ค่าอัตสหสัมพันธ์หรือค่าอัตราส่วนของ

พลังงานความถี่ต่ำต่อพลังงานที่ความถี่สูงเข้ามาช่วยพิจารณาในการแบ่งส่วนภายในเซกเมนต์ของหน่วยเสียงดังกล่าวได้



รูปที่ 6.1 ตัวอย่างเซกเมนต์หน่วยเสียง “ส”(/s/)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

รายการอ้างอิง

[1] Rabiner, L. R. A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE *77*, pp. 257-286. 1989.

[2] Maneenoi, E. An Acoustic Study of Syllable Rhymes: A Basis for Thai Continuous Speech Recognition System. Doctoral dissertation, Department of Electrical Engineering, Faculty of Engineering, Chulalongkorn University. 2003.

[3] Glass, J. R. A probabilistic framework for segment-based speech recognition. Computer Speech and Language 17. (2003): 137-152.

[4] Young, S., Jansen. J., Odell, J. Ollason, D., and Woodland, P. The HTK Book (for HTK Version 3.3). Cambridge, 2005.

[5] Becchetti, C., and Ricotti, L. P. Speech Recognition Theory and C++ Implementation. England: John Wiley & Sons, 1999.

[6] Furui, S. Digital Speech Processing, Synthesis, and Recognition. New York and Basel: Marcel Dekker, 1989.

[7] Balakrishnama, S.; and Ganapathiraju, A. Linear Discriminant Analysis - A Brief Tutorial. Institute for signal and Infomation Processing. 1998.

[8] Goodwin, M. M., and Laroche, J. Audio Segmentation by Feature-space Clustering Using Linear Discriminant Analysis and Dynamic Programming. 2003 IEEE Workshop on Applications of Signal Processing to Audio and Acoustic. New Paltz, NY, 2003.

[9] MIT, MITOpenCourseWare. Pattern Classification [Online]. (n.d.). Available from: http://ocw.mit.edu/OcwWeb/Electrical-Engineering-and-Computer-Science/6-345Automatic-Speech-RecognitionSpring2003/LectureNotes/index.htm [2006, October 14]

[10] Meng, H., and Zue, V. Signal representation romparison for phonetic classification. Proceeding of the IEEE International Conference on Acoustics, Speech and Signal Proceeding. Toronto, Canada, 1991.

[11] Leung, H. C., Chigier, B., and Glass, J. R. A comparative study of signal representations and classification techniques for speech recognition. Proceedings of the ICASSP-93. Minneapolis, MN, USA, 1993.

[12] Zahorian, S. A., Silsbee, P., and Wang, X. Phone classification with segmental features and a binary-pair partitioned neural network classifier. Proceedings of the IEEE ICASSP'97. Munich, Germany, 1997.

[13] Kasuriya, S., Kanokphara, S., and Thatphithakkul, N. Context-independent acoustic models for Thai speech recognition. International Symposium on Communications and Information Technologies, 2004.

[14] Rabiner, L., and Juang, B. H. Fundamental of Speech Recognition. Pearson Education (Singapore) Pte., 2003.

[15] Halberstadt, A., and Glass, J. R. Heterogeneous measurements and multiple classifiers for speech recognition. Proceedings of ICSLP 98, Sydney, Australia, 1998.

[16] Ström, N., Hetherington, L., Hazen, T. J., Sandness, E., and Glass, J. R. Acoustic modeling improvements in a segment-based speech recognizer. Proceedings of IEEE Automatic Speech Recognition and Understanding Workshop, Keystone, CO, 1999.

[17] Chang, J., Glass, J. R. Segmentation and modeling in segment-based recognition. Proceedings of Eurospeech'97, Rhodes, Greece, 1997.

[18] Kasuriya, S., Sornlertlamvanich, V., Cotsomrong, P., Kanokphara, S., and Thatphithakkul, N. Thai Speech Corpus for Speech Recognition. Proceeding of the COCOSDA'03, 2003.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ประวัติผู้เขียนวิทยานิพนธ์

นางสาวหนึ่งฤทัย เอกชัยวรสิน เกิดเมื่อวันที่ 4 มิถุนายน พ.ศ. 2526 ที่จังหวัดกรุงเทพมหานคร สำเร็จการศึกษาระดับประถมศึกษาถึงมัธยมศึกษาตอนต้นจากโรงเรียนซางตาครู้สคอนแวนท์ สำเร็จการศึกษาระดับมัธยมศึกษาตอนปลายจากโรงเรียนมหิดลวิทยานุสรณ์ และสำเร็จการศึกษาระดับปริญญาบัณฑิต ในสาขาวิศวกรรมคอมพิวเตอร์ จากคณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัยในปีการศึกษา 2548

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย