วิธีการวิเคราะห์ด้วยการสังเคราะห์เพื่อประมาณพื้นที่หน้าตัดช่องทางเสียงสำหรับการสังเคราะห์เสียงภาษาไทย

นางสาว ศศิวิมล ชาญวานิชบริการ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมคอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2551

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

ANALYSIS-BY-SYNTHESIS APPROACH TO VOCAL TRACT CROSS SECTIONAL AREA ESTIMATION FOR THAI SPEECH SYNTHESIS

Miss Sasivimon Chanvanichborikarn

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 2008

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	วิธีการวิเคราะห์ด้วยการสังเคราะห์เพื่อประมาณพื้นที่หน้ ตัดช่องทางเสียงสำหรับการสังเคราะห์เสียงภาษาไทย
โดย	นางสาวศศิวิมล ชาญวานิชบริการ
สาขาวิชา	วิศวกรรมคอมพิวเตอร์
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	ผู้ช่วยศาสตราจารย์ ดร.อติวงศ์ สุชาโต
อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม	ผู้ช่วยศาสตราจารย์ ดร.โปรดปราน บุณยพุกกณะ

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

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

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ
(ศาสตราจารย์ ดร.บุญเสริม กิจศิริกุล)

..... อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ผู้ช่วยศาสตราจารย์ ดร.อติวงศ์ สุชาโต)

..... อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม
(ผู้ช่วยศาสตราจารย์ ดร.โปรดปราน บุณยพุกกณะ)

..... กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.พิษณุ คนองชัยยศ)

..... กรรมการภายนอกมหาวิทยาลัย
(ดร.ชัย วุฒิวิวัฒน์ชัย)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ศศิวิมล ชาญวานิชบริการ : วิธีการวิเคราะห์ด้วยการสังเคราะห์เพื่อประมาณพื้นที่หน้าตัดช่องทางเสียงสำหรับการสังเคราะห์เสียงภาษาไทย. (ANALYSIS-BY-SYNTHESIS APPROACH TO VOCAL TRACT CROSS SECTIONAL AREA ESTIMATION FOR THAI SPEECH SYNTHESIS) อ. ที่ปรึกษาวิทยานิพนธ์หลัก : ผศ.ดร.อดิวงศ์ สุชาโต, อ.ที่ปรึกษาวิทยานิพนธ์ร่วม : ผศ.ดร.โปรดปราน บุณยพุกกณะ, 99 หน้า.

โดยปรกติแล้วการเคลื่อนที่และการวางตัวของอวัยวะในช่องปากขณะทำการเปล่งเสียงเพื่อการสื่อสารไม่สามารถถูกสังเกตเห็นได้ด้วยตาเปล่า การใช้แบบจำลองอวัยวะในช่องทางเสียงแสดงกลไกการเปลี่ยนแปลงรูปร่างพื้นที่หน้าตัดช่องทางเสียงในการเปล่งเสียงพูด จะเป็นเครื่องมือในการเรียนการเปล่งเสียงพูดอย่างถูกต้อง และสามารถสังเคราะห์เสียงจากรูปร่างช่องทางเสียงที่มีการเปลี่ยนแปลงตลอดช่วงเวลาในการเปล่งเสียงต่อเนื่อง

วิทยานิพนธ์นี้นำเสนอพารามิเตอร์ที่ใช้ควบคุมรูปร่างของช่องทางเสียงและพารามิเตอร์ที่ควบคุมลักษณะของหน่วยเสียง เพื่อการแสดงกลไกการเปล่งเสียงพูดที่ถูกต้องจึงเสนอวิธีการหารูปร่างพื้นที่หน้าตัดของช่องทางเสียงของหน่วยเสียงพยัญชนะและหน่วยเสียงสระในภาษาไทยโดยเปรียบเทียบข้อมูลของความถี่ฟอร์แมนต์อันดับที่ 1 ถึง 3 ของเสียงสังเคราะห์ที่ได้กับข้อมูลของเสียงภาษาไทยจากเสียงพูดจริง รวมทั้งเสนอวิธีการหาการเปลี่ยนแปลงของพื้นที่หน้าตัดของช่องทางเสียงขณะเปล่งเสียงโดยใช้เส้นโค้งเบซิเยร์ที่สามารถสร้างเส้นโค้งรูปแบบต่าง ๆ เปลี่ยนแปลงไปตามจุดควบคุมความโค้งเบซิเยร์ จากนั้นทำการเปรียบเทียบกับลักษณะการเปลี่ยนแปลงของความถี่ฟอร์แมนต์ที่ 1 และ 2 จุดควบคุมความโค้งคู่ที่ให้ผลสหสัมพันธ์ของแนวฟอร์แมนต์ที่ 1 และ 2 ของเสียงสังเคราะห์และเสียงจริงมากที่สุดจะถูกเลือกให้เป็นจุดควบคุมเส้นการเปลี่ยนแปลงพื้นที่หน้าตัดของช่องทางเสียงของหน่วยเสียงพยัญชนะและคู่สระนั้น ๆ และเสนอการประมาณรูปร่างช่องทางเสียงพยัญชนะที่เกิดการใช้อวัยวะออกเสียงร่วมกับหน่วยเสียงสระที่ติดกัน

ผลจากการเสนอวิธีการข้างต้นการหารูปร่างช่องทางเสียงสระด้วยการวิเคราะห์จากเสียงสังเคราะห์ที่ได้ให้ผลสอดคล้องกับรูปร่างช่องทางเสียงที่ใช้ในการออกเสียงจริง การประมาณการเปลี่ยนพื้นที่หน้าตัดของช่องทางเสียงขณะเปล่งเสียงให้ผลเฉพาะเจาะจงกับคู่พยัญชนะและสระเฉพาะตัว และเช่นเดียวกันผลวิเคราะห์การประมาณรูปร่างช่องทางเสียงพยัญชนะที่เกิดการใช้อวัยวะออกเสียงร่วมกับหน่วยเสียงสระข้างเคียงก็ได้ผลเฉพาะเจาะจงกับคู่พยัญชนะและสระคู่นั้นๆ

ภาควิชา..... วิศวกรรมคอมพิวเตอร์..... ลายมือชื่อนิสิต ศศิวิมล ชาญวานิชบริการ

สาขาวิชา..... วิศวกรรมคอมพิวเตอร์..... ลายมือชื่อ อ. ที่ปรึกษาวิทยานิพนธ์หลัก.....

ปีการศึกษา..... 2551..... ลายมือชื่อ อ. ที่ปรึกษาวิทยานิพนธ์ร่วม .....

487 04815 21 : MAJOR COMPUTER ENGINEERING

KEYWORDS: ARTICULATORY MECHANISM / HUMAN SPEECH PRODUCTION / COARTICULATION / VOCAL TRACT VISUALIZATION

SASIVIMON CHANVANICHBORIKARN : ANALYSIS-BY-SYNTHESIS APPROACH TO VOCAL TRACT CROSS SECTIONAL AREA ESTIMATION FOR THAI SPEECH SYNTHESIS. ADVISOR : ASST. PROF. ATIWONG SUCHATO, Ph.D., CO-ADVISOR : ASST. PROF. PROADPRAN PUNYABUKKANA, Ph.D., 99 pp.

Movement and alignment of human speech production organs while producing speech sound are partially visible. Using a model of vocal tract that represent the variations of the cross section area of the vocal tract is the tool for correctly learning speech production utterance and with co-articulation compensation capability.

The aims of this thesis are to present the parameters that were used to control the shape of vocal tract and the parameters that control the phoneme properties and to present the accurate speech production mechanism of vocal tract. Therefore this thesis will introduce the vocal tract shape estimation method of the consonant and vowel in Thai by comparing the F1 to F3 of the synthetic speech and real speech. This thesis also introduces the method to estimate the variations of the cross-section area of the vocal tract using bezier interpolation then compare the deviation of F1 and F2. The most dominant control points in the bezier curve that produce the correlation of F1 and F2 of the synthetic speech closet to the real speech will be chosen as control points.

Vocal tract shape estimation by analyzing synthetic speech conforms to the vocal tract shape of real speech. Cross section area estimation from synthetic speech that conforms to real speech is unique for each couple of consonant and vowel as well as the estimation with co-articulation compensation.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Department : .....Computer Engineering	Student's Signature : ศศิวิมล ชาญวนิชบริกาญ
Field of Study : ...Computer Engineering	Advisor's Signature : .....
Academic Year : .....2008.....	Co-Advisor's Signature : .....

กิตติกรรมประกาศ

ในโอกาสนี้ข้าพเจ้าใคร่ขอขอบคุณ ผู้ช่วยศาสตราจารย์ ดร. อติวงศ์ สุชาโต อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก และ ผู้ช่วยศาสตราจารย์ ดร. โปรดปราน บุณยพุกกณะ อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม ซึ่งท่านทั้งสองได้ช่วยเหลือ ให้คำแนะนำ ข้อคิดที่เป็นประโยชน์ อันเป็นส่วนสำคัญที่ทำให้วิทยานิพนธ์นี้สำเร็จลุล่วงไปได้ด้วยดี

ขอขอบคุณ ศาสตราจารย์ ดร. บุญเสริม กิจศิริกุล ผู้ช่วยศาสตราจารย์ ดร.พิษณุ คนองชัยยศ และ ดร.ชัย วุฒิวิวัฒน์ชัย ที่ให้ข้อแนะนำดี ๆ และข้อคิดที่เป็นประโยชน์ต่อการทำวิทยานิพนธ์นี้

นอกจากนี้ข้าพเจ้าขอขอบคุณ พี่ ๆ เพื่อน ๆ และน้อง ๆ ห้องปฏิบัติการ SLS ที่ได้ให้ความร่วมมือ สบับสนุน ช่วยเหลือ จนกระทั่งวิทยานิพนธ์นี้สำเร็จไปได้ด้วยดี

ท้ายที่สุดขอขอบคุณครอบครัวของข้าพเจ้า ที่ได้มอบกำลังใจ ความห่วงใย ช่วยเหลือแก่ข้าพเจ้าเสมอมา

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ง
บทคัดย่อภาษาอังกฤษ ..... จ
กิตติกรรมประกาศ ..... ฉ
สารบัญ ..... ช
สารบัญตาราง ..... ญ
สารบัญภาพ ..... ฏ
บทที่ 1 1
บทนำ 1
ความเป็นมาและความสำคัญของปัญหา 1
วัตถุประสงค์ของการวิจัย 2
ขอบเขตของการวิจัย 2
ประโยชน์ที่คาดว่าจะได้รับ 3
วิธีดำเนินการวิจัย 3
โครงสร้างวิทยานิพนธ์ 4
บทที่ 2 5
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง 5
แนวคิดและทฤษฎี 5
ทฤษฎีบทเรื่องอวัยวะในการเปล่งเสียง 5
1. ปอดและหลอดลม 5
2. กล่องเสียงและเส้นเสียง 7
3. อวัยวะที่อยู่เหนือเส้นเสียง 9
ทฤษฎีด้านอวัยวะออกเสียงร่วม 16
ทฤษฎีด้านภาษาศาสตร์ 16
1. เสียงพยัญชนะ 17
2. เสียงสระ 20
3. เสียงวรรณยุกต์ 25
ทฤษฎีในการสังเคราะห์เสียง 29
1. เทคนิคการสังเคราะห์เสียงพูด 29
2. แบบจำลองในการสังเคราะห์เสียง 29
ทฤษฎีการวิเคราะห์สัญญาณเสียง 33

หน้า

1. สเปกโตรแกรม (Spectrogram) 33
2. สมการโลคัส (Locus Equation) 34
ทฤษฎีในการวิเคราะห์ข้อมูล 34
1. ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error) 34
2. สหสัมพันธ์ (Correlation) 35
ทฤษฎีในการประมาณค่า 36
เส้นโค้งเบซิเยร์ (Bezier Curve) 36
งานวิจัยที่เกี่ยวข้อง 36
1. แบบจำลองของช่องทางเสียงและอวัยวะในการเปล่งเสียง 37
2. การศึกษาถึงผลของอวัยวะออกเสียงร่วมในการเปล่งเสียง 37
3. การสังเคราะห์เสียงด้วยวิธี Articulatory Speech Synthesis 38
บทที่ 3 39
วิธีดำเนินการวิจัย 39
ขั้นตอนการดำเนินการวิจัย 39
เครื่องมือที่ใช้ในการวิจัย 39
ข้อมูลที่ใช้พัฒนาระบบ 40
แบบจำลองช่องทางเสียง 40
การออกแบบพารามิเตอร์ควบคุมที่สอดคล้องกับหน่วยเสียง 43
การทดสอบหาค่าพารามิเตอร์ควบคุมช่องทางเสียง 45
1. การทดสอบหาค่าพารามิเตอร์ควบคุมช่องทางเสียงสระ 45
2. การทดสอบหาค่าพารามิเตอร์ควบคุมช่องทางเสียงพยัญชนะ 47
การทดลองการประมาณการเปลี่ยนแปลงของพื้นที่หน้าตัดขณะเปล่งเสียง 47
การทดลองการประมาณรูปร่างช่องทางเสียงพยัญชนะที่เกิดผลจากการใช้อวัยวะออกเสียงร่วม.. 47
การออกแบบระบบจำลองกลไกการเปล่งเสียง 50
ภาพรวมของระบบ 50
บทที่ 4 52
ผลการวิจัย 52
ผลการวิจัย 52
ผลการวิเคราะห์การกำหนดพารามิเตอร์ควบคุมช่องทางเสียง 52
ผลการวิเคราะห์การประมาณการเปลี่ยนแปลงของพื้นที่หน้าตัดขณะเปล่งเสียง 57

หน้า

ผลการวิเคราะห์การหาค่าน้ำหนักที่เหมาะสมสำหรับพยัญชนะและสระที่มีผลการใช้อวัยวะออกสียงร่วม 58
บทที่ 5 64
สรุปผลการวิจัยและข้อเสนอแนะ 64
สรุปผลการวิจัย 64
ข้อเสนอแนะ 65
รายการอ้างอิง 66
ภาคผนวก 70
ภาคผนวก ก 71
ภาคผนวก ข 90
ประวัติผู้เขียนวิทยานิพนธ์ 99

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญตาราง

หน้า

ตารางที่ 2.1 ตำแหน่งของอวัยวะในช่องทางเสียงที่สัมพันธ์กับเสียงพยัญชนะไทย.....13
ตารางที่ 2.2 เสียงพยัญชนะภาษาไทยแบ่งตามลักษณะของเสียง 19
ตารางที่ 2.3 เปรียบเทียบรูปสระในภาษาไทยกับสัทอักษรสากล.....23
ตารางที่ 2.4 เสียงสระตามระดับลิ้น ตำแหน่งลิ้น และลักษณะริมฝีปาก 25
ตารางที่ 3.1 พารามิเตอร์ที่ใช้ควบคุมช่องทางเสียง.....44
ตารางที่ 3.2 พารามิเตอร์ที่ใช้ควบคุมลักษณะของหน่วยเสียง.....44
ตารางที่ 3.3 ค่าเฉลี่ยความถี่ฟอร์แมนต์อันดับที่ 1 ถึง 3 ของชายและหญิงไทยในเสียงสระเสียงเดี่ยว 45
ตารางที่ 4.1 ผลค่าคลาดเคลื่อนกำลังสองเฉลี่ยของความถี่ฟอร์แมนต์อันดับ ที่ 1-3 ของเสียงสระเดี่ยวจากการสังเคราะห์เสียงเปรียบเทียบกับเสียงสระเดี่ยวจริง 53
ตารางที่ 4.2 ค่าความถี่ฟอร์แมนต์อันดับที่ 1 – 3 ของเสียงสังเคราะห์ที่ให้ผลรวมค่าคลาดเคลื่อนกำลังสองเฉลี่ยน้อยที่สุดในชุดข้อมูลทดลอง 53
ตารางที่ 4.3 ค่าของพารามิเตอร์ควบคุมสระเดี่ยวที่ให้ผลรวมค่าคลาดเคลื่อนกำลังสองเฉลี่ยน้อยที่สุดในชุดข้อมูลทดลอง.....54
ตารางที่ 4.4 ค่าพารามิเตอร์ควบคุมช่องทางเสียงหน่วยเสียงพยัญชนะภาษาไทย 55
ตารางที่ 4.5 ค่าพารามิเตอร์ควบคุมลักษณะของหน่วยเสียงของหน่วยเสียงพยัญชนะภาษาไทย 56
ตารางที่ 4.6 ผลของจุดควบคุมเส้นโค้งเบซิเยร์ของหน่วยเสียงพยัญชนะและสระเดี่ยว.....57
ตารางที่ 4.7 ผลการวิเคราะห์ค่าน้ำหนักของพยัญชนะและสระที่มีการใช้อวัยวะออกเสียงร่วม.....60

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญภาพ

หน้า

รูปที่ 2.1 บริเวณภายในช่องอกและคอ....6
รูปที่ 2.2 การทำงานของกล้ามเนื้อกะบังลมเวลาหายใจ....7
รูปที่ 2.3 เส้นเสียง....8
รูปที่ 2.4 ภาพตัดขวางของเส้นเสียงตามแนวตั้ง....8
รูปที่ 2.5 อวัยวะที่อยู่เหนือเส้นเสียง....10
รูปที่ 2.6 ตำแหน่งต่าง ๆ ของอวัยวะในช่องปาก....11
รูปที่ 2.7 ลิ้นลักษณะต่าง ๆ....14
รูปที่ 2.8 ริมฝีปากลักษณะต่าง ๆ....15
รูปที่ 2.9 การเคลื่อนตัวของลิ้นไก่และเพดานเพื่อควบคุมทิศทางของลม....15
รูปที่ 2.10 บริเวณการเกิดเสียงสระ....22
รูปที่ 2.11 ตำแหน่งของลิ้นในสระภาษาไทยตามแบบของแผนภาพสระ....23
รูปที่ 2.12 ความถี่ของเสียงวรรณยุกต์ที่เปลี่ยนแปลงกับเวลา....27
รูปที่ 2.13 สัทอักษรสากล....28
รูปที่ 2.14 กระบวนการการเกิดเสียงโดยการแทนด้วยฟังก์ชั่นถ่ายโอน....30
รูปที่ 2.15 ช่องทางเสียงแทนด้วยท่อตรงพื้นที่หน้าตัดต่าง ๆ กัน....31
รูปที่ 2.16 ทิศทางของความเร็วปริมาตรของอากาศในท่อ 2 ท่อที่ติดกัน....32
รูปที่ 2.17 โครงสร้างของสเปกโตรแกรม....33
รูปที่ 2.18 ภาพสเปกโตรแกรมของเสียงพูดสระอีและอาต่อเนื่องกัน....33
รูปที่ 3.1 แบบจำลองสามมิติที่ใช้ในวิทยานิพนธ์ (ก) กรามและฟันล่าง (ข) ปุ่มเหงือก เพดานแข็ง และเพดานอ่อน (ค) ลิ้น (ง) ริมฝีปาก และ (จ) กระโหลกศีรษะและฟันบน....41
รูปที่ 3.2 อวัยวะที่ต้องกำหนดรูปแบบการเคลื่อนไหว (ก) ริมฝีปาก (ข) ลิ้น และ (ค) กราม....42
รูปที่ 3.3 การกำหนดรูปแบบการเคลื่อนไหวตั้งต้นของลิ้น....43
รูปที่ 3.4 สมการโลคัสของเสียงจริงของพยัญชนะ ก[k]....48
รูปที่ 3.5 สมการโลคัสของเสียงจริงของพยัญชนะ ค[k^h]....49
รูปที่ 3.6 สมการโลคัสของเสียงจริงของพยัญชนะ ง[ŋ]....49
รูปที่ 3.7 แผนผังของระบบจำลองกลไกการออกเสียง....50
รูปที่ 4.1 สมการโลคัสของเสียงสังเคราะห์ของพยัญชนะ ก[k] (ก)ได้จากการปรับปรุงรูปร่างช่องทางเสียงตามหน่วยเสียงสระจากผลของอวัยวะออกเสียงร่วม (ข) ไม่มีการปรับปรุงรูปร่าง....61

หน้า

รูปที่ 4.2 สมการโลคัสของเสียงสังเคราะห์ของพยัญชนะ ค[k^h] (ก)ได้จากการปรับปรุงรูปร่างช่องทางเสียงตามหน่วยเสียงสระจากผลของอวัยวะออกเสียงร่วม (ข) ไม่มีการปรับปรุงรูปร่าง 62

รูปที่ 4.3 สมการโลคัสของเสียงสังเคราะห์ของพยัญชนะ ง[ŋ] (ก) ได้จากการปรับปรุงรูปร่างช่องทางเสียงตามหน่วยเสียงสระจากผลของอวัยวะออกเสียงร่วม (ข) ไม่มีการปรับปรุงรูปร่าง 63

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 1
บทนำ

ความเป็นมาและความสำคัญของปัญหา

การออกเสียงเพื่อการสื่อสารด้วยคำพูดของมนุษย์นั้นนอกจากจะมีกลไกที่ซับซ้อนแล้ว โดยทั่วไปกลไกเหล่านั้นส่วนใหญ่จะไม่สามารถมองเห็นได้โดยตรง นอกจากรูปร่างของริมฝีปาก การเคลื่อนที่ของขากรรไกร และ บางส่วนของลิ้นที่สามารถมองเห็นได้ผ่านช่องเปิดของปาก การจำลองกลไกดังกล่าวนี้ด้วยกลุ่มของวัตถุสามมิติที่มีลักษณะเหมือนกับอวัยวะต่างๆ ภายในช่องปากของมนุษย์ที่มีความสำคัญต่อการออกเสียงในภาษานั้น เป็นวิธีหนึ่งซึ่งสามารถแสดงกลไกเหล่านี้ได้โดยสะดวก การแสดงกลไกดังกล่าวทำได้โดยเริ่มจากทำการแสดงรูปร่างของช่องทางเสียง (vocal tract) จากตำแหน่งของเสียงพยัญชนะ ตามด้วยการเปลี่ยนแปลงรูปร่างของช่องทางเสียงไปเป็นรูปร่างของช่องทางเสียงสระที่ต้องการ โดยมีค่ารูปร่างของช่องทางเสียงพยัญชนะและสระที่ตายตัว แต่เมื่อมีการทำการศึกษากลับพบว่า รูปร่างของช่องทางเสียงนั้นไม่ได้กำหนดตายตัวแน่นอนทั้งหมด จะมีส่วนของต้นและปลายของหน่วยเสียงที่ทำการเชื่อมต่อกับหน่วยเสียงเสียงอื่นจะมีการเปลี่ยนแปลงไป กล่าวคือจะมีผลขึ้นอยู่กับรูปร่างของช่องทางเสียงของหน่วยเสียงที่อยู่รอบข้าง (co-articulation) และที่สังเกตเห็นได้ชัดคือการออกเสียงพยัญชนะที่มีต้นกำเนิดจากอวัยวะบริเวณเพดานอ่อน เมื่อรวมรูปร่างของช่องทางเสียงพยัญชนะที่มีแหล่งกำเนิดจากเพดานอ่อนดังกล่าวเข้ากับรูปร่างช่องทางเสียงของเสียงสระ รูปร่างของช่องทางเสียงพยัญชนะจะเปลี่ยนแปลงไปขึ้นกับเสียงสระที่นำไปผสมร่วม

การได้มาซึ่งรูปร่างพื้นที่หน้าตัดของช่องทางเสียงสามารถใช้ภาพถ่ายเอ็มอาร์ไอของอวัยวะช่องทางเสียงในการเปล่งเสียงหน่วยเสียงนั้น ๆ แต่วิธีดังกล่าวมีค่าใช้จ่ายสูง หากทำการวิเคราะห์จากเสียงสังเคราะห์เปรียบเทียบกับเสียงพูดจริงของมนุษย์วิธีนี้จะสะดวกและสามารถนำไปใช้ได้ไม่จำกัดภาษา เมื่อผนวกกับความสามารถในการรับพารามิเตอร์ที่สามารถควบคุมคุณสมบัติและพื้นที่หน้าตัดของวัตถุสามมิติเหล่านี้ให้สอดคล้องกับหน่วยเสียงที่ต้องการแบบจำลองดังกล่าวจะช่วยอำนวยความสะดวก ในการศึกษากลไกการเปล่งเสียงของมนุษย์โดยละเอียด ภาพเคลื่อนไหวแสดงการเคลื่อนที่ของอวัยวะที่เกี่ยวข้องในการเปล่งเสียงนั้นสามารถนำไปใช้เป็นเครื่องมือสำคัญในการเรียนการสอนการเปล่งเสียงพูดอย่างถูกต้อง ตัวอย่างของการเรียนการสอนประเภทนี้ ได้แก่ การสอนการเปล่งเสียงให้ถูกต้องแก่ผู้ที่มีความบกพร่องทางการได้ยิน หรือ การเรียนการสอน ภาษาต่างประเทศซึ่งอาจจะมีหน่วยเสียงที่ไม่อยู่ในระบบภาษาท้องถิ่นของผู้เรียน นอกจากประโยชน์ในแง่ของการช่วยการเรียนการสอนลักษณะดังกล่าวแล้ว การแสดง

การจำลองการออกเสียงด้วยภาพเคลื่อนไหวน่าจะมีส่วนช่วยให้นักวิจัยทางเสียงพูดและนักภาษาศาสตร์อธิบายปรากฏการณ์ทางภาษาบางอย่างได้สะดวกยิ่งขึ้น

รูปร่างของช่องทางเสียงโดยละเอียดตลอดช่วงเวลาของการออกเสียงต่อเนื่องนั้นสามารถนำไปใช้ประโยชน์ในการสังเคราะห์เสียง (Speech synthesis) พูดได้ โดยใช้สมมติฐานที่ว่าเสียงเกิดจากลมที่ถูกขับออกมาจากปอดไหลผ่านช่องทางเสียง ซึ่งเปรียบเสมือนท่อที่มีรูปร่างเปลี่ยนแปลงไปตามเวลา การสังเคราะห์เสียงเช่นนี้เป็นการสังเคราะห์เสียงแบบพาราเมตริกที่พารามิเตอร์ที่เกี่ยวข้องนั้นควบคุมรูปร่างของช่องทางเสียงโดยตรง ต่างกับการสังเคราะห์เสียงแบบพาราเมตริกแบบอื่นๆ ซึ่งพารามิเตอร์ที่ใช้นั้น บ่งบอกถึงลักษณะสำคัญของสัญญาณเสียงที่ถูกเปล่งออกมาสู่ภายนอกช่องปาก เสียงที่ได้จากการสังเคราะห์เสียงด้วยวิธีการนี้น่าจะมีความเป็นธรรมชาติมากกว่า ทั้งนี้ความเป็นธรรมชาติขึ้นอยู่กับความถูกต้องในการจำลองการเคลื่อนไหวของอวัยวะต่างๆ ในช่องทางเสียง และ ความละเอียดของแบบจำลองทางพลศาสตร์ของไหลที่ใช้ในการสังเคราะห์เสียง

วัตถุประสงค์ของการวิจัย

1. เสนอชุดของพารามิเตอร์ควบคุมที่ใช้กำหนดรูปร่างของช่องทางเสียง เพื่อสามารถนำไปใช้ในการสังเคราะห์เสียงที่ใช้การวิเคราะห์พื้นที่หน้าตัดบริเวณต่าง ๆ ของช่องทางเสียงได้
2. เสนอการประมาณการเปลี่ยนแปลงของพื้นที่หน้าตัดในบริเวณต่าง ๆ ของช่องทางเสียง ขณะเปล่งเสียงต่อเนื่องของหน่วยเสียงสองหน่วยเสียง โดยการวิเคราะห์เสียงที่ได้จากการสังเคราะห์เปรียบเทียบกับเสียงจริง
3. เสนอการประมาณรูปร่างช่องทางเสียงพยัญชนะ ก ค และ ง ที่เกิดการใช้อวัยวะออกเสียงร่วมกับหน่วยเสียงสระเดี่ยวที่ติดกัน โดยการวิเคราะห์เสียงที่ได้จากสังเคราะห์เปรียบเทียบกับเสียงจริง
4. สร้างโปรแกรมต้นแบบแสดงภาพเคลื่อนไหวสามมิติ ตามชุดพารามิเตอร์ที่กำหนดรูปร่างของช่องทางเสียงและการประมาณการเคลื่อนที่ของช่องทางเสียงระหว่างหน่วยเสียง

ขอบเขตของการวิจัย

1. ชุดของพารามิเตอร์ควบคุมใช้กำหนดรูปร่างช่องทางเสียง ชุดพารามิเตอร์ควบคุมนี้ครอบคลุมลักษณะตามหน่วยเสียงพยัญชนะภาษาไทย 21 หน่วยเสียง และหน่วยเสียงสระเดี่ยว 18 หน่วยเสียง

2. สร้างชุดเครื่องมือที่มีส่วนติดต่อกับผู้ใช้เป็นแบบกราฟิก โดยชุดเครื่องมือนี้ประกอบด้วย
    2.1. เครื่องมือสำหรับกำหนดสัญลักษณ์ของหน่วยเสียง และพารามิเตอร์ควบคุมที่สอดคล้องกับสัญลักษณ์หน่วยเสียงนั้น ๆ
    2.2. เครื่องมือสำหรับอ่านชุดของสัญลักษณ์ของหน่วยเสียง เพื่อแสดงแบบจำลองช่องทางเสียงที่สอดคล้องกับชุดของสัญลักษณ์นั้น ๆ
3. การแสดงกลไกการเปล่งเสียงจะใช้การเปลี่ยนแปลงของพื้นที่หน้าตัดที่ทำการศึกษาไว้เป็นข้อมูลพื้นฐาน
4. การแสดงกลไกการเปล่งเสียงที่มีหน่วยเสียง ก ค และ ง จะมีส่วนอวัยวะออกเสียงร่วมเพิ่มเติม โดยใช้ข้อมูลที่ได้จากการศึกษา
5. ตรวจสอบความถูกต้องโดยใช้โปรแกรมสังเคราะห์เสียงจากรูปร่างช่องทางเสียงที่เปลี่ยนแปลงโดยโปรแกรมสังเคราะห์เสียงดังกล่าวไม่ได้เป็นส่วนหนึ่งของงานวิจัยนี้

ประโยชน์ที่คาดว่าจะได้รับ

ทราบถึงวิธีการหาลักษณะรูปร่างช่องทางเสียงหน่วยเสียงสระและการเปลี่ยนแปลงของพื้นที่หน้าตัดของช่องทางเสียงในการเปล่งเสียงหน่วยเสียงสองหน่วยเสียงติดต่อกันโดยการวิเคราะห์สัญญาณเสียงแทนการใช้ภาพถ่ายเอ็มอาร์ไอ และทราบถึงวิธีการหารูปร่างของช่องทางเสียงพยัญชนะกลุ่มลิ้นส่วนหลัง-เพดานอ่อนที่มีรูปร่างเปลี่ยนแปลงไปขึ้นกับเสียงสระที่ติดกันซึ่งเกิดจากผลการใช้อวัยวะออกเสียงร่วม

สามารถนำวิธีการเหล่านี้ไปประยุกต์ใช้ในภาษาอื่นได้ เนื่องจากเป็นการวิเคราะห์สัญญาณเสียงซึ่งเป็นพื้นฐานของงานวิจัยทางเสียง และข้อมูลจากผลการทดลองยังสามารถนำไปใช้ในการสร้างโปรแกรมแบบจำลองช่องทางเสียงแสดงกลไกการเปล่งเสียงในแต่ละภาษา เพื่อนำไปใช้เป็นเครื่องมือสำคัญในการเรียนการสอนการเปล่งเสียงพูดอย่างถูกต้อง

วิธีดำเนินการวิจัย

1. ศึกษากลไกการเปล่งเสียงของหน่วยเสียงภาษาไทย เพื่อหาพารามิเตอร์ควบคุมที่เหมาะสม
2. ออกแบบรูปแบบของแฟ้มอักขระที่ใช้ในการเก็บพารามิเตอร์ควบคุม
3. เตรียมแบบจำลองสามมิติของช่องทางเสียง

4. สร้างเครื่องมือสำหรับกำหนดสัญลักษณ์ของหน่วยเสียง และพารามิเตอร์ควบคุมที่สอดคล้องกับหน่วยเสียงนั้น และเครื่องมือสำหรับอ่านชุดสัญลักษณ์ของหน่วยเสียง
5. ทดลองใช้เครื่องมือกำหนดพารามิเตอร์ของหน่วยเสียงสระและพยัญชนะภาษาไทย
6. ทดลองหาการเปลี่ยนแปลงพื้นที่หน้าตัดขณะเปล่งเสียงต่อเนื่องของหน่วยเสียงสองหน่อยเสียง วิเคราะห์ผลจากข้อมูลความถี่ฟอร์แมนต์ ของเสียงสังเคราะห์จากหน่วยเสียงสองหน่วยเสียงนั้นเปรียบเทียบกับข้อมูลความถี่ฟอร์แมนต์ของเสียงจริง
7. ทดลองหารูปร่างช่องทางเสียงของพยัญชนะ ก ค และ ง ร่วมกับหน่วเสียงสระเดี่ยว เปรียบเทียบข้อมูลความถี่ฟอร์แมนต์อันดับที่ 2 ของเสียงสังเคราะห์และเสียงจริง
8. สรุปผลการวิจัย และจัดทำรายงานวิทยานิพนธ์

โครงสร้างวิทยานิพนธ์

ในวิทยานิพนธ์ฉบับนี้จะแบ่งเนื้อหาในการนำเสนอออกเป็น 5 ส่วน คือ ในบทที่ 2 จะกล่าวถึง ทฤษฎีและงานวิจัยที่เกี่ยวข้องของวิทยานิพนธ์นี้ ทฤษฎีเรื่องอวัยวะในการเปล่งเสียงภาษาไทย การเกิดการใช้อวัยวะออกเสียงร่วม รวมทั้งทฤษฎีที่ใช้ในการวิเคราะห์ผลการวิจัย

ในบทที่ 3 จะได้กล่าวถึงวิธีการดำเนินการวิจัย ฐานข้อมูลที่ใช้ในพัฒนาระบบและกระบวนการขั้นตอนการวิจัยเรื่อง การหารูปร่างช่องทางเสียงของสระที่เหมาะสมกับเสียงสระเดี่ยวภาษาไทย การหาการประมาณการเปลี่ยนแปลงพื้นที่หน้าตัดขณะเปล่งเสียงของคู่พยัญชนะและสระเสียงเดี่ยวและการหารูปร่างช่องทางเสียงพยัญชนะที่เกิดการผลจากการใช้อวัยวะออกเสียงร่วม

ในบทที่ 5 และบทที่ 6 จะได้กล่าวถึงการวิเคราะห์ผลการทดสอบการหารูปร่างช่องทางเสียงของสระเสียงเดี่ยว ผลวิเคราะห์การประมาณการเปลี่ยนแปลงพื้นที่หน้าตัดของช่องทางเสียงขณะเปล่งเสียงพูด ผลวิเคราะห์ค่าน้ำหนักของคู่พยัญชนะและสระเสียงเดี่ยวที่เกิดจากผลการใช้อวัยวะออกเสียงร่วม และสรุปผลที่ได้จากวิทยานิพนธ์นี้ รวมทั้งให้ข้อเสนอแนะในการพัฒนางานวิจัยนี้ต่อไปในอนาคต

บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงทฤษฎีและงานวิจัยที่เกี่ยวข้อง การศึกษาทฤษฎีประกอบไปด้วย ทฤษฎีเรื่องอวัยวะในการเปล่งเสียง ศึกษาถึงหน้าที่และความสำคัญของอวัยวะในการเปล่งเสียง ทฤษฎีด้านอวัยวะออกเสียงร่วม ศึกษาลักษณะและผลของการเกิดอวัยวะออกเสียงร่วม ทฤษฎีด้านภาษาศาสตร์ ศึกษาลักษณะของเสียงภาษาไทยสัมพันธ์กับตำแหน่งการเกิดเสียงในอวัยวะการเปล่งเสียง ทฤษฎีในการสังเคราะห์เสียง ศึกษาการสังเคราะห์เสียงโดยการใช้พื้นที่หน้าตัดของช่องทางเสียง ทฤษฎีการวิเคราะห์สัญญาณเสียง ศึกษาการใช้สมการโลคัสในการชี้วัดหน่วยเสียงพยัญชนะที่เกิดการใช้อวัยวะออกเสียงร่วม ศึกษาการใช้ภาพสเปกโตรแกรมในการวิเคราะห์คุณลักษณะของเสียง ทฤษฎีในการวิเคราะห์ข้อมูล ศึกษาการใช้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยและค่าสหสัมพันธ์ในการวิเคราะห์ความถี่ฟอร์แมนต์ และทฤษฎีการประมาณค่า ศึกษาการใช้เส้นโค้งเบซิเยร์ในการประมาณแนวการเปลี่ยนแปลงพื้นที่หน้าตัดของช่องทางเสียงขณะเปล่งเสียง

ในการศึกษางานวิจัยที่เกี่ยวข้อง ศึกษางานวิจัยเกี่ยวกับแบบจำลองสามมิติของช่องทางเสียงและอวัยวะในการเปล่งเสียง งานวิจัยเกี่ยวกับการศึกษาผลของการใช้อวัยวะออกเสียงร่วมในการเปล่งเสียงและงานวิจัยเกี่ยวกับการสังเคราะห์เสียงด้วยวิธี Articulatory speech synthesis

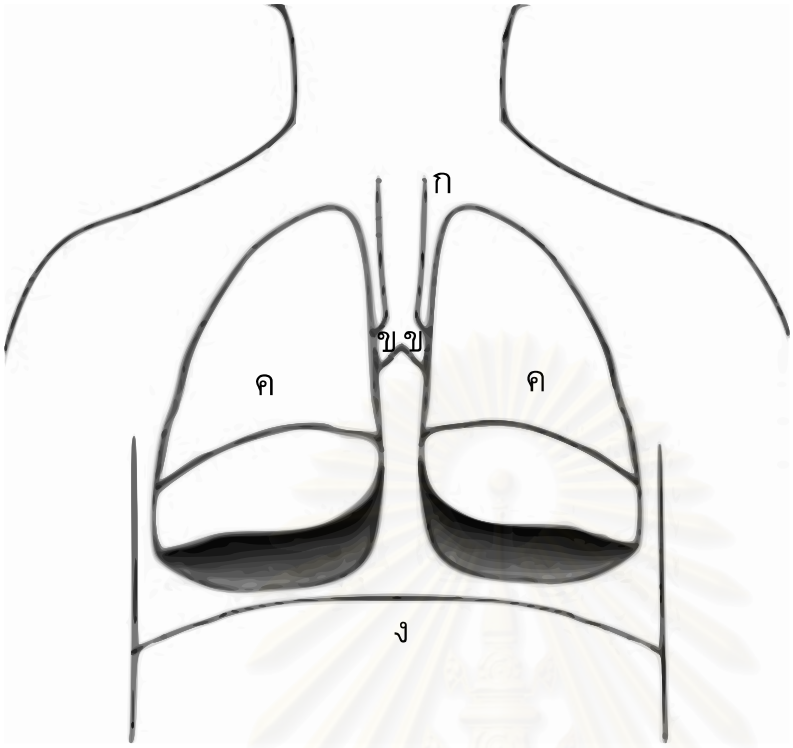
แนวคิดและทฤษฎี

ทฤษฎีบทเรื่องอวัยวะในการเปล่งเสียง

ในการศึกษากลไกการเปล่งเสียงจำเป็นต้องทำการศึกษาถึงกระบวนการของอวัยวะต่าง ๆ ที่ทำให้เกิดเสียง อวัยวะที่ใช้ในการเปล่งเสียงประกอบด้วยส่วนสำคัญแบ่งออกเป็น 3 กลุ่ม[1,2] คือ

1. ปอดและหลอดลม

ปอดทั้งสองข้างมีลักษณะเป็นถุงอากาศประกอบด้วยถุงลมขนาดเล็กจำนวนมาก ปอดมีลักษณะเหมือนฟองน้ำ ภายในปอดมีหลอดลมฝอยอยู่มากมาย หลอดลมฝอยนี้จะมารวมกันเป็นหลอดลมแขนง และหลอดลมแขนงของปอดทั้ง 2 ข้างจะมารวมกันเป็นหลอดลม ดังรูปที่ 2.1



ก หลอดลม

ข หลอดลมแขนง

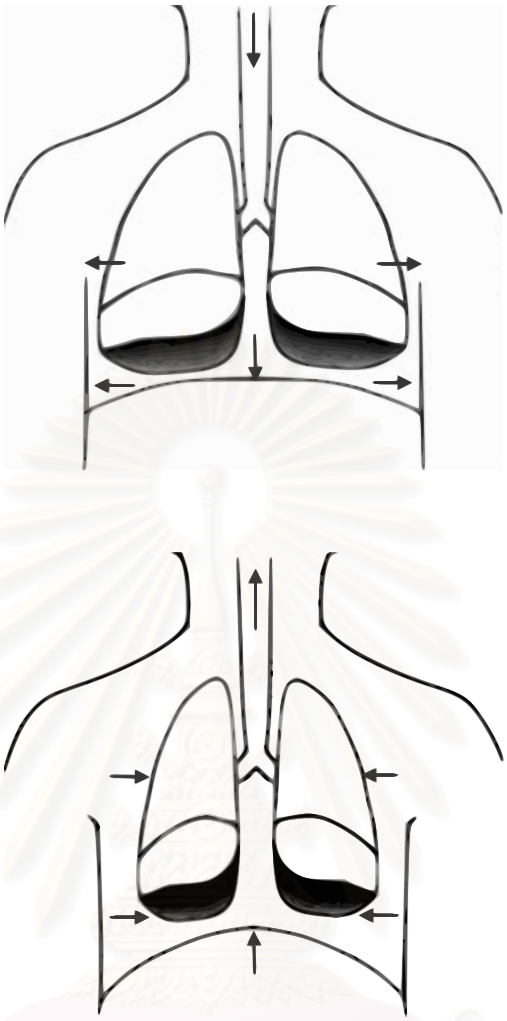
ค ปอด

ง กะบังลม

รูปที่ 2.1 บริเวณภายในช่องอกและคอ

การขับลมของปอดจะถูกควบคุมโดยกล้ามเนื้อที่ปอดและกระบังลมซึ่งเป็นแผ่นกล้ามเนื้อรูปโค้งอยู่บริเวณด้านล่างติดกับปอด แผ่นกระบังลมนี้จะเป็นตัวช่วยกดดันให้ช่องอกใหญ่ขึ้นหรือเล็กลงเวลาหายใจ กล่าวคือถ้าแผ่นกระบังลมลดตัวลง กระดูกซี่โครงจะถูกกล้ามเนื้อดึงให้ขยายตัวออก ขนาดของช่องอกจะใหญ่ขึ้น ความดันของอากาศในช่องอกจะต่ำกว่าความดันของอากาศภายนอกช่องอก อากาศก็จะไหลจากภายนอกร่างกายเข้ามาในปอดเป็นลมหายใจเข้า เมื่อสูดลมหายใจเข้าแล้ว ช่องอกจะอยู่ในสภาพพักแผ่นกระบังลมจะดันตัวโค้งเล็กน้อย กระดูกซี่โครงจะถูกกล้ามเนื้อระหว่างซี่โครงดึงให้เข้ามาหากันขนาดของช่องจะเล็กลง ความดันของอากาศในช่องอกจะสูงกว่าความดันภายนอกช่องอกอากาศก็จะไหลจากปอดไปนอกร่างกายเป็นลมหายใจออก ดังรูปที่ 2.2

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



ก ลมหายใจเข้า

ข ลมหายใจออก

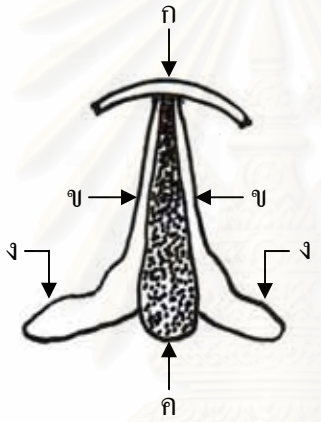
รูปที่ 2.2 การทำงานของกล้ามเนื้อกะบังลมเวลาหายใจ

ในภาษาอื่นอาจพบการใช้ลมหายใจเข้าหรือใช้ลมในคอในการเปล่งเสียงแต่ลมหายใจออกเป็นสิ่งสำคัญที่ช่วยให้เกิดเสียงพูดในภาษาไทย โดยที่ลมจะไหลผ่านหลอดลมเข้าสู่กล่องเสียง ผ่านสู่อวัยวะในช่องปากหรือช่องจมูก อวัยวะเหล่านี้จะดัดแปลงลมให้เป็นเสียงพูดต่าง ๆ กัน

2. กล่องเสียงและเส้นเสียง

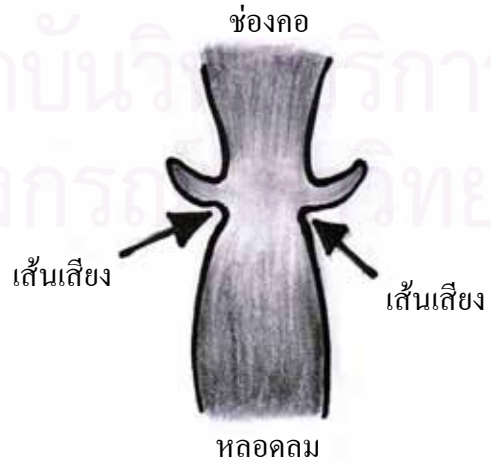
อวัยวะที่สำคัญที่สุดในการกล่องเสียงคือเส้นเสียง โดยเส้นเสียงจะอยู่ภายในกล่องเสียงอีกชั้นหนึ่ง กล่องเสียงประกอบขึ้นด้วยกระดูก กระดูกอ่อน และกล้ามเนื้อหลายชิ้น ส่วนหนึ่งของกล่องเสียงคือ สิ่งที่เรียกว่าลูกกระเดือกซึ่งเห็นได้ชัดเจนในเพศชาย กล่องเสียงอยู่เหนือหลอดลม ด้านล่างของกล่องเสียงประกอบด้วยกระดูกอ่อนชื่อ ไครคอยด์ (Cricoid) เป็นกระดูกอ่อนรูปคล้ายแหวน ด้านกว้างหันมาทางด้านหลังของร่างกาย ทางด้านหน้าของกล่องเสียงประกอบด้วย

กระดูกอ่อนชื่อ ไทรอยด์ (Thyroid) เป็นกระดูกอ่อนแผ่นเดียวรูปโค้งครึ่งวงกลม กระดูกชิ้นนี้เป็นส่วนหนึ่งของกล่องเสียงที่สัมผัสได้และเห็นได้บริเวณคอ นอกจากนี้กล่องเสียงยังประกอบด้วยกระดูกอ่อนรูปปิรามิดอีก 1 คู่ ชื่อ อาริตินอยด์ (Arytenoid) กระดูกอ่อนคู่นี้อยู่ด้านหลังของกล่องเสียงบนกระดูกอ่อนรูปวงแหวน ส่วนประกอบที่สำคัญที่สุดของกล่องเสียงได้แก่ เส้นเสียงซึ่งเปรียบเหมือนลิ้นของกล่องเสียงที่ทำหน้าที่ปิดปรือเปิดทางเดินลมหายใจ มีลักษณะเป็นแผ่นบาง ๆ 2 แผ่นประกอบด้วยเส้นเอ็นและกล้ามเนื้อ เส้นเสียงทั้งคู่พาดตัวตามแนวนอนอยู่ตรงกลางของกล่องเสียง ปลายด้านหนึ่งของเส้นเสียงทั้ง 2 เส้นอยู่ติดกันและเชื่อมอยู่กับกระดูกอ่อนไทรอยด์ทางด้านหน้าของกล่องเสียงปลายอีกด้านหนึ่งของเส้นเสียงแต่ละเส้นเชื่อมอยู่กับกระดูกอ่อนอาริตินอยด์แต่ละอัน ดังรูปที่ 2.3 และ 2.4



ก. กระดูกอ่อนไทรอยด์
ข. เส้นเสียง
ค. ช่องระหว่างเส้นเสียง
ง. กระดูกอ่อนอาริตินอยด์

รูปที่ 2.3 เส้นเสียง



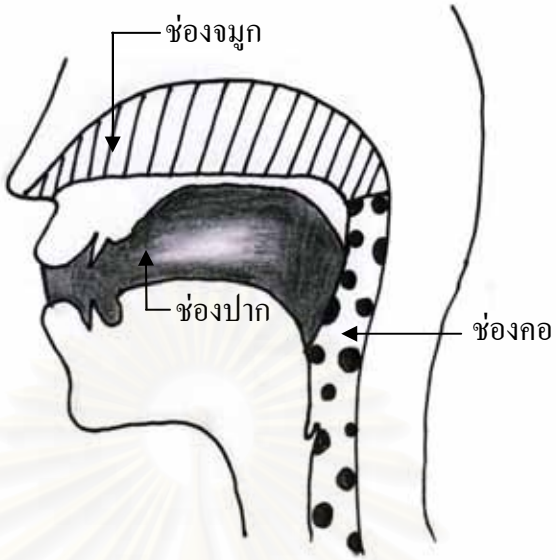
รูปที่ 2.4 ภาพตัดขวางของเส้นเสียงตามแนวตั้ง

ในการเปล่งเสียงหน้าที่สำคัญของเส้นเสียง คือ เปลี่ยนลมจากปอดให้เป็นคลื่นเสียง ลักษณะของเส้นเสียงที่สำคัญในการเปล่งเสียงแบ่งเป็น 4 ลักษณะ[1]

1) เส้นเสียงเปิดกว้าง (Open) เกิดขึ้นเมื่อเส้นเสียงแยกตัวออกจากกัน ทำให้เกิดระแสลมผ่านออกมาจากปอด โดยไม่มีการกีดขวางใด ๆ เป็นลักษณะของเส้นเสียงในการหายใจตามปกติ เสียงพูดที่เกิดขึ้นจากเส้นเสียงลักษณะนี้เป็นเสียงประเภท อโฆษะหรือเสียงไม่ก้อง (Voiceless) เช่น เสียง ซ[ʂ] ฟ[f] และ ฮ[h]
2) เส้นเสียงปิดสนิท (Close) เกิดขึ้นเมื่อเส้นเสียงเลื่อนมาชิดกัน ลมไม่สามารถผ่านออกมาจากปอดได้ เป็นลักษณะที่เกิดการกลั้นหายใจ จะเกิดเสียงพูดได้ก็ต่อเมื่อลมที่ถูกกักไว้ที่ดันออกมาทันที เสียงพูดที่เกิดจากการกักลมไว้นี้เรียกว่าเสียงกักฐานเส้นเสียง (Glottal Stop) เช่น อ[ʔ]
3) เส้นเสียงสั่น (Vibrating) เกิดขึ้นเมื่อเส้นเสียงถูกดึงให้เลื่อนมาติดกัน แต่เมื่อมีลมผ่านจากปอดดันขึ้นมาอย่างแรงทำให้เส้นเสียงนี้เปิดออกในลักษณะพลิ้วไหวและสั่น เสียงพูดที่เกิดจากเส้นเสียงลักษณะนี้เรียกว่าเสียงโฆษะหรือเสียงก้อง (Voiced) ได้แก่เสียงสระและพยัญชนะบางส่วน เช่น ด[d] บ[b] ม[m] น[n] ร[r] ล[l] ว[w]
4) เส้นเสียงปิดไม่สนิท เกิดขึ้นเมื่อเส้นเสียงส่วนหน้าแผ่ตัวปิดเข้าหากัน แต่ส่วนหลังเปิดออกเล็กน้อย ขณะออกเสียงเส้นเสียงตอนหน้าที่อยู่ชิดกันเกิดการสั่นสะเทือนด้วย ลักษณะของเส้นเสียงชนิดนี้เกิดขึ้นเมื่อพูดเสียงกระซิบ

3. อวัยวะที่อยู่เหนือเส้นเสียง

อวัยวะที่มีส่วนในการเปลี่ยนแปลงลักษณะของเสียงเป็นอวัยวะที่อยู่เหนือเส้นเสียงทั้งหมด ลมจากปอดที่ผ่านเส้นเสียงแล้วจะผ่านขึ้นมายังช่องคอ และจะออกสู่ภายนอกทางช่องปากหรือช่องจมูก ช่องใดช่องหนึ่งหรือทั้ง 2 ช่องพร้อม ๆ กัน ดังแสดงในรูปที่ 2.5

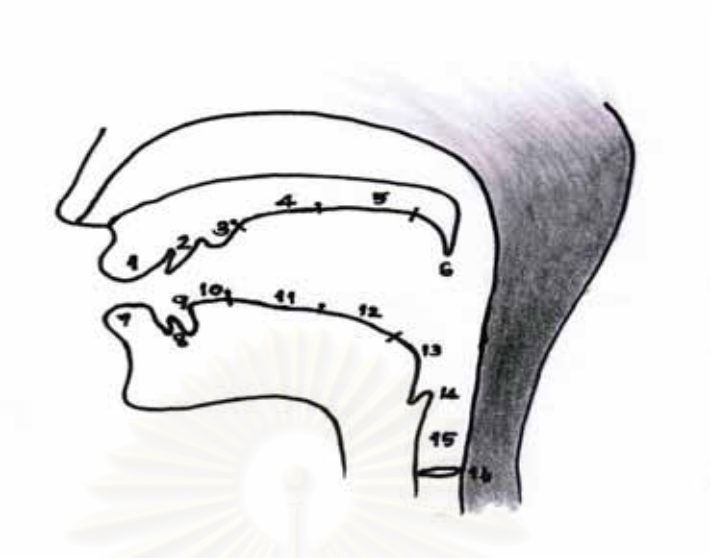


รูปที่ 2.5 อวัยวะที่อยู่เหนือเส้นเสียง

อวัยวะที่อยู่เหนือเส้นเสียงซึ่งมีองค์ประกอบของอวัยวะต่าง ๆ ดังแสดงในรูปที่ 2.6 เหล่านี้อาจจำแนกตามหน้าที่ในการช่วยให้เกิดเสียงเป็น 2 ประเภท[1] คือ

1) อวัยวะที่เป็นตำแหน่งที่เกิดเสียงต่าง ๆ (Passive Articulator) เรียกอีกชื่อหนึ่งว่า ฐานกรณ์บน เป็นอวัยวะที่ไม่เคลื่อนที่ในขณะที่ออกเสียงอยู่ทางด้านบนของช่องปาก ได้แก่ ริมฝีปากบน ฟันบน ปุ่มเหงือก เพดานแข็ง เพดานอ่อน และ ลิ้นไก่
2) อวัยวะที่เป็นส่วนกระทำการ (Active Articulator) เรียกอีกชื่อหนึ่งว่า ฐานกรณ์ล่าง เป็นอวัยวะที่เคลื่อนที่ขณะออกเสียง อยู่ด้านล่างของช่องปาก ได้แก่ ริมฝีปากล่าง ฟันล่าง ปลายลิ้น ลิ้นส่วนปลาย ลิ้นส่วนหน้าและลิ้นส่วนหลัง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 2.6 ตำแหน่งต่าง ๆ ของอวัยวะในช่องปาก

1. ริมฝีปากบน (Upper lip)
2. ฟันบน (Upper teeth)
3. ปุ่มเหงือก (Alveolar ridge)
4. เพดานแข็ง (Hard palate)
5. เพดานอ่อน (Soft palate)
6. ลิ้นไก่ (Uvula)
7. ริมฝีปากล่าง (Lower lip)
8. ฟันล่าง (Lower teeth)
9. ปลายลิ้น (The tip of the tongue)
10. ลิ้นส่วนปลาย (The blade of the tongue)
11. ลิ้นส่วนหน้า (The front of the tongue)
12. ลิ้นส่วนหลัง (The back of the tongue)
13. โคนลิ้น (The root of the tongue)
14. แผ่นเนื้อปากหลอดลม (Epiglottis)
15. กรวยคอ (Pharynx)
16. เส้นเสียง (Vocal cord)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

คำอธิบายรูปที่ 2.6

- หมายเลข 1,7 ริมฝีปากบนและริมฝีปากล่าง เป็นอวัยวะที่เคลื่อนไหวมาก ริมฝีปากคู่นี้จะดัดแปลงลมให้เป็นเสียงต่าง ๆ กัน โดยปิดริมฝีปาก เปิดเล็กน้อย เปิดกว้าง รูปปากแผ่ หรือรูปปากห่อ หมายเลข 2,8 ฟัน ช่วยกักลมทำให้เกิดเสียงหลายประเภท เช่น เมื่อฟันบนกดลงบนฟันล่างและปล่อยให้ลมเสียดแทรกช่องฟันออกมาเป็นเสียง ฟ[f] หรือใช้ลิ้นส่วนปลายแตะหลังฟันบน ทำให้เกิดเสียงต่าง ๆ เช่น ด[d] ต[t] ฯลฯ
- หมายเลข 3 ปุ่มเหงือก คือ ส่วนนูนออกมาจากแผ่นเพดานบน อยู่หลังฟันบน เป็นส่วนสำคัญในการเปล่งเสียงพยัญชนะอีกส่วนหนึ่ง เช่น ร[r] ล[l] ฯลฯ
- หมายเลข 4 เพดานแข็ง คือ ส่วนโค้งที่เป็นกระดูกแข็งอยู่ถัดจากปุ่มเหงือกเข้าไป เป็นบริเวณที่เกิดเสียงพยัญชนะอีกส่วนหนึ่ง เช่น จ[tɕ] ช[tɕ^h] ฯลฯ
- หมายเลข 5 เพดานอ่อน คือ กระดูกที่อยู่ต่อกับเพดานแข็ง สามารถเคลื่อนขยับขึ้นลงได้เล็กน้อย เป็นบริเวณที่เกิดเสียงพยัญชนะอีกส่วนหนึ่ง เช่น ก[k] ค[k^h] ฯลฯ
- หมายเลข 6 ลิ้นไก่ เป็นก้อนเนื้อเล็ก ๆ อยู่ปลายเพดานอ่อนตรงกลางช่องปาก สามารถขยับขึ้นลงได้เป็นอวัยวะที่สำคัญที่ช่วยในการเปล่งเสียงทางจมูกและเสียงที่ไม่ใช่เสียงทางจมูก
- หมายเลข 9 -13 ลิ้น เป็นอวัยวะที่เคลื่อนไหวได้มากในการเปล่งเสียงพูด แบ่งออกเป็น 5 ส่วนตามตำแหน่งที่ช่วยให้เกิดเสียงต่าง ๆ กัน ได้แก่
    - หมายเลข 9 ปลายลิ้น คือ ส่วนปลายสุดของลิ้น สามารถยกขึ้นไปแตะอวัยวะต่าง ๆ ในปากส่วนบนได้ง่าย ช่วยให้เกิดเสียงพยัญชนะ ร[r] ล[l] ฯลฯ
    - หมายเลข 10 ลิ้นส่วนปลาย คือ ส่วนที่ติดกับปลายลิ้น สามารถขึ้นไปแตะอวัยวะในปากส่วนบนได้ ช่วยให้เกิดเสียงพยัญชนะ ต[t] ด[d] ฯลฯ
    - หมายเลข 11 ลิ้นส่วนหน้า คือ ส่วนที่อยู่ตรงกับเพดานแข็ง เมื่อวางลิ้นราบ ช่วยให้เกิดเสียงพยัญชนะ จ[tɕ] ช[tɕ^h] และสระหน้า ฯลฯ

- หมายเลข 12 ลิ้นส่วนหลัง คือ ส่วนของลิ้นที่อยู่ตรงข้ามกับเพดานอ่อน เมื่อวางลิ้นราบ ช่วยในการเกิดเสียงพยัญชนะ ก[k] ค[k^h] และสระหลัง ฯลฯ
- หมายเลข 13 โคนลิ้น คือ ส่วนที่อยู่ต่อจากลิ้นส่วนหลังลงไปในลำคอ

- หมายเลข 14 แผ่นเนื้อปากหลอดลม เป็นก้อนเนื้อเล็ก ๆ คล้ายลิ้นไก่อยู่ต่อจากโคนลิ้นลงไปในช่องคอ มีหน้าที่ปิดช่องลมเมื่อกลืนอาหารและเปิดช่องลมเมื่อพูด
- หมายเลข 15 กรวยคอ คือ โพรงคอที่อยู่ถัดจากช่องปากลงไปถึงเส้นเสียง
- หมายเลข 16 เส้นเสียง เป็นส่วนที่สำคัญที่ช่วยให้เกิดเสียงพูด

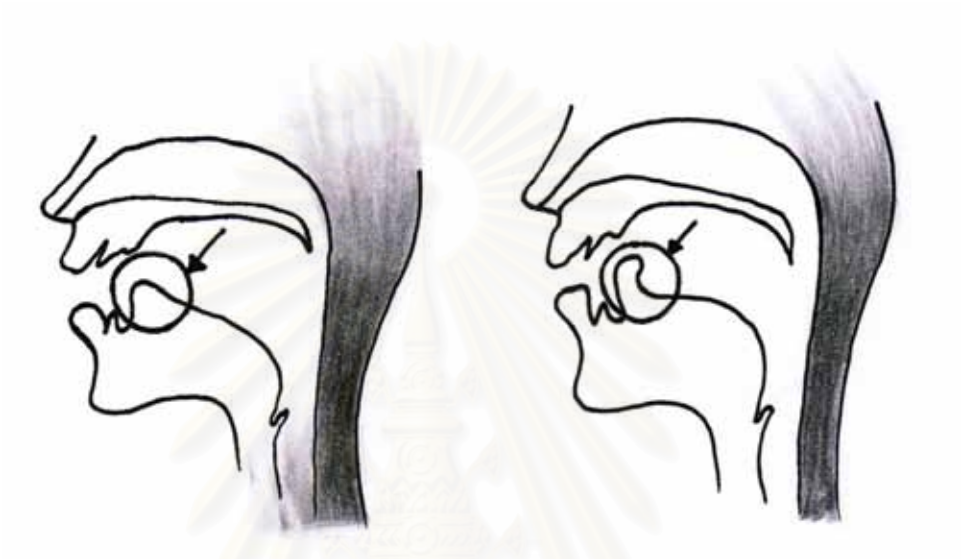
ตารางที่ 2.1 ตำแหน่งของอวัยวะในช่องทางเสียงที่สัมพันธ์กับเสียงพยัญชนะไทย

ฐานกรณ์ล่าง	ฐานกรณ์บน	พยัญชนะ
ริมฝีปากล่าง	ริมฝีปากบน	พ ภ ผ ป บ
	ริมฝีปากบน-ลิ้นไก่	ม
	ฟันบน	ฟ ฝ
ปลายลิ้น	ปุ่มเหงือก	ถ ท ธ ฒ ฐ ฑ ต ฏ ด ฎ ร ล ศ
	ปุ่มเหงือก-ลิ้นไก่	น ณ
	หลังปุ่มเหงือก	ซ ศ ษ ส
ลิ้นส่วนหน้า	หลังปุ่มเหงือก	จ ฉ ช ฌ
กลางลิ้น	เพดานแข็ง	ย ญ
ลิ้นส่วนหลัง	เพดานอ่อน	ก ข ค ฆ
	เพดานอ่อน-ลิ้นไก่	ง
	เพดานอ่อน-ริมฝีปากบน	ว
เส้นเสียง		ห อ ฮ

อวัยวะต่าง ๆ ภายในบริเวณช่องปากนี้ ในการเปล่งเสียงนั้นอวัยวะสำคัญที่ทำหน้าที่ดัดแปลงลมให้เกิดเป็นเสียงต่าง ๆ ได้แก่ ลิ้น ริมฝีปาก ลิ้นไก่ และเพดานอ่อน

1) ลิ้น เป็นอวัยวะที่เคลื่อนไหวได้อย่างคล่องแคล่วในหลายทิศทาง ทั้งกระดกขึ้น กระดกลง แลบยื่นไปข้างหน้า ดึงไปข้างหลัง นอกจากนี้ปลาย

ลิ้นยังเคลื่อนตัวได้มากกว่าส่วนอื่น อาจกระดกขึ้นไปในทิศทางของฟันบนและปุ่มเหงือก หรืองอปลายลิ้นกลับเข้าไปด้านในและด้านล่างของลิ้นกระดกขึ้นไปในทิศทางของฟันบน ปุ่มเหงือก เพดานแข็งแสดงในรูปที่ 2.7 (ก) – (ข) ทำให้เกิดเสียงต่าง ๆ กันได้มาก นอกจากนี้ลักษณะของลิ้นยังแบ่งออกเป็น 3 ลักษณะ ดังแสดงในรูปที่ 2.7 (ค) – (จ)



(ก) ปลายลิ้นกระดกขึ้นโดยไม่งอกลับ (ข) ปลายลิ้นกระดกขึ้นและงอกลับ

(ค) กลม (ง) แบนราบ (จ) มีร่องตรงกลาง

รูปที่ 2.7 ลิ้นลักษณะต่าง ๆ

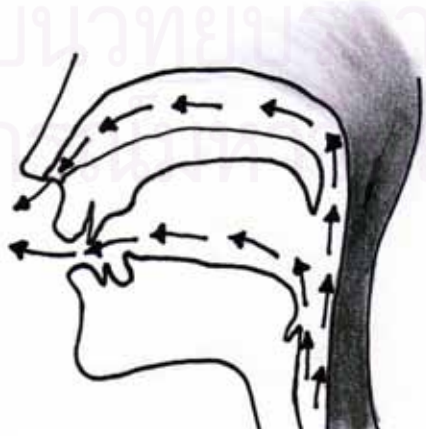
2) ริมฝีปาก มีส่วนในการดัดแปลงลมให้เป็นเสียงต่าง ๆ กัน โดยริมฝีปากอาจปิดสนิท เปิดเล็กน้อย เปิดกว้าง ห่อลม หรือทำเป็นรูปรีก็ได้ดังแสดงในรูป 2.8

(ก) ห่อ (ข) เปิดกว้าง (ค) เปิดธรรมดา

รูปที่ 2.8 ริมฝีปากลักษณะต่าง ๆ

3) ลิ้นไก่และเพดานอ่อน เป็นส่วนที่เคลื่อนไหวได้เล็กน้อย โดยอาจยกขึ้นไปติดกับผนังคอทำให้ลมจากปอดเคลื่อนไปสู่ช่องจมูกไม่ได้ หรืออาจลดตัวลงปล่อยให้ลมจากปอดเข้าสู่ช่องจมูก

(ก) ลมออกทางช่องปาก (ข) ลมออกทางช่องจมูก



(ค) ลมออกทางช่องปากและช่องจมูก

รูปที่ 2.9 การเคลื่อนตัวของลิ้นไก่และเพดานเพื่อควบคุมทิศทางของลม

ทฤษฎีด้านอวัยวะออกเสียงร่วม

จากความรู้ในทฤษฎีบทเรื่องอวัยวะในการเปล่งเสียงข้างต้นทำให้ทราบถึงรูปแบบลักษณะของอวัยวะในการเปล่งเสียง ในการเปล่งเสียงพูดทั้งพยัญชนะและสระภาษาไทย แต่ในกลไกการเปล่งเสียงนั้นยังต้องมีการทำการศึกษาถึงเรื่องการใช้อวัยวะออกเสียงร่วมด้วย

อวัยวะออกเสียงร่วมคือการที่อวัยวะในการเปล่งเสียงมีการจัดรูปร่างของช่องทางเสียงของหน่วยเสียงพยัญชนะและสระขณะเวลาเดียวกันเพื่อที่จะเปล่งเสียงพยางค์หนึ่งๆ รูปร่างของช่องทางเสียงพยัญชนะจะมีการเปลี่ยนแปลงไปขึ้นกับหน่วยเสียงสระที่อยู่ติดกัน[4]

ผลของการใช้อวัยวะออกเสียงร่วมจะเห็นได้ชัดเจนพยัญชนะกลุ่มลิ้นส่วนหลัง-เพดานอ่อน[5,6] เมื่อพิจารณาถึงรูปร่างของช่องเสียงพยัญชนะต้นเดียวกันที่มีการผสมกับรูปร่างของช่องทางเสียงสระที่ต่างกัน ตัวอย่างเช่น กา กับ กี หากทดลองจัดรูปร่างของช่องทางเสียงเพื่อเตรียมที่จะเปล่งเสียงพยางค์ทั้ง 2 จะพบว่าตำแหน่งของลิ้นที่ยกขึ้นไปแตะเพดานปากจะมีการเลื่อนตำแหน่งมาด้านหน้าของช่องปาก เนื่องจากผลของตำแหน่งลิ้นของเสียงสระ อา และ อี ตามลำดับ ดังนั้นจึงสามารถกล่าวได้ว่ารูปร่างของช่องทางเสียงพยัญชนะต้นถูกเปลี่ยนแปลงไปจากผลของรูปร่างช่องทางเสียงสระที่ติดกัน ในกลไกการเปล่งเสียงพยางค์หนึ่ง ๆ จึงต้องพิจารณาถึงผลของการใช้อวัยวะออกเสียงร่วมด้วย

ทฤษฎีด้านภาษาศาสตร์

หลักการที่ใช้ในการสังเคราะห์แบบจำลองสามมิติเพื่อแสดงกระบวนการการเปล่งเสียงหนึ่ง ๆ มีความเกี่ยวข้องกับทฤษฎีด้านภาษาศาสตร์และทฤษฎีที่เกี่ยวข้องกับเสียงและอวัยวะที่ใช้ในการเปล่งเสียง ซึ่งทฤษฎีทางภาษาศาสตร์และทฤษฎีที่เกี่ยวข้องกับเสียงและอวัยวะที่ใช้ในการเปล่งเสียงใช้ในการทำความเข้าใจลักษณะของการเปล่งเสียงภาษาไทย เช่น อวัยวะที่เกี่ยวข้องกับการเปล่งเสียงนั้น ๆ ลักษณะของลิ้นและริมฝีปาก เป็นต้น เมื่อเข้าใจหลักการการเปล่งเสียงแล้วก็สามารถนำความรู้ส่วนนี้ไปหาค่าพารามิเตอร์ที่เหมาะสมแก่แบบจำลองสามมิติ

หน่วยเล็กที่สุดของการพูดภาษาไทย[1,7]คือ พยางค์ ในการเปล่งเสียงแต่ละพยางค์ประกอบด้วยส่วนที่สำคัญได้แก่ เสียงสระ 1 เสียง เสียงพยัญชนะ 1 เสียงหรือเสียงพยัญชนะควบกล้ำ และเสียงวรรณยุกต์ 1 เสียง ส่วนเสียงพยัญชนะท้ายอาจมีหรือไม่มีก็ได้ เช่น “กิน” ประกอบด้วยเสียงพยัญชนะต้น คือ “ก” มีเสียงสระคือ “อิ” เสียงพยัญชนะท้ายคือ “น” และเสียงวรรณยุกต์เป็นเสียงสามัญ

ภาษาไทยมีตัวพยัญชนะทั้งหมด 44 รูป แต่เปล่งเสียงได้ต่างกัน 21 เสียง ทั้ง 21 เสียงสามารถเกิดต้นพยางค์หรือเป็นพยัญชนะต้นได้ แต่เสียงพยัญชนะสะกดมีเพียง 9 หน่วยเสียง ส่วนเสียงสระประสม 3 เสียง เสียงสระเดี่ยวมีความแตกต่างกันในเรื่องความสั้น-ยาวของเสียง ดังนั้นจึงสามารถแบ่งได้เป็นสระเสียงสั้น 9 เสียง และสระเสียงยาว 9 เสียง ส่วนวรรณยุกต์ในภาษาไทยมี 4 รูป แต่มี 5 เสียง[1,2,7] ดังรายละเอียดต่อไปนี้

1. เสียงพยัญชนะ

เสียงพยัญชนะเกิดจากอวัยวะในการเปล่งเสียงมีลักษณะแตกต่างกันออกไปดังที่ได้กล่าวในหัวข้ออวัยวะในการเปล่งเสียง การบอกลักษณะของเสียงจะต้องบอกถึงลักษณะทั้ง 4 แบบ[1] ดังนี้

1) เสียงก้องหรือไม่ก้อง (Voiced or Voiceless) หรือในภาษาไทยเรียกว่า เสียงโฆษะและอโฆษะ โดยพิจารณาจากเส้นเสียงเกิดการสั่นหรือไม่ หากเส้นเสียงสั่นเสียงที่เปล่งออกมาคือเสียงโฆษะ และในทางกลับกัน หากเส้นเสียงไม่มีการสั่น เส้นเสียงเปิดหรือปิดอยู่เฉย ๆ เสียงที่เปล่งออกมาก็คือเสียงอโฆษะ

2) พ่นลมหรือไม่พ่นลม (Aspirated or Unaspirated) หรือในภาษาไทยเรียกว่า เสียงธนิตและเสียงสิถิล โดยพิจารณาจากกลุ่มลมที่พ่นออกมาจากอวัยวะในช่องปากหลังจากมีการเปิดช่องปิดในช่องปาก หากมีกลุ่มลมพ่นตามออกมาเสียงที่เปล่งออกมาก็คือเสียงพ่นลม หากไม่มีกลุ่มลมพ่นออกมาก้จะเป็นเสียงไม่พ่นลม

3) ตำแหน่งการเกิดเสียง (Place of Articulation) พิจารณาถึงฐานกรณ์ล่างและฐานกรณ์บนคู่ใดทำให้เกิดเสียง ดังนี้

3.1) ริมฝีปากล่างกับฟันบน (Labio-dental)

3.2) ริมฝีปากล่างกับริมฝีปากบน (Bilabial)

3.3) ปลายลิ้นกับฟันบน (Apico-dental)

3.4) ปลายลิ้นกับปุ่มเหงือก (Apico-alveolar)

3.5) ปลายลิ้นกับหลังปุ่มเหงือก (Apico-prepalatal)

3.6) ลิ้นส่วนหน้ากับปุ่มเหงือก (Lamino- alveolar)

3.7) ลิ้นส่วนหน้ากับหลังปุ่มเหงือก (Lamino-prepalatal)

3.8) กลางลิ้นกับเพดานแข็ง (Fronto-palatal)

3.9) ปลายลิ้นและลิ้นส่วนหน้ากับปุ่มเหงือกและหลังปุ่มเหงือก (Palato-alveolar)

3.10) ลิ้นส่วนหลังกับเพดานอ่อน (Dorso-velar)

3.11) ลิ้นส่วนหลังกับลิ้นไก่ (Dorso-uvular)

3.12) เส้นเสียงกับเส้นเสียง (Glottal)

3.13) ริมฝีปากล่างและริมฝีปากบนกับลิ้นส่วนหลังและเพดานอ่อน (Labio-velar)

3.14) ปลายลิ้นและหลังปุ่มเหงือก โดยลิ้นงอม้วนกลับ (Retroflex)

4) ลักษณะการออกเสียง (Manner of Articulation) พิจารณาการบังคับลมในช่องทางเสียงว่ามีรูปแบบใด ได้แก่ เสียงกัก(Stop) เสียงเสียดแทรก (Fricative) เสียงกึ่งเสียดสี(Affricate) เสียงนาสิก(Nasal) เสียงข้างลิ้น (Lateral) เสียงลิ้นกระดก(Flap) เสียงรัว(Trill)และเสียงเปิด(Approximant)

เสียงพยัญชนะ เสียงสำคัญที่ใช้ในภาษาไทยอีกพวกหนึ่งคือเสียงพยัญชนะ เป็นเสียงที่เกิดจากลมซึ่งผ่านเส้นเสียง แล้วมาถูกดัดแปลงด้วยอวัยวะเปล่งเสียงส่วนต่าง ๆ ในปาก ทำให้เกิดเสียงขึ้น เสียงพยัญชนะมีหลายประเภทมีลักษณะการเปล่งเสียงแตกต่างกันหลายแบบ ความแตกต่างของเสียงจะมีความสำคัญยิ่งขึ้นเมื่อความแตกต่างนั้นทำให้ความหมายของคำในภาษาต่างกันหรือเรียกได้ว่า เป็นความแตกต่างที่ทำให้เสียงสองเสียงทำหน้าที่เป็นสองหน่วยเสียงต่างกันในภาษานั้นเพื่อให้ทราบว่าในภาษามีหน่วยเสียงกี่หน่วย หน่วยเสียงพยัญชนะในภาษาไทยมี 21 หน่วยเสียง

หน่วยเสียงพยัญชนะในภาษาไทยจะปรากฏในตำแหน่งต่าง ๆ ได้ 4 ตำแหน่ง คือ

1) เกิดต้นคำ นำหน้าเสียงสระในพยางค์หนึ่ง ๆ เป็นหน่วยเสียงพยัญชนะต้น
2) เกิดนำเสียงพยัญชนะอื่นอีกเสียงหนึ่งในตำแหน่งต้นคำเรียกว่า พยัญชนะควบ
3) เกิดตามเสียงพยัญชนะอื่นในต้นคำ คือ เกิดควบกับพวกที่ 2
4) เกิดตามหลังเสียงสระ เป็นเสียงพยัญชนะสะกด

ตารางที่ 2.2 เสียงพยัญชนะภาษาไทยแบ่งตามลักษณะของเสียง[1]

ลักษณะการออกเสียง (Manner of Articulation)			ตำแหน่งการเกิดเสียง (Place of Articulation)								
			Bilabial	Labio-dental	Apico-alveolar	Apico-prepalatal	Lamino-prepalatal	Fronto-palatal	Dorso-velar	Labio-velar	Glottal
เสียงกัก (Stop)	พ่นลม (Aspirated)	เสียงไม่ก้อง (Voiceless)	พ ภ ผ [p^h]		ฐ ฑ ฒ ถ ท ธ [t^h]				ข ฃ ฅ ค ฆ [k^h]		อ [ʔ]
	ไม่พ่นลม (Unaspirated)	เสียงไม่ก้อง (Voiceless)	ป [p]		ต ฏ [t]				ก [k]		
		ก้อง(Voiced)	บ [b]		ด ฎ [d]						
เสียงเสียดแทรก (Fricative)	เสียงไม่ก้อง (Voiceless)			ฟ ฝ [f]		ซ ศ ษ ส [s̩]					ห ฮ[h]
เสียงกึ่งเสียดสี (Affricate)	พ่นลม (Aspirated)	เสียงไม่ก้อง (Voiceless)					ฉ ช ฌ [$tɕ^h$]				
	ไม่พ่นลม (Unaspirated)						จ [tɕ]				
เสียงนาสิก (Nasal)	เสียงก้อง (Voiced)		ม [m]		ณ น [n]				ง [ŋ]		
เสียงข้างลิ้น (Lateral)	เสียงก้อง (Voiced)				ล ฬ [l]						
เสียงรัว (Trill)	เสียงก้อง (Voiced)				ร [r]						
เสียงเปิด (Approximant)	เสียงก้อง (Voiced)							ญ ย [j]		ว [w]	

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

2. เสียงสระ

หน่วยเสียงที่สำคัญในภาษาทุกภาษาคือหน่วยเสียงสระ เสียงสระเป็นเสียงที่เกิดจากลมที่ผ่านเส้นเสียงในตำแหน่งปิดเกือบสนิท และลมที่ต้องดันตัวออกมาทำให้เส้นเสียงสั่นเกิดเป็นเสียงดังขึ้นเรียกว่าเสียงก้อง เสียงที่เรียกว่าเสียงสระจะต้องถูกเปล่งออกมาทางปากโดยที่ไม่มีอวัยวะส่วนใดในปากมาปิดกั้นทางลมไว้ แต่อวัยวะในปากอาจจะอยู่ในท่าและตำแหน่งต่าง ๆ ที่ทำให้โพรงปากมีลักษณะต่างกันได้หลายแบบ ลมที่ผ่านออกมาจึงเกิดเป็นเสียงต่าง ๆ กัน การห่อริมฝีปากหรือไม่ห่อริมฝีปาก การยกลิ้นส่วนหนึ่งส่วนใด การยกลิ้นสูงต่ำต่าง ๆ ย่อมมีส่วนในการทำให้เกิดเสียงสระต่าง ๆ กันทั้งสิ้น หน่วยเสียงสระในภาษาไทยมีทั้งหมด 21 หน่วย แบ่งเป็น 2 ประเภทคือ สระเดี่ยว และสระเลื่อนหรือสระประสม เป็นหน่วยเสียงสระเสียงสั้น 9 หน่วย หน่วยเสียงสระเดี่ยวเสียงยาว 9 หน่วย และหน่วยเสียงสระประสม 3 หน่วย

1) สระเดี่ยว (Monophthongs) ได้แก่ สระที่ขณะเปล่งเสียงลักษณะของลิ้นและริมฝีปากไม่มีการเปลี่ยนแปลงหรือมีการเปลี่ยนแปลงน้อยมาก ตำแหน่งจุดสูงสุดของลิ้นจะมีตำแหน่งเดียวแบ่งเป็น 2 ชนิดดังนี้
   1.1) หน่วยเสียงสระเดี่ยวเสียงสั้น 9 หน่วยมี อะ[a] อิ[i]อึ[ɯ]อุ[u] เอะ[e] แอะ[æ] เออะ[ɤ] โอะ[o]และ เอาะ[ɔ]
   1.2) หน่วยเสียงสระเดี่ยวเสียงยาว 9 หน่วยมี อา[aː] อี[iː] อือ[ɯː] อู[uː] เอ[eː] แอ[æː] เออ[ɤː] โอ[oː] และ ออ[ɔː]
2) สระเลื่อนหรือสระประสม ได้แก่ สระที่ขณะเปล่งเสียง ลักษณะของลิ้นมีการเปลี่ยนแปลงอย่างชัดเจน ส่วนลักษณะริมฝีปากจะเปลี่ยนแปลงหรือไม่ก็ได้ สระประเภทนี้ตำแหน่งจุดสูงสุดของลิ้นมีมากกว่า 1 จุด แบ่งเป็น 2 ประเภท คือ
   2.1) สระเลื่อนสองส่วน (Diphthongs) ได้แก่ เสียงสระซึ่งเลื่อนไปในทิศทางเดียว ซึ่งเสียงสระประสมภาษาไทยอยู่ในประเภทนี้ ได้แก่ หน่วยเสียงสระประสม 3 หน่วย มีเสียงย่อยหน่วยละ 2 เสียง เป็นสระเสียงสั้นและสระเสียงยาว คือ เอียะ[ia] เอีย[iːa] เอือะ[ɯa] เอือ[ɯːa] อัวะ[ua] และ อัว[uːa]
   2.2) สระเลื่อนสามส่วน (Triphthongs) ได้แก่ เสียงสระซึ่งเลื่อนไปในทิศทางหนึ่งแล้วเลื่อนต่อไปยังอีกทิศทางหนึ่ง

เสียงสระเป็นเสียงก้องทุกเสียงที่ดังกระจายไปได้ไกลกว่าเสียงพยัญชนะ เสียงสระจึงเป็นแกนของพยางค์ซึ่งทำให้เสียงอื่น ๆ ในพยางค์นั้นได้ยินได้ หน่วยเสียงสระในภาษาไทยสามารถเกิดร่วมกับวรรณยุกต์ได้ทุกหน่วย แต่เกิดตามหลังหน่วยเสียงพยัญชนะต้นและนำหน่วยเสียงพยัญชนะสะกดไม่ได้ทุกหน่วย

เนื่องจากเสียงสระเป็นเสียงก้อง เกิดจากการสั่นของเส้นเสียงผ่านอวัยวะในช่องทางเสียงแล้วออกจากปากโดยไม่มีอวัยวะอื่นใดมาขวางทางออกของเสียง ถ้าดูรูปคลื่นของเสียงสระจะพบว่าเป็นรูปคลื่นที่เป็นรายคาบ โดยมีคลื่นเสียงที่มีความถี่หลักมูลเท่ากับอัตราที่เส้นเสียงสั่น และเนื่องจากเกิดการกำทอนจากช่องทางเสียงคือช่องปากและลำคอทำให้มีค่าความถี่ฮาร์มอนิกส์อื่นผ่านออกมาได้มากน้อยต่าง ๆ กัน ปนเข้ากับคลื่นเสียงจากเส้นเสียงด้วย ทำให้เกิดเป็นคลื่นเสียงที่ซับซ้อน

เสียงสระแต่ละตัวจะมีลักษณะเฉพาะตัวของความถี่เสียงที่ต่างกันออกไป ถ้านำสระต่าง ๆ มาวิเคราะห์หาองค์ประกอบทางความถี่ดู จะพบว่า สระต่าง ๆ จะมีรูปแบบความถี่ฟอร์แมนท์ (Formant frequency) หรือความถี่การกำทอนที่ต่างกันออกไป เนื่องจากรูปร่างของช่องทางเสียงที่จะเปลี่ยนไปตามแต่ละสระ นั่นคือจะทำให้ค่าความถี่ฟอร์แมนท์ของสระต่าง ๆ มีค่าแตกต่างกัน

การศึกษาเรื่องสระมีสิ่งที่ควรพิจารณาคือ อวัยวะในช่องปากที่เกี่ยวข้องกับการดัดแปลงลมจากปอดวางตัวห่างจากกัน ลมจากปอดจะผ่านบริเวณกลางของช่องปากออกมาได้โดยสะดวกไม่ถูกปิดกั้น อวัยวะที่มีบทบาทในการดัดแปลงลมจากปอดมาเป็นเสียงสระต่าง ๆ คือ ลิ้นและริมฝีปาก[1,2]

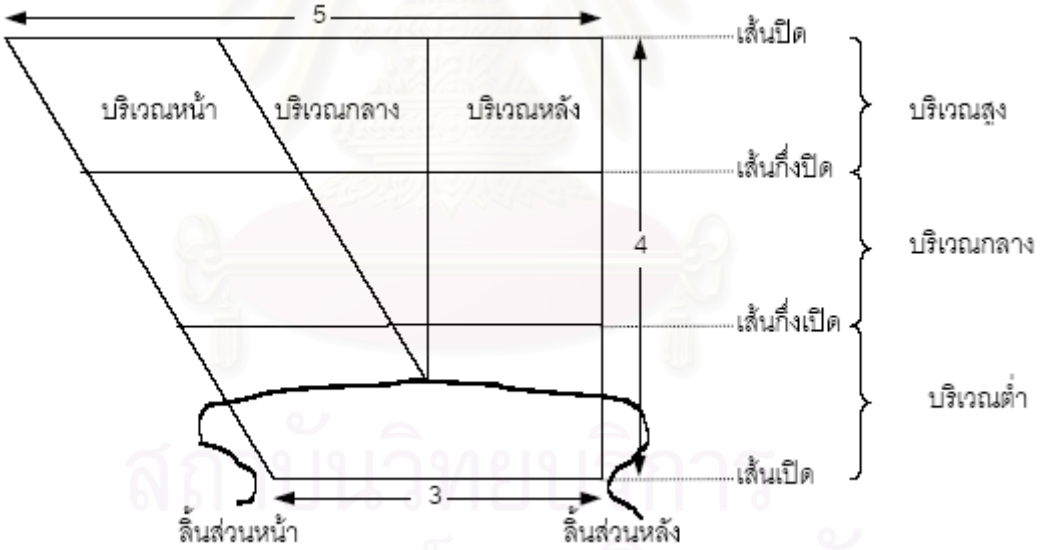
1) ลักษณะของลิ้น นักสัทศาสตร์ได้ใช้วิธีถ่ายภาพเอกซเรย์และภาพยนตร์เอกซเรย์ช่องปาก ในขณะที่ผู้พูดเปล่งเสียงสระต่าง ๆ พบว่าจุดสูงสุดของลิ้นจะอยู่ต่างกันเสมอ ดังนั้นสิ่งที่ต้องพิจารณาในการเปล่งสระ คือ ตำแหน่งของเสียงสระหรือที่เกิดของสระ การเคลื่อนไหวของลิ้นในการเปล่งเสียงสระเป็นดังนี้

1.1) ส่วนของลิ้น ขณะที่เปล่งเสียงจะใช้ลิ้นส่วนต่าง ๆ คือใช้ลิ้นส่วนหน้า ส่วนกลางและส่วนหลัง ยกขึ้นไปในระดับต่าง ๆ กัน จึงเรียกสระตามส่วนต่าง ๆ ของลิ้นที่ยกขึ้น ถ้าเปล่งเสียงสระโดยยกลิ้นส่วนหน้าขึ้นเรียกว่า สระหน้า ถ้ายกลิ้นไปส่วนหลังก็เรียกว่า สระหลัง

1.2) ระดับของลิ้น ความสูงต่ำของลิ้นที่ยกขึ้นใกล้เพดานมีส่วนสำคัญในการทำให้เสียงสระต่างกัน ถ้าลิ้นยกในระดับสูงช่องระหว่างลิ้นกับเพดานก็จะแคบ สระนั้นก็เรียกว่า สระปิด (Close vowel) ถ้าลิ้นยกในระดับต่ำช่องระหว่างลิ้นกับเพดานก็จะสูง สระนั้นก็เรียกว่า สระเปิด (Open vowel)

2) ลักษณะริมฝีปาก ในการเปล่งเสียงสระริมฝีปากจะอยู่ในลักษณะต่าง ๆ กัน คือ ริมฝีปากห่อกลม แผ่ หรืออยู่ในลักษณะปกติ

นักสัทศาสตร์ได้สร้างแผนภูมิสระขึ้นดังแสดงในรูป 2.10 โดยดัดแปลงรูปมาจาก บริเวณที่เกิดเสียงสระในปากประกอบด้วยเส้นนอนขนาน 4 เส้น เส้นตั้งที่อยู่ขวามือสุดจะตั้งฉากกับเส้นทั้ง 4 สัดส่วนระหว่างเส้นนอนเส้นบนสุดกับเส้นนอนเส้นล่างสุด และเส้นที่ตั้งฉากทางขวามือคือ 5:4:3 เรียกเส้นนอนทั้ง 4 จากบนลงล่างว่าเส้นปิด เส้นกึ่งปิด เส้นกึ่งเปิด และเส้นเปิด เรียกบริเวณที่อยู่ระหว่างเส้นทั้ง 4 จากบนลงล่างว่า บริเวณสูง บริเวณกลาง และบริเวณต่ำ และเรียกบริเวณของแผนภูมิสระจากซ้ายไปขวาว่า บริเวณหน้า บริเวณกลาง และบริเวณหลังตามลำดับ



รูปที่ 2.10 บริเวณการเกิดเสียงสระ[2]

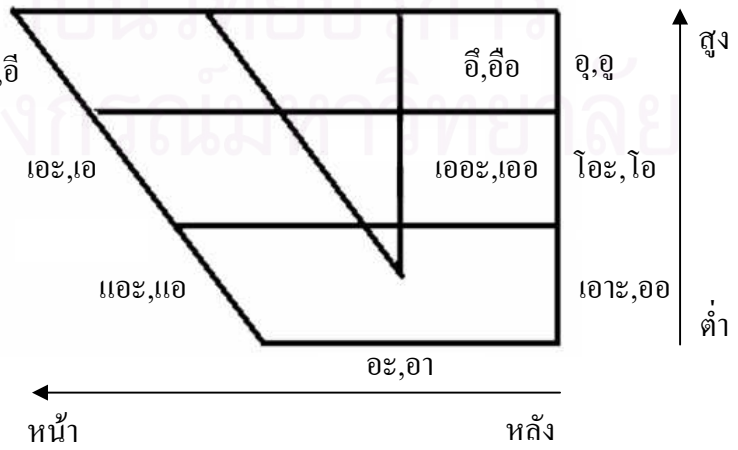
ก่อนที่จะพิจารณาการแทนตำแหน่งเสียงสระภาษาไทยในแผนภาพสระ จำเป็นที่จะต้องเข้าใจสัญลักษณ์ที่ใช้แทนสัทอักษรในภาษาสากลเสียก่อน ซึ่งสามารถแสดงได้ด้วยตารางที่ 2.4 โดยเลือกแสดงเฉพาะเสียงสระเท่านั้น สำหรับรายละเอียดของตารางเป็นการแสดงทั้งสระเดี่ยวและสระประสม ซึ่งสระเสียงสั้นสามารถแทนได้ด้วยสัทอักษรตัวเดียว ส่วนสระเสียงยาวที่มีบริเวณ

การเกิดเสียงตำแหน่งเดียวกัน แทนได้ด้วยสัทอักษรของคู่สระเสียงสั้นนั้น ๆ 2 ตัวติดกัน หรืออาจจะแทนได้ด้วยสัทอักษรของสระเสียงสั้นยาวด้วยเครื่องหมาย “ ː ”

ตารางที่ 2.3 เปรียบเทียบรูปสระในภาษาไทยกับสัทอักษรสากล

สระ	สัทอักษรสากล (IPA)
อะ,อา	[a],[aː]
อิ,อี	[i],[iː]
อึ,อือ	[ɯ],[ɯː]
อุ,อู	[u],[uː]
เอะ,เอ	[e],[eː]
แอะ,แอ	[æ],[æː]
โอะ,โอ	[o],[oː]
อัวะ,อัว	[ua],[uːa]
เอาะ,ออ	[ɔ],[ɔː]
เอียะ,เอีย	[ia],[iːa]
เอือะ,เอือ	[ɯa],[ɯːa]
เออะ,เออ	[ɤ],[ɤː]

เสียงสระในภาษาไทยสามารถแสดงตำแหน่งการเปล่งเสียงของสระภาษาไทยลงบนแผนภาพสระได้ดังรูปที่ 2.11



รูปที่ 2.11 ตำแหน่งของลิ้นในสระภาษาไทยตามแบบของแผนภาพสระ[8]

คำอธิบาย

1) หน้า กลาง หลัง เป็นชื่อเรียกสระตามส่วนของลิ้นที่ยกขึ้นใกล้เพดานดังนี้
   1.1) สระหน้า เกิดจากการที่เปล่งเสียงสระโดยปล่อยลมจากปอดให้ผ่านขึ้นมาในช่องปาก ในขณะที่ยกลิ้นส่วนหน้าขึ้นใกล้เพดานแข็ง สระหน้าในภาษาไทยมี 3 คู่ คือ [i],[iː],[e],[eː],[æ],[æː]
   1.2) สระกลาง เกิดจากการที่เปล่งเสียงสระโดยปล่อยลมออกจากปอดให้ผ่านขึ้นมาในช่องปาก ในขณะที่ยกลิ้นส่วนหน้ากับลิ้นส่วนหลังขึ้นใกล้บริเวณเพดานแข็งต่อกับบริเวณเพดานอ่อน สระกลางในภาษาไทยมี 1 คู่ คือ [a],[aː]
   1.3) สระหลัง เกิดจากการที่เปล่งเสียงสระโดยปล่อยลมออกจากปอดให้ผ่านขึ้นมาในช่องปาก ในขณะที่ยกลิ้นส่วนหลังขึ้นใกล้เพดานอ่อน สระหลังมี 5 คู่ คือ [ɯ],[ɯː],[ɤ],[ɤː],[u],[uː],[o],[oː],[ɔ],[ɔː]
2) สูง กลาง ต่ำ เป็นชื่อเรียกสระตามระดับความสูง-ต่ำ ของลิ้นดังนี้
   2.1) สระสูง เป็นสระที่เปล่งเสียงโดยยกลิ้นส่วนต่าง ๆ ขึ้นในระดับสูงสุดเท่าที่จะเกิดเสียงสระได้ สระสูงในภาษาไทยมี 3 คู่ คือ [i],[iː],[ɯ],[ɯː],[u],[uː]
   2.2) สระกลาง เป็นสระที่เปล่งเสียงโดยยกลิ้นส่วนต่าง ๆ ขึ้นในระดับกึ่งกลางระหว่างการเปล่งเสียงสระสูงกับสระต่ำ สระกลางในภาษาไทยมี 3 คู่ คือ [e],[eː],[ɤ],[ɤː],[o],[oː]
   2.3) สระต่ำ เป็นสระที่เปล่งเสียงโดยยกลิ้นส่วนต่าง ๆ ขึ้นในระดับต่ำสุดเท่าที่จะเกิดเสียงสระได้ สระต่ำในภาษาไทยมี 3 คู่ คือ [æ],[æː],[a],[aː],[ɔ],[ɔː]

สระเดี่ยวดังกล่าวอธิบายเสียงแต่ละเสียงได้ด้วยตารางที่ 2.4 ซึ่งให้ข้อมูลเกี่ยวกับระดับลิ้นในขณะที่เปล่งเสียงนั้น ๆ รวมถึงตำแหน่งของลิ้นที่ยกขึ้นสูงสุดระหว่างการเปล่งเสียงและลักษณะของริมฝีปากแบบห่อรมฝีปากหรือไม่ห่อริมฝีปาก

ตารางที่ 2.4 เสียงสระตามระดับลิ้น ตำแหน่งลิ้น และลักษณะริมฝีปาก

ตำแหน่งลิ้น	หน้า	กลาง	หลัง	
ริมฝีปาก / ระดับลิ้น	แผ่-รี	แผ่-รี	แผ่-รี	ห่อ-กลม
สูง	[i],[iː]		[ɯ],[ɯː]	[u],[uː]
กลาง	[e],[eː]		[ɤ],[ɤː]	[o],[oː]
ต่ำ	[æ],[æː]	[a],[aː]		[ɔ],[ɔː]

สระเลื่อนในภาษาไทยเป็นสระเลื่อนสองส่วนและเป็นสระเลื่อนลง (Opening diphthong) มี 6 เสียง คือ เอียะ[ia], เอีย[iːa], เอือะ[ɯa], เอือ[ɯːa], อัวะ[ua], อัว[uːa] ลักษณะการเลื่อนของทั้ง 6 สระเป็นดังนี้[1,2]

1) เอียะ เป็นสระที่เริ่มต้นที่สระสูงหน้า ริมฝีปากไม่ห่อ แล้วเลื่อนลิ้นไปยังตำแหน่งสระต่ำ กลาง ริมฝีปากไม่ห่อ ระยะเวลาการเปล่งเสียงสั้น
2) เอีย มีลักษณะเหมือนเอียะ แต่ระยะเวลาการเปล่งเสียงนานกว่า เป็นสระเสียงยาว
3) เอือะ เป็นสระที่เริ่มต้นที่สระสูงกลาง ริมฝีปากไม่ห่อ แล้วเลื่อนลิ้นไปยังตำแหน่งสระต่ำ กลาง ริมฝีปากไม่ห่อ ระยะเวลาการเปล่งเสียงนานพอ ๆ กับ เอียะ
4) เอือ มีลักษณะเหมือนเอือะ แต่ระยะเวลาการเปล่งเสียงนานกว่า เป็นสระเสียงยาว
5) อัวะ เป็นสระที่เริ่มต้นที่สระสูงหลัง ริมฝีปากไม่ห่อ แล้วเลื่อนลิ้นไปยังตำแหน่งสระต่ำ กลาง ริมฝีปากไม่ห่อ ระยะเวลาการเปล่งเสียงสั้น
6) อัว มีลักษณะเหมือนอัวะ แต่ระยะเวลาการเปล่งเสียงนานกว่า เป็นสระเสียงยาว

3. เสียงวรรณยุกต์

ในการเปล่งเสียงพูดในภาษาหนึ่ง ๆ นั้น ไม่ว่าผู้พูดจะเปล่งเสียงสระ พยัญชนะเรียงต่อกันออกมาเป็นคำวลีหรือเป็นประโยคก็ตาม ในการเปล่งเสียงนั้นผู้พูดจะดัดแปลงเสียงสระเสียงพยัญชนะนั้น ๆ ให้อยู่ในระดับเสียงสูง ต่ำ หรือเปลี่ยนแปลงเสียงขึ้นลงต่าง ๆ ได้อีก เสียงสูงต่ำในภาษาพูดเกิดได้ด้วยการสั่นสะเทือนของเส้นเสียงในอัตราต่าง ๆ กัน แต่เสียงที่เปล่งออกมาในขณะที่เส้นเสียงนั้นต้องเป็นเสียงก้องจึงเป็นเสียงที่ช่วยทำให้เกิดระดับสูงต่ำได้ โดยมองว่าระดับ

เสียงสูงต่ำของสระในพยางค์หนึ่ง ๆ ก็เป็นระดับเสียงสูงต่ำของพยางค์นั้นด้วย เพราะเมื่อเสียงสระซึ่งเป็นแกนของพยางค์มีระดับใด เสียงอื่น ๆ ในพยางค์นั้นก็ย่อมมีระดับเสียงนั้นตามไปด้วยเช่นกัน หรืออีกนัยหนึ่งคือ เสียงสูงต่ำในภาษา คือการที่ค่าความถี่มูลฐานของเสียงสระมีการเปลี่ยนแปลงไปตามเวลา

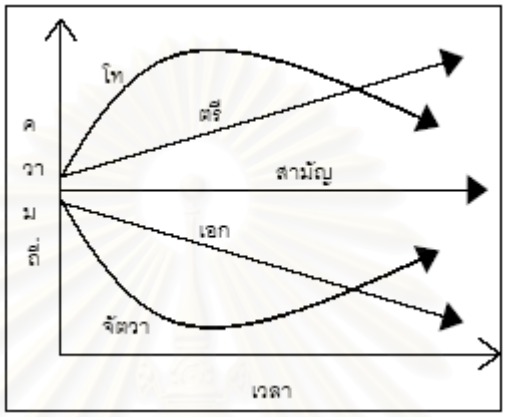
ในภาษาไทยระดับเสียงสูงต่ำของคำเรียกว่า วรรณยุกต์ เป็นส่วนสำคัญที่ทำให้คำตั้งแต่ 2 คำขึ้นไปซึ่งมีส่วนประกอบอื่น ๆ คือ พยัญชนะ สระ และตัวสะกดอย่างเดียวกันมีความหมายแตกต่างกันได้ วรรณยุกต์ในภาษาไทยจึงจัดเป็นหน่วยเสียงเรียกว่าหน่วยเสียงวรรณยุกต์ หน่วยเสียงวรรณยุกต์จัดว่าเป็นหน่วยเสียงซ้อน เสียงวรรณยุกต์ในภาษาไทยมีลักษณะเค้าโครงและความถี่ของเสียงโดยประมาณดังรูปที่ 2.12

หน่วยเสียงวรรณยุกต์ทั้ง 5 หน่วยนี้สามารถจัดกลุ่มได้เป็น 2 กลุ่ม[1,2,9] คือ

1) วรรณยุกต์คงระดับ มีระดับความถี่ของเสียงค่อนข้างคงที่ตลอดพยางค์ จัดได้ว่ามีวรรณยุกต์คงระดับในการออกเสียงพูดนั้น โดยปกติต้นพยางค์และท้ายพยางค์มักจะไม่อยู่ในระดับเดียวกัน ในทางสัทศาสตร์ถือว่าระดับเสียงที่แตกต่างกันไปเล็กน้อยเมื่อเปรียบเทียบกับวรรณยุกต์อีกพวกหนึ่ง เสียงวรรณยุกต์คงระดับในภาษาไทยมี 3 หน่วยเสียง คือ
   1.1) หน่วยเสียงวรรณยุกต์ระดับต่ำ (Low tone) คือ วรรณยุกต์เอก วรรณยุกต์ระดับต่ำจะปรากฏในพยางค์ได้ทุกแบบ
   1.2) หน่วยเสียงวรรณยุกต์ระดับกลาง (Mid tone) คือ วรรณยุกต์สามัญ วรรณยุกต์มีระดับเสียงกลาง ๆ ในปลายพยางค์จะมีความถี่ลดลง วรรณยุกต์สามัญไม่ปรากฏในพยางค์ที่สะกดด้วยพยัญชนะกัก
   1.3) หน่วยเสียงวรรณยุกต์ระดับสูง (High tone) คือ วรรณยุกต์ตรี เสียงวรรณยุกต์นี้มีลักษณะเด่นที่มีระดับสูง โดยจะค่อย ๆ สูงขึ้นที่ละน้อยจากต้นเสียง ในภาษาไทยหน่วยเสียงวรรณยุกต์ตรีไม่ปรากฏในพยางค์ที่ประสมด้วยสระเสียงยาวที่ตัวสะกดเป็นเสียงหลัก
2) วรรณยุกต์เปลี่ยนระดับ เป็นเสียงซึ่งมีระดับความถี่ของการเปล่งเสียงเปลี่ยนแปลงมากในช่วงพยางค์หนึ่ง ๆ เช่น ต้นพยางค์เปล่งเสียงให้มีระดับสูงแล้วลดระดับเสียงลงอย่างรวดเร็วไปสู่ระดับต่ำที่ปลายพยางค์ หรือต้นพยางค์เสียงมีระดับต่ำแล้วเสียงเพิ่มระดับเสียงอย่างรวดเร็วเป็นระดับสูงที่ปลายพยางค์ หรือเปลี่ยนสูงแล้วต่ำแล้วสูงอีกครั้งก็ได้ ในภาษาไทยวรรณยุกต์เปลี่ยนระดับมี 2 หน่วย คือ

2.1) หน่วยเสียงวรรณยุกต์เปลี่ยนตก (Falling tone) คือ เสียงวรรณยุกต์โท วรรณยุกต์โทจะไม่ปรากฏในพยางค์ที่มีสระเสียงสั้นและสะกดด้วยพยัญชนะกัก

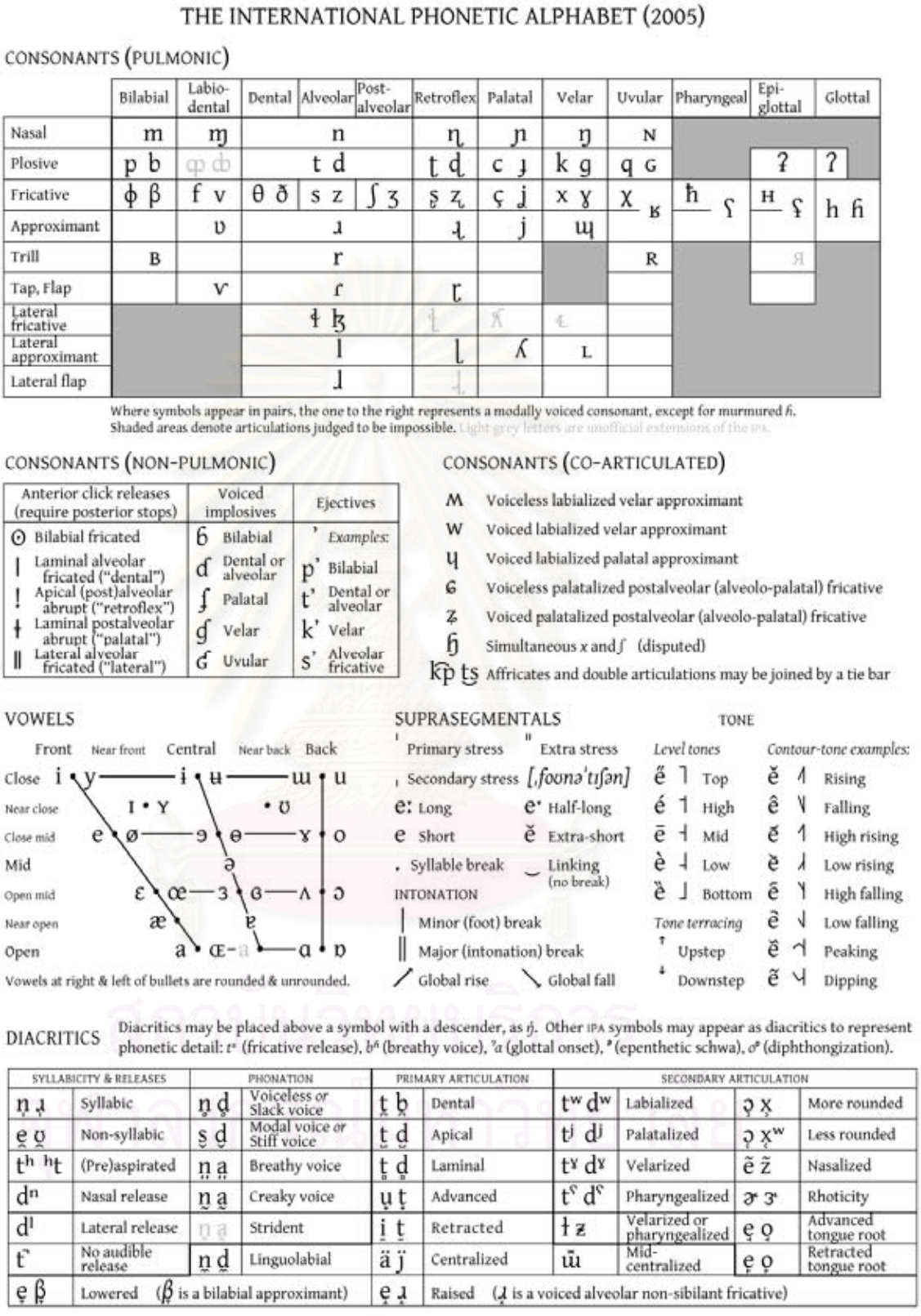
2.2) หน่วยเสียงวรรณยุกต์เปลี่ยนขึ้น (Rising tone) คือ เสียงวรรณยุกต์จัตวา จะไม่ปรากฏในพยางค์ที่มีพยัญชนะสะกดเป็นพยัญชนะกัก



รูปที่ 2.12 ความถี่ของเสียงวรรณยุกต์ที่เปลี่ยนแปลงกับเวลา[9]

นักภาษาศาสตร์ได้ศึกษาและบรรยายการเปล่งเสียงอย่างเป็นระบบและกำหนดสัญลักษณ์ทางภาษาเพื่อแสดงคุณสมบัติของเสียงแต่ละเสียง สัญลักษณ์ที่ใช้เรียกว่าสัทอักษร (Phonetic symbol) สัทอักษรที่ใช้เป็นภาษาสากลเรียกว่า IPA (International Pronunciation Alphabet) สัทอักษรแต่ละตัวทำหน้าที่แทนหน่วยเสียงซึ่งแสดงให้เห็นว่าแต่ละเสียงทำหน้าที่ต่างกัน ดังนั้นทุกภาษาในโลกก็สามารถจัดให้อยู่ในสัทอักษรสากล ในบางภาษาอาจมีหน่วยเสียงที่มีสัทอักษรตรงกัน มีการใช้อวัยวะในการเปล่งเสียงเดียวกัน

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



THE INTERNATIONAL PHONETIC ALPHABET (2005)

CONSONANTS (PULMONIC)

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

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

CONSONANTS (NON-PULMONIC)

Anterior click releases (require posterior stops)	Voiced implosives	Ejectives
ʘ Bilabial fricated	ɓ Bilabial	' *Examples:*
ǀ Laminal alveolar fricated ("dental")	ɗ Dental or alveolar	pʼ Bilabial
ǃ Apical (post)alveolar abrupt ("retroflex")	ʄ Palatal	tʼ Dental or alveolar
ǂ Laminal postalveolar abrupt ("palatal")	ɠ Velar	kʼ Velar
ǁ Lateral alveolar fricated ("lateral")	ʛ Uvular	sʼ Alveolar fricative

CONSONANTS (CO-ARTICULATED)

ʍ Voiceless labialized velar approximant
w Voiced labialized velar approximant
ɥ Voiced labialized palatal approximant
ɕ Voiceless palatalized postalveolar (alveolo-palatal) fricative
ʑ Voiced palatalized postalveolar (alveolo-palatal) fricative
ɧ Simultaneous *x* and ʃ (disputed)
k͡p t͡s Affricates and double articulations may be joined by a tie bar

VOWELS

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

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

SUPRASEGMENTALS

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

INTONATION

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

TONE

Level tones			*Contour-tone examples:*	
e̋	˥	Top	ě	˩˧ Rising
é	˦	High	ê	˥˩ Falling
ē	˧	Mid	e᷌	˧˥ High rising
è	˨	Low	e᷄	˩˧ Low rising
ȅ	˩	Bottom	e᷈	˦˩ High falling
		Tone terracing	e᷅	˧˩ Low falling
	ꜛ	Upstep	e᷉	˥˧ Peaking
	ꜜ	Downstep	e᷇	˩˦ Dipping

DIACRITICS Diacritics may be placed above a symbol with a descender, as ŋ̊. Other IPA symbols may appear as diacritics to represent phonetic detail: tˣ (fricative release), bʱ (breathy voice), ʔa (glottal onset), ᵊ (epenthetic schwa), oᵊ (diphthongization).

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

รูปที่ 2.13 สัทอักษรสากล [8]

ทฤษฎีในการสังเคราะห์เสียง

1. เทคนิคการสังเคราะห์เสียงพูด

การสังเคราะห์เสียงพูดในปัจจุบันมีเทคนิคแตกต่างกันออกไปขึ้นกับลักษณะของข้อมูลและวิธีการสร้างใหม่ของเสียงที่ต้องการ แบ่งออกเป็น 3 เทคนิคใหญ่ ๆ คือ

1. Concatenative synthesis เป็นการนำหน่วยเสียงที่ทำการบันทึกเอาไว้มาเล่นใหม่โดยใช้วิธีการนำหน่วยเสียงที่บันทึกมาต่อกัน ตัวอย่างของหน่วยเสียงที่ใช้ เช่น คำ (Words) พยางค์ (Syllables) หน่วยคู่เสียง (Diaphones) และหน่วยเสียงอนุภาค (Micro phonemes) นอกจากนี้อาจจะเก็บแต่ละหน่วยเสียงให้อยู่ในรูปของพารามิเตอร์ ด้วยวิธีการวิเคราะห์และสังเคราะห์ออกมา (Analysis-synthesis method) เมื่อต้องการสังเคราะห์เสียงออกมาก็นำค่าพารามิเตอร์เหล่านี้มาต่อกัน[10-12]

2. Formant synthesis เป็นการวิเคราะห์หาค่าฟอร์แมนต์เพื่อเป็นตัวแทนของหน่วยเสียงในภาษาเพื่อนำค่าที่ได้นี้ไปเป็นอินพุทให้กับ ฟังก์ชั่นถ่ายโอนของช่องทางเสียง เพื่อสังเคราะห์เป็นเสียงออกมา สามารถแบ่งแบบจำลองได้เป็น 3 แบบ คือ แบบอนุกรม แบบขนานและแบบอนุกรมขนาน[13]

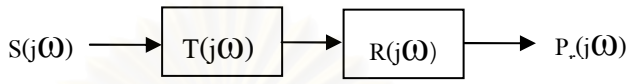
3. Articulatory synthesis เป็นวิธีจำลองระบบของการเกิดเสียงพูดของมนุษย์ ประกอบด้วยแบบจำลองของ 3 ส่วนหลักคือ แหล่งกำเนิดเสียง (Sound source production) รูปร่างของช่องทางเสียง และ การแพร่เสียง (Radiation) [11-19]

2. แบบจำลองในการสังเคราะห์เสียง

มนุษย์สร้างเสียงพูดขึ้นมาได้โดยมนุษย์จะมีอวัยวะสำคัญที่ทำให้เกิดเสียงหรือที่เรียกว่า เส้นเสียง ซึ่งอยู่ในกล่องเสียง และจากการสั่นของเส้นเสียงจะทำให้เกิดคลื่นเสียงความถี่ต่าง ๆ ผ่านเข้าสู่ลำคอ ผ่านจากลำคอเข้าสู่ช่องปากหรือช่องจมูกออกไปภายนอก ซึ่งขนาดและรูปร่างของอวัยวะภายในช่องปากนี้จะเป็นสิ่งกำหนดว่าคลื่นความถี่ไหนปรากฏออกมาให้ได้ยิน หรือคลื่นความถี่ไหนจะถูกดูดซับไว้ไม่ให้ปรากฏออกมา ปรากฏการณ์ที่ทำให้เกิดเสียงพูดออกมาได้เนื่องจากรูปร่างของช่องคอและช่องปากของมนุษย์ ที่สามารถเปลี่ยนแปลงไปมาได้อย่างอิสระ ทำ

ให้ผู้พูดมีความสามารถในการกำหนดคลื่นเสียงความถี่ต่าง ๆ ให้ปรากฏออกมาหรือดูดซับไว้ได้ตามความต้องการ การเรียงลำดับชั้นของการคัดเลือกความถี่คลื่นเสียงระดับต่าง ๆ เหล่านี้นั่นเองที่ทำให้เกิดเสียงพูดออกมา

สามารถเขียนกระบวนการการเกิดเสียงในรูปของฟังก์ชันถ่ายโอน (Transfer function) ได้ดังรูปที่ 2.14



รูปที่ 2.14 กระบวนการการเกิดเสียงโดยการแทนด้วยฟังก์ชั่นถ่ายโอน

$S(j\omega)$ ฟังก์ชั่นถ่ายโอนของแหล่งกำเนิดสัญญาณเสียง

$T(j\omega)$ ฟังก์ชั่นถ่ายโอนของรูปร่างช่องทางเสียง

$R(j\omega)$ ฟังก์ชั่นถ่ายโอนของการแพร่ของคลื่นเสียง

$P_r(j\omega)$ ฟังก์ชั่นถ่ายโอนของความดันอากาศที่แพร่ออกมาจากการเปล่งเสียง

สามารถเขียนออกมาให้อยู่ในรูปของสมการคุณสมบัติของการเกิดเสียงในมนุษย์ได้ว่า

$$P_r(j\omega) = S(j\omega)T(j\omega)R(j\omega) \qquad (2.1)$$

จากพจน์ $P_r(j\omega)$,$S(j\omega)$ และ $R(j\omega)$ จะมีคุณสมบัติคงตัว มีเพียงพจน์ $T(j\omega)$ ที่จะเปลี่ยนแปลงไปตามรูปร่างของช่องทางเสียงในการเปล่งเสียงหนึ่ง ๆ ซึ่งสามารถหาได้จากสมการ

$$T(j\omega) = U_L(j\omega)/U_G(j\omega) \qquad (2.2)$$

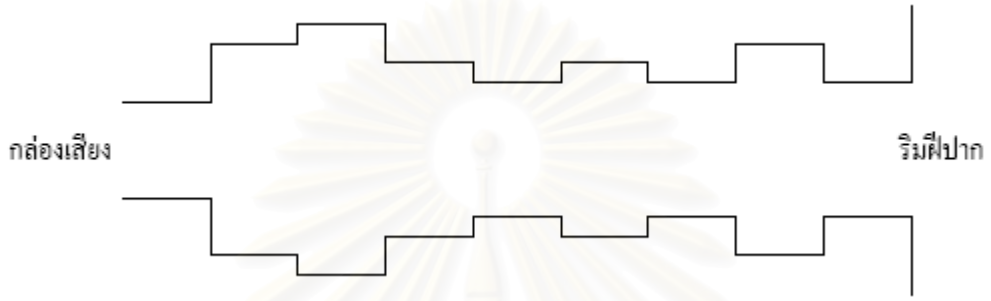
$U_L(j\omega)$ ฟังก์ชั่นถ่ายโอนของความเร็วปริมาตรของอากาศที่บริเวณริมฝีปาก

$U_G(j\omega)$ ฟังก์ชั่นถ่ายโอนของความเร็วปริมาตรของอากาศที่บริเวณเส้นเสียง

ระบบของเสียงพูดสามารถที่จะพิจารณาได้ว่าประกอบด้วยลำดับของท่อต่อออกมาจากปอดซึ่งมีความยาวประมาณ 17 เซนติเมตร[19,20] รูปแบบของแบบจำลองของช่องเสียงอย่างง่ายอาจมองได้เป็นลักษณะของท่อทรงกลมที่มีต้นกำเนิดของเสียงที่อยู่ปลายข้างหนึ่ง (ส่วนของกล่องเสียง) และปลายอีกข้างหนึ่งจะเปิด (ส่วนของช่องปาก) ดังนั้นตามหลักการทางฟิสิกส์จะเกิดการกำทอน (Resonance) ภายในท่อที่ความยาวคลื่นเท่ากับ 4L, 4L/3, 4L/5 ,… โดยที่ L คือความยาวท่อที่ใช้แทนช่องทางเสียง ซึ่งถ้าคิดเป็นความถี่ที่เกิดการกำทอนก็จะได้ความถี่ที่ c/4L, 3c/4L, 5c/4L,… Hz โดยที่ c คือความเร็วของเสียงในอากาศ ถ้าแทนความเร็วของเสียงในอากาศด้วยค่า 340

เมตรต่อวินาที และ L เท่ากับ 17 เซนติเมตร ก็จะมีการกำทอนที่ความถี่ประมาณ 500 Hz, 1500Hz, 2500Hz,… ตามลำดับ

แบบจำลองของช่องทางเสียง (Vocal tract model) แทนช่องทางเสียงด้วยลำดับของท่อตรงที่มีพื้นที่หน้าตัดขนาดต่าง ๆ กัน ดังรูปที่ 2.15 ด้านซ้ายคือ ส่วนที่ต่อกับกล่องเสียง ด้านขวาแทนริมฝีปาก



รูปที่ 2.15 ช่องทางเสียงแทนด้วยท่อตรงพื้นที่หน้าตัดต่าง ๆ กัน

แบ่งความยาวของท่อเป็นส่วน ๆ ซึ่งจะคิดว่าช่องทางเสียงมีความยาวโดยตลอดเป็น L และใช้ทฤษฎีคลื่นในการพิจารณาเฉพาะส่วน ในการพิจารณาขนาดความยาวของแต่ละส่วน

$$\Delta x = \mathrm{L}/\mathrm{N} \tag{2.3}$$

$$\tau = \Delta x / \mathrm{c} \tag{2.4}$$

Δx ความยาวของท่อในแต่ละส่วน

L ความยาวของท่อทั้งหมด

N จำนวนส่วนที่ทำการแบ่ง

τ เวลาที่ใช้ในการเดินทางของคลื่นในท่อแต่ละส่วน

c ของเร็วของเสียงในอากาศ

A พื้นที่หน้าตัดของท่อในแต่ละส่วน

สามารถเขียนสมการคลื่นของท่อแต่ละส่วนได้ว่า

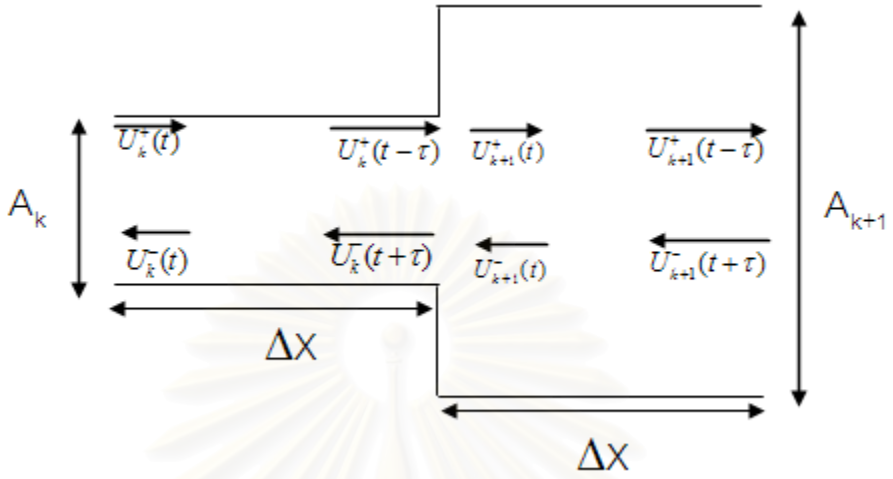
$$p_k(x,t) = \frac{\rho c}{A_k}\left[U_k^+\left(t-\frac{x}{c}\right)+U_k^-\left(t+\frac{x}{c}\right)\right] \tag{2.5}$$

$$U_k(x,t) = U_k^+\left(t-\frac{x}{c}\right)-U_k^-\left(t+\frac{x}{c}\right) \tag{2.6}$$

p_k ความดันอากาศภายในท่อที่ k

U_k ความเร็วปริมาตรของอากาศภายในท่อที่ k

หากพิจารณาเป็นท่อ 2 ท่อที่ติดกันแต่มีพื้นที่หน้าตัดแตกต่างกัน (A_n) ค่าของความเร็วปริมาตรของอากาศก็จะต้องแตกต่างกัน



รูปที่ 2.16 ทิศทางของความเร็วปริมาตรของอากาศในท่อ 2 ท่อที่ติดกัน[21]

จะได้สมการความเร็วปริมาตรของอากาศได้ว่า

$$U_{k+1}^{+}(t) = (1+r_k)U_k^{+}(t-\tau) + r_k U_{k+1}^{-}(t) \tag{2.7}$$

$$U_k^{-}(t+\tau) = -r_k U_k^{+}(t-\tau) + (1-r_k)U_{k+1}^{-}(t) \tag{2.8}$$

r_k สัมประสิทธิ์การสะท้อน (Reflection coefficient) ระหว่างรอยต่อพื้นที่หน้าตัดของท่อทั้ง 2 ส่วนและจะได้ค่าสัมประสิทธิ์การสะท้อนจากสมการ

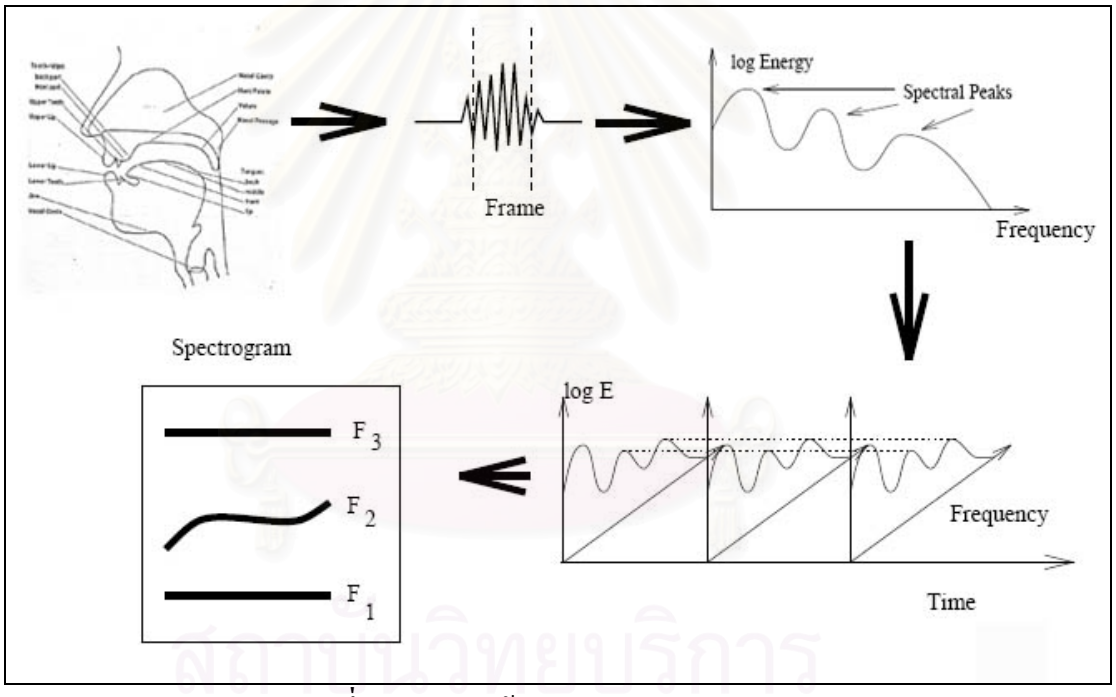
$$r_k = \frac{A_{k+1} - A_k}{A_{k+1} + A_k} \tag{2.9}$$

ก็จะสามารถแก้สมการโดยแทนค่า r_k ในสมการ (2.7) และ (2.8) เพื่อให้ได้ค่าความเร็วปริมาตรของอากาศ U_k ของท่อในแต่ละส่วน ไล่ไปเรื่อยจากความเร็วปริมาตรของอากาศที่บริเวณเส้นเสียง $U_{Glottal}(t)$ จนถึงความเร็วปริมาตรของอากาศที่บริเวณริมฝีปาก $U_{Lip}(t)$ แปลงกลับให้อยู่ในรูปของโดเมนความถี่ จากนั้นแทนในสมการที่ (2.2) เพื่อให้ได้ฟังก์ชั่นถ่านโอนของรูปร่างของช่องทางเสียง เพื่อที่จะนำไปแทนในสมการ (2.1) ให้ได้ฟังก์ชั่นถ่ายโอนของความดันอากาศที่แพร่ออกมาจากการเปล่งเสียง[22]

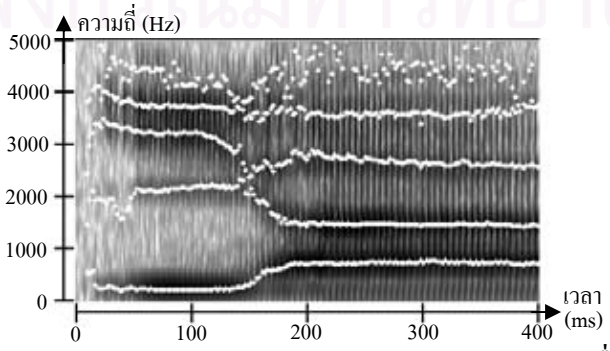
ทฤษฎีการวิเคราะห์สัญญาณเสียง

1. สเปกโตรแกรม (Spectrogram)

สัญญาณเสียงถูกสร้างจากอวัยวะในการออกเสียงต่าง ๆ ที่กล่าวไว้ข้างต้น สัญญาณเสียงประกอบด้วยค่าคุณลักษณะทางเสียง ความถี่ และแอมปลิจูด จากรูป 2.17 เมื่อเส้นเสียงสั่นสร้างความถี่ผ่านมาในช่องทางเสียงเกิดการเปลี่ยนแปลงของความถี่กำทอนอันดับต่าง ๆ และพลังงาน ในการวิเคราะห์สัญญาณเสียงไม่เพียงแค่ข้อมูลของพลังงาน ณ ความถี่ต่าง ๆ ยังต้องการทราบถึงการเปลี่ยนตามเวลาด้วย การแสดงคุณลักษณะให้เห็นรายละเอียดของฟอร์แมนต์และแอมปลิจูดเปลี่ยนแปลงตามเวลาจะช่วยให้สามารถระบุว่าเป็นหน่วยเสียงใด รูปร่างช่องทางเสียงเป็นอย่างไร จึงนิยมแปลงให้อยู่ในรูปแบบของภาพสเปกโตรแกรม



รูปที่ 2.17 โครงสร้างของสเปกโตรแกรม[23]



รูปที่ 2.18 ภาพสเปกโตรแกรมของเสียงพูดสระอีและอาต่อเนื่องกัน

รูปร่างของช่องทางเสียงมีการเปลี่ยนแปลงได้หลายรูปแบบขึ้นอยู่กับหน่วยเสียง การแสดงผลทางด้านคุณสมบัติทางเสียงจึงนิยมใช้ภาพสเปกโตรแกรมดังรูปที่ 2.18 โดยภาพจะให้รายละเอียดค่าความถี่ฟอร์แมนต์ลำดับต่าง ๆ (แนวเส้นสีขาว) ตามแกนแนวตั้ง และบอกแอมปลิจูดว่ามากน้อยจากการดูความเข้มของภาพ โดยแกนนอนจะบอกการเปลี่ยนแปลงทางเวลา

ในวิทยานิพนธ์นี้จะใช้ข้อมูลของภาพสเปกโตรแกรมของเสียงพูดภาษาไทยในการวิเคราะห์ร่วมกับข้อมูลของภาพสเปกโตรแกรมของเสียงพูดภาษาไทยที่ได้จากการสังเคราะห์ขึ้น

2. สมการโลคัส (Locus Equation)[24]

สมการโลคัส เป็นเครื่องมือที่ใช้ในงานวิจัยการใช้อวัยวะออกเสียงร่วม[25] เป็นตัวชี้วัดถึงการเกิดการเปลี่ยนแปลงของอวัยวะในช่องทางเสียงของหน่วยเสียงพยัญชนะที่มีการเปล่งเสียงร่วมกับหน่วยเสียงสระต่าง ๆ สมการโลคัสเป็นสมการถดถอยเชิงเส้นชนิดหนึ่ง โดยจะได้จากการลงจุดของกราฟสองมิติ โดยให้แกน Y เป็นค่าของความถี่ฟอร์แมนต์อันดับที่ 2 ของเสียงสระในจุดเริ่มต้น (ช่วงเริ่มต้นที่ติดกับรอยต่อของหน่วยเสียงพยัญชนะ) และแกน X เป็นค่าของความถี่ฟอร์แมนต์อันดับที่ 2 ของเสียงสระในจุดคงที่ (ใช้ตำแหน่งจุดกึ่งกลางของความยาวเสียงสระทั้งหมด) จากนั้นหาเส้นตรงที่ตัดผ่านจุดในเส้นกราฟ เส้นตรงที่ได้จะจัดในรูปของสมการ (2.10)

$$y = mx + c \tag{2.10}$$

ค่า m ที่ได้จะบอกความชันของเส้นตรง โดยค่าความชันยิ่งมากแสดงให้เห็นถึงผลจากสระที่ทำให้เกิดการเปลี่ยนแปลงของพยัญชนะที่มาก หรือสรุปได้ว่าพยัญชนะดังกล่าวมีการใช้อวัยวะออกเสียงร่วมเด่นชัด และเช่นเดียวกันหากค่า c มีค่าใกล้เคียงหรือเท่ากับ 0 แสดงให้เห็นถึงพยัญชนะดังกล่าวมีรูปร่างช่องทางเสียงที่ไม่มีการเปลี่ยนแปลงตามหน่วยเสียงสระ[24,26]

ทฤษฎีในการวิเคราะห์ข้อมูล

1. ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error)[27]

ค่าคลาดเคลื่อนกำลังสองเฉลี่ยเป็นการศึกษาผลรวมของผลต่างของข้อมูล 2 ชุด โดยไม่สนใจเครื่องหมายหรือทิศทางของข้อมูล จะพิจารณาเพียงค่าผลต่างเท่านั้น การหาค่าคลาดเคลื่อนกำลังสองเฉลี่ยคำนวณจากสมการ (2.11)

$$MSE = \frac{\sum_{i=1}^{n} (X_i - Y_i)^2}{n} \tag{2.11}$$

X ข้อมูลชุด (X)

Y ข้อมูลชุด (Y)

n ขนาดของกลุ่มตัวอย่าง

การบอกระดับความสัมพันธ์จะใช้ค่าคลาดเคลื่อนกำลังสองเฉลี่ย โดยค่าคลาดเคลื่อนกำลังสองเฉลี่ยมีค่าใกล้เคียงหรือเท่ากับ 0 แสดงถึงการมีความสัมพันธ์กันในระดับสูง แต่อาจมีความคลาดเคลื่อนได้บ้างเนื่องจากเป็นการพิจารณาผลรวมของค่าคลาดเคลื่อนกำลังสองเฉลี่ย

2. สหสัมพันธ์ (Correlation)[28]

สหสัมพันธ์เป็นการศึกษาความสัมพันธ์ของชุดข้อมูลตั้งแต่ 2 ชุดขึ้นไป ในการพิจารณาความสัมพันธ์ระหว่างชุดข้อมูลว่ามีมากน้อยเพียงใด จะใช้ค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation coefficient, r) เป็นค่าที่วัดความสัมพันธ์ คำนวณจากสมการ (2.12)

$$r_{xy} = \frac{n\sum_{i=1}^{n} X_i Y_i - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{\sqrt{n\sum_{i=1}^{n} X^2{}_i - (\sum_{i=1}^{n} X_i)^2}\sqrt{n\sum_{i=1}^{n} Y^2{}_i - (\sum_{i=1}^{n} Y_i)^2}} \qquad (2.12)$$

r_{xy} ค่าสัมประสิทธิ์สหสัมพันธ์

$\sum X$ ผลรวมของข้อมูลที่วัดได้จากข้อมูลชุด (X)

$\sum Y$ ผลรวมของข้อมูลที่วัดได้จากข้อมูลชุด (Y)

$\sum XY$ ผลรวมของผลคูณของข้อมูลชุด (X) และ (Y)

$\sum X^2$ ผลรวมของกำลังสองของข้อมูลที่วัดได้จากข้อมูลชุด (X)

$\sum Y^2$ ผลรวมของกำลังสองของข้อมูลที่วัดได้จากข้อมูลชุด (Y)

n ขนาดของกลุ่มตัวอย่าง

การบอกระดับหรือขนาดความสัมพันธ์จะใช้ตัวเลขของค่าสัมประสิทธิ์สหสัมพันธ์ โดยค่าสัมประสิทธิ์สหสัมพันธ์มีค่าใกล้เคียงหรือเท่ากับ 1 แสดงถึงการมีความสัมพันธ์กันในระดับสูง

ทฤษฎีในการประมาณค่า

เส้นโค้งเบซิเยร์ (Bezier Curve)[29]

เส้นโค้งเบซิเยร์ เป็นเส้นโค้งชนิดหนึ่งที่นิยมใช้ในงานของคอมพิวเตอร์กราฟฟิก ในการสร้างเส้นโค้งใด ๆ เส้นโค้งเบซิเยร์ใช้ในการสร้างความราบลื่นของเส้นโค้งในซอฟต์แวร์ประเภทการจัดแต่งภาพและกลุ่มแฟลช โดยเส้นโค้งเบซิเยร์มีรูปสมการทั่วไปดังสมการ (2.13)

$$B(t) = \sum_{i=0}^{n} \binom{n}{i} (1-t)^{n-i} t^{i} P_i \qquad (2.13)$$

n กำหนดดีกรีของเส้นโค้งเบซิเยร์

t เวลา กำหนด 0 ถึง 1

P_i พิกัดจุดตำแหน่งที่ i

ในวิทยานิพนธ์นี้ต้องการได้แนวเส้นโค้งเพื่อเป็นแนวการเปลี่ยนแปลงพื้นที่หน้าตัดของแบบจำลองสามมิติ จึงใช้เส้นโค้งเบซิเยร์ในดีกรีที่ 3 ได้ดังสมการ (2.14) จากการกำหนดสมการเบซิเยร์ดีกรี 3 ทำให้เกิดการเปลี่ยนแปลงจุดควบคุมความโค้งที่อิสระต่อกัน 2 จุด

$$B(t) = (1-t)^3 P_0 + 3t(1-t)^2 P_1 + 3t^2(1-t)P_2 + t^3 P_3 \qquad (2.14)$$

จากสมการ 2.14 สามารถอธิบายได้ว่า จุด P_0, P_3 เป็นจุดอ้างอิงตำแหน่งเริ่มต้นและตำแหน่งสิ้นสุดของเส้นโค้ง โดยมีจุด P_1, P_2 เป็นจุดควบคุมแนวความโค้งของเส้นโค้ง เมื่อกำหนดจุด P_0, P_3 คงที่และให้จุด P_1, P_2 เปลี่ยนแปลงเป็นจุดใด ๆ ก็จะสามารถสร้างเส้นโค้งที่มีรูปแบบเส้นโค้งใด ๆ ได้โดยใช้เพียงสมการ (2.14) สมการเดียว

งานวิจัยที่เกี่ยวข้อง

วิทยานิพนธ์ฉบับนี้ศึกษาการใช้แบบจำลองสามมิติในการแสดงกลไกการเปล่งเสียงภาษาไทย งานวิจัยที่เกี่ยวข้องที่ต้องทำการศึกษาจึงแบบออกเป็น 3 ส่วน คืองานวิจัยเกี่ยวกับแบบจำลองสามมิติของช่องทางเสียงและอวัยวะในการเปล่งเสียง งานวิจัยเกี่ยวกับการศึกษาถึงผลของอวัยวะออกเสียงร่วมในการเปล่งเสียงและงานวิจัยเกี่ยวกับการสังเคาะห์เสียงด้วยวิธี Articulatory Speech Synthesis

1. แบบจำลองของช่องทางเสียงและอวัยวะในการเปล่งเสียง

ในการเปล่งเสียงของมนุษย์จะใช้องค์ประกอบหลาย ๆ ส่วนด้วยกัน หากแบ่งออกเป็นส่วนใหญ่ ๆ ก็จะสามารถแบ่งออกเป็น 3 ส่วนคือ ส่วนของปอดที่ขับลมผ่านเส้นเสียง กล่องเสียง และส่วนสุดท้ายบริเวณช่องปาก ริมฝีปากรวมถึงช่องจมูก การสร้างแบบจำลองขององค์ประกอบดังกล่าวมีการสร้างในรูปของแบบจำลองสองมิติ [14,30] แต่เนื่องจากข้อจำกัดในมุมมองที่มีเพียง 2 ระนาบ จึงมีการพัฒนาแบบจำลองในยุคต่อมาให้เป็นแบบจำลองสามมิติ [31] ใช้วิธี MRI[32-34] และ X-Ray ในการจับภาพช่องทางเสียงและลิ้นในขณะออกเสียง อีกทั้งยังมีการจำลองการเคลื่อนไหวของลิ้นโดยการคำนวณแบบ Finite Element Method (FEM) และควบคุมการเคลื่อนไหวโดยวิธี Electromyographic (EMG) และ Equilibrium Point Hypothesis (EPH) การคำนวณหาพื้นที่หรือรูปร่างของ ช่องทางเสียง ก็มีความสำคัญเนื่องจากสามารถนำไปเป็นส่วนหนึ่งของการคำนวณหารูปแบบของเสียงที่จะเปล่งออกมา [34] นำภาพถ่าย MRI ที่ถ่ายตัดส่วนต่าง ๆ ของ ช่องทางเสียง จำนวน 14 จุดมาทำการหาพื้นที่หน้าตัดของจุดทั้ง 14 จุด จากนั้นนำไปขึ้นเป็นแบบจำลองของ ช่องทางเสียง นอกจากการคำนวณหาพื้นที่แล้วยังมี [35] การคำนวณหาเส้นกึ่งกลาง ช่องทางเสียง เพื่อหารูปแบบการเปลี่ยนแปลงของ ช่องทางเสียง ด้วยวิธี Kohonen Algorithm

โปรแกรมตัวอย่างที่เป็น 3 มิติ คือ ArtiSynth [16] เป็นแพลตฟอร์มสำหรับการจำลองด้านชีวจักรกลแบบ 3 มิติ ที่เกี่ยวข้องกับช่องทางอากาศในปากและทางเดินอากาศส่วนบน มีระบบการพัฒนาแบบเปิด (Open-source) และสามารถทำงานข้ามแพลทฟอร์ม(Platform) ผู้ทำงานวิจัยสามารถสร้างและเชื่อมต่อระหว่างแบบจำลองแต่ละชนิด เพื่อที่จะประกอบขึ้นเป็นระบบชีวจักรกลที่มีความสมบูรณ์ ซึ่งมีความสามารถในการสังเคราะห์เสียงแบบอวัยวะออกเสียง มีส่วนติดต่อผู้ใช้แบบกราฟิก และมีส่วนให้ผู้ใช้ทำการจำลองและปรับเปลี่ยน เคลื่อนย้ายตำแหน่งของช่องทางเสียงได้ แต่ข้อจำกัดคือในการปรับเปลี่ยนช่องทางเสียงสำหรับหน่วยเสียงอื่น ๆ จะต้องมีข้อมูลของภาพถ่ายเอ็มอาร์ไอในการเปรียบเทียบกำหนดรูปร่างของช่องทางเสียง อีกโปรแกรมคือ TractSyn [15] สามารถทำการกำหนดค่าพารามิเตอร์ต่าง ๆ และจำลองการเคลื่อนไหวตำแหน่งต่างๆ ในช่องปากได้เช่นเดียวกับ ArtiSynth แต่แบบจำลองเป็นลักษณะรูปทรงเรขาคณิตเพื่อให้ได้ค่าจากพารามิเตอร์ที่กำหนดนำไปสังเคราะห์เสียง

2. การศึกษาถึงผลของอวัยวะออกเสียงร่วมในการเปล่งเสียง

การศึกษาผลของอวัยวะออกเสียงร่วมมีการทำการศึกษาในภาษาต่าง ๆ เช่น ภาษาอังกฤษ [36,37] ภาษาญี่ปุ่น [3] ภาษาสวีเดน[25] ภาษาในกลุ่มคนในประเทศปากีสถาน [38]

เป็นต้น ข้อมูลที่นำมาวิเคราะห์มีหลายรูปแบบ เช่น ข้อมูลจากภาพ EMA (Electromagnetic articulagraphic) [3] ข้อมูลจากภาพ ultrasound แต่ข้อมูลที่นิยมนำมาวิเคราะห์คือค่าความถี่ฟอร์แมนต์อันดับที่ 2 (F2) โดยใช้การวิเคราะห์ของสมการโลคัส [37-40] ที่ได้กล่าวไว้ในทฤษฎีการวิเคราะห์สัญญาณเสียง ผลจากการสำรวจในงานวิจัย [5,6] กลุ่มพยัญชนะลิ้นส่วนหลัง-เพดานอ่อนเป็นกลุ่มพยัญชนะที่มีผลจากอวัยวะออกเสียงร่วมในการเปลี่ยนแปลงรูปร่างขึ้นกับสระมากที่สุด

3. การสังเคราะห์เสียงด้วยวิธี Articulatory Speech Synthesis

GNU Speech [19] เป็นระบบที่ถูกพัฒนาโดย Leonard Manzara และคณะพัฒนาขึ้นบนระบบปฏิบัติการ NeXTSTEP ประกอบด้วยส่วนหลักคือ ระบบ MONET และ ส่วน Tube Resonance Model (TRM)[41] โดยส่วนแรกเป็นส่วนติดต่อกับผู้ใช้ ทำหน้าที่กำหนดคุณลักษณะของสัญลักษณ์สัมพันธ์กับค่าพารามิเตอร์ที่กำหนด ส่วนที่ 2 จะทำหน้าที่อ่านค่าจากระบบ MONET รับพารามิเตอร์แล้วแปลงเป็นสัญญาณเสียง

จากการศึกษางานวิจัยข้างต้น ยังมีบางส่วนที่ยังไม่สมบูรณ์โดยไม่มีส่วนของการประเมินหน่วยเสียงที่เชื่อมต่อกัน เพื่อที่จะจัดตำแหน่งลิ้น (Place of articulation) ของหน่วยเสียงที่ต้องการจะออกเสียงให้สอดคล้องกับหน่วยเสียงที่อยู่ติดกัน ในงานวิจัยนี้มุ่งเน้น การจำลองการเคลื่อนที่ของแบบจำลองสามมิติของอวัยวะต่างๆ ขึ้นมาจากหน่วยเสียงพื้นฐานที่ผู้ใช้กำหนดขึ้นและปรับให้สอดคล้องกับหน่วยเสียงรอบข้างเพื่อให้การจำลองการเปล่งเสียงออกมาถูกต้อง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 3
วิธีดำเนินการวิจัย

ในบทนี้จะกล่าวถึงขั้นตอนการดำเนินการวิจัย เริ่มจากเครื่องมือที่ใช้ในการวิจัย ข้อมูลที่ใช้ในการทดลอง แบบจำลองช่องทางเสียงและการปรับแต่ง การออกแบบพารามิเตอร์ที่ใช้ควบคุมลักษณะสำคัญของหน่วยเสียง การทดลองการประมาณการเปลี่ยนแปลงของพื้นที่หน้าตัดขณะเปล่งเสียงโดยใช้สมการเส้นโค้งเบซิเยร์ การหาค่าน้ำหนักที่เหมาะสมสำหรับพยัญชนะและสระที่มีผลการใช้อวัยวะออกเสียงร่วมกันวิเคราะห์ผลโดยใช้สมการโลคัส และสุดท้ายการออกแบบระบบจำลองกลไกการเปล่งเสียง

ขั้นตอนการดำเนินการวิจัย

1. ออกแบบพารามิเตอร์ควบคุมที่สอดคล้องกับหน่วยเสียง
2. เตรียมแบบจำลองสามมิติของช่องทางเสียง
3. ทดลองหาค่าพารามิเตอร์ควบคุมช่องทางเสียงของหน่วยเสียงภาษาไทย
4. ทดลองหาการเปลี่ยนแปลงพื้นที่หน้าตัดขณะเปล่งเสียงต่อเนื่องของหน่วยเสียงสองหน่อยเสียง วิเคราะห์ผลจากข้อมูลความถี่ฟอร์แมนต์ ของเสียงสังเคราะห์จากหน่วยเสียงสองหน่วยเสียงนั้นเปรียบเทียบกับข้อมูลความถี่ฟอร์แมนต์ของเสียงจริง
5. ทดลองหารูปร่างช่องทางเสียงของพยัญชนะ ก ค และ ง ร่วมกับหน่วเสียงสระเดี่ยว เปรียบเทียบข้อมูลความถี่ฟอร์แมนต์อันดับที่ 2 ของเสียงสังเคราะห์และเสียงจริง
6. สร้างเครื่องมือสำหรับกำหนดสัญลักษณ์ของหน่วยเสียง และพารามิเตอร์ควบคุมที่สอดคล้องกับหน่วยเสียงนั้น และเครื่องมือสำหรับอ่านชุดสัญลักษณ์ของหน่วยเสียง
7. สรุปผลการวิจัย

เครื่องมือที่ใช้ในการวิจัย

1) Praat [42] เป็นเครื่องมือใช้ในการวิเคราะห์และศึกษาคุณลักษณะของเสียงโดยเฉพาะโดเมนทางความถี่สัญญาณเสียง เครื่องมือนี้ถูกพัฒนาจากสถาบัน The Institute of Phonetic Science ของมหาวิทยาลัยอัมสเตอร์ดัม

2) Blender [43] เป็นเครื่องมือใช้สร้างวัตถุสามมิติ กำหนดรูปแบบพื้นผิว จัดสภาพแสง กำหนดรูปแบบการเคลื่อนไหวของวัตถุต่างๆในฉากที่สร้างขึ้น ในวิทยานิพนธ์นี้ใช้ Blender เป็นเครื่องมือในการสร้างโมเดลจำลองสามมิติของอวัยวะในการเปล่งเสียง

3) Orge3D [44] เป็นกราฟิกเอนจินสามมิติ ที่มีประสิทธิภาพสูง ทำงานได้รวดเร็วและเหมาะกับการสร้างโปรแกรมจำลองแบบสามมิติ เช่น การสนับสนุนการทำภาพเคลื่อนไหวสามมิติด้วยเทคนิค Morph target และการทำ pose animation ผู้พัฒนาสามารถใช้ภาษา C/C++ ในการเขียนโปรแกรม และเป็นโครงการโอเพนซอร์ส

4) TRM [41] เป็นเครื่องมือสำหรับใช้ในการสังเคราะห์เสียง โดยเป็นส่วนหนึ่งของระบบ GnuSpeech [19]

ข้อมูลที่ใช้พัฒนาระบบ

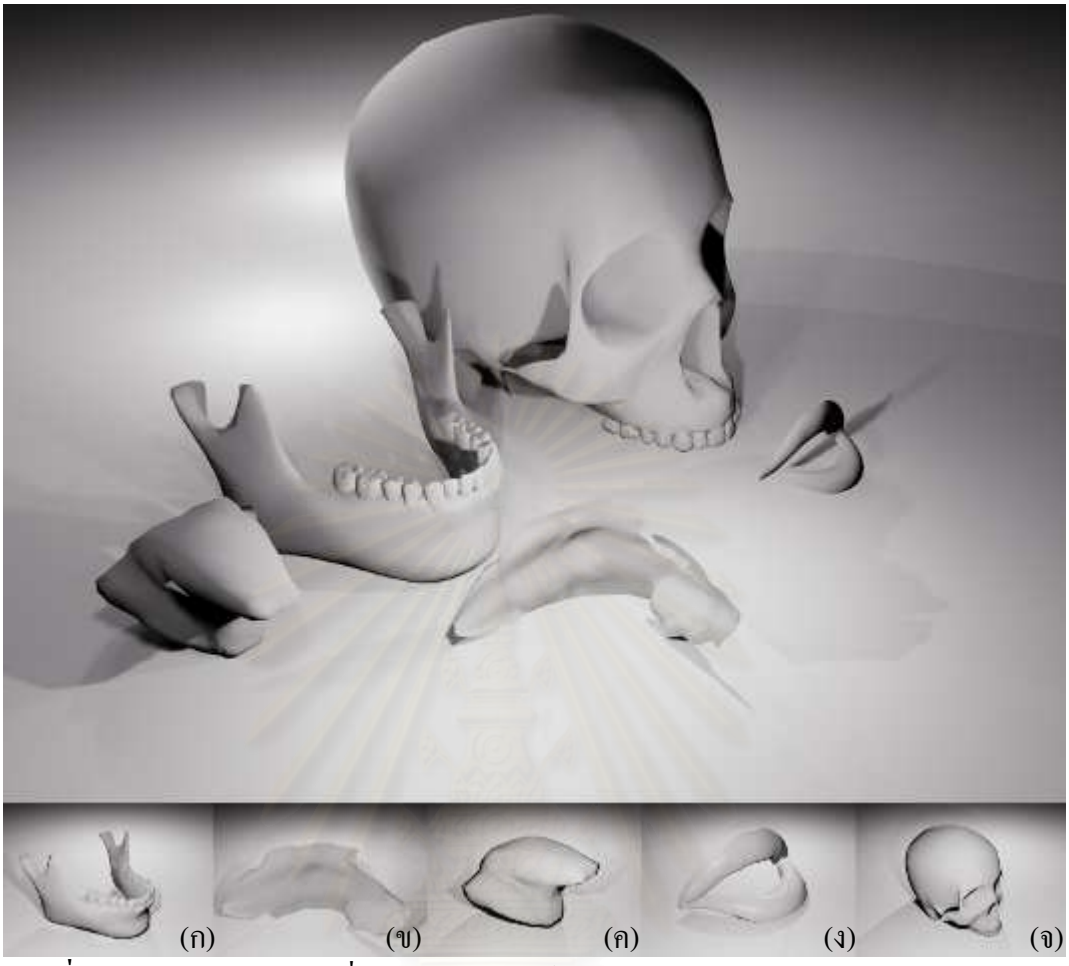
ข้อมูลที่ต้องการใช้ในการพัฒนาระบบแบ่งออกเป็น 2 ส่วนคือ

1) ฐานข้อมูลเสียงภาษาไทย ประกอบด้วยเสียงสระเดี่ยวเสียงยาว 9 เสียงและเสียงพยางค์ที่เกิดจาก พยัญชนะที่เกิดจาก ลิ้นส่วนหลัง-เพดานอ่อน ได้แก่ ก ค และ ง ร่วมกับสระ สระเดี่ยวเสียงยาว 9 เสียง รวมเป็นทั้งหมด 27 หน่วยเสียง

2) ฐานข้อมูลรูปร่างช่องทางเสียง [41] ในรูปแบบของพื้นที่หน้าตัดแต่ละส่วนของช่องทางเสียงในเสียงภาษาอังกฤษ เพื่อใช้อ้างอิงเป็นเสียงในภาษาไทยโดยพิจารณาจากสัทอักษรสากลที่ตรงกัน

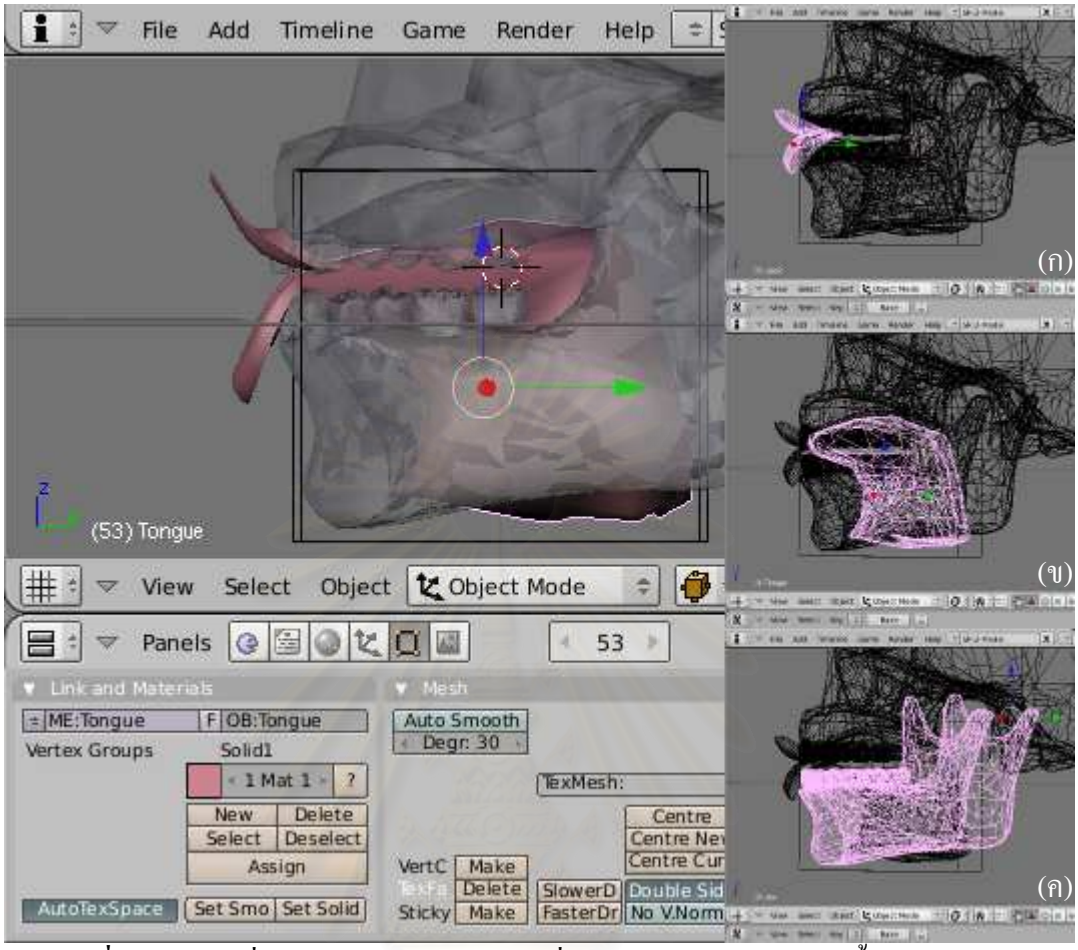
แบบจำลองช่องทางเสียง

จากงานวิจัยที่เกี่ยวข้อง[16] มีแบบจำลองสามมิติของอวัยวะที่เกี่ยวข้องกับการเปล่งเสียงที่เปิดให้นักวิจัยทั่วไปสามารถนำมาพัฒนาต่อได้ โดยเป็นแบบจำลองของโครงการ ArtiSynth 2.0 ที่อยู่ในรูปแบบของ แฟ้มข้อมูล Object. ซึ่งเป็นรูปแบบแฟ้มที่จัดเก็บโครงสร้างข้อมูลของวัตถุสามมิติ



รูปที่ 3.1 แบบจำลองสามมิติที่ใช้ในวิทยานิพนธ์ (ก) กรามและฟันล่าง (ข) ปุ่มเหงือก เพดานแข็ง และเพดานอ่อน (ค) ลิ้น (ง) ริมฝีปาก และ (จ) กระโหลกศีรษะและฟันบน

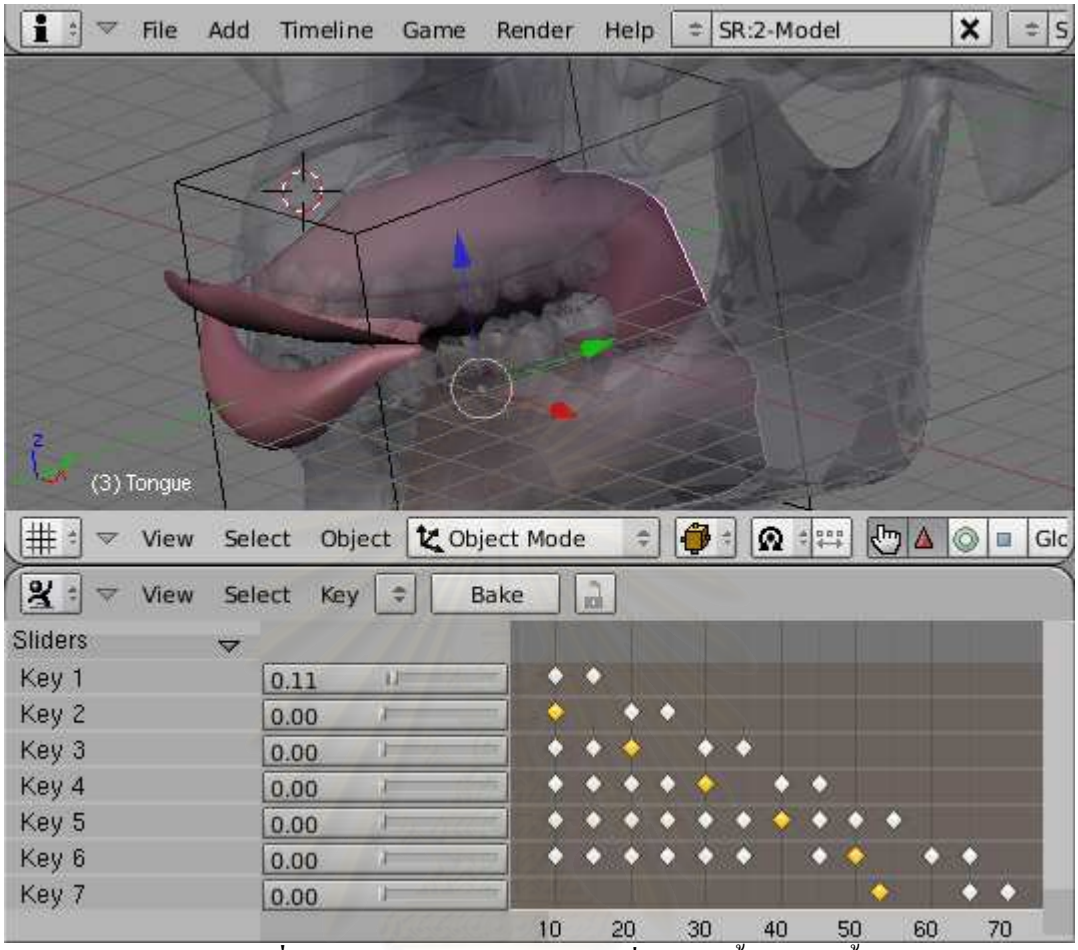
แบบจำลองที่ใช้ในงานวิจัยนี้ประกอบด้วย แบบจำลองของอวัยวะต่างๆที่เกี่ยวข้องกับการออกเสียง ดังรูป 3.1 แบบจำลองแต่ละชิ้น จะถูกนำมาเชื่อมโยงกันเพื่อสร้างแบบจำลองช่องทางเสียงที่สามารถกำหนดรูปแบบการเคลื่อนไหว ดังรูป 3.2 เพื่อจำลองการเคลื่อนไหวของอวัยวะในการเปล่งเสียงได้ การกำหนดตำแหน่งและรูปแบบการเคลื่อนไหว จะถูกกำหนดค่าตั้งต้นด้วยโปรแกรม Blender ซึ่งเป็นโปรแกรมทางด้านคอมพิวเตอร์กราฟิกสามมิติ



รูปที่ 3.2 อวัยวะที่ต้องกำหนดรูปแบบการเคลื่อนไหว (ก) ริมฝีปาก (ข) ลิ้น และ (ค) กราม

การกำหนดรูปแบบการเคลื่อนไหวตั้งต้น เป็นการกำหนดว่าชิ้นส่วนแต่ละชิ้นจะสามารถเคลื่อนไหวไปในทิศทางใดมีค่าเท่าใดได้บ้าง โดยการเคลื่อนที่ระยะน้อยสุดจะถูกกำหนดเป็นค่า 0 และการเคลื่อนที่ระยะมากสุดจะถูกกำหนดค่าเป็น 1 จากรูปที่ 3.3 จะพบว่ามีการกำหนดการเคลื่องไหวของอวัยวะ 3 ส่วนด้วยกัน คือ ริมฝีปาก ลิ้น และกรามซึ่งจะเชื่อมโยงกับพารามิเตอร์ ต่าง ๆ ซึ่งพารามิเตอร์เหล่านี้จะถูกควบคุมโดยโปรแกรมที่สร้างขึ้นสำหรับงานวิจัยนี้ โดยอ้างอิงจากค่าพารามิเตอร์ของหน่วยเสียงที่ถูกเก็บไว้ในฐานข้อมูลตั้งต้น หากผู้ใช้ปรับพารามิเตอร์ของหน่วยเสียงหนึ่งก็จะมีผลต่อการเคลื่อนไหวของแบบจำลองด้วย

เมื่อทำการกำหนดค่าตั้งต้นการเคลื่อนไหวของอวัยวะต่างๆเสร็จแล้ว แบบจำลองดังกล่าวจะต้องถูกแปลงให้อยู่ในรูปแบบข้อมูล Mesh ซึ่งเป็นรูปแบบข้อมูลที่ OGRE3D Engine ที่ใช้ในงานวิจัยนี้จะสามารถนำไปใช้ในโปรแกรมจำลองต้นแบบในงานนี้ได้



รูปที่ 3.3 การกำหนดรูปแบบการเคลื่อนไหวตั้งต้นของลิ้น

การออกแบบพารามิเตอร์ควบคุมที่สอดคล้องกับหน่วยเสียง

จากการศึกษาถึงอวัยวะที่ใช้ในการออกเสียง ตำแหน่งของช่องทางเสียงที่สามารถเปลี่ยนแปลงได้แบ่งออกเป็น 8 ส่วน[41] แต่ละส่วนมีความยาวเท่ากับ 1.7 เซนติเมตร ยกเว้นส่วนลิ้นส่วนหลังและเพดานอ่อน (r4) และส่วนกลางลิ้นและเพดานแข็ง (r5) มีความยาวเป็น 2 เท่าของส่วนปกติ หรือเท่ากับ 3.4 เซนติเมตร รายละเอียดของพารามิเตอร์แสดงดังตารางที่ 3.1

พารามิเตอร์ r1-r8 มีค่าระหว่าง 0.1-3.0 เป็นค่าพื้นที่หน้าตัดของช่องทางเสียง (ตารางเซนติเมตร)ในส่วนนั้น ๆ พารามิเตอร์ Mouth มีค่า 0 และ 1 โดย 0 หมายถึงรูปร่างของปากลักษณะห่อกลม และ 1 หมายถึงรูปร่างของปากลักษณะแผ่แบน นอกจากพารามิเตอร์สำหรับควบคุมรูปร่างของช่องทางเสียงแล้ว ยังมีพารามิเตอร์ที่ใช้ในการควบคุมลักษณะของหน่วยเสียงอีกด้วย โดยยึดพารามิเตอร์บางส่วนที่จำเป็นต่อโปรแกรมสังเคราะห์เสียง [41] แสดงในตารางที่ 3.2

ตารางที่ 3.1 พารามิเตอร์ที่ใช้ควบคุมช่องทางเสียง

ชื่อพารามิเตอร์	ช่วงค่าน้อย-มาก	ตำแหน่ง
r1	0.1 ถึง 3.0	ช่องลมเหนือเส้นเสียงถึงเพดานอ่อน
r2	0.1 ถึง 3.0	
r3	0.1 ถึง 3.0	
r4	0.1 ถึง 3.0	ลิ้นส่วนหลังและเพดานอ่อน
r5	0.1 ถึง 3.0	กลางลิ้นและเพดานแข็ง
r6	0.1 ถึง 3.0	ลิ้นส่วนหน้าและหลังปุ่มเหงือก
r7	0.1 ถึง 3.0	ปลายลิ้นและฟันหรือปุ่มเหงือก
r8	0.1 ถึง 3.0	ริมฝีปาก
Mouth	0 และ1	0 หมายถึง ปากห่อ 1 หมายถึง ปากแบน

ตารางที่ 3.2 พารามิเตอร์ที่ใช้ควบคุมลักษณะของหน่วยเสียง

ชื่อพารามิเตอร์	ช่วงค่าน้อย-มาก	คำอธิบาย
microInt	-22 ถึง 22	ใช้ควบคุมโทนเสียงโดย -12 หมายถึง เสียงสามัญ -17 หมายถึง เสียงเอก -2 หมายถึง เสียงโท -7 หมายถึง เสียงตรี -22 หมายถึง เสียงจัตวา
glotVol	0 ถึง 60	ใช้ควบคุมระดับความสั่นของเส้นเสียง โดย 0 หมายถึง เส้นเสียงไม่สั่น 60 หมายถึง เส้นเสียงสั่นมากที่สุด
aspVol	0 ถึง 60	เสียงพ่นลม
fricVol	0 ถึง 60	เสียงเสียดแทรก
fricPos	0 ถึง 7	ตำแหน่งการเกิดการเสียดแทรกในช่องปาก ค่าแทน ตำแหน่งของพารามิเตอร์ r#
fricCF	100 ถึง 20,000	ความถี่กลางของเสียงเสียดแทรก
fricBW	250 ถึง 20,000	แถบความถี่ของเสียงเสียดแทรก
Velum	0 ถึง 1.5	0 หมายถึง เพดานอ่อนยกปิด ไม่มีลมผ่านช่องจมูก

		การเพิ่มค่าหมายถึงการเปิดให้ลมผ่านช่องจมูกมากขึ้น
step	1 ถึง 500	ความยาวในการเปล่งเสียง

การทดสอบหาค่าพารามิเตอร์ควบคุมช่องทางเสียง

รูปร่างของช่องทางเสียงในวิทยานิพนธ์ฉบับนี้แบ่งออกเป็น รูปร่างช่องทางเสียงพยัญชนะและรูปร่างช่องเสียงสระ

1. การทดสอบหาค่าพารามิเตอร์ควบคุมช่องทางเสียงสระ

ใช้ข้อมูล [41,45] ของเสียงในภาษาอังกฤษอ้างอิงจากสัทอักษรสากลเดียวกัน ได้พารามิเตอร์ควบคุมช่องทางเสียงสระเป็นสระเดี่ยว เนื่องจากรูปร่างของช่องทางเสียงสระเดี่ยวเสียงยาวและเสียงสั้นมีรูปร่างช่องทางเสียงเหมือนกัน ในที่นี้จึงเลือกเฉพาะรูปร่างของช่องทางเสียงสระเดี่ยวเสียงยาว ได้สระเดี่ยวเสียงยาวทั้งหมด 4 เสียง คือ อา อี แอ และ ออ

ส่วนสระเดี่ยวเสียงยาวอีก 5 เสียง ใช้การเทียบตำแหน่งของลิ้นตามแผนภาพสระในรูป 2.11 จากนั้นเมื่อได้พารามิเตอร์ควบคุมรูปร่างของช่องทางเสียงสระทั้ง 9 เสียง ทำการตรวจสอบความถูกต้องกับเสียงสระภาษาไทย โดยใช้โปรแกรมสังเคราะห์เสียง TRM [41] รับค่าพารามิเตอร์ควบคุมทั้งหมดแล้วสร้างเสียงสังเคราะห์ขึ้นมา จากนั้นพิจารณาค่าความถี่ฟอร์แมนต์ที่ 1 ถึง 3 ของเสียงสังเคราะห์เปรียบเทียบกับข้อมูลค่าความถี่ฟอร์แมนต์อันดับ 1 ถึง 3 ในเสียงสระเสียงเดี่ยว จากฐานข้อมูล LOTUS ชุดสมดุล (Phonetically distributes : PD)โดยเลือกเสียงพูดชายและหญิงอย่างละ 5 คนเป็นข้อมูลอ้างอิงเสียงสระในภาษาไทย แสดงค่าเฉลี่ยของสระเสียงเดี่ยวทั้ง 9 ในตารางที่ 3.3

เพื่อความครอบคลุมรูปร่างของช่องทางเสียงที่ได้มาให้มากขึ้น จึงทำการเพิ่มขอบเขตขนาดของช่องทางเสียงแต่ละส่วน เริ่มจาก r1 ถึง ส่วน r8 อีกจากค่าที่ได้มาจากข้างต้น เพื่อเพิ่มความเป็นไปได้ที่จะได้พารามิเตอร์ที่เหมาะสมกับหน่วยเสียงสระภาษาไทย การเพิ่มความเป็นไปได้ทำโดยลดขนาดพื้นที่หน้าตัดลงจากขนาดเดิมครั้งละ 0.1 ตารางเซนติเมตร จำนวน 5 ครั้ง และเพิ่มขนาดพื้นที่หน้าตัดขึ้นจากขนาดเดิม 0.1 ตารางเซนติเมตร จำนวน 5 ครั้งเช่นกัน

ตารางที่ 3.3 ค่าเฉลี่ยความถี่ฟอร์แมนต์อันดับที่ 1 ถึง 3 ของชายและหญิงไทยในเสียงสระเสียงเดี่ยว

สระ	ชาย			หญิง		
	ความถี่ฟอร์แมนต์ที่ 1 (เฮิร์ต)	ความถี่ฟอร์แมนต์ที่ 2 (เฮิร์ต)	ความถี่ฟอร์แมนต์ที่ 3 (เฮิร์ต)	ความถี่ฟอร์แมนต์ที่ 1 (เฮิร์ต)	ความถี่ฟอร์แมนต์ที่ 2 (เฮิร์ต)	ความถี่ฟอร์แมนต์ที่ 3 (เฮิร์ต)
อา	804.43	1576.35	2431.82	957.92	1769.28	2384.12
อี	342.44	2221.82	2972.63	400.141	2447.67	2986.18
อือ	402.82	1756.83	2548.74	450.65	1809.18	2776.34
อู	440.70	897.61	2584.23	461.19	950.83	2809.03
เอ	488.98	2061.82	2717.90	579.51	2302.94	2691.59
แอ	709.55	1765.67	2507.89	926.16	1934.08	2421.07
โอ	528.04	1011.87	2649.28	643.34	1144.03	2662.44
ออ	714.85	1166.53	2586.87	860.70	1281.79	2437.44
เออ	533.63	1594.59	2609.50	658.03	1678.12	2742.82

ดังนั้นแต่ละส่วนของช่องทางเสียงจะมีค่าที่เป็นไปได้เพิ่มขึ้น 10 ค่า จำนวนส่วนของช่องทางเสียงทั้งหมด 8 ส่วน ทำให้ได้รูปร่างช่องทางเสียงสระที่เป็นไปได้ทั้งสิ้น 10^8 ค่าต่อหนึ่งหน่วยเสียงสระ ทำการตรวจสอบค่าเหมาะสมจากการวิเคราะห์ข้อมูลด้วยค่าคลาดเคลื่อนกำลังสองเฉลี่ย เพื่อหาค่าพารามิเตอร์ของช่องทางเสียงของสระเสียงเดี่ยวที่มีผลรวมค่าคลาดเคลื่อนกำลังสองเฉลี่ยของฟอร์แมนต์ทั้ง 3 น้อยที่สุด ได้พารามิเตอร์ควบคุมช่องทางเสียงที่มีค่าความถี่ฟอร์แมนต์อันดับที่ 1 ถึง 3 ใกล้เคียงที่สุดกับข้อมูลในตารางที่ 3.3 ซึ่งเป็นข้อมูลความถี่ฟอร์แมนต์อันดับที่ 1 ถึง 3 ของเสียงสระเดี่ยวภาษาไทย

2. การทดสอบหาค่าพารามิเตอร์ควบคุมช่องทางเสียงพยัญชนะ

ใช้ข้อมูลจาก [41,45] โดยอ้างอิงจากสัทอักษรสากลเดียวกัน ได้พารามิเตอร์ควบคุมช่องทางเสียงพยัญชนะทั้งสิ้น 17 หน่วยเสียง ในกลุ่มหน่วยเสียงที่มีการพ่นลม ได้แก่ [p^h] [t^h] [k^h] และ [$tɕ^h$] สามารถใช้พารามิเตอร์ควบคุมรูปร่างช่องทางเสียงของหน่วยเสียง [p] [t] [k] และ [tɕ] ตามลำดับแต่มีการเพิ่มส่วนการพ่นลมด้วยพารามิเตอร์ aspVol.

การทดลองการประมาณการเปลี่ยนแปลงของพื้นที่หน้าตัดขณะเปล่งเสียง

ใช้สมการเส้นโค้งเบซิเยร์เป็นหลักในการหาแนวการเปลี่ยนแปลงรูปร่างช่องทางเสียงจากหน่วยเสียงต้นไปเป็นรูปร่างช่องทางเสียงปลาย โดยจะกำหนดรูปร่างช่องทางเสียงต้นเป็นจุดอ้างอิง P_0 และกำหนดรูปร่างช่องทางเสียงปลายเป็นจุดอ้างอิง P_3

ส่วนจุด P_1, P_2 ที่ใช้เป็นจุดควบคุมแนวความโค้ง ทำการวนซ้ำจุด P_1, P_2 โดยจุดทั้งสองจะอยู่ในสเปซระหว่างจุดอ้างอิง P_0, P_3 โดยกำหนดให้มีจุดที่ต้องวนซ้ำ 100 จุด ดังนั้นจะได้เส้นโค้งที่เป็นไปได้ทั้งสิ้น 100^2 เส้น ใช้แนวเส้นโค้งดังกล่าวเป็นแนวการเปลี่ยนแปลงรูปร่างของช่องทางเสียงหน่วยเสียงต้นไปยังหน่วยเสียงปลาย จากนั้นนำพื้นที่หน้าตัดของช่องทางเสียงที่มีการเปลี่ยนแปลงไปใช้ในการสังเคราะห์เสียงแล้วทำการวิเคราะห์ด้วยวิธีความสัมพันธ์สหสัมพันธ์ เพื่อดูการเปลี่ยนแปลงของความถี่ฟอร์แมนต์ที่ 1 และ 2 ของเสียงสังเคราะห์และเสียงจริง แนวเส้นโค้งที่ทำให้ค่าสหสัมพันธ์มีค่าน้อยที่สุดชุดหนึ่งในจำนวนเส้นโค้งทั้งหมดจะนำไปเป็นส่วนการคำนวณการเปลี่ยนแปลงของช่องทางเสียงต่อไป

ในการทดลองเลือกทำการทดลองกับหน่วยเสียงพยัญชนะโดยเลือกจากตำแหน่งการเกิดเสียงอย่างละ 1 หน่วยเสียง จะได้หน่วยเสียงพยัญชนะทั้งสิ้น 9 หน่วยเสียงได้แก่ บ ฟ ด ซ จ ย ก ว และ ฮ จับคู่ร่วมกับหน่วยเสียงสระเดี่ยวทั้ง 9 หน่วยเสียง

การทดลองการประมาณรูปร่างช่องทางเสียงพยัญชนะที่เกิดผลจากการใช้อวัยวะออกเสียงร่วม

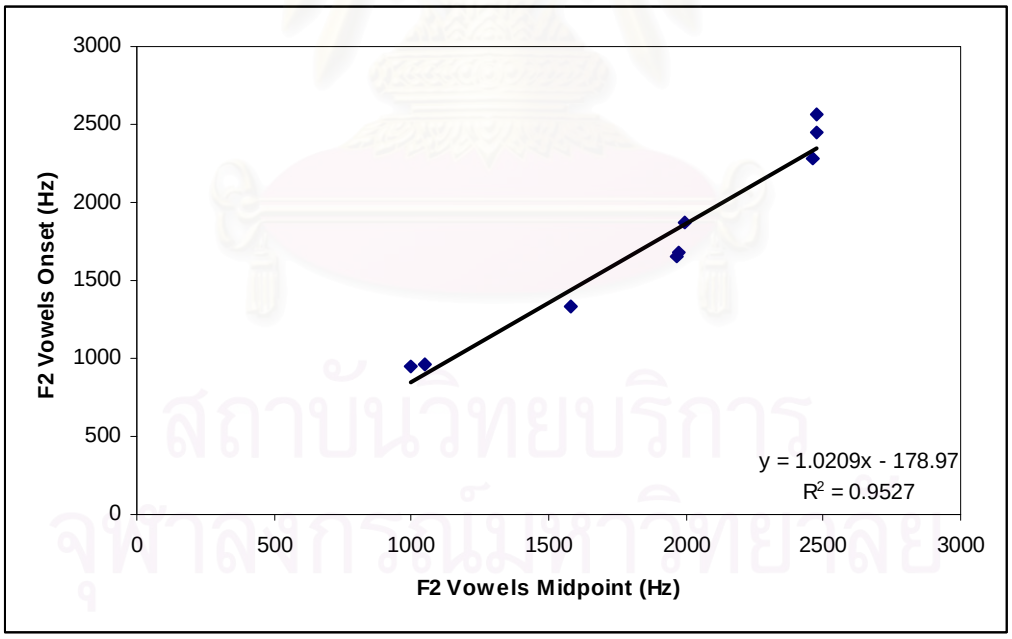
เมื่อกำหนดพารามิเตอร์ควบคุมสำหรับพยัญชนะและสระแล้ว ยังต้องพิจารณาถึงการเปลี่ยนแปลงที่เกิดขึ้นได้จากการใช้อวัยวะออกเสียงร่วมในการเปล่งเสียงจริง จึงทำการทดลองในกลุ่มพยัญชนะลิ้นส่วนหลัง-เพดานอ่อนร่วมกับหน่วยเสียงสระเดี่ยวทั้งหมดยกเว้น สระอา เนื่องจากการกำหนดพารามิเตอร์ของพยัญชนะข้างต้นกำหนดให้อิงกับรูปร่างช่องทางเสียงสระอาแล้ว กำหนดการเปลี่ยนแปลงของพื้นที่หน้าตัดตามผลที่ได้จากการทดลองในหัวข้อ"การทดลองการประมาณการเปลี่ยนแปลงของพื้นที่หน้าตัดขณะเปล่งเสียง"

จะมีการกำหนดค่าน้ำหนักให้แก่คู่พยัญชนะและสระโดยค่าน้ำหนักนี้จะใช้ในการคำนวณรูปร่างช่องทางเสียงของหน่วยเสียงพยัญชนะใหม่ดังสมการ (3.1)

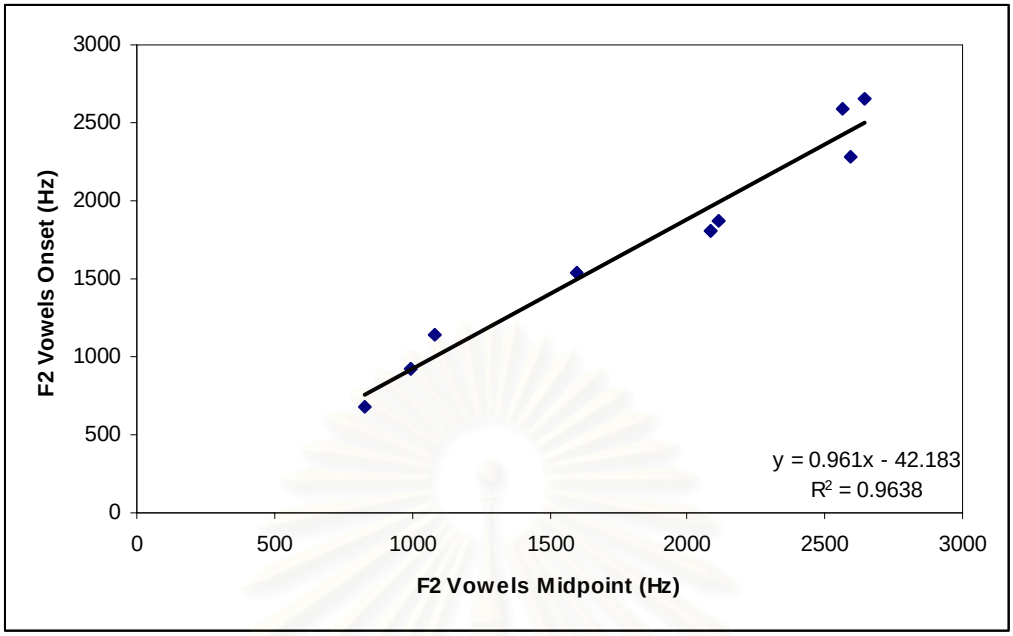
$$VTC_{new}(n) = VTC_{old}(n)W(n) + (1 - W(n))VTV(n) \quad (3.1)$$

โดย VTC_{new}, VTC_{old} และ VTV คือ รูปร่างของช่องทางเสียงพยัญชนะใหม่ที่ได้จากการคำนวณ รูปร่างของช่องทางเสียงพยัญชนะก่อนการคำนวณ และ รูปร่างของช่องทางเสียงสระตามลำดับ W เป็นค่าน้ำหนักที่จะกำหนดว่าต้องการให้รูปร่างของช่องทางเสียงใหม่โน้มเอียงไปมีรูปร่างตามช่องทางเสียงสระมากเท่าไร โดยค่านี้มีค่าระหว่าง 0 ถึง 1 โดยการคำนวณจะทำการแยกคิดตามส่วนต่าง ๆ ทั้ง 8 ส่วนของช่องทางเสียง

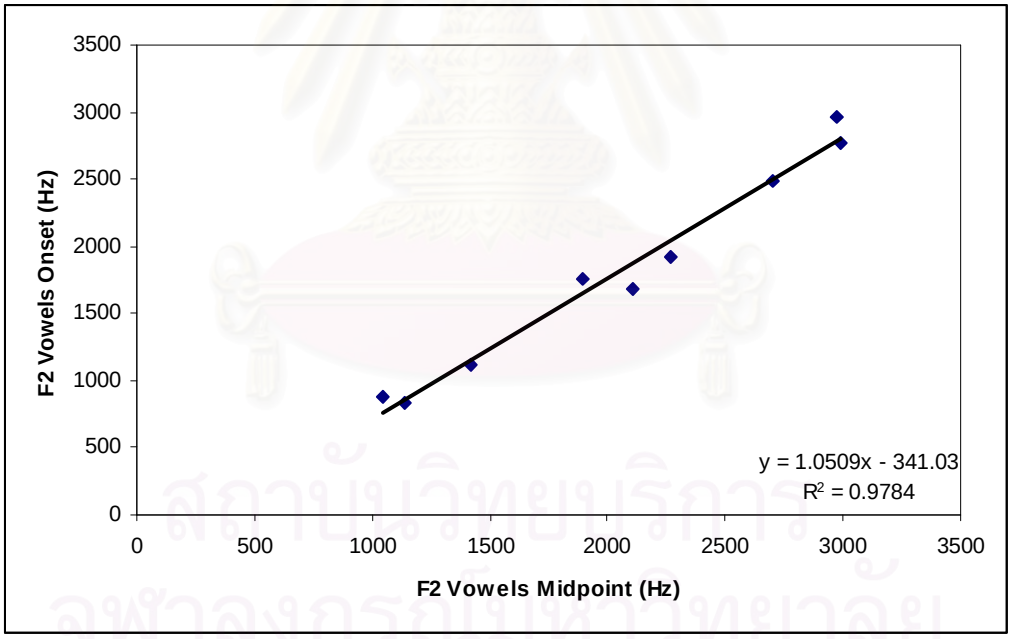
ในการทดลองจะหาค่าน้ำหนักที่เหมาะสม โดยทำการวนซ้ำค่าน้ำหนักในแต่ละส่วนของช่องทางเสียงจาก 0 ถึง 1 โดยเพิ่มค่าครั้งละ 0.1 ทำให้ได้ค่าของชุดน้ำหนักทั้งสิ้น 10^8 ค่า ในการวิเคราะห์จะวิเคราะห์หาชุดน้ำหนักที่จะทำให้ค่า m และ c ในสมการโลคัสของเสียงสังเคราะห์มีค่าแตกต่างจากเสียงจริงน้อยที่สุดในจำนวนชุดน้ำหนัก 10^8 ชุด รูปที่ 3.4-3.6 แสดงสมการโลคัสของเสียงพยัญชนะ ก ค และ ง ร่วมกับสระเสี่ยงเดี่ยว 9 เสียง



รูปที่ 3.4 สมการโลคัสของเสียงจริงของพยัญชนะ ก[k]



รูปที่ 3.5 สมการโลคัสของเสียงจริงของพยัญชนะ ค[k^h]

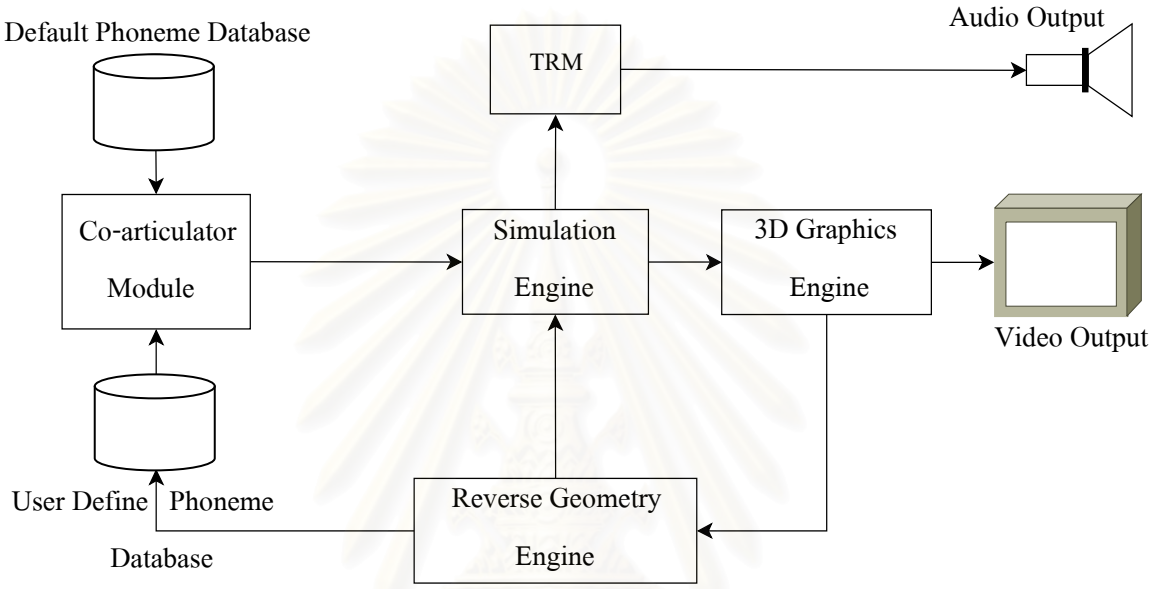


รูปที่ 3.6 สมการโลคัสของเสียงจริงของพยัญชนะ ง [ŋ]

การออกแบบระบบจำลองกลไกการเปล่งเสียง

ภาพรวมของระบบ

ระบบจำลองกลไกการออกเสียงของงานวิจัยนี้ สามารถอธิบายภาพรวมของระบบดังรูปที่ 3.7 และแยกอธิบายที่ละส่วนดังนี้



รูปที่ 3.7 แผนผังของระบบจำลองกลไกการออกเสียง

1) ฐานข้อมูลหน่วยเสียง (Phoneme Database) ข้อมูลที่ใช้สำหรับการสังเคราะห์และการจำลองลักษณะการออกเสียงในงานวิจัยนี้ถูกเก็บข้อมูลเริ่มต้นประกอบด้วยข้อมูลสำคัญสองชุด คือ ข้อมูลชุดแรก พารามิเตอร์ควบคุมลักษณะทางกายภาพของช่องทางเสียง ได้แก่ ขนาดของพื้นที่หน้าตัดของช่องทางเสียง ซึ่งจะถูกแบ่งเป็น 8 ส่วนตามตาราง 3.1 ข้อมูลชุดที่สองจะเป็นพารามิเตอร์ที่จะใช้ควบคุม Simulation Engine ได้แก่ ความถี่ในการสร้างลำดับข้อมูล ความยาวของแต่ละหน่วยเสียง คุณลักษณะของแหล่งกำเนิดเสียง โดยที่ส่วน Default Phoneme Database จะเก็บค่าที่ถูกตั้งค่ามาโดยซอฟต์แวร์ และส่วน User Define Phoneme Database จะเก็บที่ถูกแก้ไขโดยผู้ใช้
2) หน่วยคำนวณอวัยวะออกเสียงร่วม (Co-articulator Module) ส่วนประมวลผลข้อมูลของหน่วยเสียงตั้งต้น จาก Phoneme database ทำการคำนวณรูปร่างของช่องทางเสียงใหม่โดยคำนวณจากค่าน้ำหนักที่กำหนดไว้ระหว่างพยัญชนะและสระคู่นั้น ๆ

3) หน่วยประมวลผลการจำลอง (Simulation Engine) มีหน้าที่สำคัญของระบบสองส่วน คือ ส่วนแรกจะควบคุมการเคลื่อนไหวของแบบจำลองสามมิติ โดยมีการคำนวณการเคลื่อนไหวของอวัยวะออกเสียงให้เป็นไปอย่างราบรื่น โดยใช้การประมาณค่าที่มีการควบคุมอัตราการเปลี่ยนแปลงตามเวลา ด้วยสมการเส้นโค้งเบซิเยร์ ค่าที่ใช้ในการคำนวณ คือ รูปร่างของช่องทางเสียงส่วนต่าง ๆ ทั้ง 8 ส่วน ถูกอ่านเข้ามาจาก Co-articulator Module ที่ได้รูปร่างของช่องทางเสียงที่ปรับแต่งให้สอดคล้องกับเสียงสระข้างเคียงแล้ว และส่วนที่สองคือแปลงข้อมูลที่ได้จากการจำลองเพื่อส่งไปยัง TRM ต่อไป
4) หน่วยคำนวณรูปร่างย้อนกลับ (Reverse Geometry Engine) คำนวณค่าพารามิเตอร์ลักษณะทางกายภาพของอวัยวะในช่องทางเสียงใหม่ จากรูปร่างของช่องทางเสียง ของหน่วยเสียงที่ได้จากการปรับแต่งโดยผู้ใช้ โดยจะทำแก้ไขค่าพารามิเตอร์ที่เกี่ยวข้องกับช่องทางเสียง เพื่อนำกลับไปใช้เป็นค่าพารามิเตอร์ตั้งต้นในการจำลองอีกครั้ง ที่จะถูกะบันทึกสู่ฐานข้อมูลเพื่อนำกลับมาใช้ใหม่ภายหลังได้
5) หน่วยแสดงผลสามมิติ (3D Graphics Engine) OGRE3D graphic engine [44] คือ ส่วนแสดงผลภาพสามมิติ ของแบบจำลองอวัยวะในช่องทางเสียง การเคลื่อนไหวของแบบจำลองจะถูกควบคุมโดยคำสั่งจากหน่วยประมวลผลการจำลอง ที่ใช้ในการแสดงผลสนับสนุนทั้ง OpenGL และ Direct 3D
6) TRM(Tube Resonance Model) เป็นส่วนสังเคราะห์เสียงโดยรับค่าลักษณะทางกายภาพจากช่องทางเสียง ทำหน้าที่คล้ายกับเครื่องดนตรีประเภทเครื่องเป่า ที่จะกำเนิดเสียงความถี่ต่างๆ แตกต่างกันออกไป ตามคุณสมบัติของช่องทางเสียง ซึ่งเปรียบได้กับช่องทางเสียงของมนุษย์

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 4
ผลการวิจัย

ผลการวิจัย

วิทยานิพนธ์นี้ได้เสนอการจำลองสามมิติกลไกการเปล่งเสียงพูดภาษาไทยโดยเปรียบเทียบกับข้อมูลหน่วยเสียงในภาษาไทย ซึ่งในการสร้างระบบจำลองการเปล่งเสียงพูดเพื่อให้ได้ความถูกต้องในรูปร่างของช่องทางเสียงหน่วยเสียงภาษาไทยและความถูกต้องในการเปลี่ยนแปลงพื้นที่หน้าตัดของช่องทางเสียง ต้องทำการวิเคราะห์พารามิเตอร์ควบคุมช่องทางเสียงเพื่อกำหนดค่าที่เหมาะสม วิเคราะห์หาค่าน้ำหนักที่เหมาะสมสำหรับพยัญชนะและสระที่มีผลการใช้อวัยวะออกเสียงร่วม และวิเคราะห์การประมาณการเปลี่ยนแปลงของพื้นที่หน้าตัดขณะเปล่งเสียง ดังนั้นวิทยานิพนธ์นี้จึงรายงานผลการวิเคราะห์ของแต่ละขั้นตอนดังนี้

1. ผลการวิเคราะห์กำหนดพารามิเตอร์ควบคุมช่องทางเสียง

2. ผลการวิเคราะห์การประมาณการเปลี่ยนแปลงของพื้นที่หน้าตัดขณะเปล่งเสียง

3. ผลการวิเคราะห์การหาค่าน้ำหนักที่เหมาะสมสำหรับพยัญชนะและสระที่มีผลการใช้อวัยวะออกเสียงร่วม

มีผลการวิเคราะห์รายงานอย่างละเอียดดังนี้

ผลการวิเคราะห์การกำหนดพารามิเตอร์ควบคุมช่องทางเสียง

จากการทดลองการทดสอบหาค่าพารามิเตอร์ควบคุมช่องทางเสียง ในส่วนของหน่วยเสียงสระ เลือกผลที่มีค่าคลาดเคลื่อนกำลังสองเฉลี่ยของความถี่ฟอร์แมนต์อันดับ ที่ 1-3 น้อยที่สุดในกลุ่มค่าการทดลองทั้งหมดของหน่วยเสียงสระนั้น ๆ เพราะค่าคลาดเคลื่อนกำลังสองเฉลี่ยของความถี่ฟอร์แมนต์อันดับ ที่ 1-3 น้อยที่สุดหมายถึงหน่วยเสียงสระดังกล่าวมีคุณสมบัติของความถี่ฟอร์แมนต์อันดับ ที่ 1-3 ใกล้เคียงกับข้อมูลของเสียงสระจริงเป็นผลให้มีเสียงสังเคราะห์ที่คล้ายและใกล้เคียงเสียงจริงที่สุดในกลุ่มการทดลอง ผลของค่าคลาดเคลื่อนกำลังสองเฉลี่ยของความถี่ฟอร์แมนต์อันดับ ที่ 1-3 ของหน่วยเสียงสระเปรียบเทียบกับเสียงสระเดี่ยวจริงดังตารางที่ 4.1

ตารางที่ 4.1 ผลค่าคลาดเคลื่อนกำลังสองเฉลี่ยของความถี่ฟอร์แมนต์อันดับ ที่ 1-3 ของเสียงสระเดี่ยวจากการสังเคราะห์เสียงเปรียบเทียบกับเสียงสระเดี่ยวจริง

หน่วยเสียง	ค่า MSE			
	ความถี่ฟอร์แมนต์ที่ 1	ความถี่ฟอร์แมนต์ที่ 2	ความถี่ฟอร์แมนต์ที่ 3	ผลรวมความถี่ฟอร์แมนทั้ง 3
อา [aː]	136.07	13.80	1.80	151.67
อี [iː]	1.48	6.82	6.57	14.87
อือ [ɯː]	3.61	0.88	0.23	4.74
อู [uː]	1.37	1.46	23.29	26.13
เอ [eː]	1.19	5.79	2.15	9.14
แอ [æː]	4.23	3.85	0.85	8.94
โอ [oː]	2.55	1.83	5.33	9.73
ออ [ɔː]	59.06	6.60	4.27	69.64
เออ [ɤː]	6.10	0.98	3.59	10.69

ตารางที่ 4.2 ค่าความถี่ฟอร์แมนต์อันดับที่ 1 – 3 ของเสียงสังเคราะห์ที่ให้ผลรวมค่าคลาดเคลื่อนกำลังสองเฉลี่ยน้อยที่สุดในชุดข้อมูลทดลอง

หน่วยเสียง	ความถี่ฟอร์แมนต์ที่ 1 (เฮิร์ต)	ความถี่ฟอร์แมนต์ที่ 2 (เฮิร์ต)	ความถี่ฟอร์แมนต์ที่ 3 (เฮิร์ต)
อา [aː]	668.36	1562.54	2430.01
อี [iː]	340.96	2228.65	2979.20
อือ [ɯː]	399.20	1755.94	2548.97
อู [uː]	439.32	896.14	2560.94
เอ [eː]	487.79	2056.03	2715.74
แอ [æː]	713.78	1761.81	2507.03
โอ [oː]	530.60	1013.71	2643.94
ออ [ɔː]	655.79	1173.13	2582.59
เออ [ɤː]	527.53	1595.58	2605.90

ผลของความถี่ฟอร์แมนต์ 1 – 3 ของเสียงสังเคราะห์จากค่าพารามิเตอร์ควบคุมช่องทางเสียงที่ให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของความถี่ฟอร์แมนต์อันดับ ที่ 1-3 ของเสียงสระเดี่ยวจากการสังเคราะห์เสียงเปรียบเทียบกับเสียงสระเดี่ยวจริงน้อยที่สุด ดังตารางที่ 4.2

ค่าของพารามิเตอร์ควบคุมรูปร่างของหน่วยเสียงสระ ดังตารางที่ 4.3 พบว่าในการทดลองหารูปร่างหรือพื้นที่หน้าตัดของช่องทางเสียงจากการพิจารณาค่าความถี่ฟอร์แมนต์ให้ผลที่สอดคล้องกับรูปร่างช่องทางเสียงจริง เช่น รูปร่างช่องทางเสียงของสระเดี่ยวปากห่อ อู โอ และ ออ จะมีค่าพารามิเตอร์ r8 ที่กำหนดขนาดพื้นที่หน้าตัดของปากมีค่าน้อยกว่าหน่วยเสียงสระอื่น ๆ และมีขนาดเปลี่ยนแปลงตามรูปร่างของปากจริง คือ อู และ โอ จะมีขนาดพื้นที่หน้าตัดน้อยกว่า ออ และ อู มีขนาดพื้นที่หน้าตัดน้อยที่สุด

ตารางที่ 4.3 ค่าของพารามิเตอร์ควบคุมสระเดี่ยวที่ให้ผลรวมค่าคลาดเคลื่อนกำลังสองเฉลี่ยน้อยที่สุดในชุดข้อมูลทดลอง

หน่วยเสียง	r1	r2	r3	r4	r5	r6	r7	r8	Mouth
อา [aː]	0.80	0.82	0.85	1.09	1.19	1.57	1.32	2.09	1
อี [iː]	0.80	1.17	1.81	1.69	0.71	0.61	1.03	1.99	1
อือ [ɯː]	0.80	0.88	1.22	2.00	1.89	1.17	0.75	1.56	1
อู [uː]	0.80	1.41	1.14	0.70	1.52	1.83	1.66	0.85	0
เอ [eː]	0.80	1.08	1.62	1.80	1.09	1.17	1.15	2.56	1
แอ [æː]	0.80	1.15	1.14	1.44	1.61	1.89	2.09	2.51	1
โอ [oː]	0.80	0.90	1.16	0.61	1.39	1.69	1.50	1.14	0
ออ [ɔː]	0.80	0.71	0.64	0.62	1.39	1.37	1.48	1.55	0
เออ [ɤː]	0.80	0.79	1.49	0.91	1.06	0.94	1.33	1.62	1
ทุกสระ	microInt	glotVol	aspVol	fricVol	fricPos	fricCF	fricBW	Velum	step
	หมายเหตุ	60	0	0	5.5	2500	500	0.1	320

หมายเหตุ : ค่าเปลี่ยนตามโทนเสียงวรรณยุกต์

สำหรับหน่วยเสียงสระประสมเนื่องจากมีการเปลี่ยนแปลงภายในระหว่างการเปล่งเสียง จึงไม่สามารถกำหนดรูปร่างช่องทางเสียงตั้งต้นให้กับสระประสมได้ แต่หากต้องการออกเสียงสระประสมดังกล่าวก็สามารถทำได้โดยจัดเรียงสระเดี่ยวที่เป็นต้นกำเนิดของสระประสม

นั้น ๆ ตามด้วยหน่วยเสียงสระอา เช่น หน่วยเสียงสระ เอีย เลือกกำหนดพารามิเตอร์ตามสระอี จากนั้นตามด้วยพารามิเตอร์สระอา เช่นเดียวกับสระ เอือและอัว เลือกกำหนดพารามิเตอร์สระ อือ อู ตามด้วยสระอา ตามลำดับ

สำหรับค่าพารามิเตอร์สำหรับช่องเสียงพยัญชนะได้จากการเทียบสัทอักษรสากล (สัญลักษณ์ภายในเครื่องหมาย “[]”) ของหน่วยเสียงพยัญชนะไทยกับสัทอักษรสากลของหน่วยเสียงในภาษาสากล แสดงค่าพารามิเตอร์ควบคุมช่องทางเสียงในตารางที่ 4.4 และค่าพารามิเตอร์ควบคุมลักษณะของหน่วยเสียงในตารางที่ 4.5

ตารางที่ 4.4 ค่าพารามิเตอร์ควบคุมช่องทางเสียงหน่วยเสียงพยัญชนะภาษาไทย

หน่วยเสียง	r1	r2	r3	r4	r5	r6	r7	r8	Mouth
ก [k]	0.80	1.70	1.30	0.99	0.10	1.07	0.73	1.49	1
ข ฃ ค ฅ ฆ [k^h]	0.80	1.70	1.30	0.99	0.10	1.07	0.73	1.49	1
ง [ŋ]	0.80	1.70	1.30	0.99	0.10	1.07	0.73	1.49	1
จ [tɕ]	0.80	1.36	1.74	1.87	0.94	0.1	0.79	0.79	1
ฉ ช ฌ [$tɕ^h$]	0.80	1.36	1.74	1.87	0.94	0.1	0.79	0.79	1
ซ ศ ษ ส [s̪]	0.80	1.31	1.49	1.25	0.90	0.20	0.40	1.31	1
ญ ย [j]	0.80	1.36	1.74	1.87	0.94	0.10	0.79	0.79	1
ด ฎ [d]	0.80	1.31	1.49	1.25	0.76	0.10	1.44	1.30	1
ต ฏ [t]	0.80	1.31	1.49	1.25	0.76	0.10	1.44	1.31	1
ฐ ฑ ฒ ถ ท ธ [t^h]	0.80	1.31	1.49	1.25	0.76	0.10	1.44	1.31	1
ณ น [n]	0.80	1.31	1.49	1.25	1.00	0.05	1.44	1.31	1
บ [b]	0.80	0.89	0.76	1.28	1.80	0.99	0.84	0.10	1
ป [p]	0.80	0.89	0.76	1.28	1.80	0.99	0.84	0.10	1
พ ภ ผ [p^h]	0.80	0.89	0.76	1.28	1.80	0.99	0.84	0.10	1
ฟ ฝ [f]	0.80	0.89	0.99	0.81	0.76	0.89	0.84	0.50	1
ม [m]	0.80	0.89	0.76	1.28	1.80	0.99	0.84	0.10	1
ร [r]	0.80	1.31	0.73	1.07	2.12	0.47	1.78	0.65	1
ล ฬ [l]	0.80	0.89	1.10	0.97	0.89	0.34	0.29	1.12	1
ว [w]	0.80	1.91	1.44	0.60	1.02	1.33	1.56	0.55	0
ห ฮ [h]	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	1

อ [ʔ]	หมายเหตุ ค่าพารามิเตอร์เปลี่ยนแปลงตามสระ

ตารางที่ 4.5 ค่าพารามิเตอร์ควบคุมลักษณะของหน่วยเสียงของหน่วยเสียงพยัญชนะภาษาไทย

หน่วยเสียง	microInt	glotVol	aspVol	fricVol	fricPos	fricCF	fricBW	Velum	step
ก [k]	-12	0	0	0	4.7	2000	2000	0.1	70
ข ฃ ค ฅ ฆ [k^h]	-12	0	10	0	4.7	2000	2000	0.1	70
ง [ŋ]	-12	60	0	0	5.5	2500	500	0.5	70
จ [tɕ]	-12	0	10	0	5.6	2500	500	0.1	70
ฉ ช ฌ [$tɕ^h$]	-12	0	10	0	5.6	2500	2600	0.1	70
ซ ศ ษ ส [ʂ]	-12	0	0	0.8	5.8	5500	500	0.1	70
ญ ย [j]	-12	30	0	0	5.6	2500	2500	0.1	70
ด ฎ [d]	-12	30	0	0	6.7	4500	2000	0.1	70
ต ฏ [t]	-12	0	0	0	7.0	1500	2000	0.1	70
ฐ ฑ ฒ ถ ท ธ [t^h]	-12	0	10	0	7.0	4500	2000	0.1	70
ณ น [n]	-12	30	0	0	5.5	2500	500	0.5	70
บ [b]	-12	30	0	0	7.0	2000	700	0.1	70
ป [p]	-12	0	0	0	7.0	2000	700	0.1	70
พ ภ ผ [p^h]	-12	0	10	0	7.0	2000	700	0.1	70
ฟ ฝ [f]	-12	0	0	0.5	6.0	3300	1000	0.1	70
ม [m]	-12	30	0	0	5.5	2500	500	0.5	70
ร [r]	-12	30	0	0	5.5	2500	500	0.1	70
ล ฬ [l]	-12	30	0	0	5.5	2500	500	0.1	70
ว [w]	-12	30	0	0	5.5	2500	500	0.1	70
ห ฮ [h]	-12	0	10	0	5.5	2500	500	0.1	70
อ [ʔ]	หมายเหตุ ค่าพารามิเตอร์เปลี่ยนแปลงตามหน่วยเสียงสระ								

ผลการวิเคราะห์การประมาณการเปลี่ยนแปลงของพื้นที่หน้าตัดขณะเปล่งเสียง

ผลการทดลองของการประมาณการเปลี่ยนแปลงของพื้นที่หน้าตัด โดยจุด P_1, P_2 ที่คัดเลือกจะเป็นจุดที่ทำให้เส้นโค้งที่ใช้เป็นการเปลี่ยนแปลงพื้นที่หน้าตัดขณะเปล่งเสียง ให้เสียงสังเคราะห์ที่มีผลรวมสหสัมพันธ์ของค่าความถี่ฟอร์แมนต์ที่ 1 และ 2 มากที่สุดจากชุดข้อมูลในการทดลองทั้งหมด ผลที่ได้พบว่าจุด P_1, P_2 ที่ได้ในแต่ละหน่วยเสียงสระไม่มีความสัมพันธ์กัน และในแต่ละหน่วยเสียงพยัญชนะเมื่อมีหน่วยเสียงสระต่างกัน จุด P_1, P_2 ก็ไม่มีความสัมพันธ์เช่นเดียวกัน สรุปได้ว่า การประมาณการเปลี่ยนแปลงของพื้นที่หน้าตัดมีค่าเป็นรูปแบบเฉพาะของคู่หน่วยเสียงพยัญชนะและหน่วยเสียงสระนั้นๆ รายละเอียดของผลการวิเคราะห์จุด P_1, P_2 แสดงในภาคผนวก ก

ตารางที่ 4.6 ผลของจุดควบคุมเส้นโค้งเบซิเยร์ของหน่วยเสียงพยัญชนะและสระเดี่ยว

	P_1 (x,y) P_2 (x,y) ผลรวมค่าสหสัมพันธ์ของถวามถี่ฟอร์แมนต์ที่ 1 และ 2				
หน่วยเสียง	บ	ฟ	ด	ซ	จ
อา [aː]	(0,3) (2,6) 1.96	(0,8) (0,9) 1.80	(3,7) (0,9) 1.98	(1,6) (8,5) 1.93	(0,9) (1,9) 1.96
อี [iː]	(6,9) (2,9) 1.67	(0,8) (0,9) 1.33	(3,9) (4,8) 1.95	(0,9) (0,6) 1.65	(0,7) (1,7) 1.59
อือ [ɯː]	(9,2) (9,0) 1.33	(9,8) (5,0) 1.87	(1,9) (3,9) 1.98	(0,9) (0,0) 1.30	(0,9) (0,5) 1.84
อู [uː]	(9,0) (9,1) 1.37	(0,8) (0,6) 1.65	(0,7) (1,9) 1.94	(3,9) (2,7) 1.98	(0,9) (0,4) 1.60
เอ [eː]	(3,9) (9,2) 1.74	(7,7) (9,1) 1.57	(4,9) (0,9) 1.88	(0,9) (1,6) 1.78	(0,9) (0,6) 1.33
แอ [æː]	(7,1) (9,0) 1.72	(8,4) (5,5) 1.28	(2,8) (7,1) 1.92	(0,9) (0,6) 1.06	(0,8) (1,9) 1.93
โอ [oː]	(9,2) (9,0) 1.10	(1,9) (4,8) 1.64	(0,2) (1,9) 1.96	(0,8) (0,9) 1.72	(0,9) (2,3) 1.75
ออ [ɔː]	(0,8) (0,9) 1.16	(0,9) (1,3) 1.33	(2,8) (1,9) 1.95	(0,9) (1,9) 1.82	(0,9) (4,9) 1.98
เออ [ɤː]	(0,9) (6,1) 1.36	(9,9) (9,1) 1.17	(0,9) (1,5) 1.88	(0,8) (3,5) 1.96	(0,8) (0,9) 1.97

	P_1 (x,y) P_2 (x,y) ผลรวมค่าสหสัมพันธ์ของถวามถี่ฟอร์แมนต์ที่ 1 และ 2			
หน่วยเสียง	ย	ก	ว	ฮ
อา [aː]	(3,9) (3,7) 1.98	(1,1) (0,9) 1.9	(2,7) (0,9) 1.99	(7,4) (0,2) 1.56
อี [iː]	(0,9) (5,1) 0.67	(1,9) (2,6) 1.93	(2,9) (9,0) 1.88	(5,9) (3,0) 1.82
อือ [ɯː]	(5,9) (2,9) 1.99	(4,2) (0,8) 1.24	(9,8) (9,0) 1.36	(0,9) (0,1) 0.68
อู [uː]	(5,8) (0,8) 1.97	(2,2) (0,2) 1.975	(0,9) (9,0) 1.91	(9,0) (9,0) 1.45
เอ [eː]	(0,8) (5,5) 1.87	(0,9) (3,5) 1.275	(1,9) (8,1) 1.93	(8,0) (9,7) 0.95
แอ [æː]	(1,9) (4,8) 1.98	(4,9) (4,9) 1.91	(1,9) (7,8) 1.97	(7,1) (9,2) 1.30
โอ [oː]	(1,8) (2,9) 1.99	(0,9) (0,9) 1.48	(9,1) (9,0) 1.12	(0,8) (2,9) 1.62
ออ [ɔː]	(5,7) (1,8) 1.99	(0,9) (4,9) 1.95206	(2,9) (0,9) 1.91824	(0,8) (2,2) 1.17
เออ [ɤː]	(4,8) (0,9) 1.99	(3,7) (0,7) 1.99	(1,7) (0,9) 1.98	(1,8) (6,5) 1.93

ผลการวิเคราะห์การหาค่าน้ำหนักที่เหมาะสมสำหรับพยัญชนะและสระที่มีผลการใช้อวัยวะออกสียงร่วม

ผลการวิเคราะห์การหาค่าน้ำหนักทดลองโดยเปลี่ยนค่าน้ำหนักในแต่ละส่วนของพารามิเตอร์ควบคุมรูปร่างช่องทางเสียง จากนั้นทำการสังเคราะห์เสียงของหน่วยเสียงพยัญชนะที่มีรูปร่างช่องทางเสียงเปลี่ยนแปลงไปตามค่าน้ำหนักที่กำหนดให้กับหน่วยเสียงสระ จากนั้นพิจารณาค่าความถี่ฟอร์แมนต์อันดับที่ 2 ของเสียงสระที่ได้จากการสังเคราะห์ในตำแหน่งเริ่มต้นและช่วงกลางที่มีความความถี่คงที่เพื่อสร้างสมการโลคัส ค่าน้ำหนักในชุดที่ทำให้ค่า m และ c ในสมการ

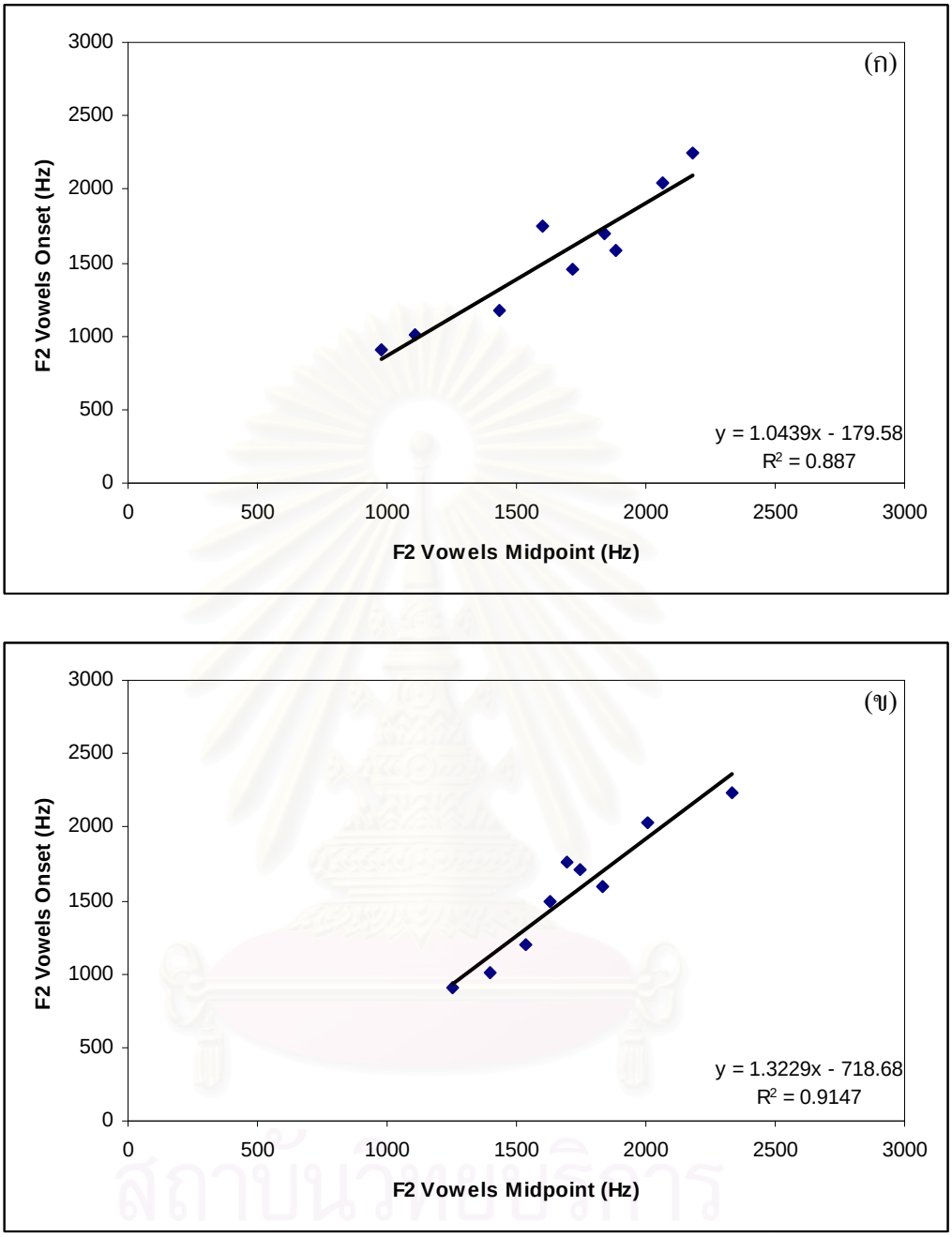
โลคัสมีความใกล้เคียงกับสมการโลคัสของเสียงจริงที่ทำการเก็บข้อมูลไว้ก่อนหน้านี้ ถือว่าเป็นชุดค่าน้ำหนักที่เหมาะสมที่สุดในชุดการทดลองทั้งหมด ดังตารางที่ 4.7 และรูปสมการโลคัสของหน่วยเสียงพยัญชนะทั้ง 3 ที่ทำการทดลอง เปรียบเทียบระหว่างสมการโลคัสที่ได้จากหน่วยเสียงพยัญชนะที่มีการปรับปรุงรูปร่างช่องทางเสียงด้วยค่าน้ำหนัก และสมการโลคัสที่ได้จากพยัญชนะที่ไม่มีการปรับปรุงรูปร่างตามหน่วยเสียงสระ ดังรูป 4.1-4.3

ผลจากการทดลองพบว่าค่าน้ำหนักที่ได้ของพยัญชนะ ก ค และ ง ซึ่งเป็นหน่วยเสียงที่ตำแหน่งการเกิดเสียงที่เดียวกัน คือ ลิ้นส่วนหลัง-เพดานอ่อน มีรูปร่างช่องทางเสียงที่เหมือนกัน แต่คุณลักษณะในการออกเสียงแตกต่างกัน เช่น ค มีการพ่นลม หรือ ง เป็นเสียงนาสิก เสียงออกทางช่องจมูก ค่าน้ำหนักที่ได้มีความแตกต่างกัน แต่มีความสัมพันธ์กันในหน่วยเสียงสระเดียวกัน ตัวอย่างที่สังเกตได้ชัดเจนเช่น ในหน่วยเสียงสระปากห่อ ค่าน้ำหนักในส่วนของพารามิเตอร์ที่ควบคุมรูปร่างริมฝีปากจะมีค่าน้ำหนักที่น้อย ซึ่งหมายถึงรูปร่างของหน่วยเสียงพยัญชนะจะมีการเปลี่ยนแปลงในส่วนนี้เอนเอียงไปมีรูปร่างคล้ายกับหน่วยเสียงสระค่อนข้างมาก

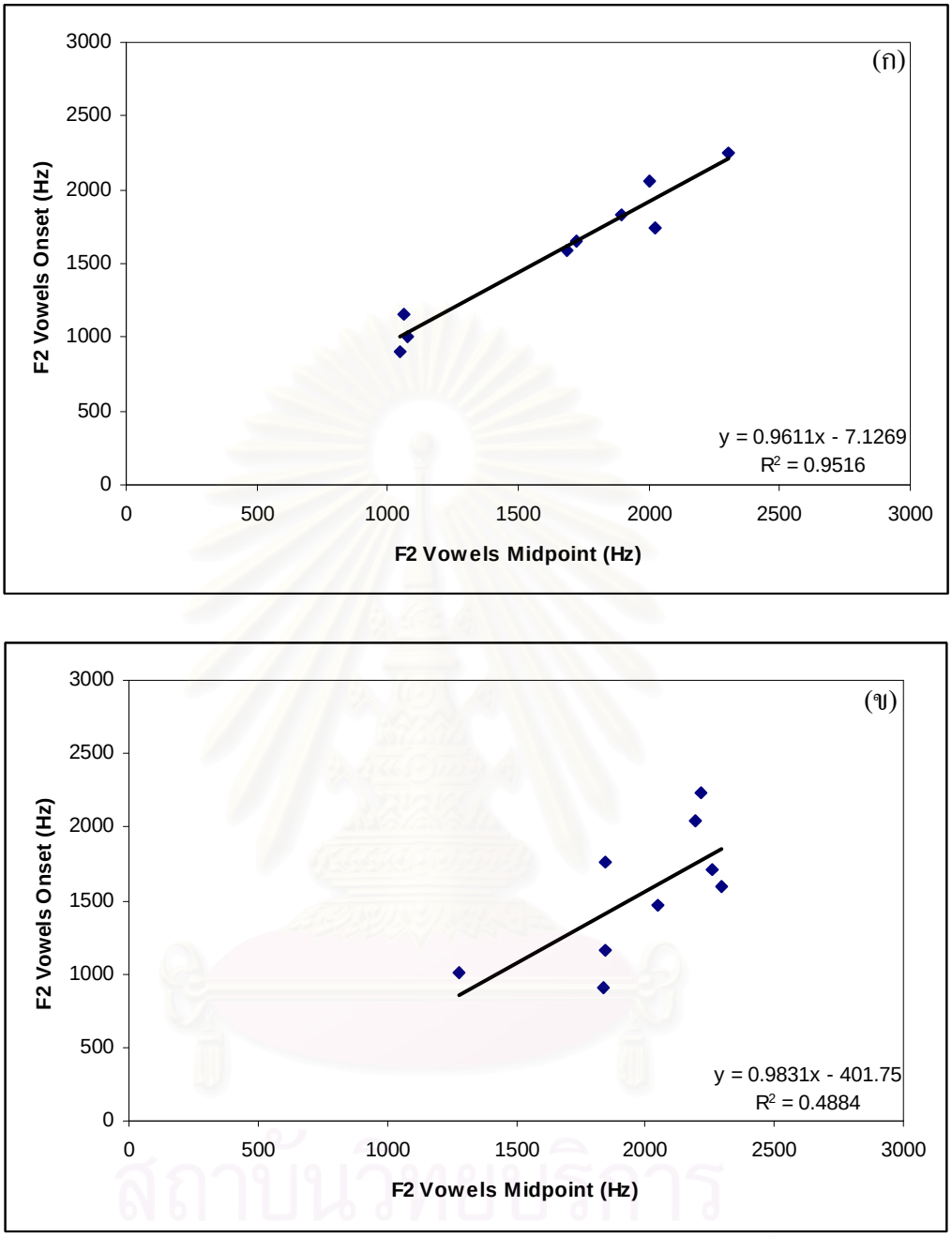
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.7 ผลการวิเคราะห์ค่าน้ำหนักของพยัญชนะและสระที่มีการใช้อวัยวะออกเสียงร่วม

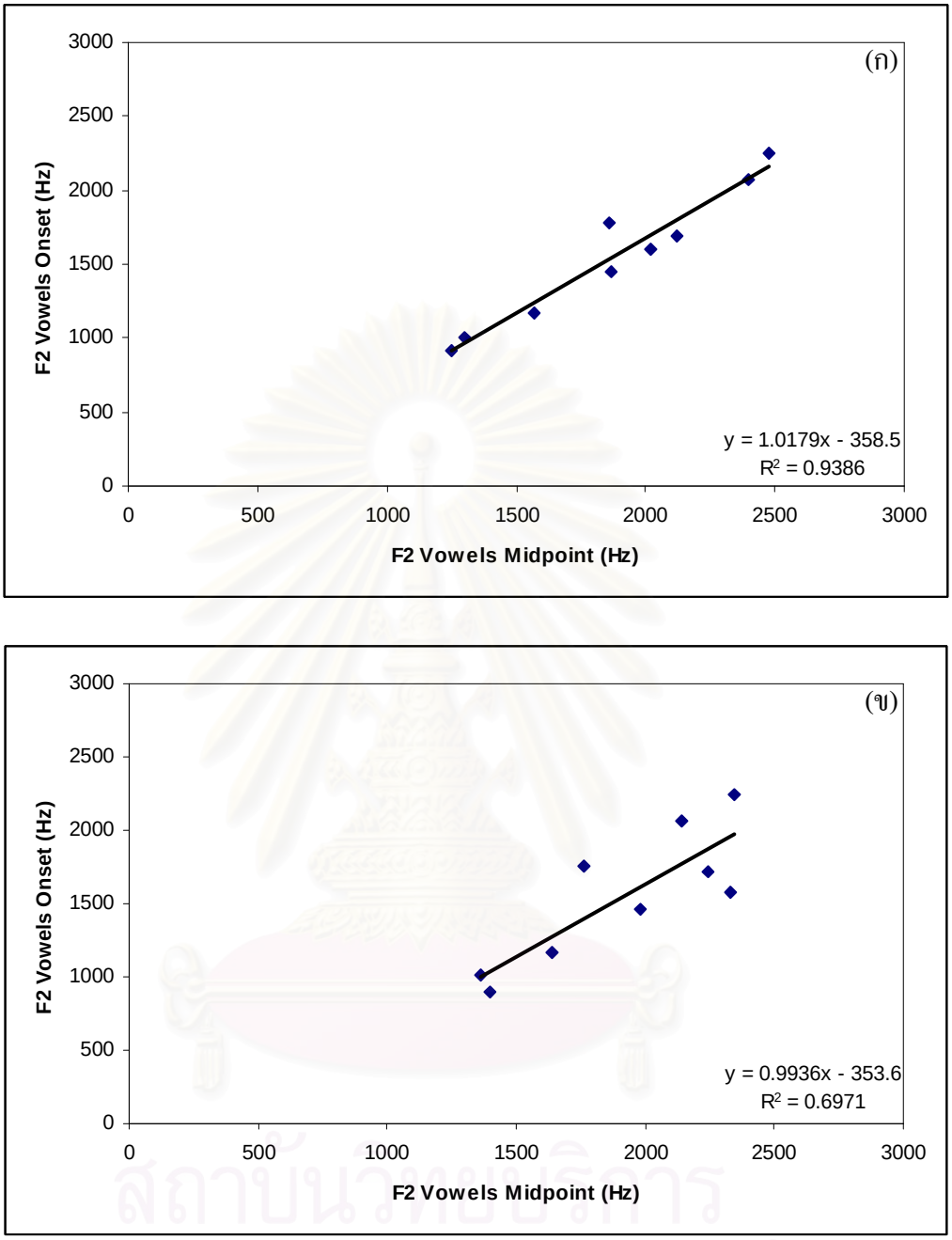
หน่วยเสียง	r#	อา [aː]	อี [iː]	อือ [ɯː]	อู [uː]	เอ [eː]	แอ [æː]	โอ [oː]	ออ [ɔː]	เออ [ɤː]
ก	1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	2	0.6	0.8	0.9	0.7	0.7	0.8	0.3	0.9	0.2
	3	0.9	0.7	0.9	0.7	0.9	0.2	0.4	0.7	0.4
	4	0.8	0.9	0.7	0.7	0.2	0.9	0.3	0.9	0.9
	5	0.9	0.7	0.2	0.7	0.5	0.7	0.5	0.7	0.9
	6	0.7	0.9	0.8	0.1	0.9	0.3	0.3	0.2	0.1
	7	0.1	0.4	0.1	0.3	0.3	0.3	0.9	0.7	0.3
	8	0.3	0.8	0.2	0.1	0.7	0.9	0.3	0.7	0.1
ค	1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	2	0.3	0.7	0.3	0.8	0.7	0.9	0.7	0.1	0.9
	3	0.7	0.9	0.9	0.1	0.9	0.5	0.3	0.9	0.9
	4	0.5	0.2	0.7	0.9	0.9	0.5	0.7	0.1	0.1
	5	0.9	0.9	0.7	0.9	0.3	0.9	0.1	0.1	0.1
	6	0.1	0.9	0.1	0.9	0.1	0.5	0.1	0.1	0.1
	7	0.9	0.2	0.1	0.8	0.1	0.1	0.9	0.3	0.9
	8	0.1	0.7	0.7	0.2	0.5	0.5	0.5	0.9	0.1
ง	1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	2	0.3	0.6	0.7	0.1	0.3	0.3	0.1	0.5	0.7
	3	0.1	0.1	0.3	0.9	0.1	0.7	0.9	0.9	0.3
	4	0.7	0.7	0.1	0.9	0.7	0.1	0.9	0.9	0.1
	5	0.9	0.9	0.7	0.9	0.9	0.7	0.9	0.9	0.7
	6	0.1	0.7	0.7	0.9	0.3	0.9	0.9	0.7	0.7
	7	0.7	0.5	0.9	0.9	0.1	0.7	0.9	0.9	0.9
	8	0.3	0.6	0.3	0.1	0.5	0.5	0.4	0.1	0.3



รูปที่ 4.1 สมการโลคัสของเสียงสังเคราะห์ของพยัญชนะ ก[k] (ก)ได้จากการปรับปรุงรูปร่างช่องทางเสียงตามหน่วยเสียงสระจากผลของอวัยวะออกเสียงร่วม (ข) ไม่มีการปรับปรุงรูปร่าง



รูปที่ 4.2 สมการโลคัสของเสียงสังเคราะห์ของพยัญชนะ ค[k^h] (ก)ได้จากการปรับปรุงรูปร่างช่องทางเสียงตามหน่วยเสียงสระจากผลของอวัยวะออกเสียงร่วม (ข) ไม่มีการปรับปรุงรูปร่าง



รูปที่ 4.3 สมการโลคัสของเสียงสังเคราะห์ของพยัญชนะ ง[ŋ] (ก) ได้จากการปรับปรุงรูปร่างช่องทางเสียงตามหน่วยเสียงสระจากผลของอวัยวะออกเสียงร่วม (ข) ไม่มีการปรับปรุงรูปร่าง

บทที่ 5
สรุปผลการวิจัยและข้อเสนอแนะ

สรุปผลการวิจัย

วิทยานิพนธ์นี้เสนอการจำลองสามมิติกลไกการเปล่งเสียงพูดภาษาไทยโดยเปรียบเทียบกับข้อมูลหน่วยเสียงในภาษาไทย ซึ่งในการสร้างระบบจำลองการเปล่งเสียงพูดเพื่อให้ได้ความถูกต้องในรูปร่างของช่องทางเสียงหน่วยเสียงภาษาไทยและความถูกต้องในการเปลี่ยนแปลงพื้นที่หน้าตัดของช่องทางเสียง ต้องทำการวิเคราะห์พารามิเตอร์ควบคุมช่องทางเสียงเพื่อกำหนดค่าที่เหมาะสม วิเคราะห์หาค่าน้ำหนักที่เหมาะสมสำหรับพยัญชนะและสระที่มีผลการใช้อวัยวะออกเสียงร่วม และวิเคราะห์การประมาณการเปลี่ยนแปลงของพื้นที่หน้าตัดขณะเปล่งเสียง

วิธีการหารูปร่างหรือพื้นที่หน้าตัดช่องทางเสียง แบ่งออกเป็น 2 กลุ่มคือ หน่วยเสียงสระใช้การเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ยของค่าความถี่ฟอร์แมนต์อันดับที่ 1 – 3 ของเสียงสังเคราะห์เปรียบเทียบกับเสียงจริง หน่วยเสียงพยัญชนะใช้การเปรียบเทียบสัทอักษรสากลของพยัญชนะภาษาไทยกับพยัญชนะภาษาอังกฤษ ผลที่ได้จากการวิเคราะห์ค่าความถี่ฟอร์แมนต์ให้รูปร่างของช่องทางเสียงสระที่มีความสัมพันธ์กับรูปร่างของช่องทางเสียงในการเปล่งเสียงจริง

วิธีการหาการเปลี่ยนแปลงพื้นที่หน้าตัดของช่องทางเสียงขณะเปล่งเสียงพูดใช้เส้นโค้งเบซิเยร์ที่กำหนดจุดควบคุมความโค้งคือจุด P_1, P_2 ทำการวิเคราะห์เสียงสังเคราะห์ที่ได้ว่า ที่จุดใดให้ผลรวมสหสัมพันธ์ของค่าความถี่ฟอร์แมนต์ที่ 1 และ 2 มากที่สุดจากชุดข้อมูลในการทดลองทั้งหมด ผลที่ได้พบว่าจุด P_1, P_2 ที่ได้ในแต่ละหน่วยเสียงสระไม่มีความสัมพันธ์กัน และในแต่ละหน่วยเสียงพยัญชนะเมื่อมีหน่วยเสียงสระต่างกัน จุด P_1, P_2 ก็ไม่มีความสัมพันธ์เช่นเดียวกัน สรุปได้ว่า การประมาณการเปลี่ยนแปลงของพื้นที่หน้าตัดมีค่าเป็นรูปแบบเฉพาะของคู่หน่วยเสียงพยัญชนะและหน่วยเสียงสระนั้น ๆ

วิธีการการหาค่าน้ำหนักที่เหมาะสมสำหรับพยัญชนะและสระที่มีผลการใช้อวัยวะออกเสียงร่วม ทดลองโดยเปลี่ยนค่าน้ำหนักในแต่ละส่วนพารามิเตอร์รูปร่างช่องทางเสียง จากนั้นทำการสังเคราะห์เสียงของหน่วยเสียงพยัญชนะที่มีรูปร่างช่องทางเสียงเปลี่ยนแปลงไปตามค่าน้ำหนักที่กำหนดให้กับหน่วยเสียงสระ ทำการพิจารณาค่าความถี่ฟอร์แมนต์อันดับที่ 2 ของเสียงสระที่ได้จากการสังเคราะห์ในตำแหน่งเริ่มต้นและช่วงกลางที่มีความความถี่คงที่ เพื่อสร้างสมการโลคัส ค่าน้ำหนักในชุดที่ทำให้ค่า m และ c ในสมการโลคัสมีความใกล้เคียงกับสมการโลคัสที่ทำ

การเก็บข้อมูลเสียงจริงก่อนหน้านี้ ถือว่าเป็นชุดค่าน้ำหนักที่เหมาะสมที่สุดในชุดการทดลองทั้งหมด ผลจากการทดลองได้ค่าน้ำหนักในแต่ละคู่หน่วยเสียงพยัญชนะและสระที่แตกต่างกัน

ประโยชน์จากการศึกษาของวิทยานิพนธ์นี้สามารถนำวิธีการเหล่านี้ไปประยุกต์ใช้ในภาษาอื่นได้ เนื่องจากเป็นการวิเคราะห์สัญญาณเสียงซึ่งเป็นพื้นฐานของงานวิจัยทางเสียง ในการปรับพื้นที่หน้าตัดของรูปร่างช่องทางเสียงก็สามารถใช้ข้อมูลจากการวิเคราะห์ผลแทนการใช้ภาพถ่ายเอ็มอาร์ไอแตกต่างจากงานวิจัย Artisynth ที่ต้องใช้ข้อมูลดังกล่าวในการปรับอวัยวะในช่องปาก และข้อมูลจากผลการทดลองยังสามารถนำไปใช้ในการสร้างโปรแกรมแบบจำลองช่องทางเสียงแสดงกลไกการเปล่งเสียงในแต่ละภาษา เพื่อนำไปใช้เป็นเครื่องมือสำคัญในการเรียนการสอนการเปล่งเสียงพูดอย่างถูกต้อง

ข้อเสนอแนะ

1. ผลที่ได้จากการทดลองทั้งหมดอาจมีการเปลี่ยนแปลงได้ขึ้นกับข้อมูลตั้งต้นที่ใช้ในการเปรียบเทียบ ในงานวิจัยนี้ผู้ทำการวิจัยเป็นผู้ทำการเก็บข้อมูลจากฐานข้อมูล LOTUS หากมีการเก็บข้อมูลค่าความถี่ฟอร์แมนต์ในสระภาษาไทยและข้อมูลสมการโลคัสของหน่วยเสียงพยัญชนะภาษาไทยที่เป็นมาตรฐานกลาง ผลที่ได้จากวิธีการทดลองในวิทยานิพนธ์นี้จะมีความน่าเชื่อถือมากขึ้น

2. การหารูปร่างช่องทางเสียงสระสามารถทำให้ครอบคลุมมากกว่าในการทดลองโดยทำการวนซ้ำค่าพารามิเตอร์ r1-r8 และทำการเพิ่มค่าครั้งละ 0.01 ตารางเซนติเมตร ถ้าหากต้องการรูปร่างช่องทางเสียงที่ให้เสียงสังเคราะห์ที่มีค่าความถี่ฟอร์แมนต์ที่มีค่าคลาดเคลื่อนกำลังสองเฉลี่ยใกล้ศูนย์ ก็สามารถใช้วิธีนี้ได้ แต่มีการใช้เวลาในการคำนวณที่ใช้ระยะเวลานาน ต้องทำการศึกษาเพิ่มเติมในเรื่องของแบบจำลอง TRM ว่าแต่ละส่วนของแบบจำลองมีผลต่อความถี่ฟอร์แมนต์อันดับต่าง ๆ อย่างไร

3. การหาค่าการเปลี่ยนแปลงพื้นที่หน้าตัด เช่นเดียวกับข้อ 2 หากต้องการค่าสหสัมพันธ์ใกล้หนึ่ง ก็สามารถเพิ่มเติมความละเอียดของจุด P_1, P_2 ได้เช่นกัน

4. การหาค่าน้ำหนักที่เหมาะสมสำหรับพยัญชนะและสระ ในการทดลองในวิทยานิพนธ์นี้ใช้การเพิ่มค่าน้ำหนักครั้งละ 0.1 สามารถทำการวิเคราะห์ที่ละเอียดกว่าได้โดยการเพิ่มค่าน้ำหนักให้น้อยลง แต่ต้องใช้เวลาในการคำนวณมากขึ้น

รายการอ้างอิง

[1] เทียนมณี บุญจุน. สัทศาสตร์: ระบบเสียงในภาษาอังกฤษและภาษาไทย. กรุงเทพมหานคร: โอเดียนสโตร์, 2548.

[2] จินดา เฮงสมบูรณ์. ภาษาศาสตร์เบื้องต้น. กรุงเทพมหานคร: สุวิริยาสาส์น, 2541.

[3] Jianwu, D., Masaaki, H., and Kiyoshi, H. Investigation of coarticulation in continuous speech of Japanese. Proceedings of the 25th Acoustic. Sci. & Tech (February 2004): pp.318-329.

[4] John, C. Multiple Articulation and Coarticulation[Online]. 2006. Available from: http://www.phon.ox.ac.uk/~jcoleman/MULTART.htm [2006, September 12]

[5] Löfqvist, A. Interarticulator phasing locus equations and degree of coarticulation. Journal of the 106th Acoustical Society of America (1999): pp. 2022-2030.

[6] Gibson, T., and Ohde, R.N. F2 Locus Equations:Phonetic Descriptors of Coarticulation in 17-to-22-Month-Old Children. Journal of Speech, Language, and Hearing Research. (February 2007): pp. 97-108

[7] กำชัย ทองหล่อ. หลักภาษาไทย. กรุงเทพมหานคร: รวมสาส์น, 2539.

[8] The International Phonetic Association, Reproduction of The International Phonetic Alphabet [Online]. 2005. Available from: http://www2.arts.gla.ac.uk/IPA/fullchart.html[2006,June 16]

[9] อาทร นันทิยกุล. การสังเคราะห์เสียงพูดจากข้อความภาษาไทย. วิทยานิพนธ์ปริญญามหาบัณฑิต. ภาควิชาวิศวกรรมคอมพิวเตอร์ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 2533

[10] Alan, W.B., and Nick, C. Optimising Selection of Units from Speech Databases for Concatenative Synthesis. Proceedings of EUROSPEECH '95 (1995):

[11] Vepa, J., King, S., and Taylor, P. Objective distance measures for spectral discontinuities in concatenative speech synthesis. Proceedings of 7th ICSLP (September 2002): pp. 2605-2608.

[12] Hisashi, K., et al. Ximera: a concatenative speech synthesis system with large scale corpora. IEICE Transactions on Information and Systems (2006): pp. 2688-2698.

[13] Fushikida, K. A formant extraction method using autocorrelation domain inverse filtering and focusing method. Proceedings of the 4th Acoustics, Speech, and Signal Processing, ICASSP (1988): pp. 2260-2263.

[14] Rubin, P., Bear, T., and Mermelstein, P. An Articulatory synthesizer for perceptual research. Proceedings of the 70th JASA (August 1981): pp. 321-328.

[15] Birkholz, P., and Jackel, D. A Three-dimensional model of the vocal tract for speech synthesis. International Congress of Phonetic Sciences (2003): pp. 2597-2600.

[16] Fels, S., et al. ArtiSynth: A Biomechanical Simulation Platform for the Vocal Tract and Upper Airway 2006. Technical Report TR-2006-10 Computer Science Dept., University of British Columbia (October 2006)

[17] Engwall, O. A 3D vocal tract model for articulatory and visual speech synthesis. Proceedings of Fonetik, The Swedish Phonetics Conference (1998): pp. 196-199.

[18] Engwall, O. Modeling of the vocal tract in three dimensions. Proceedings of Eurospeech (1999): pp. 113-116.

[19] Hill, D., Manzara, L., and Schock, C. Real-time articulatory speech-synthesis-by-rules. Proceedings of AVIOS (September 1995): pp. 27-44.

[20] Fant, G. Acoustic Theory of Speech Production. 2nd ed. The Hague: Mouton, 1970.

[21] Rabiner, L.R., and Schafer, R.W. Digital in speech signals. Prentice-Hall, 1978.

[22] Charles, S. B. Articulatory Methods for Speech Production and Recognition. Ph.D. thesis, Department of Engineering, Trinity College, 1996.

[23] Schmi, P. Explicit N-Best Formant Features for Segment-Based Speech Recognition. Poctoral dissertation, Department of Electrical Engineering and Computer Science, University of Bern, Switzerland. (1990).

[24] Caleb, E. Locus Equation Analysis as a Tool for Linguistic Fieldwork. Language Documentation & Conservation vol. 2 no.2 (December 2008): pp. 185-221.

[25] Lindblom, B. On vowel reduction. Report of The 29th Royal Institute of Technology, Speech Transmission Laboratory, Stockholm, Sweden (1963).

[26] Dang, J., Honda, M., and Honda, K. Investigation of coarticulation in continuous speech of Japanese. Proceedings of the 25th Acoustical Science and Technology no.5 (2004): pp. 318-329.

[27] Wikipedia, the free encyclopedia[Online]. 2008. Available from: http://en.wikipedia.org/wiki/Mean_squared_error [2008, December 7]

[28] Wikipedia, the free encyclopedia[Online]. 2008. Available from: http://en.wikipedia.org/wiki/Correlation [2008, December 7]

[29] Wikipedia, the free encyclopedia[Online]. 2008. Available from: http://en.wikipedia.org/wiki/Bézier_curve [2008, August 3]

[30] Coker, C. A Model of articulatory dynamics and control. Proceedings of the 40th IEEE (1976): pp. 452-460.

[31] Dang, J., and Honda, K. Estimation of vocal tract shapes from speech sound with a physiological articulatory model. In Journal of Phonetics (October 2002): pp. 511-532.

[32] Yehia, H., and Tiede, M. A parametric three-dimensional of the vocal-tract based on MRI data. Proceedings of ICASSP (1997): pp. 1619-1622.

[33] Badin, P., et al. A three-dimensional linear articulatory model based on MRI data. Proceedings of ICSLP (1998): pp. 14-20.

[34] Demolin, D., Metens, T., and Soquet, A. Three-Dimensional Measurement of the Vocal Tract by MRI. Proceedings of Spoken Language Processing (October 1996): pp. 272-275.

[35] Johannes, B., and Axel, W. A Segmentation and Analysis Method for MRI Data of the Human Vocal Tract. Proceeding of FIPKM (2001): pp. 179-189.

[36] Neary, T.M., and Sherrie, E.S. Formant transitions as partly distinctive invariant properties in the identification of voiced stop. Proceedings of the 15th Canadian Acoustics (1987): pp.17-24.

[37] Whiteside, S.P., and Rixon, E. Speech characteristics of monozygotic twins and a same-sex sibling: an acoustic case study of coarticulation patterns in read speech. Proceedings of the 60th Phonetica (2003): pp. 273-297.

[38] Sussman, H., Kathryn, A.H., and Fahan, A. A cross-linguistic investication of locus equations as a phonetic descriptor for place of articulation. Journal of the 94th Acoustical Society of America (1993): pp.1256-1268.

[39] Sussman, H., Helen, A.M., and Sandra A.M. An investigation of locus equations as a source of relational invariance for stop place categorization. Journal of the 90th Acoustical Society of America (1991): pp. 1309-1325.

[40] Jennifer, C., Hansook, C., and Heejin, K. Acoustic evidence for the effect of accent on CV coarticulation in radio news speech. Proceeding of the 2003 Texas Linguistics Society Conference (2004): pp. 62-72.

[41] Hill, D. Manual for the Synthesizer application -- part of the GnuSpeech text-to-speech toolkit [Online]. 2006. Available from: http://pages.cpsc.ucalgary.ca/~hill/papers/synthesizer/index.html [2006, August 16]

[42] Boersma, P., Wennik, D., and Praat, a system for doing phonetics by computer [Online]. Available from: http://www.praat.org [1996]

[43] BlenderWiki[Online]. 2008. Available from: http://wiki.blender.org/ [2005, May 22]

[44] Junker, G. Pro OGRE 3D Programming. New York: Apress, 2006.

[45] Hill, D. Pronunciation guide [Online]. 2006. Available from: www.cpsc.ucalgary.ca/~hill/papers/pronguid.htm [2006, August 16]

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

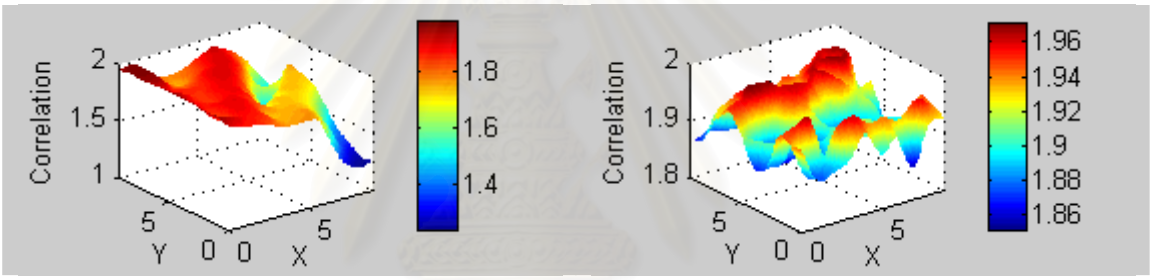
ภาคผนวก ก

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

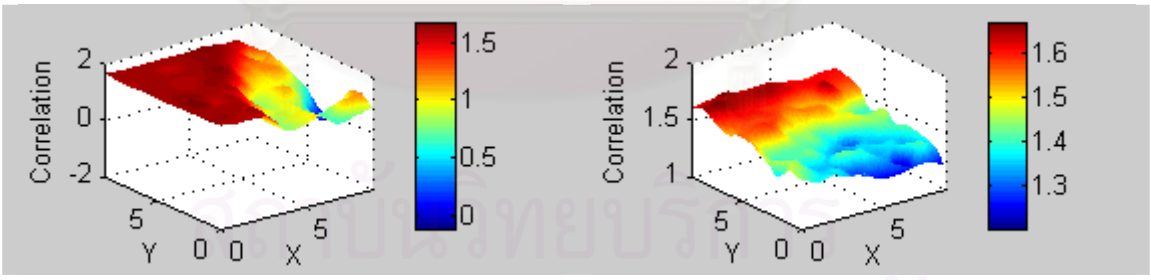
ภาคผนวก ก

ผลการวิเคราะห์การประมาณการเปลี่ยนแปลงของพื้นที่หน้าตัดช่องทางเสียงขณะเปล่งเสียง

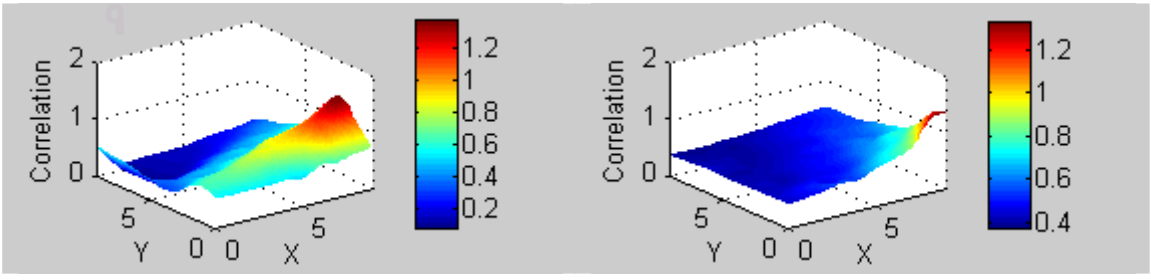
ภาคผนวก ก แสดงผลของการวิเคราะห์การประมาณการเปลี่ยนแปลงของพื้นที่หน้าตัดช่องทางเสียง วิเคราะห์เปรียบเทียบระหว่างเสียงจริงและเสียงสังเคราะห์ โดยรูปกราฟแสดงค่าสหสัมพันธ์ของผลรวมความถี่ฟอร์แมนต์ที่ 1 และ 2 ที่ได้จากจุดควบคุมความโค้งของเส้นโค้งเบซิเยร์ตามแกน x และ y ค่าสหสัมพันธ์ของจุดพิกัดที่มีค่าระหว่าง -2 ถึง 2 โดยค่าสหสัมพันธ์ที่มากแสดงถึงแนวการเปลี่ยนแปลงฟอร์แมนต์ที่ 1 และ 2 ของเสียงสังเคราะห์มีความใกล้เคียงกับการเปลี่ยนแปลงในเสียงจริง ชื่อใต้รูประบุถึงเสียงที่วิเคราะห์ รูปกราฟทางด้านซ้ายแสดงผลจุดควบคุมความโค้งเบซิเยร์ P_1 และรูปกราฟทางด้านขวาแสดงผลจุดควบคุมความโค้งเบซิเยร์ P_2



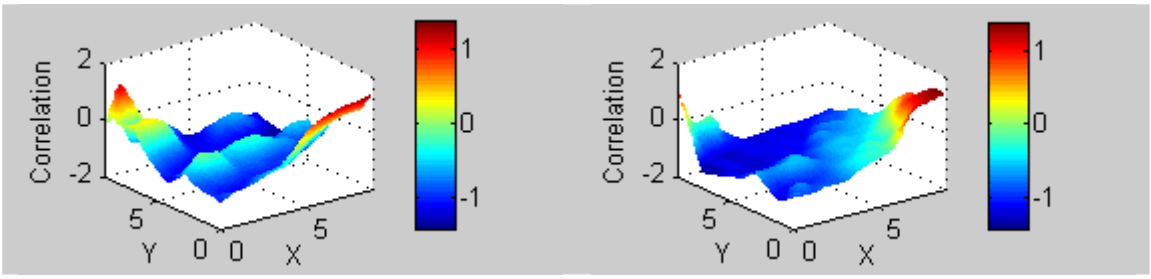
“บา”



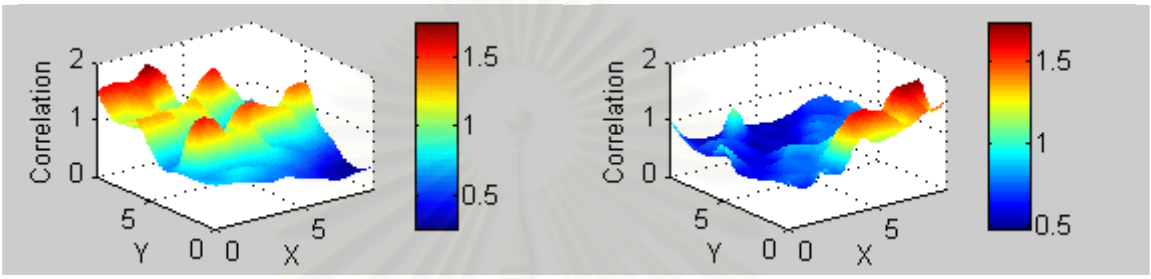
“ปี”



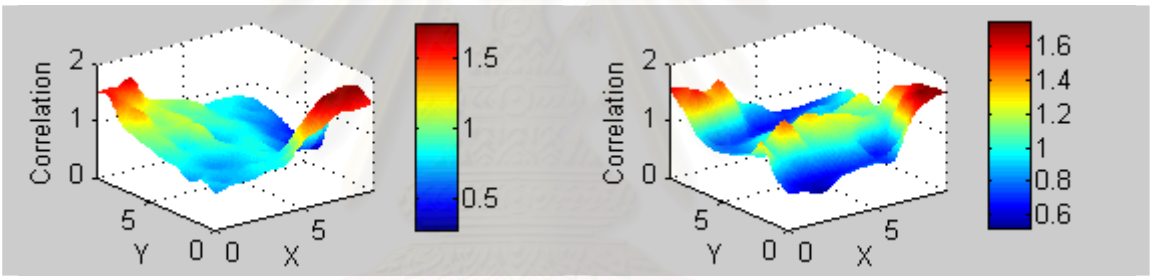
“ปือ”


Correlation
2
0
-2
5
Y
0
0
X
5
1
0
-1
Correlation
2
0
-2
5
Y
0
0
X
5
1
0
-1

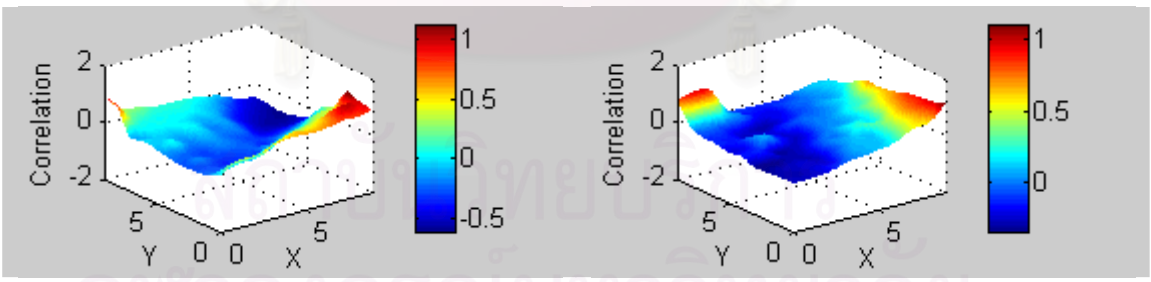
“ปู”


Correlation
2
1
0
5
Y
0
0
X
5
1.5
1
0.5
Correlation
2
1
0
5
Y
0
0
X
5
1.5
1
0.5

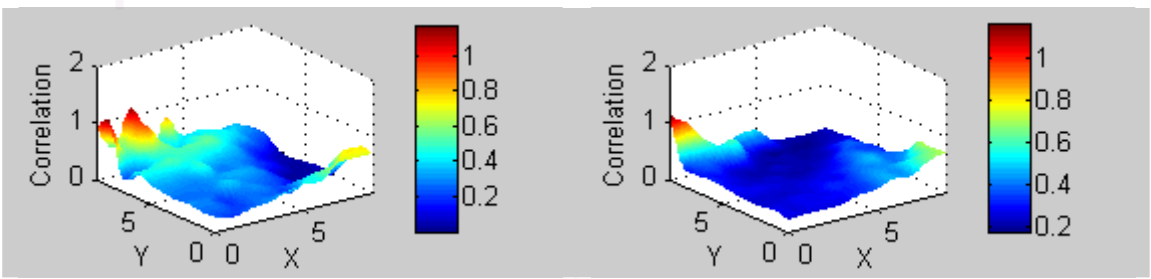
“เป”


Correlation
2
1
0
5
Y
0
0
X
5
1.5
1
0.5
Correlation
2
1
0
5
Y
0
0
X
5
1.6
1.4
1.2
1
0.8
0.6

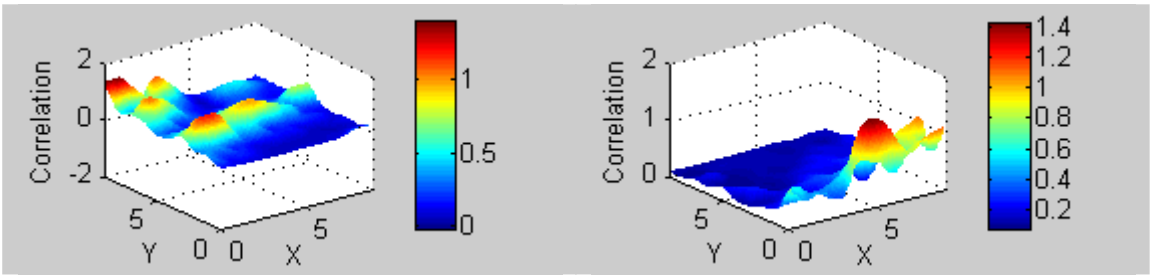
“แป”


Correlation
2
0
-2
5
Y
0
0
X
5
1
0.5
0
-0.5
Correlation
2
0
-2
5
Y
0
0
X
5
1
0.5
0

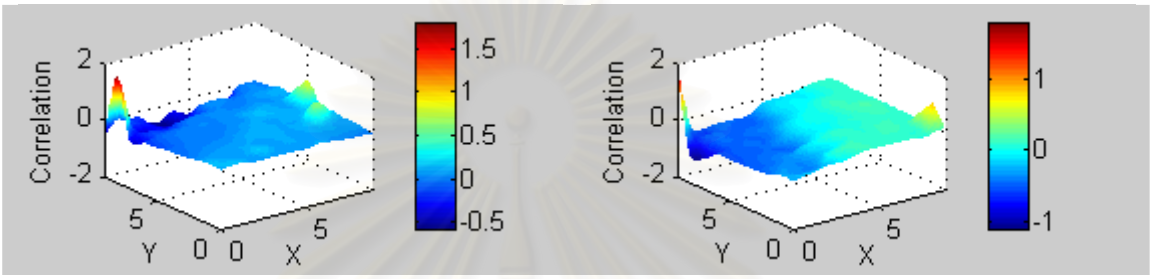
“โป”


Correlation
2
1
0
5
Y
0
0
X
5
1
0.8
0.6
0.4
0.2
Correlation
2
1
0
5
Y
0
0
X
5
1
0.8
0.6
0.4
0.2

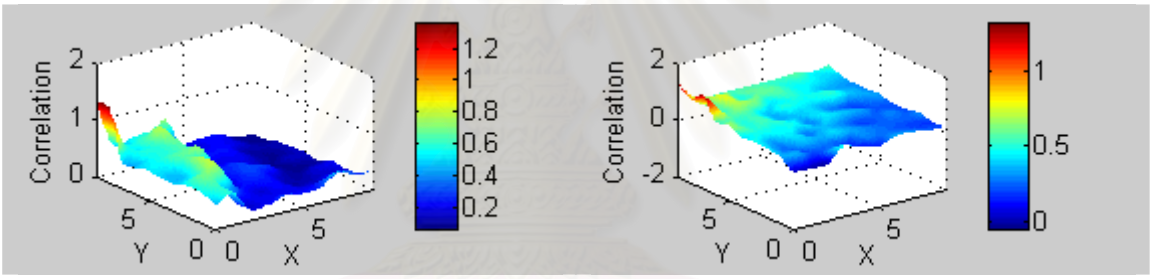
“บอ”


Correlation
2
0
-2
5
Y
0
0
X
5
1
0.5
0
Correlation
2
1
0
5
Y
0
0
X
5
1.4
1.2
1
0.8
0.6
0.4
0.2

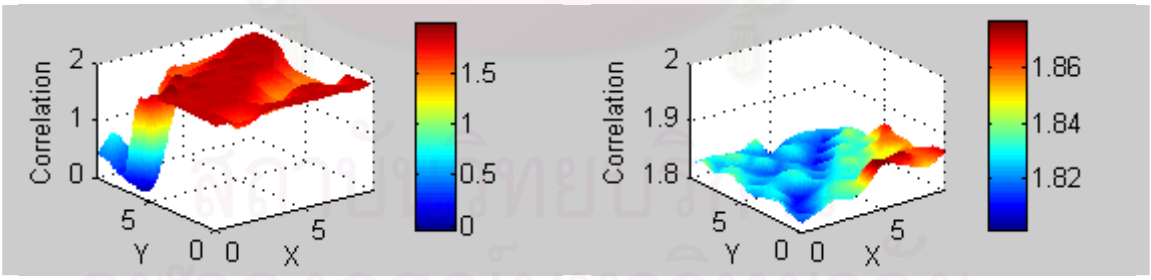
“เบอ”


Correlation
2
0
-2
5
Y
0
0
X
5
1.5
1
0.5
0
-0.5
Correlation
2
0
-2
5
Y
0
0
X
5
1
0
-1

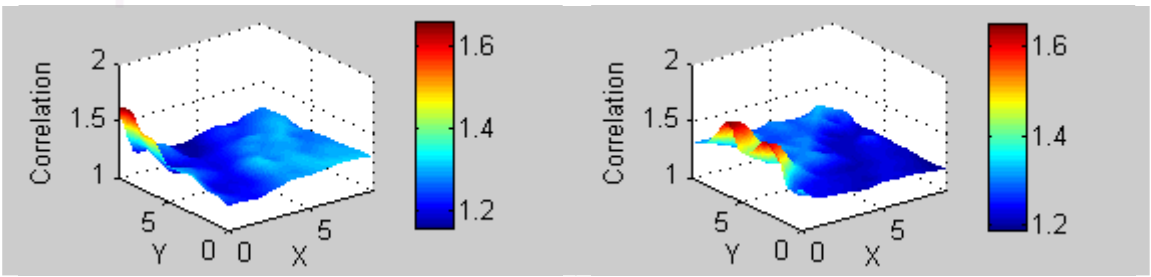
“ฟา”


Correlation
2
1
0
5
Y
0
0
X
5
1.2
1
0.8
0.6
0.4
0.2
Correlation
2
0
-2
5
Y
0
0
X
5
1
0.5
0

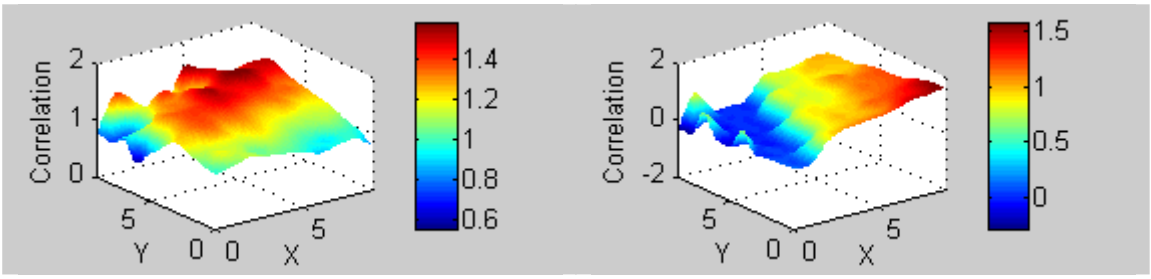
“ฟี”


Correlation
2
1
0
5
Y
0
0
X
5
1.5
1
0.5
0
Correlation
2
1.9
1.8
5
Y
0
0
X
5
1.86
1.84
1.82

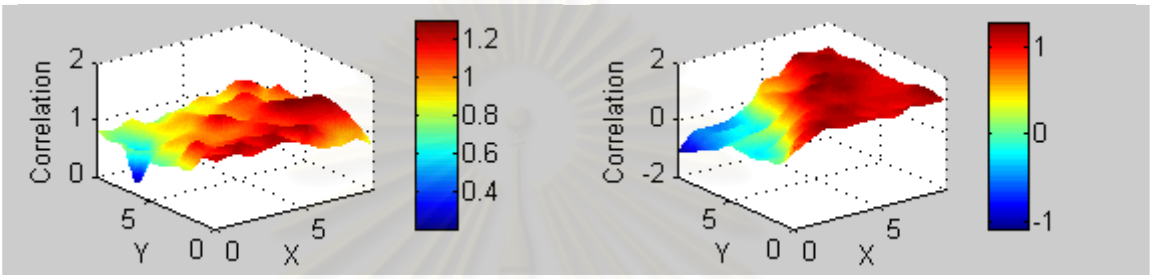
“ฟือ”


Correlation
2
1.5
1
5
Y
0
0
X
5
1.6
1.4
1.2
Correlation
2
1.5
1
5
Y
0
0
X
5
1.6
1.4
1.2

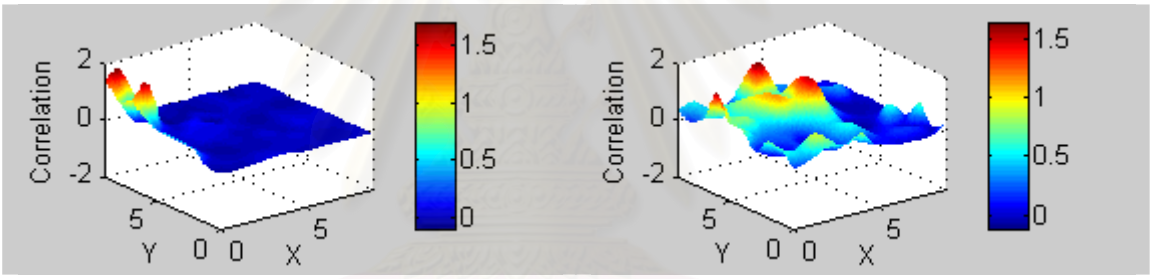
“ฟู”



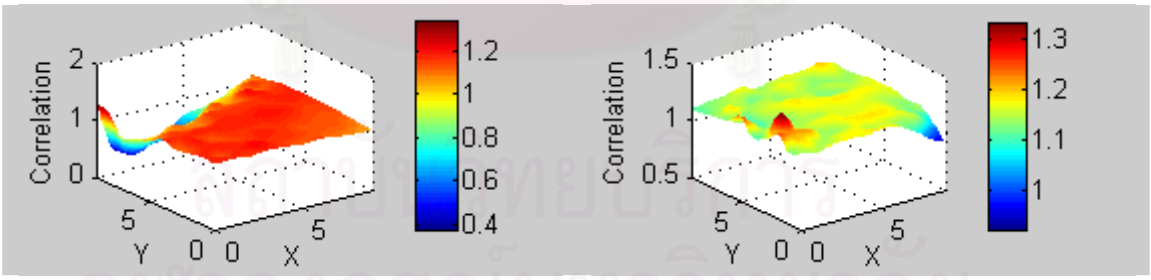
“เฟ”



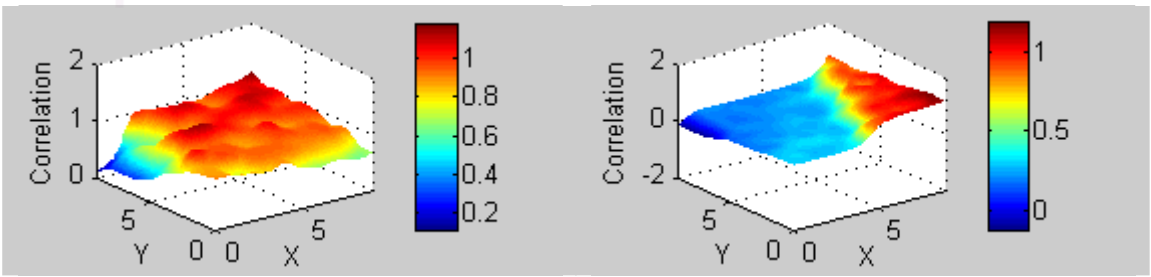
“แฟ”



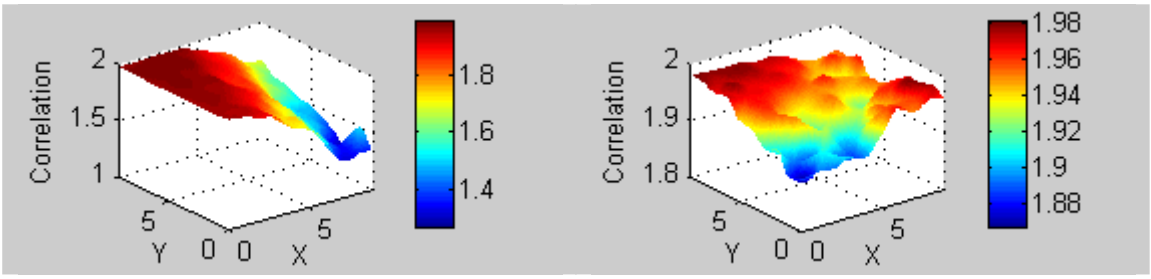
“โฟ”



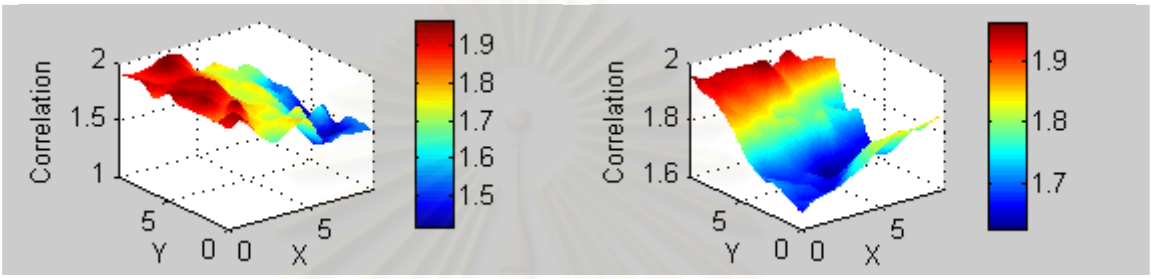
“ฟอ”



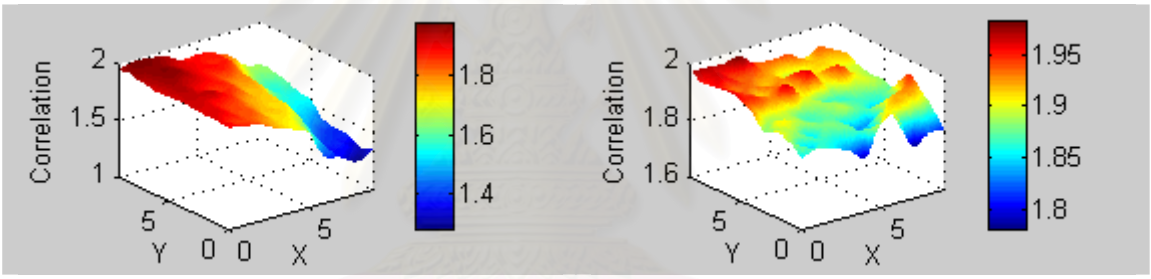
“เฟอ”


Correlation
Y
X
Correlation
Y
X

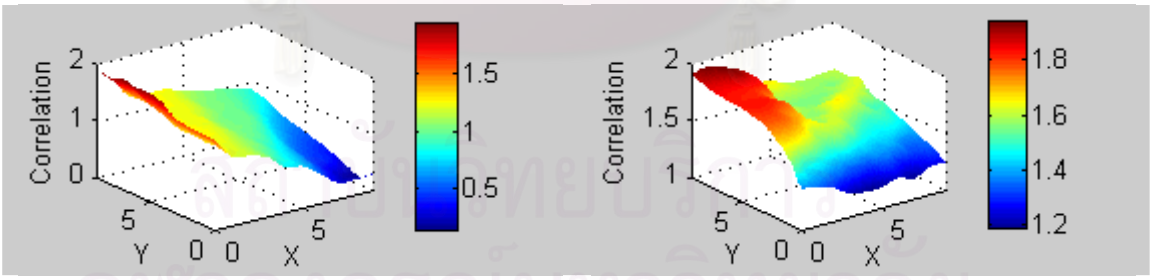
“ดา”


Correlation
Y
X
Correlation
Y
X

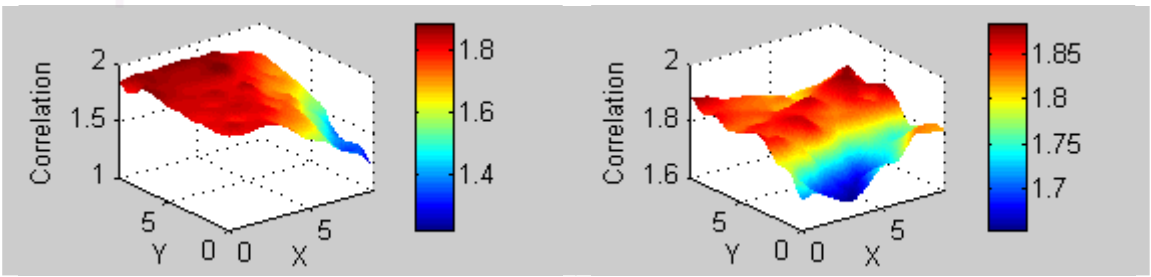
“ดี”


Correlation
Y
X
Correlation
Y
X

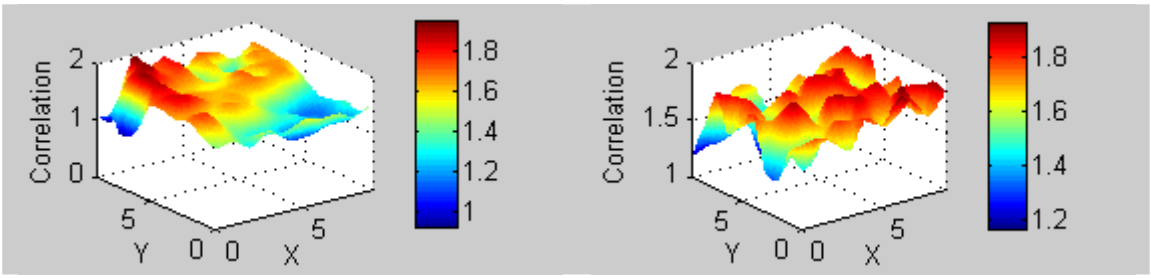
“ดือ”


Correlation
Y
X
Correlation
Y
X

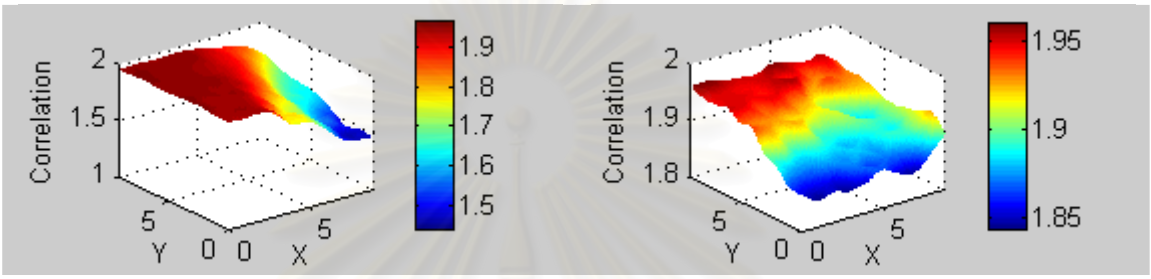
“ดู”


Correlation
Y
X
Correlation
Y
X

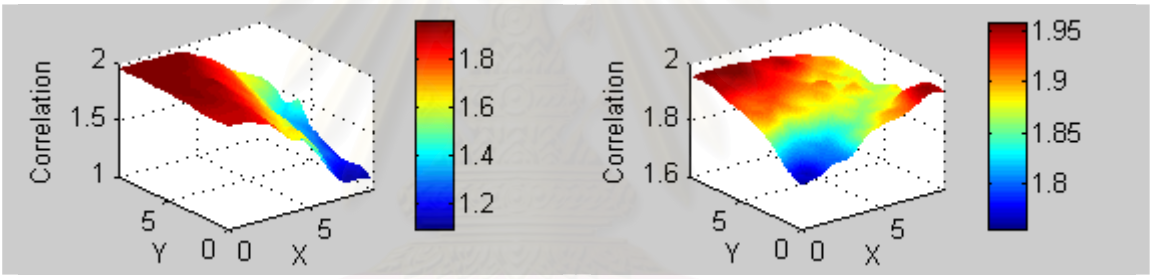
“เด”


Correlation
Y
X
Correlation
Y
X

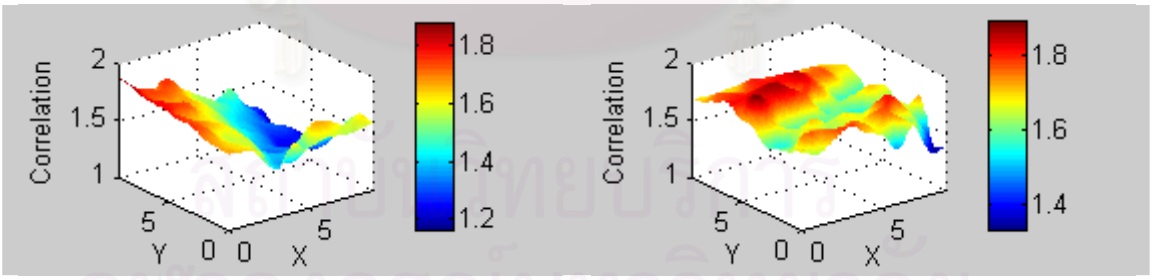
“แด”


Correlation
Y
X
Correlation
Y
X

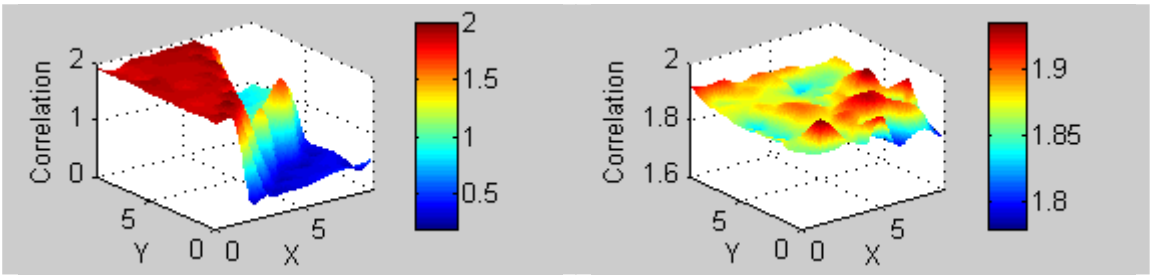
“โด”


Correlation
Y
X
Correlation
Y
X

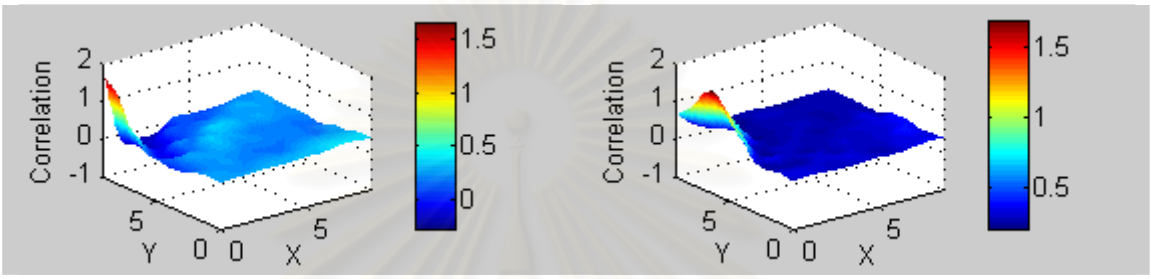
“ดอ”


Correlation
Y
X
Correlation
Y
X

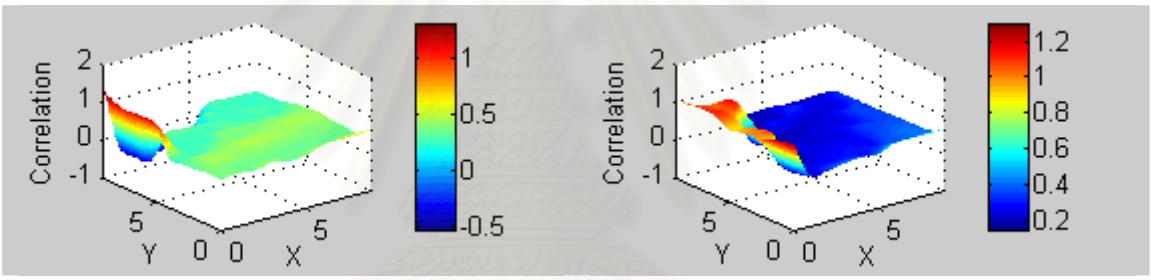
“เดอ”


Correlation
Y
X
Correlation
Y
X

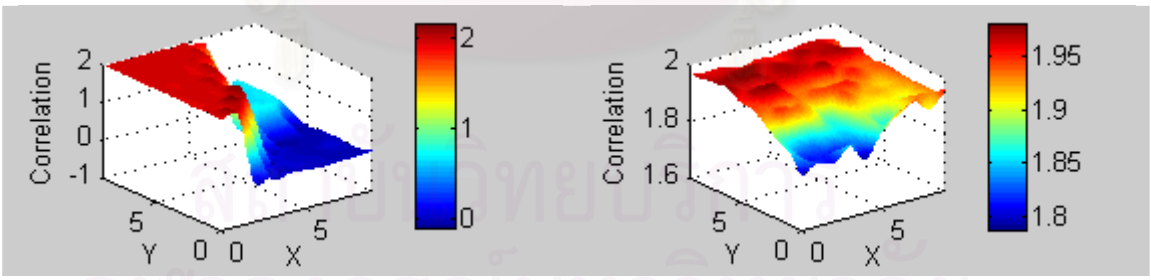
“ซา”


Correlation
Y
X
Correlation
Y
X

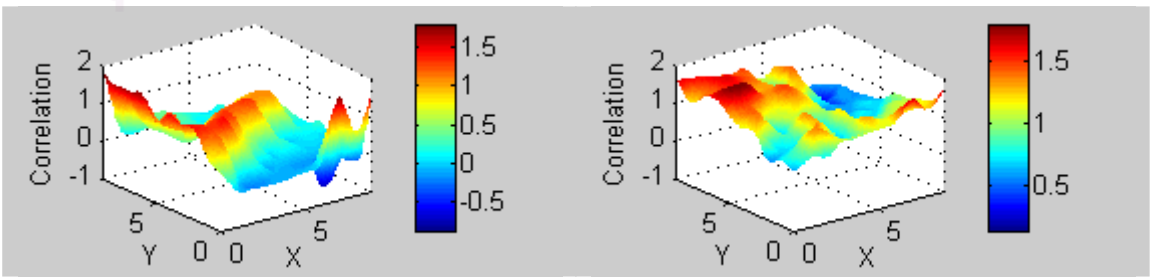
“ซี”


Correlation
Y
X
Correlation
Y
X

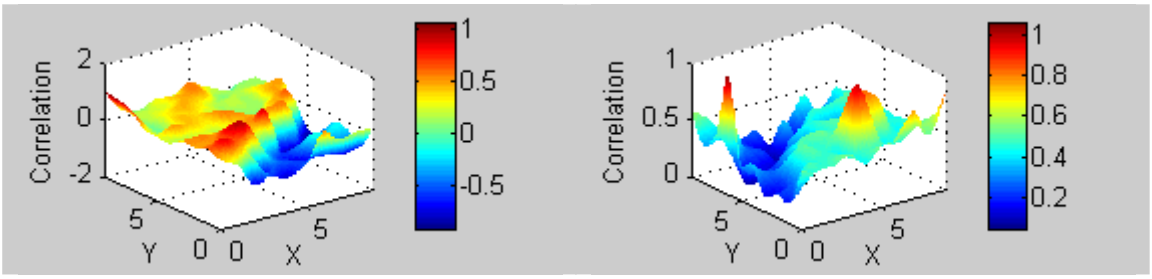
“ซือ”


Correlation
Y
X
Correlation
Y
X

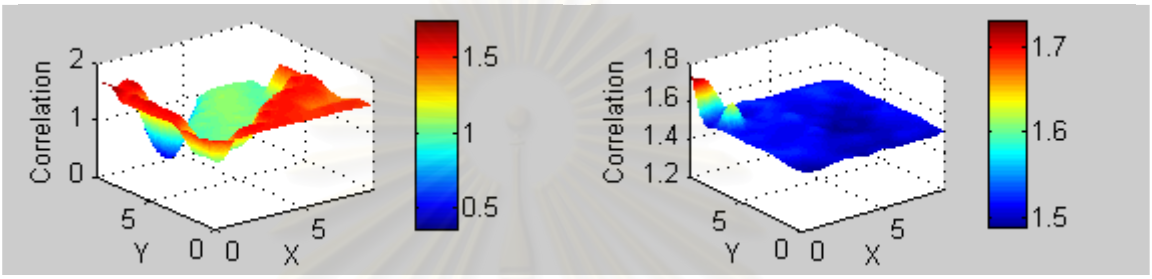
“ซู”


Correlation
Y
X
Correlation
Y
X

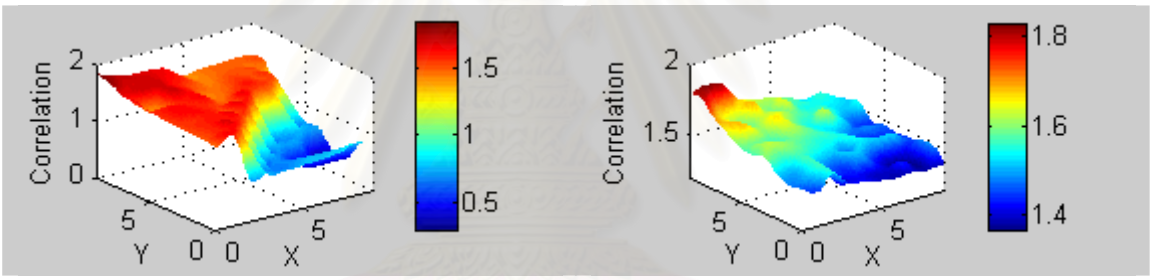
“เซ”


Correlation
2
0
-2
5
Y
0
0
X
5
1
0.5
0
-0.5
Correlation
1
0.5
0
5
Y
0
0
X
5
1
0.8
0.6
0.4
0.2

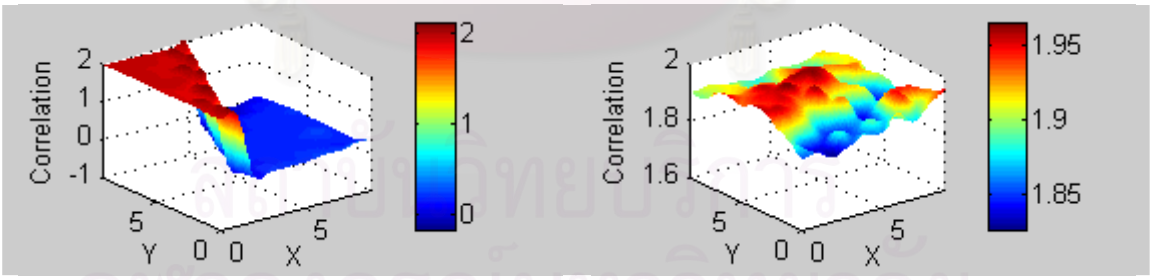
“แซ”


Correlation
2
1
0
5
Y
0
0
X
5
1.5
1
0.5
Correlation
1.8
1.6
1.4
1.2
5
Y
0
0
X
5
1.7
1.6
1.5

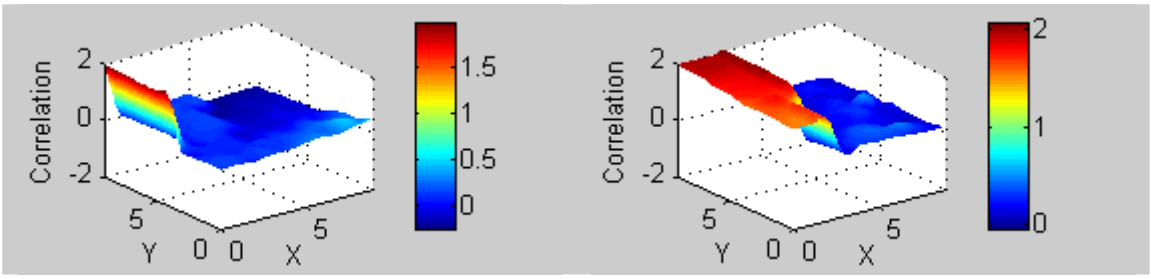
“โซ”


Correlation
2
1
0
5
Y
0
0
X
5
1.5
1
0.5
Correlation
2
1.5
5
Y
0
0
X
5
1.8
1.6
1.4

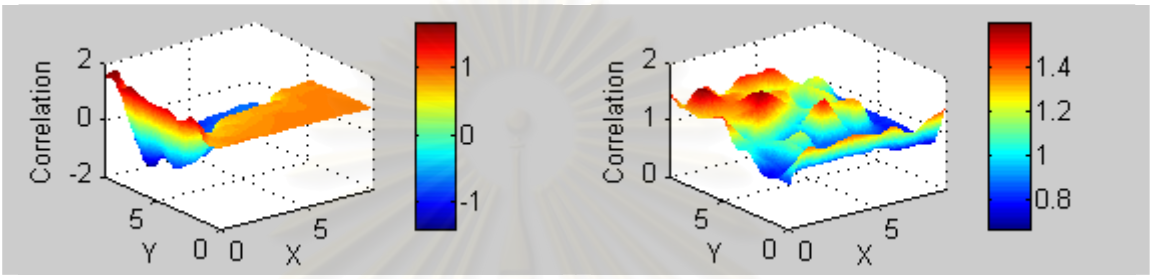
“ซอ”


Correlation
2
1
0
-1
5
Y
0
0
X
5
2
1
0
Correlation
2
1.8
1.6
5
Y
0
0
X
5
1.95
1.9
1.85

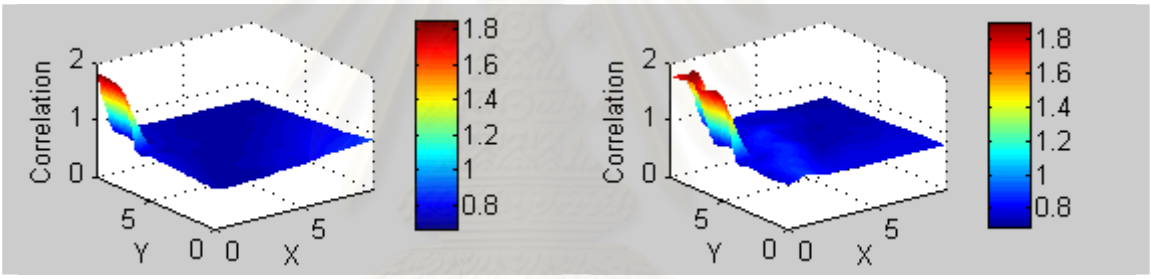
“เซอ”


Correlation
2
0
-2
5
Y
0
0
X
5
1.5
1
0.5
0
Correlation
2
0
-2
5
Y
0
0
X
5
2
1
0

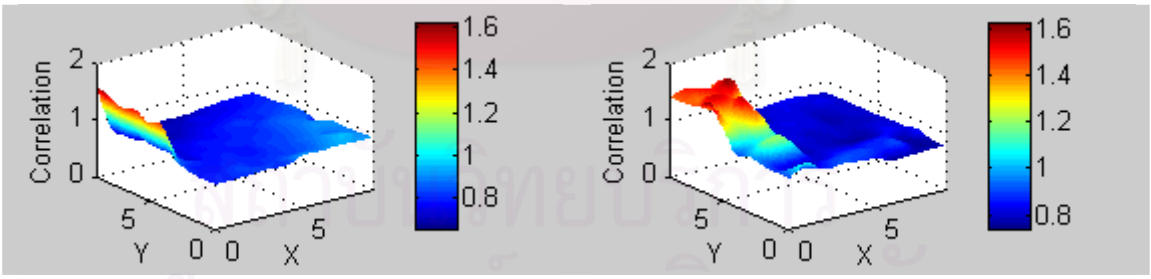
“จา”


Correlation
2
0
-2
5
Y
0
0
X
5
1
0
-1
Correlation
2
1
0
5
Y
0
0
X
5
1.4
1.2
1
0.8

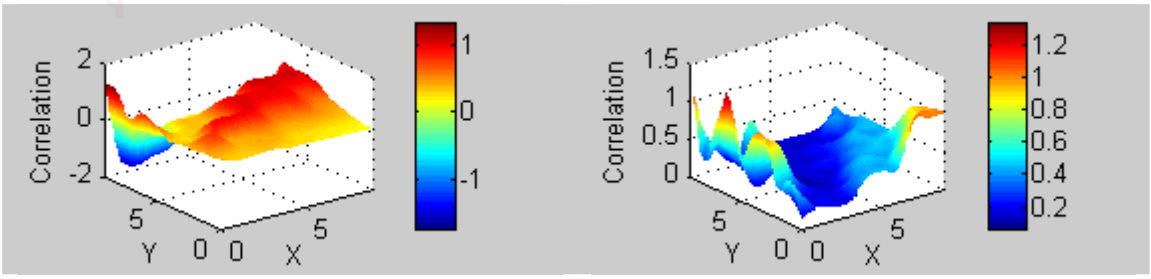
“จี”


Correlation
2
1
0
5
Y
0
0
X
5
1.8
1.6
1.4
1.2
1
0.8
Correlation
2
1
0
5
Y
0
0
X
5
1.8
1.6
1.4
1.2
1
0.8

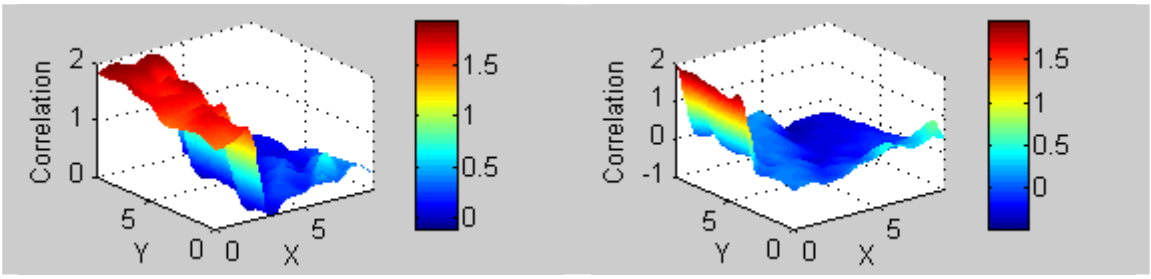
“จือ”


Correlation
2
1
0
5
Y
0
0
X
5
1.6
1.4
1.2
1
0.8
Correlation
2
1
0
5
Y
0
0
X
5
1.6
1.4
1.2
1
0.8

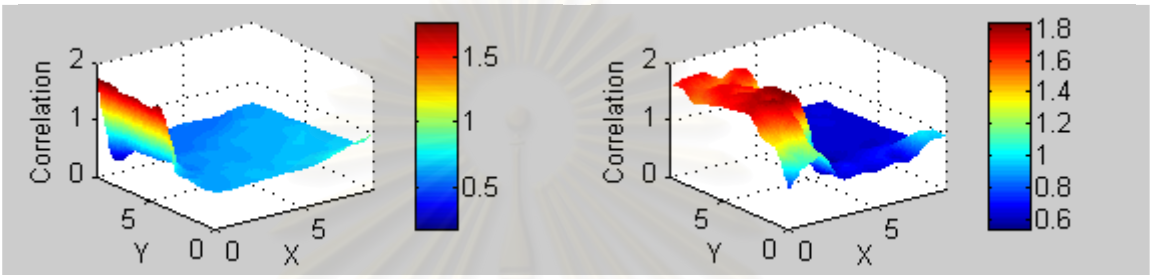
“จู”


Correlation
2
0
-2
5
Y
0
0
X
5
1
0
-1
Correlation
1.5
1
0.5
0
5
Y
0
0
X
5
1.2
1
0.8
0.6
0.4
0.2

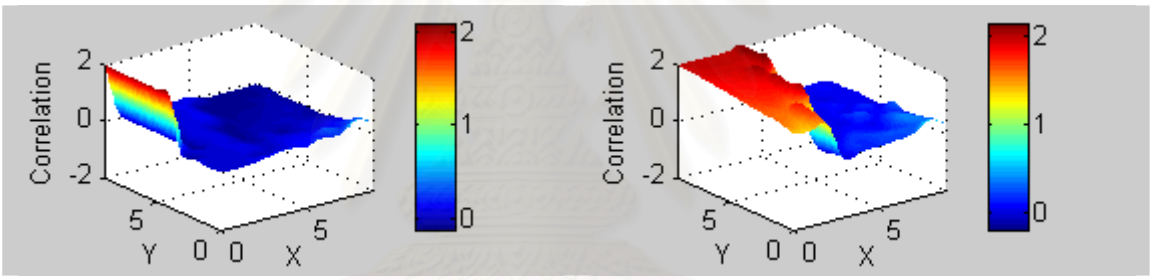
“เจ”


Correlation
2
1
0
5
Y
0
0
X
5
1.5
1
0.5
0
Correlation
2
1
0
-1
5
Y
0
0
X
5
1.5
1
0.5
0

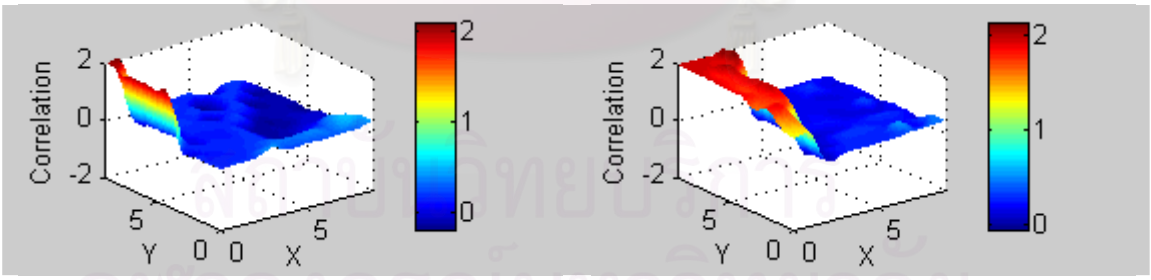
“แจ”


Correlation
2
1
0
5
Y
0
0
X
5
1.5
1
0.5
Correlation
2
1
0
5
Y
0
0
X
5
1.8
1.6
1.4
1.2
1
0.8
0.6

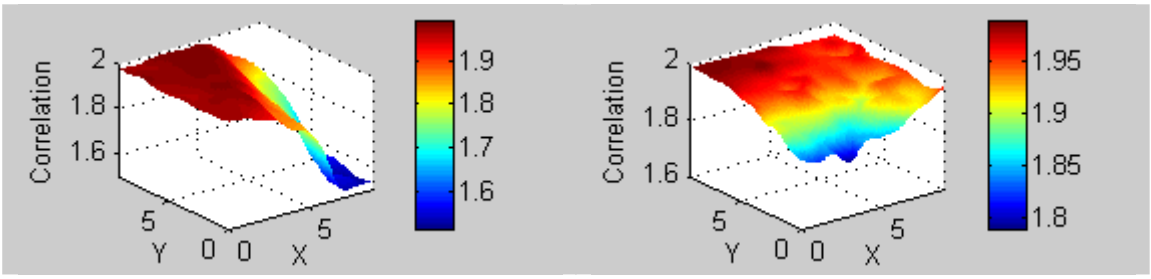
“โจ”


Correlation
2
0
-2
5
Y
0
0
X
5
2
1
0
Correlation
2
0
-2
5
Y
0
0
X
5
2
1
0

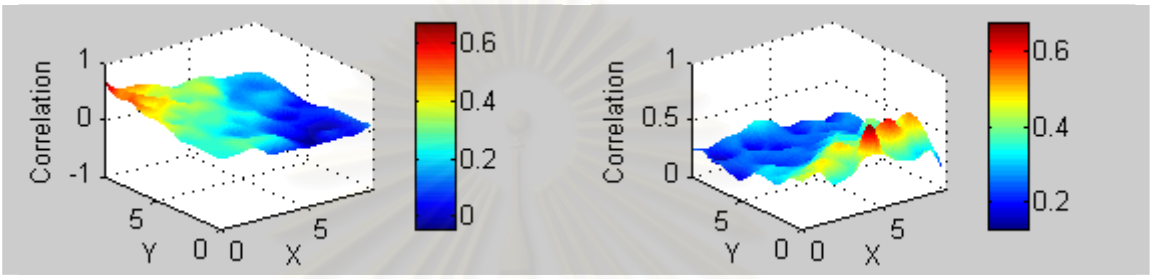
“จอ”


Correlation
2
0
-2
5
Y
0
0
X
5
2
1
0
Correlation
2
0
-2
5
Y
0
0
X
5
2
1
0

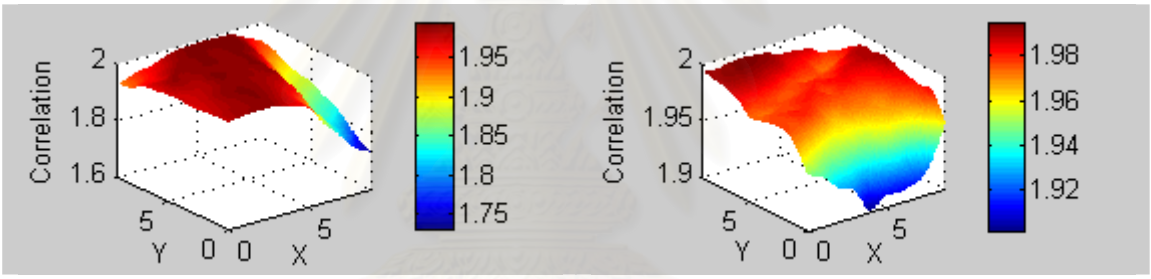
“เจอ”


Correlation
2
1.8
1.6
5
Y
0
0
X
5
1.9
1.8
1.7
1.6
Correlation
2
1.8
1.6
5
Y
0
0
X
5
1.95
1.9
1.85
1.8

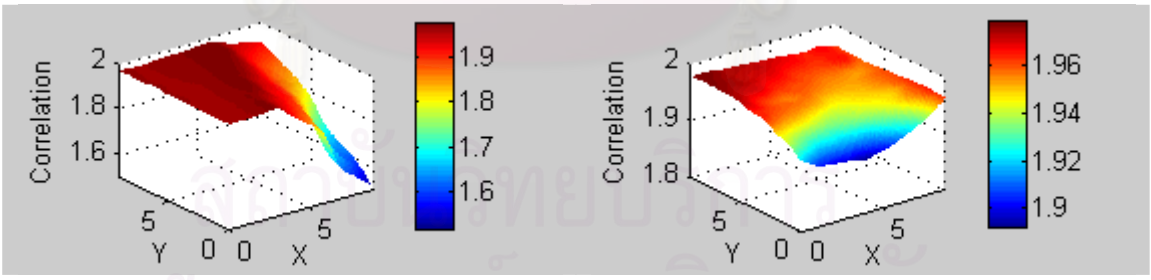
“ยา”


Correlation
1
0
-1
5
Y
0
0
X
5
0.6
0.4
0.2
0
Correlation
1
0.5
0
5
Y
0
0
X
5
0.6
0.4
0.2

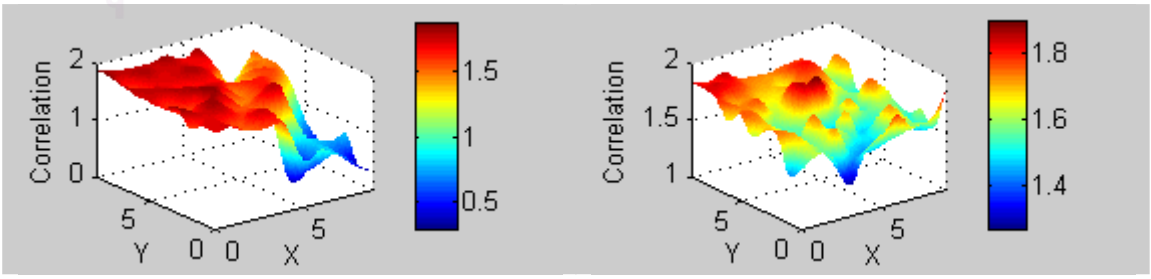
“ยี”


Correlation
2
1.8
1.6
5
Y
0
0
X
5
1.95
1.9
1.85
1.8
1.75
Correlation
2
1.95
1.9
5
Y
0
0
X
5
1.98
1.96
1.94
1.92

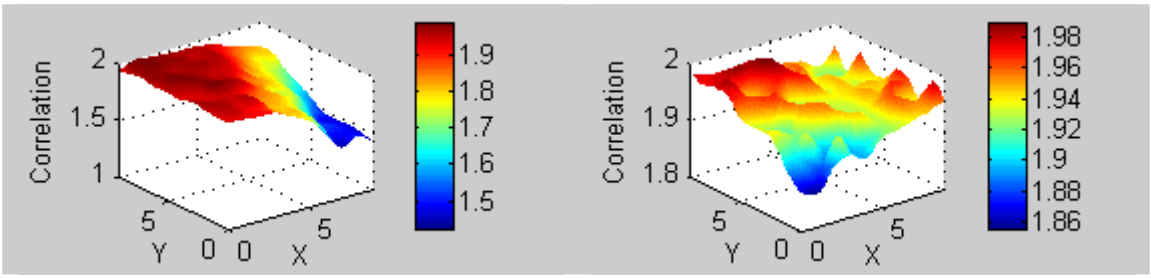
“ยือ”


Correlation
2
1.8
1.6
5
Y
0
0
X
5
1.9
1.8
1.7
1.6
Correlation
2
1.9
1.8
5
Y
0
0
X
5
1.96
1.94
1.92
1.9

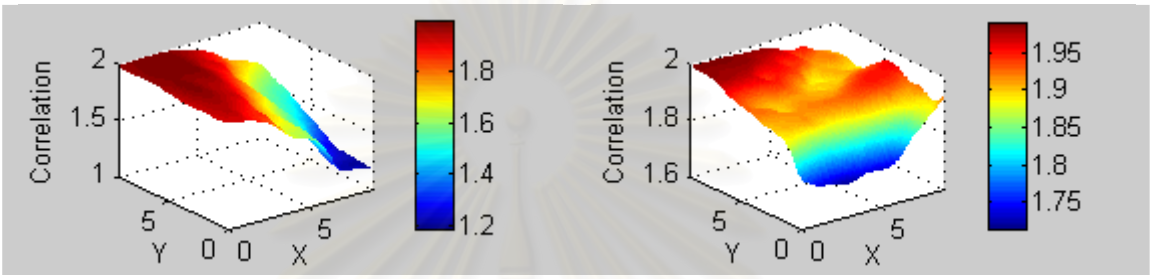
“ยู”


Correlation
2
1
0
5
Y
0
0
X
5
1.5
1
0.5
Correlation
2
1.5
1
5
Y
0
0
X
5
1.8
1.6
1.4

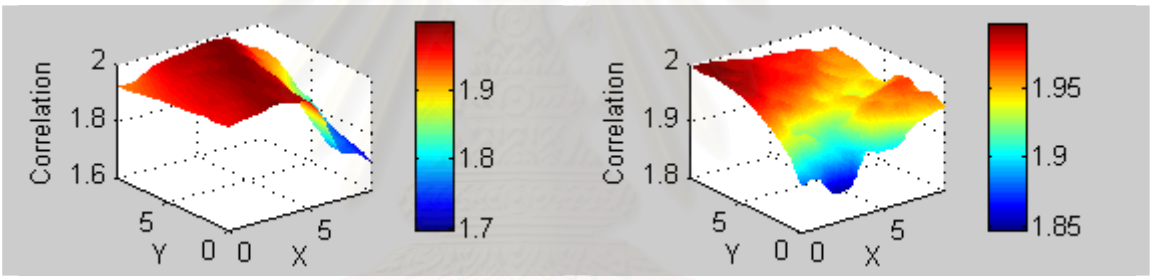
“เย”


Correlation
2
1.5
1
5
Y
0
0
X
5
1.9
1.8
1.7
1.6
1.5
Correlation
2
1.9
1.8
5
Y
0
0
X
5
1.98
1.96
1.94
1.92
1.9
1.88
1.86

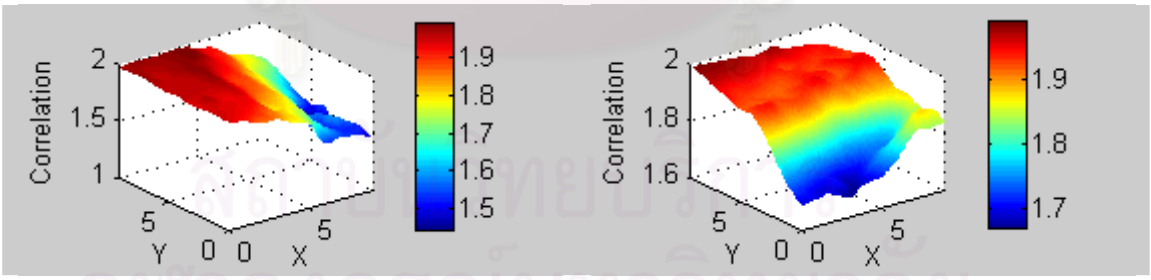
“แย”


Correlation
2
1.5
1
5
Y
0
0
X
5
1.8
1.6
1.4
1.2
Correlation
2
1.8
1.6
5
Y
0
0
X
5
1.95
1.9
1.85
1.8
1.75

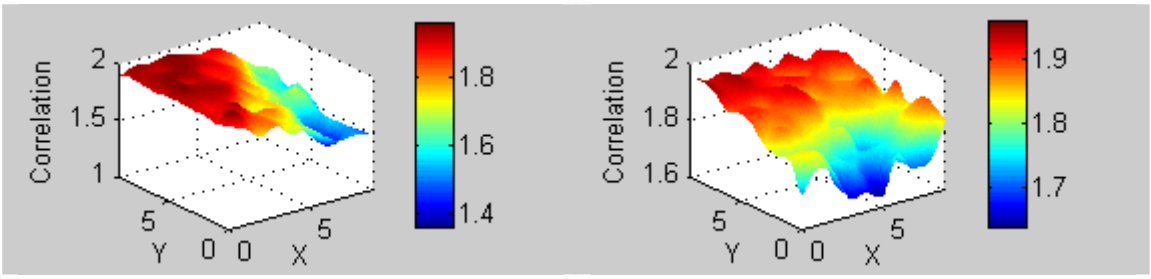
“โย”


Correlation
2
1.8
1.6
5
Y
0
0
X
5
1.9
1.8
1.7
Correlation
2
1.9
1.8
5
Y
0
0
X
5
1.95
1.9
1.85

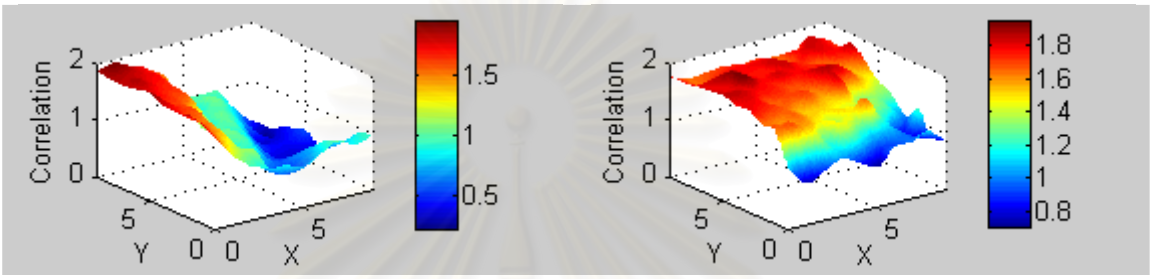
“ยอ”


Correlation
2
1.5
1
5
Y
0
0
X
5
1.9
1.8
1.7
1.6
1.5
Correlation
2
1.8
1.6
5
Y
0
0
X
5
1.9
1.8
1.7

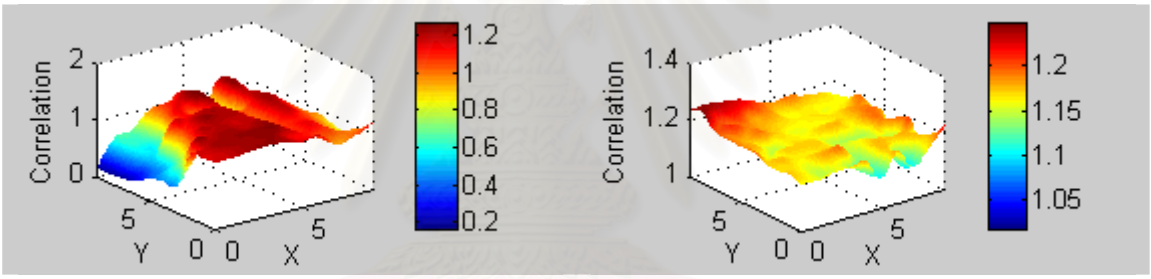
“เยอ”


Correlation
Y
X
Correlation
Y
X

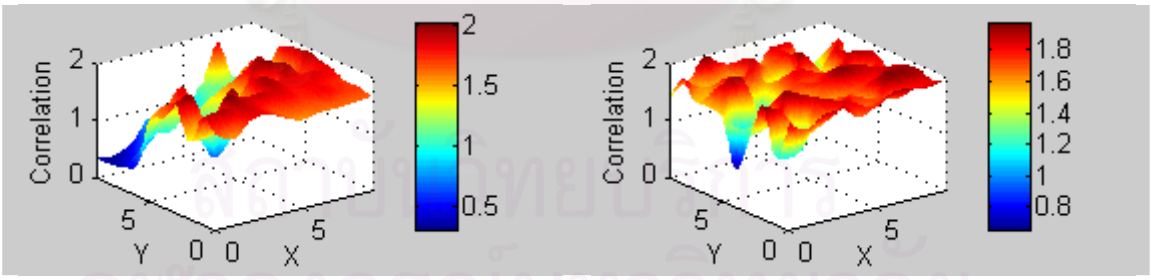
“กา”


Correlation
Y
X
Correlation
Y
X

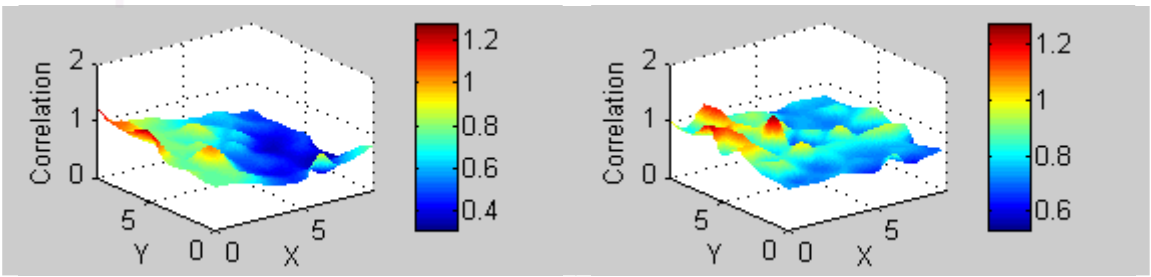
“กี”


Correlation
Y
X
Correlation
Y
X

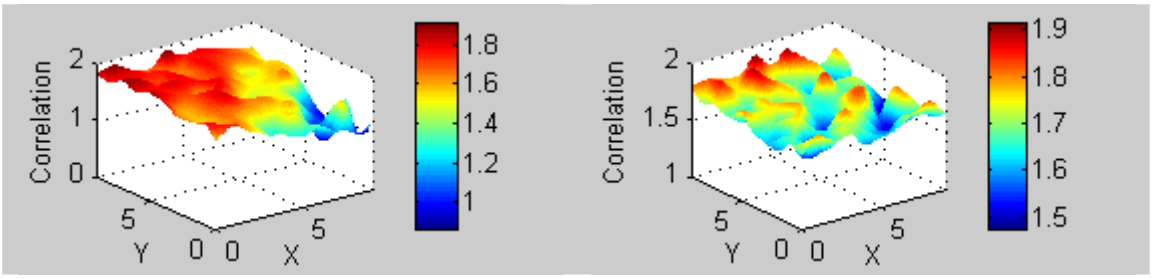
“กือ”


Correlation
Y
X
Correlation
Y
X

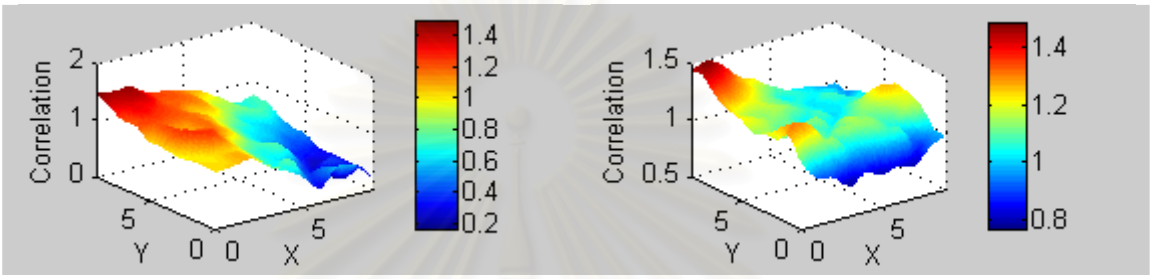
“กู”


Correlation
Y
X
Correlation
Y
X

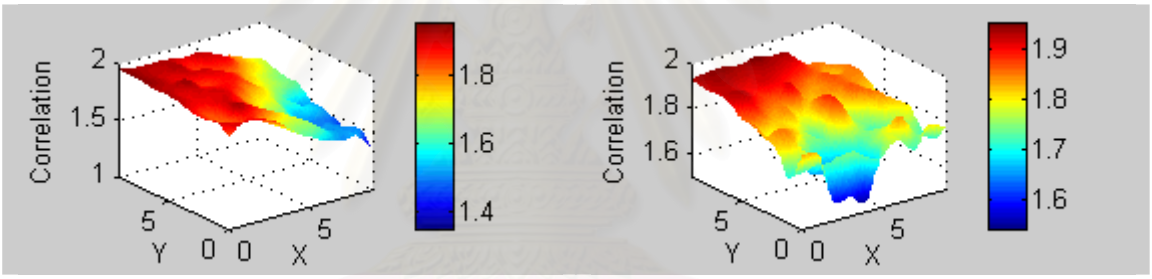
“เก”


Correlation
Y
X
Correlation
Y
X

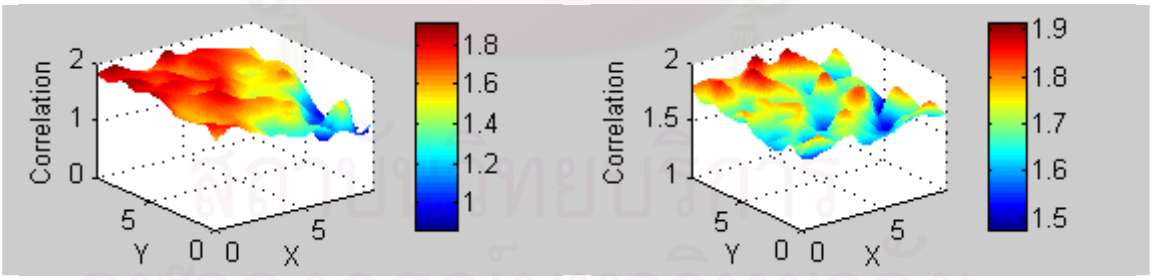
“แก”


Correlation
Y
X
Correlation
Y
X

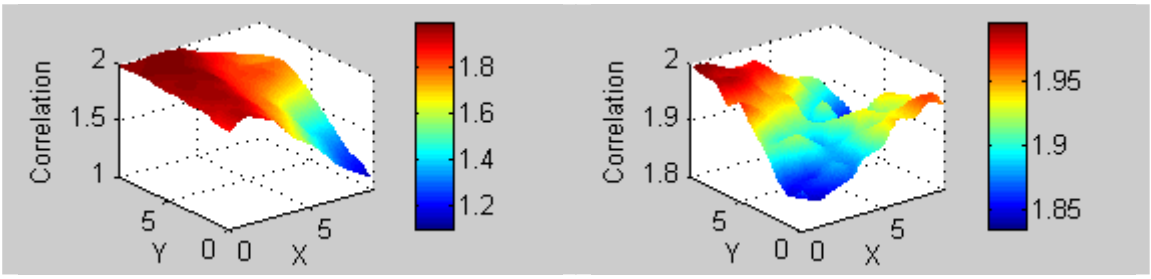
“โก”


Correlation
Y
X
Correlation
Y
X

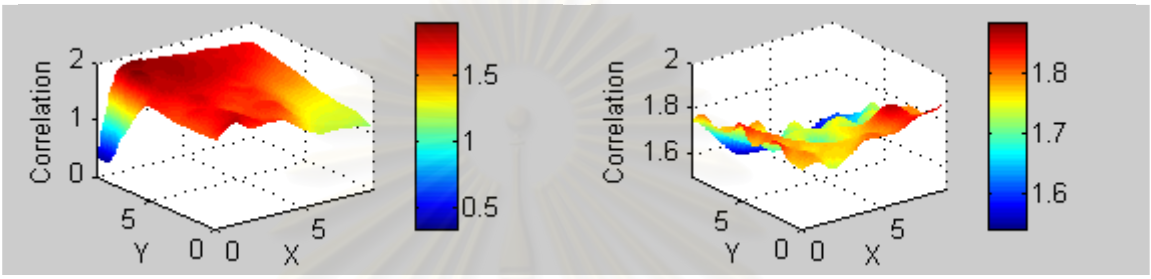
“กอ”


Correlation
Y
X
Correlation
Y
X

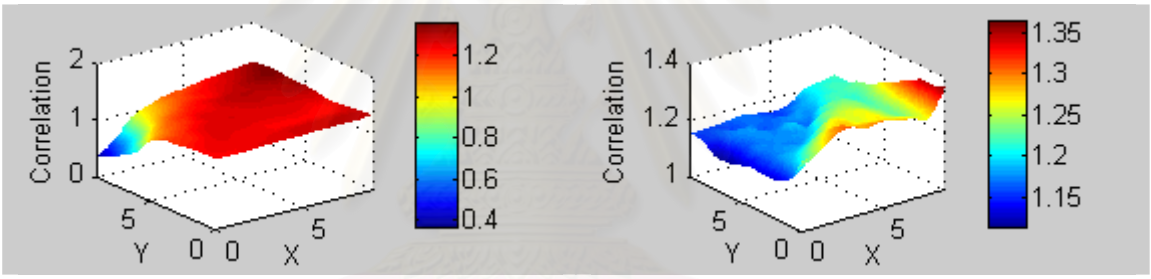
“เกอ”


Correlation
Y
X
Correlation
Y
X

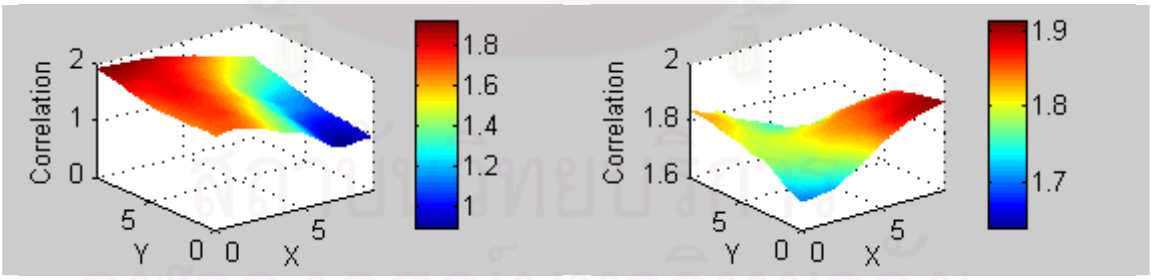
“วา”


Correlation
Y
X
Correlation
Y
X

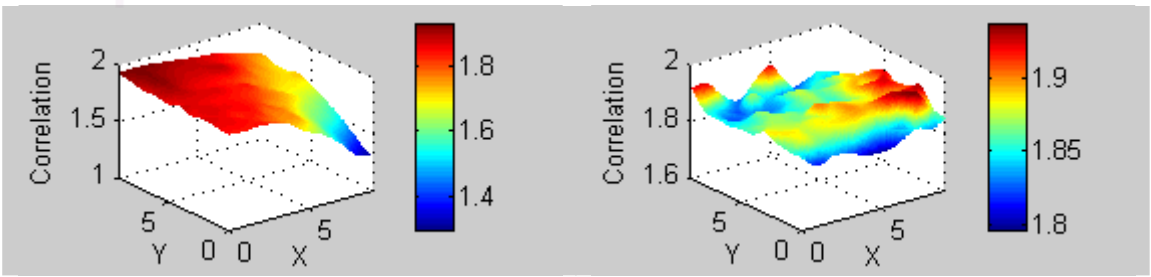
“สู”


Correlation
Y
X
Correlation
Y
X

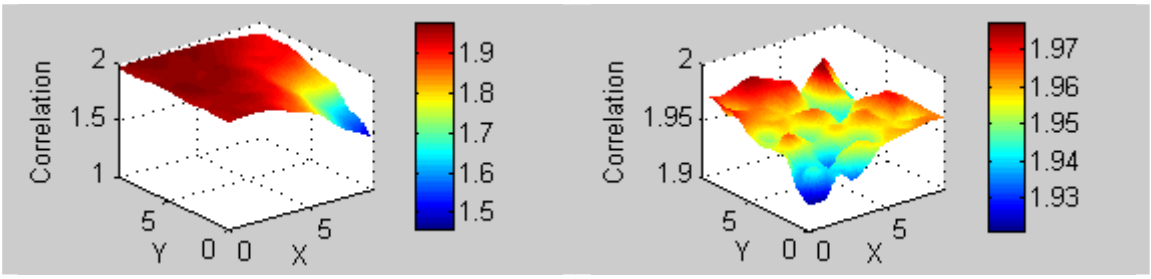
“เอ”


Correlation
Y
X
Correlation
Y
X

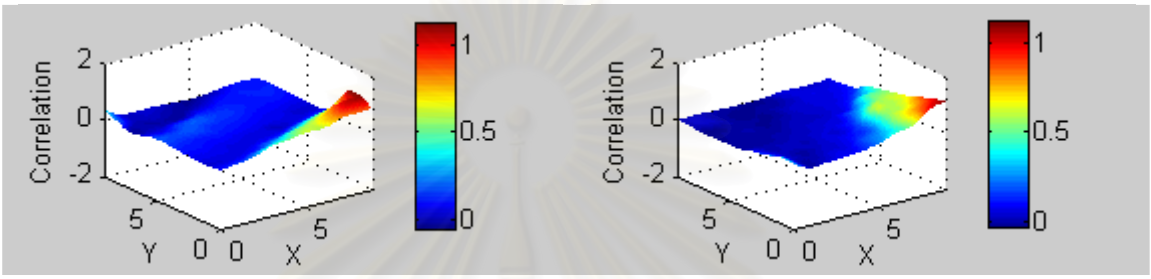
“อู”


Correlation
Y
X
Correlation
Y
X

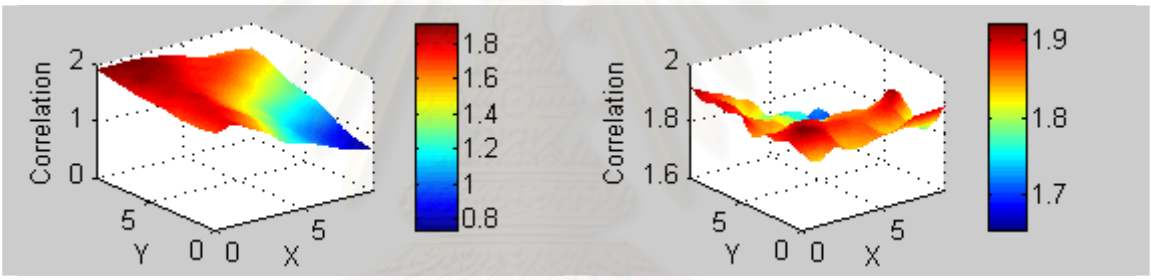
“เว”


Correlation
2
1.5
1
5
Y
0
0
X
5
1.9
1.8
1.7
1.6
1.5
Correlation
2
1.95
1.9
5
Y
0
0
X
5
1.97
1.96
1.95
1.94
1.93

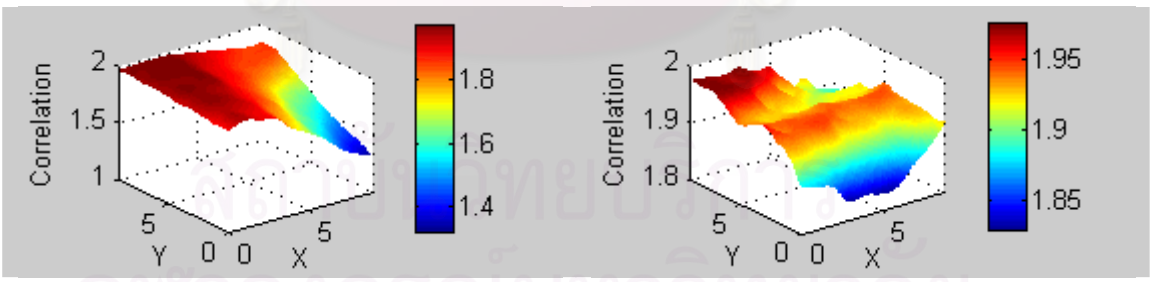
“แว”


Correlation
2
0
-2
5
Y
0
0
X
5
1
0.5
0
Correlation
2
0
-2
5
Y
0
0
X
5
1
0.5
0

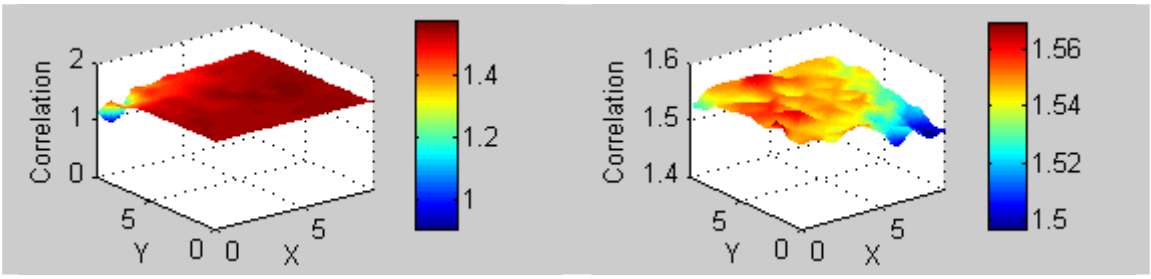
“โว”


Correlation
2
1
0
5
Y
0
0
X
5
1.8
1.6
1.4
1.2
1
0.8
Correlation
2
1.8
1.6
5
Y
0
0
X
5
1.9
1.8
1.7

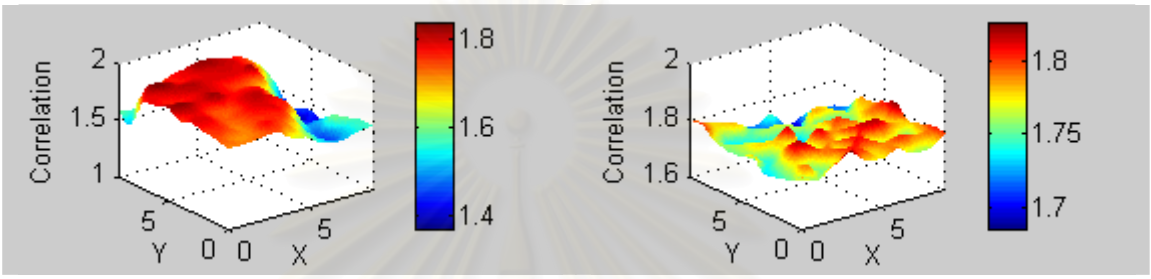
“วอ”


Correlation
2
1.5
1
5
Y
0
0
X
5
1.8
1.6
1.4
Correlation
2
1.9
1.8
5
Y
0
0
X
5
1.95
1.9
1.85

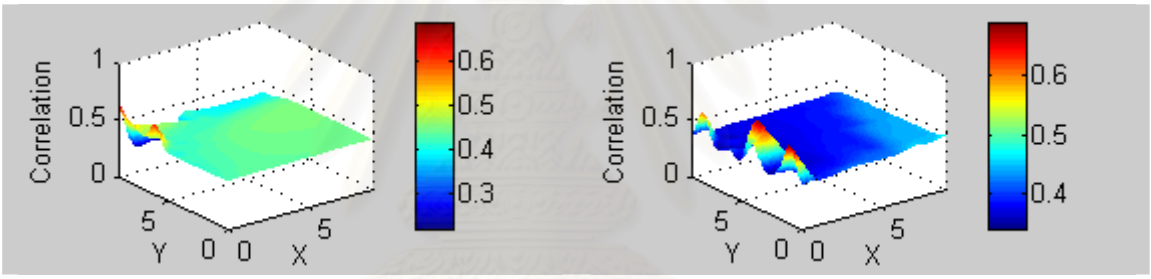
“เวอ”


Correlation
2
1
0
5
Y
0
0
X
5
1.4
1.2
1
Correlation
1.6
1.5
1.4
5
Y
0
0
X
5
1.56
1.54
1.52
1.5

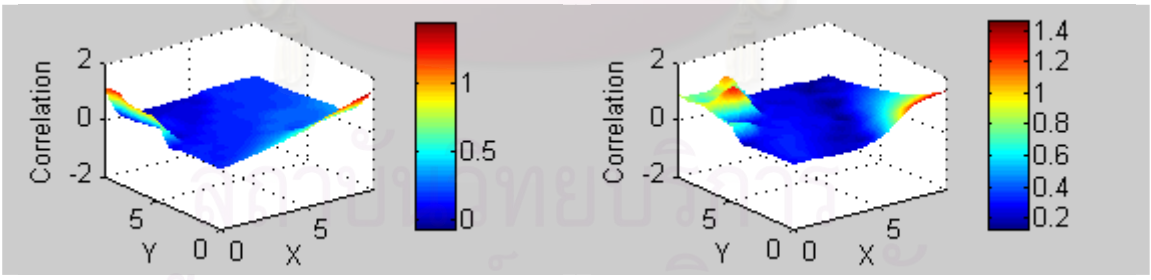
“ฮา”


Correlation
2
1.5
1
5
Y
0
0
X
5
1.8
1.6
1.4
Correlation
2
1.8
1.6
5
Y
0
0
X
5
1.8
1.75
1.7

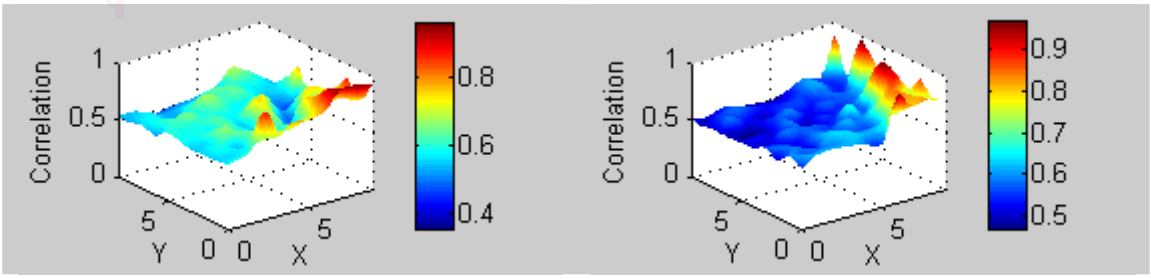
“ฮี


Correlation
1
0.5
0
5
Y
0
0
X
5
0.6
0.5
0.4
0.3
Correlation
1
0.5
0
5
Y
0
0
X
5
0.6
0.5
0.4

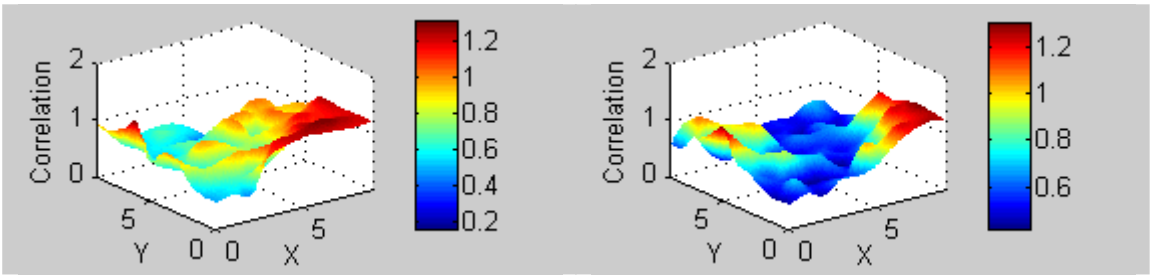
“ฮือ”


Correlation
2
0
-2
5
Y
0
0
X
5
1
0.5
0
Correlation
2
0
-2
5
Y
0
0
X
5
1.4
1.2
1
0.8
0.6
0.4
0.2

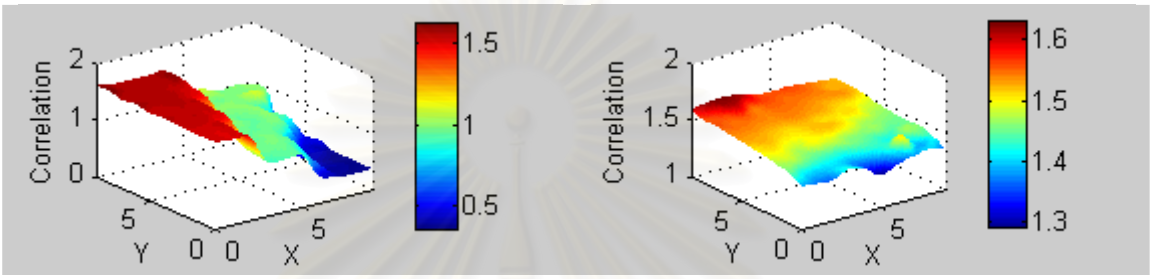
“ฮุ”


Correlation
1
0.5
0
5
Y
0
0
X
5
0.8
0.6
0.4
Correlation
1
0.5
0
5
Y
0
0
X
5
0.9
0.8
0.7
0.6
0.5

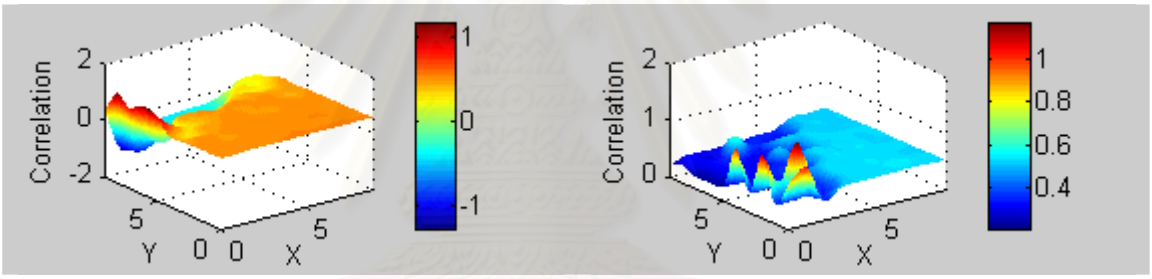
“เฮ”


Correlation
Y
X
1.2
1
0.8
0.6
0.4
0.2
Correlation
Y
X
1.2
1
0.8
0.6

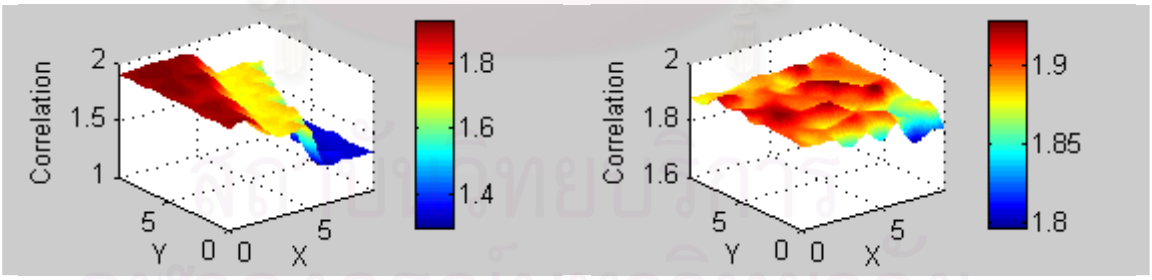
“แฮ”


Correlation
Y
X
1.5
1
0.5
Correlation
Y
X
1.6
1.5
1.4
1.3

“โฮ”


Correlation
Y
X
1
0
-1
Correlation
Y
X
1
0.8
0.6
0.4

“ฮอ”


Correlation
Y
X
1.8
1.6
1.4
Correlation
Y
X
1.9
1.85
1.8

“เฮอ”

ภาคผนวก ข

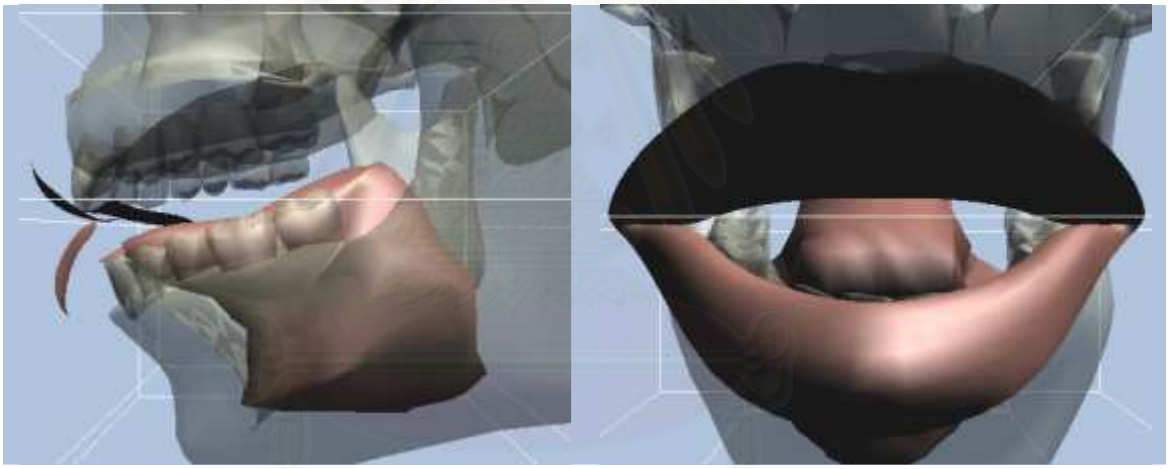
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ข

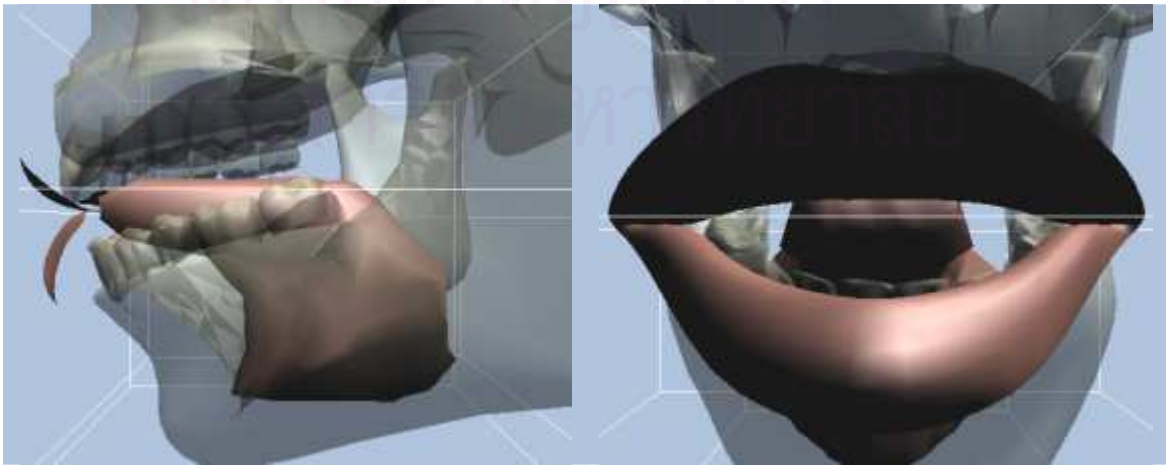
ผลการทดสอบหาค่าพารามิเตอร์ควบคุมช่องทางเสียง

ภาคผนวก ข แสดงผลของการทดสอบหาค่าพารามิเตอร์ควบคุมช่องทางเสียงสระและหน่วยเสียงพยัญชนะภาษาไทย แสดงผลค่าพารามิเตอร์ควบคุมช่องทางเสียงด้วยรูปแบบจำลองสามมิติของอวัยวะเปล่งเสียง รูปด้านซ้ายแสดงภาพด้านข้างของแบบจำลองช่องทางเสียง รูปด้านขวาแสดงภาพริมฝีปากของแบบจำลองช่องทางเสียง ชื่อใต้รูประบุหน่วยเสียงของแบบจำลองนั้น ๆ สำหรับหน่วยเสียงพยัญชนะมีบางหน่วยเสียงที่มีเสียงแตกต่างกัน เนื่องจากความแตกต่างในส่วนของพารามิเตอร์ควบคุมลักษณะหน่วยเสียง แต่มีรูปร่างช่องทางเสียงเหมือนกัน จึงรวมแสดงผลเป็นรูปเดียวกัน

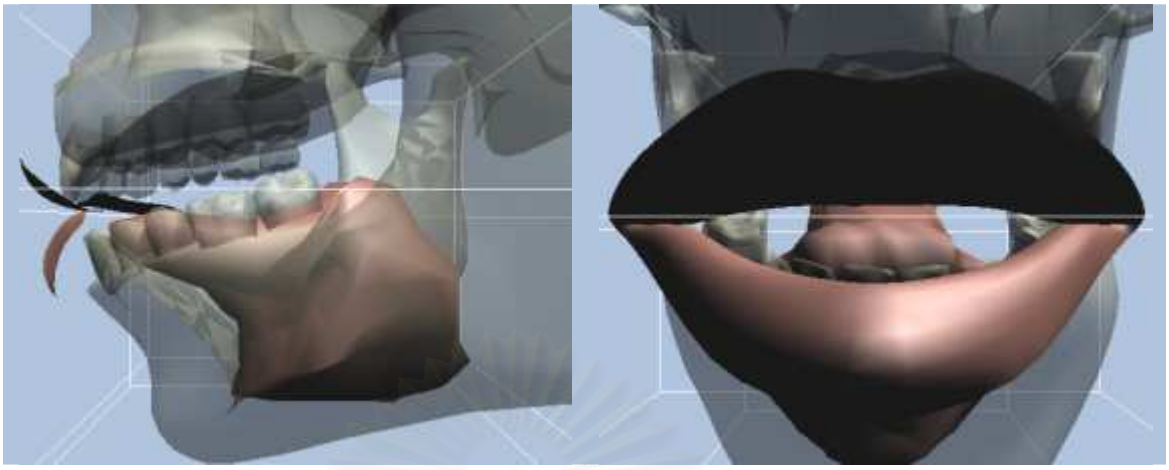
หน่วยเสียงสระ



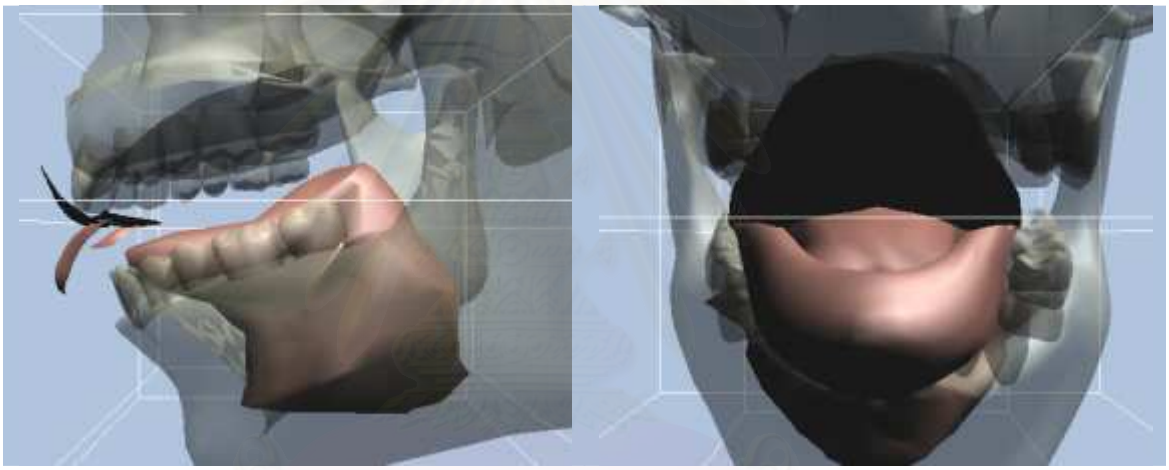
“อา”



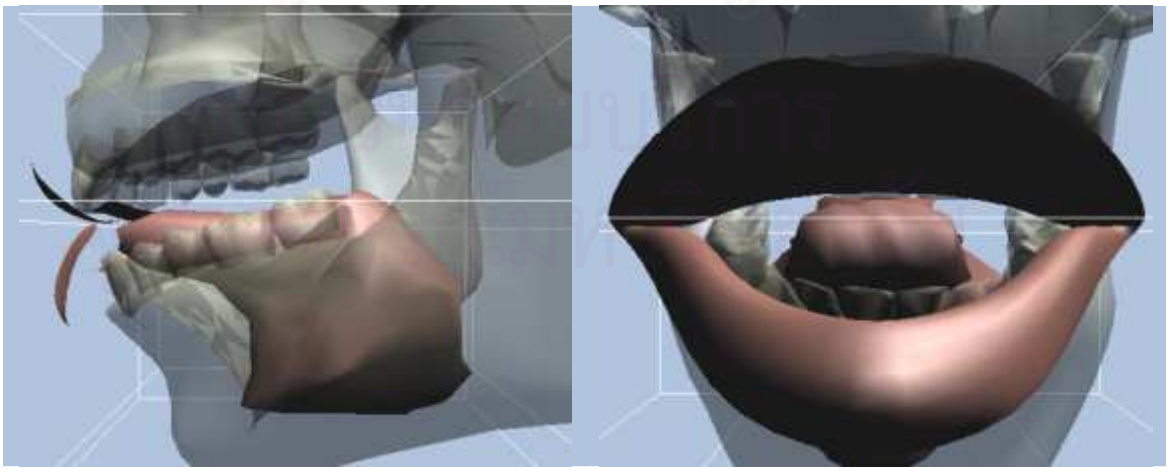
“อี”



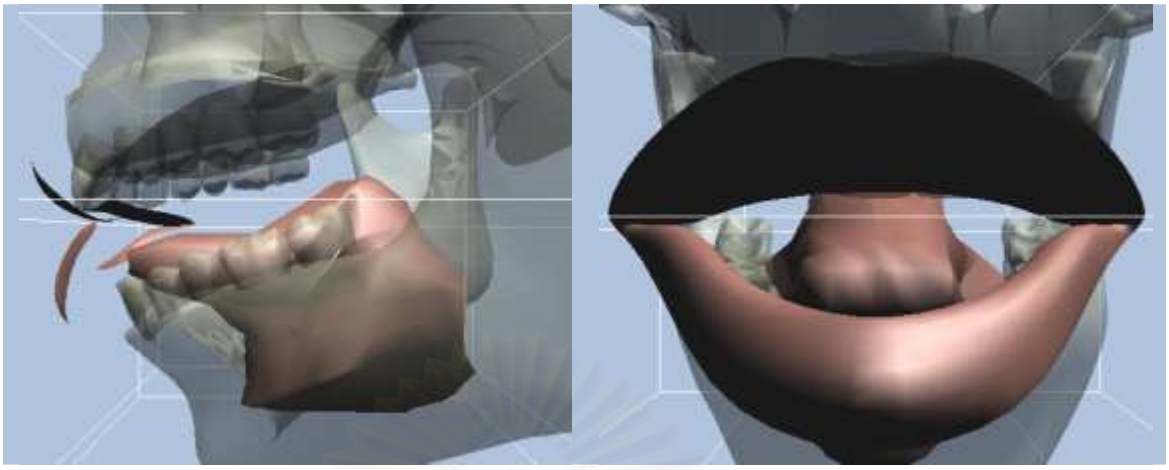
“อือ”



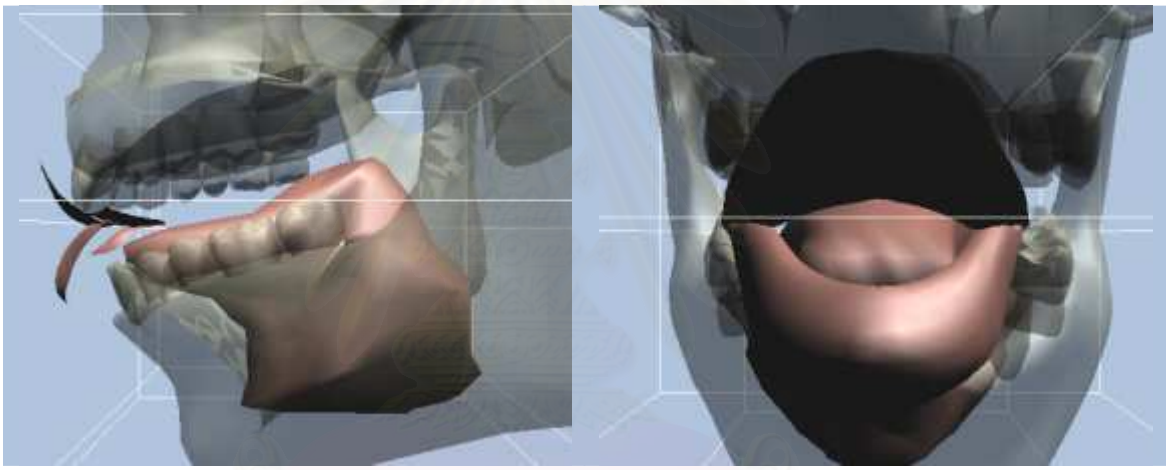
“อู”



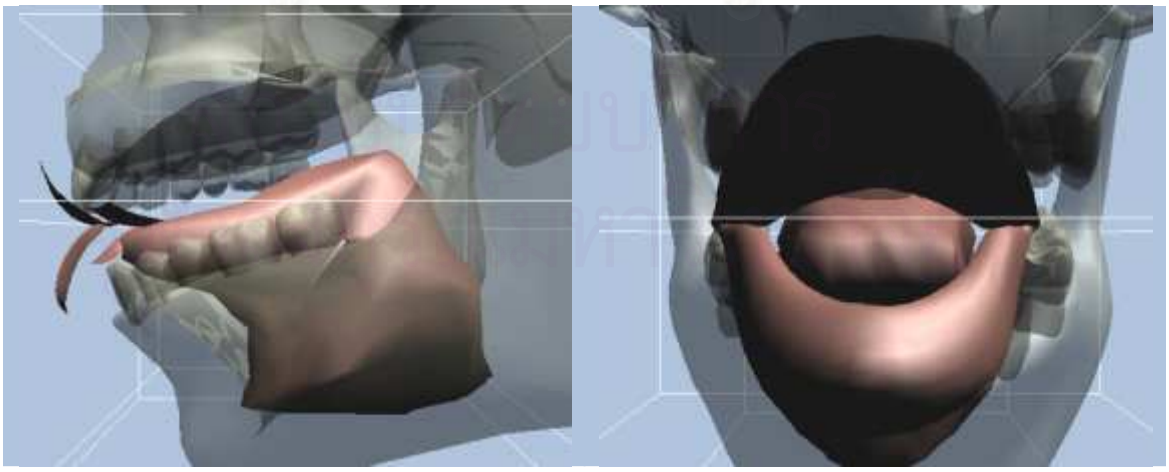
“เอ”



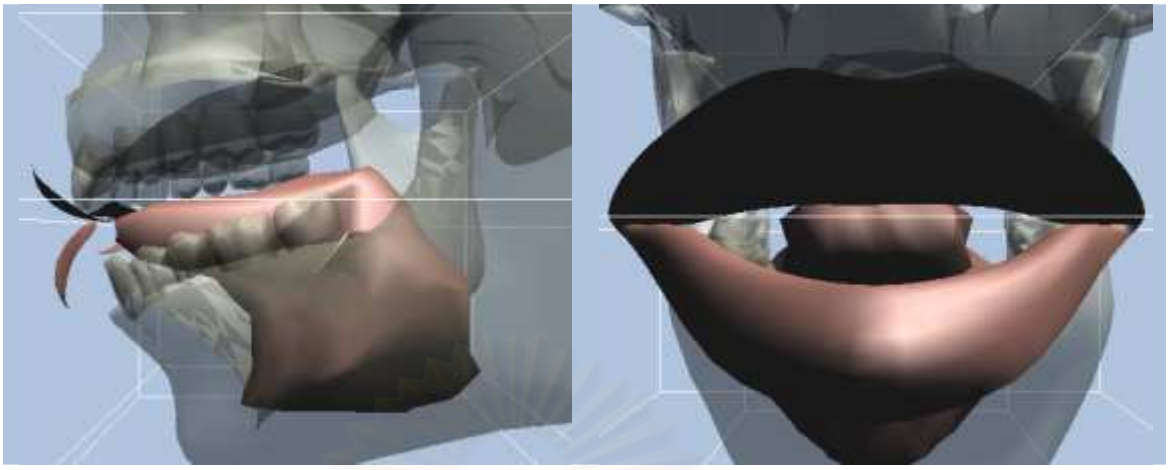
“แอ”



“โอ”

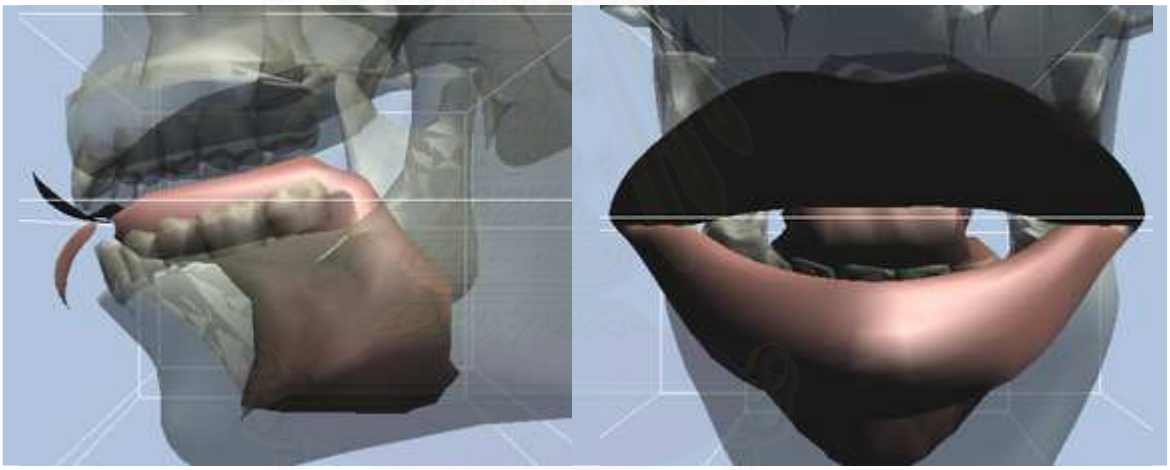


“ออ”

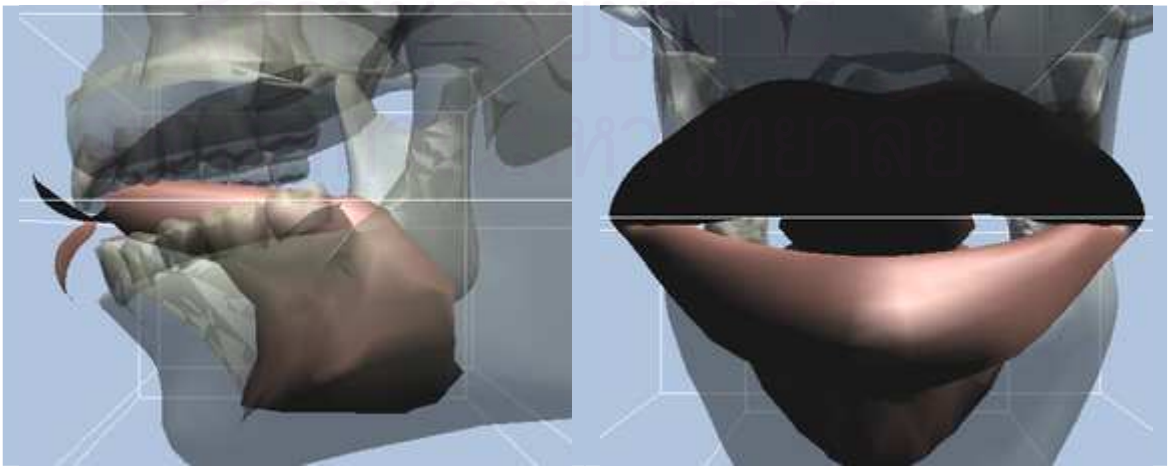


“เออ”

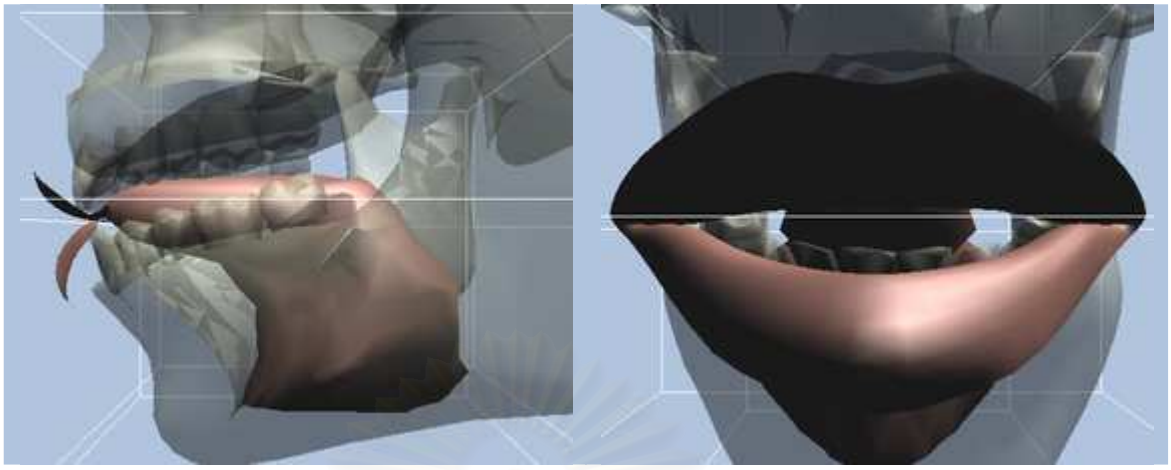
หน่วยเสียงพยัญชนะ



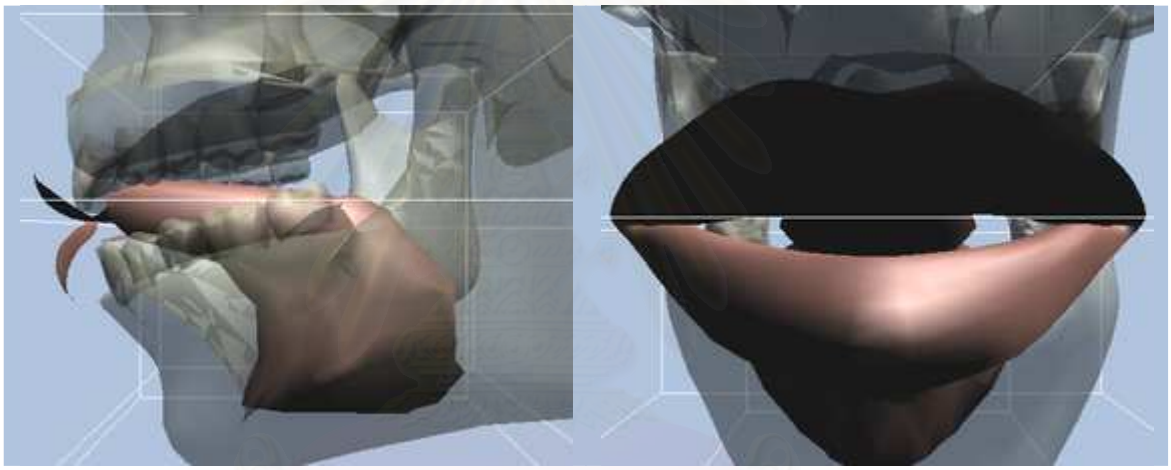
“ก ข ศ ค ศ ฆ ง”



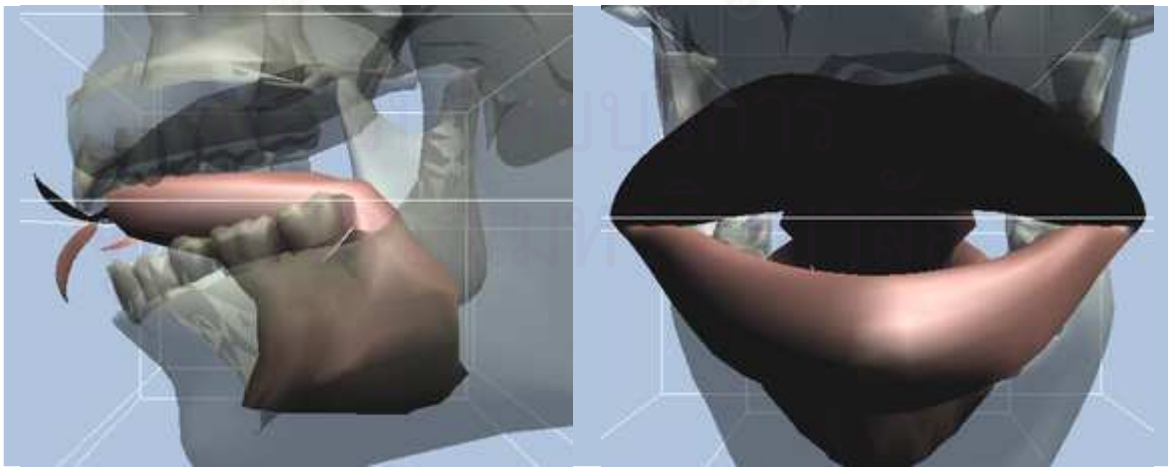
“จ ฉ ช ฌ”



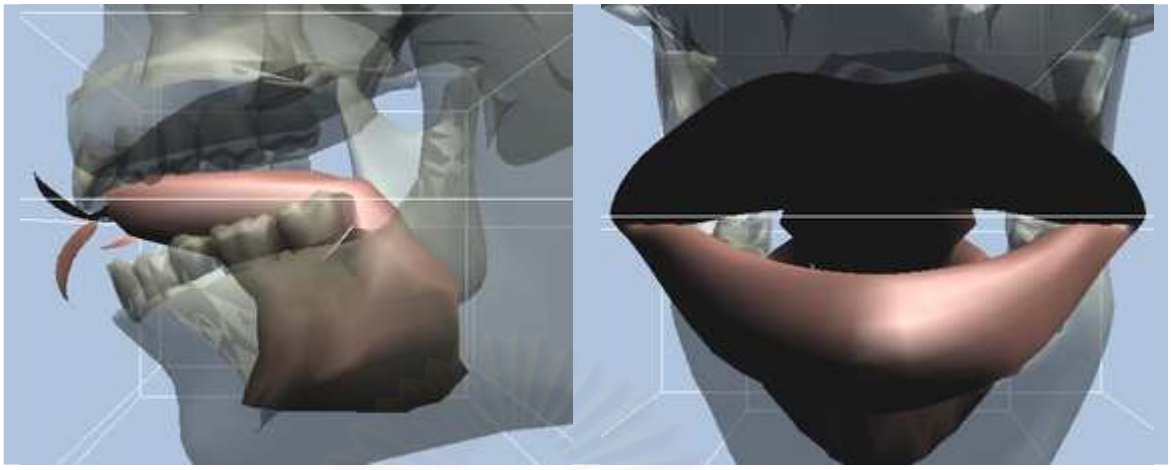
“ช ศ ษ ส”



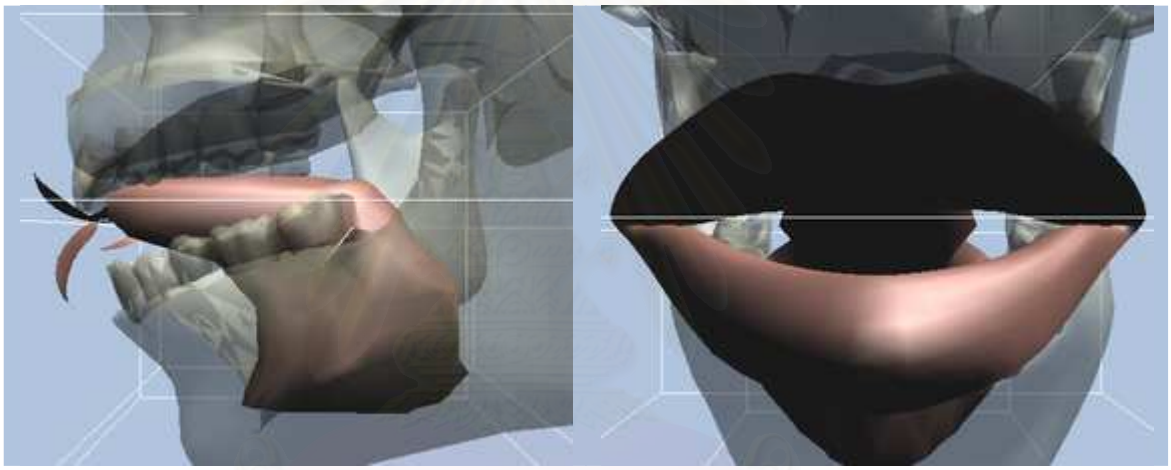
“ญ ย”



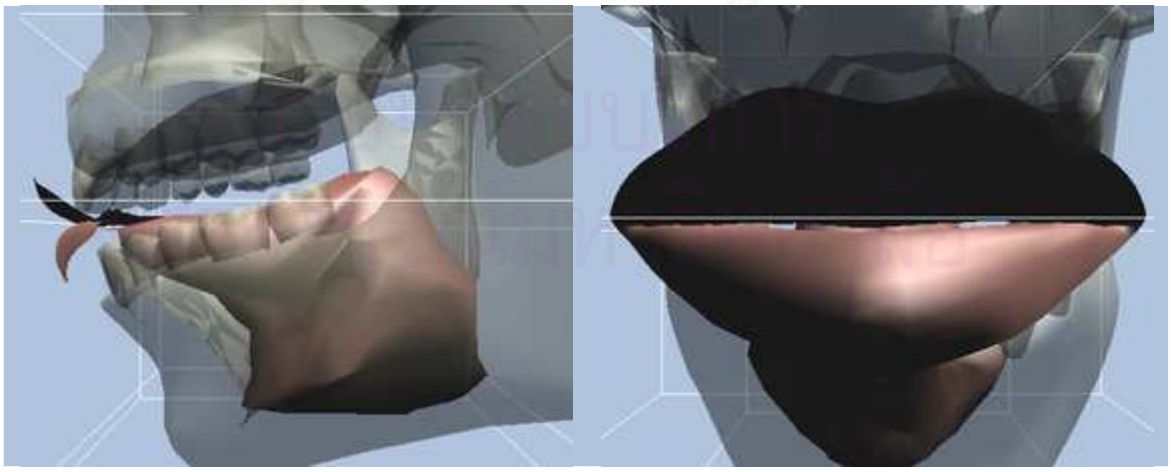
“ด ฎ”



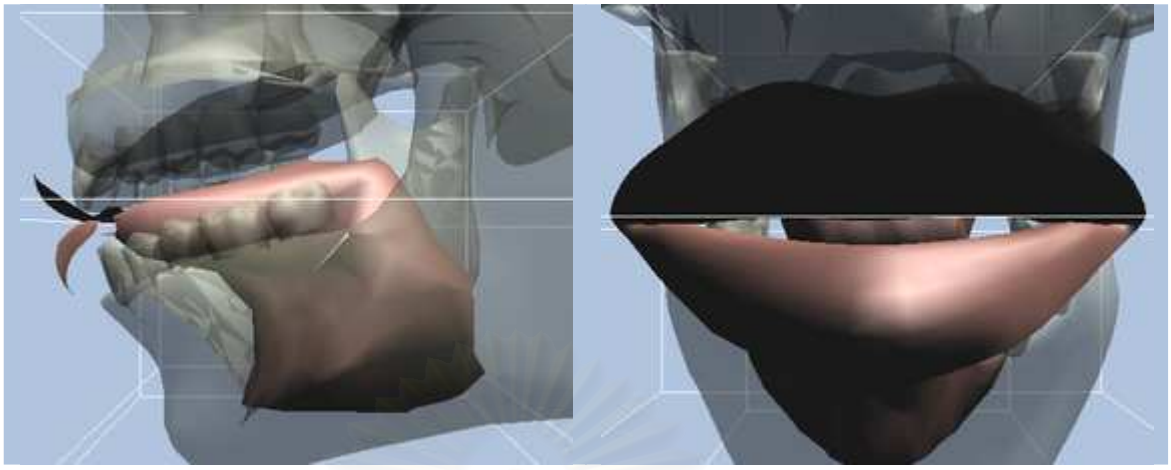
“ต ฏ ฐ ฑ ฒ ฌ ถ ท ธ”



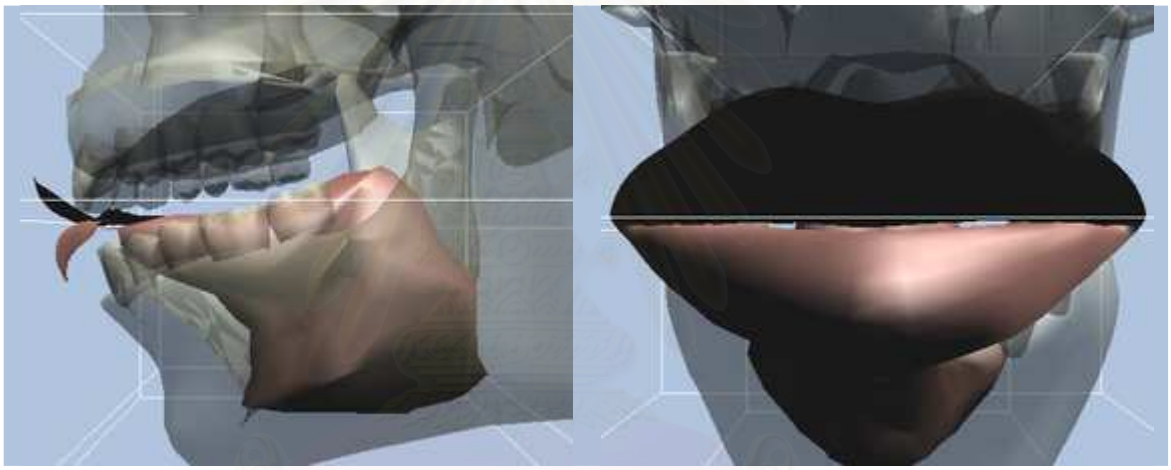
“ณ น”



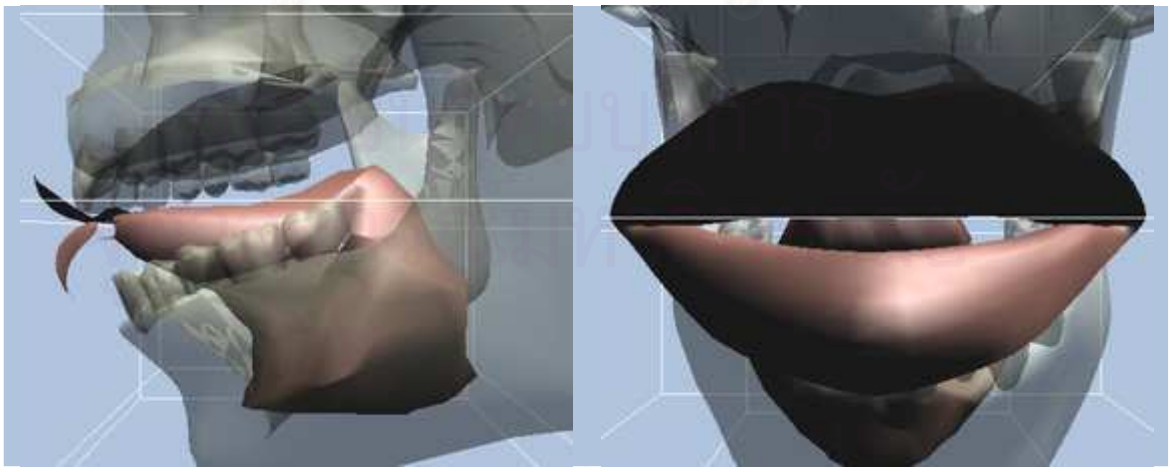
“บ ป พ ภ ผ”



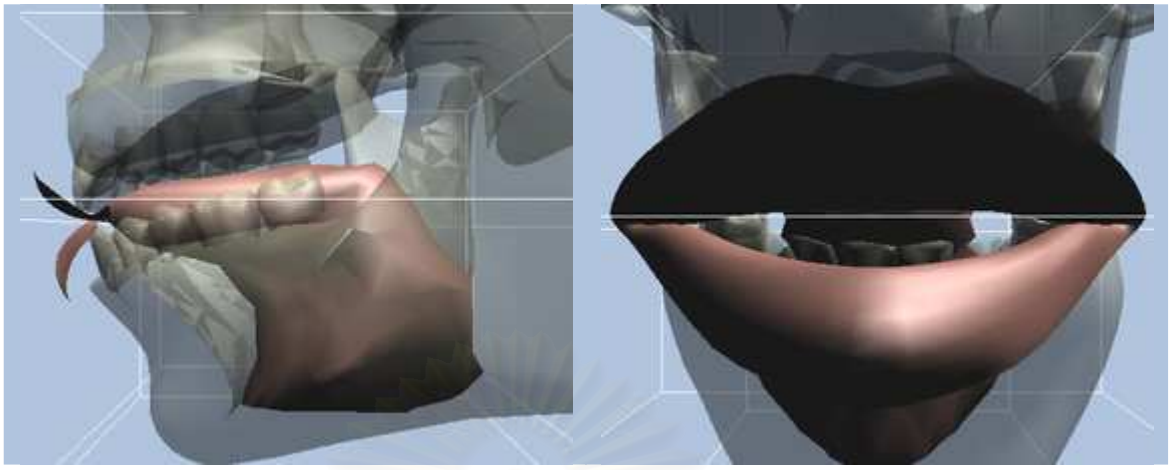
“ฟ ฝ”



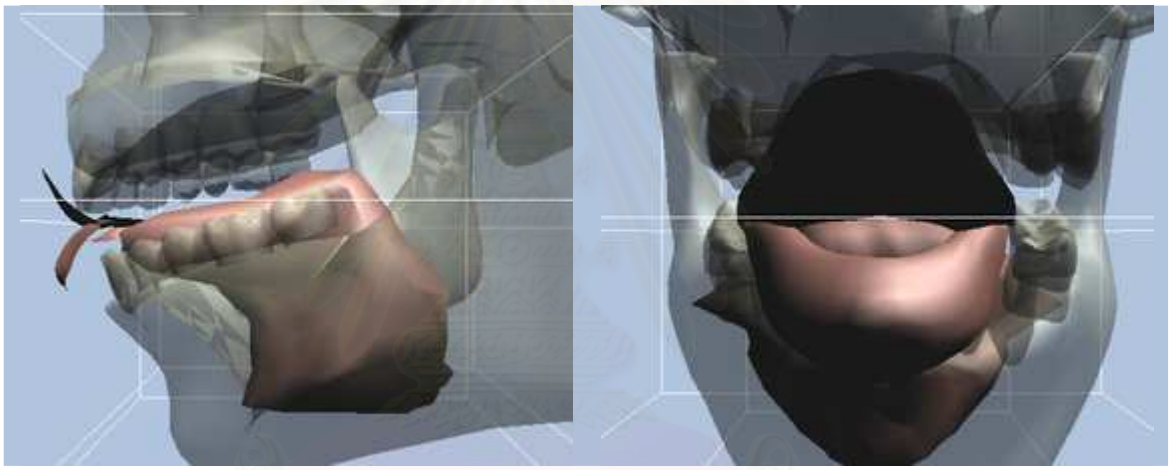
“ม”



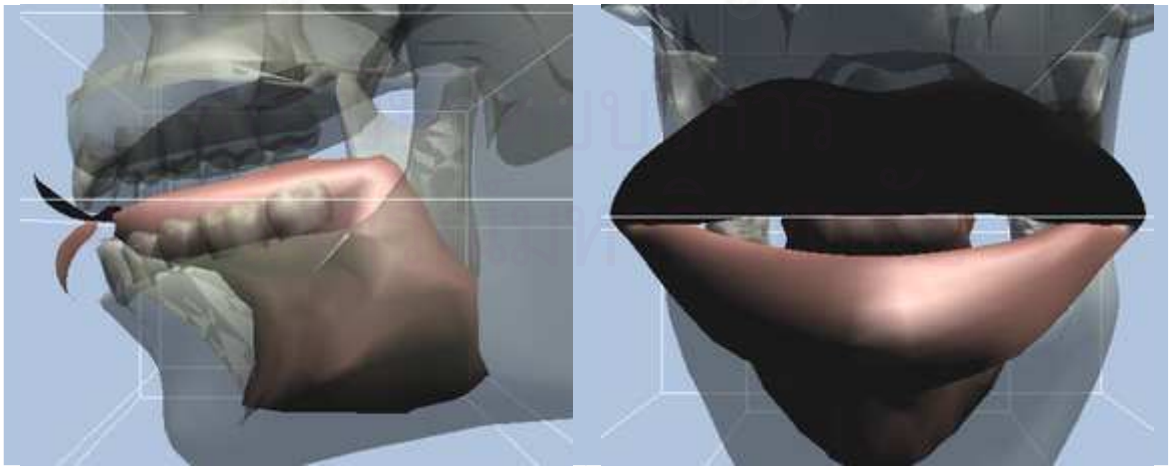
“ร”



“ล ฬ”



“ว”



“ห ฮ”

ประวัติผู้เขียนวิทยานิพนธ์

นางสาว ศศิวิมล ชาญวานิชบริการ เกิดเมื่อวันที่ 6 เมษายน พ.ศ.2526 สำเร็จการศึกษาระดับปริญญาบัณฑิตในสาขาวิศวกรรมคอมพิวเตอร์จากคณะวิศวกรรมศาสตร์ มหาวิทยาลัยสงขลานครินทร์ ในปีการศึกษา 2548

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย