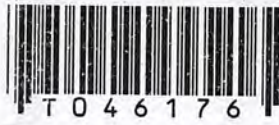


การสังเคราะห์เสียงภาษาไทย
THAI SPEECH SYNTHESIS



โดย

นายกัณนัท ริยะใจ
นายเมธี แต่งงาม



เลขที่.....
เลขทะเบียน... 46176
วัน, เดือน, ปี 20 ส.ค. 2546

b.....
i.....

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต
ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ปีการศึกษา 2544

ปริญญานิพนธ์ปีการศึกษา 2544

ภาควิชา วิศวกรรมคอมพิวเตอร์

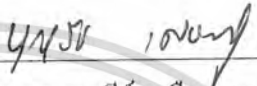
คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง การสังเคราะห์เสียงภาษาไทย

THAI SPEECH SYNTHESIS

ผู้จัดทำ

1. นายภักนันท์ ธิยะใจ รหัสประจำตัว 41014326
2. นายเมธี แด่งงาม รหัสประจำตัว 41014342


(ผศ. ดร. นุชธีร์ เกียรติราษฎร์) อาจารย์ที่ปรึกษา



การสังเคราะห์เสียงภาษาไทย

- | | | |
|---|---------|----------|
| 1. นายภคนันท์ | ธิชะใจ | 41014326 |
| 2. นายเมธี | แต่งงาม | 41014342 |
| ผศ. ดร. บุญธีร์ เครือตราชู อาจารย์ที่ปรึกษา | | |

บทคัดย่อ

ในปัจจุบันการศึกษาวิจัยทางด้านเสียง เพื่อนำมาประยุกต์ใช้ร่วมกับระบบคอมพิวเตอร์กำลังได้รับความนิยมเป็นอย่างมากในวงวิชาการ การสังเคราะห์เสียงภาษาไทยด้วยคอมพิวเตอร์ก็เป็นหัวข้อหนึ่งที่กำลังได้รับความสนใจ ซึ่งการสังเคราะห์เสียงภาษาไทยด้วยคอมพิวเตอร์จะประกอบด้วยขั้นตอนที่สำคัญ คือ การตัดคำออกจากประโยคภาษาไทย, การสร้างสัญลักษณ์แทนเสียง, การสร้างเสียง และการพัฒนาคุณภาพเสียงให้ได้ใกล้เคียงกับธรรมชาติมากที่สุด

ปริญญานิพนธ์ฉบับนี้ ตัดคำโดยใช้วิธี WordBucket ซึ่งเป็นการตัดคำที่ยาวที่สุดที่จะเป็นไปได้ โดยใช้การเปรียบเทียบกับพจนานุกรมคำศัพท์ การสร้างสัญลักษณ์แทนเสียง ใช้การสร้างสัญลักษณ์แทนเสียง โดยวิธี MPT-Networks ซึ่งจะได้สัญลักษณ์แทนหน่วยเสียงพื้นฐานในระดับที่ 3 แล้วจึงนำแฟ้มเสียงที่แทนด้วยสัญลักษณ์แทนเสียงที่ได้มารวมเป็นเสียงภาษาไทย 1 เสียง ในส่วนของการพัฒนาคุณภาพเสียง ได้มีการทดลองบันทึกและตัดเสียงในรูปแบบต่าง ๆ แล้วนำมาสังเคราะห์เพื่อเปรียบเทียบหาวิธีการตัดเสียงให้สังเคราะห์ออกมาได้ใกล้เคียงกับการพูดของมนุษย์มากที่สุด

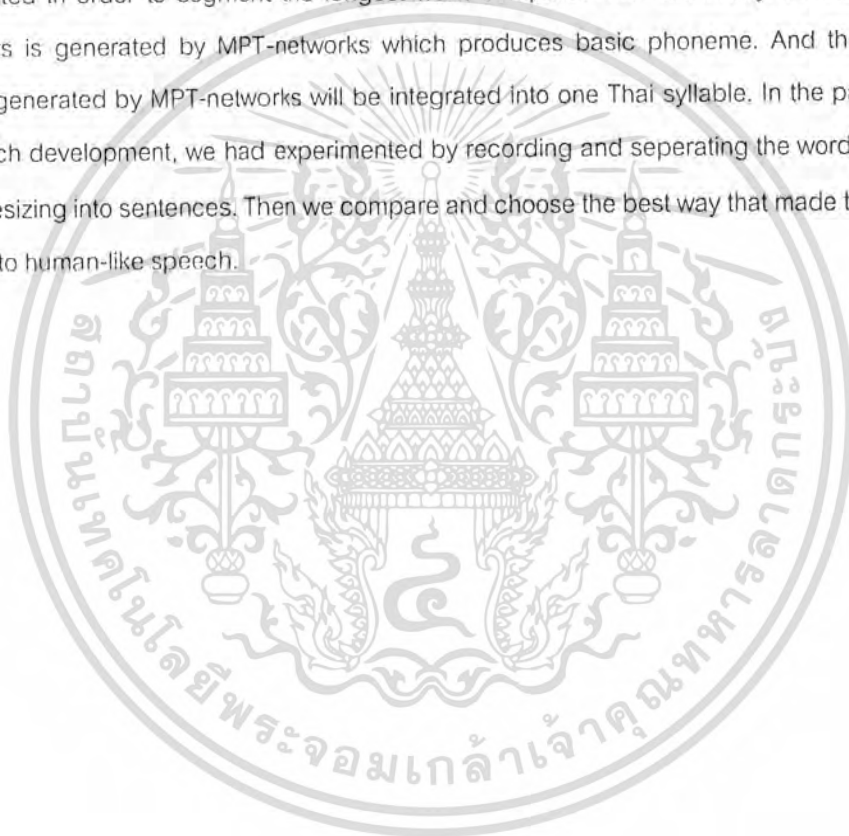
Thai Speech Synthesis

Pukkanan	Thiyajai
Methee	Taengngam
Boonthee	Kreatrachoo Adviser

ABSTRACT

At present, the computer system application of speech research is the most interesting topic in technology sphere. Also Thai Speech Synthesis is one of the interesting subjects. Thai Speech Synthesis consists of four major parts : Word Segmentation in Thai sentence, Phonetic symbols generator, Voice generator and Human-like speech development.

This research presents the method of Word Segmentation "Wordbucket Algorithm" will be implemented in order to segment the longest word compared with dictionary. Phonetics symbols generators is generated by MPT-networks which produces basic phoneme. And then phonetics symbols generated by MPT-networks will be integrated into one Thai syllable. In the part of human-like speech development, we had experimented by recording and seperating the word in many way for synthesizing into sentences. Then we compare and choose the best way that made the sentences close up to human-like speech.



กิตติกรรมประกาศ

ปริญญานิพนธ์ฉบับนี้คงไม่อาจเสร็จได้ด้วยดี หากไม่ได้รับความช่วยเหลือ และร่วมมือจากหลาย ๆ ฝ่ายด้วยกัน บุคคลแรกที่ต้องกล่าวถึงเพราะเป็นส่วนสำคัญที่ทำให้วิทยานิพนธ์นี้เสร็จลงได้ก็คือ อาจารย์ บุญธีร์ เครือตราฐ อาจารย์ที่ปรึกษาปริญญานิพนธ์ ที่ให้ความเอาใจใส่ แนะนำ และช่วยเหลือเสมอมา ซึ่งต้องขอขอบพระคุณเป็นอย่างมาก

และต้องขอขอบพระคุณบุคคลสำคัญที่สุดที่ทำให้ข้าพเจ้ามีวันนี้ ก็คือ บิดา มารดา อันเป็นที่เคารพรักยิ่ง ซึ่งได้เลี้ยงดูผู้เขียนมาเป็นอย่างดี พร้อมทั้งให้โอกาสในการศึกษาอย่างเต็มที่ และยังให้กำลังใจ เอาใจใส่เสมอมา ในทุก ๆ ด้านอันหาที่เปรียบมิได้ ข้าพเจ้าขอระลึกในพระคุณอันสุดประมาณ และขอกราบขอบพระคุณมา ณ ที่นี้

ภักนันท์ ธิยะใจ

เมธี แต่งงาม



สารบัญ

	หน้า
บทคัดย่อภาษาไทย	I
บทคัดย่อภาษาอังกฤษ	II
กิตติกรรมประกาศ	III
สารบัญ	IV
สารบัญภาพ	VI
สารบัญตาราง	VIII
บทที่ 1 บทนำ	1
1.1 ความสำคัญและที่มา	1
1.2 วัตถุประสงค์ของงานวิจัย	1
1.3 ขอบเขตของงานวิจัย	1
1.4 วิธีการดำเนินงาน	2
บทที่ 2 ระบบเสียงภาษาไทย	3
2.1 บทนำ	3
2.2 ความหมายของเสียง	3
2.3 การศึกษาเกี่ยวกับเสียง	3
2.4 อวัยวะที่ใช้ในการออกเสียงพูด	4
2.5 เสียงในภาษาไทย	5
2.6 เสียงอักษรของไทย	5
2.7 สรุป	21
บทที่ 3 การกำหนดหน่วยเสียงเพื่อการสังเคราะห์เสียง	22
3.1 บทนำ	22
3.2 คลื่นเสียง	22
3.3 หน่วยเสียงพื้นฐาน	25
3.4 การจัดเก็บหน่วยเสียง	28
บทที่ 4 การบันทึกเสียงและการแบ่งเสียงออกเป็นไบโโฟน	34
4.1 บทนำ	34
4.2 รูปแบบของไฟล์	34
4.3 การบันทึกเสียง	34
4.4 การแบ่งเสียงออกเป็นไบโโฟน	34
4.5 สรุปหลักในการบันทึกเสียง	35
4.6 สรุปหลักในการแบ่งไฟล์เสียง	35
บทที่ 5 การสังเคราะห์เสียง	37

5.1 บทนำ	37
5.2 วิธีการผสมคำในการทดลอง	37
5.3 การเปรียบเทียบระหว่างเสียงที่ได้จากการสังเคราะห์กับเสียงธรรมชาติ	41
5.4 การผสมคำเป็นประโยค	42
5.5 สรุปผลที่ได้จากการทดลอง	44
บทที่ 6 การปรับปรุงคุณภาพเสียง	45
6.1 บทนำ	45
6.2 วิธีการบันทึกเสียง	45
6.3 การปรับปรุงการแบ่งไบโโฟน	47
6.4 การแทรก Short Pause	49
6.5 สรุป	49
บทที่ 7 การตัดคำ	50
7.1 บทนำ	50
7.2 พจนานุกรมคำศัพท์	50
7.3 การตัดคำโดยใช้ WordBucket	51
7.4 สรุป	53
บทที่ 8 การแปลงหน่วยเสียงเป็นสัญลักษณ์แทนเสียง โดยวิธี MPT-networks	54
8.1 บทนำ	54
8.2 การวิเคราะห์ห้วงคำประกอบโครงสร้างพยางค์ เพื่อหารูปแบบกลาง	54
8.3 การวิเคราะห์รูปแบบกลางเพื่อหาเสียงเป้าหมาย	61
8.4 การวิเคราะห์สมการแทนเสียงเพื่อกำหนดสัญลักษณ์แทนเสียง	63
8.5 ตัวอย่างการทำงานของระบบการแปลงหน่วยคำเป็นสัญลักษณ์แทนเสียง	65
8.6 สรุป	68
บทที่ 9 การออกแบบและการสร้างโปรแกรม	69
9.1 บทนำ	69
9.2 การออกแบบโปรแกรม	69
9.3 การสร้างโปรแกรม	75
9.4 สรุป	84
บทที่ 10 บทวิจารณ์และสรุป	85
10.1 บทวิจารณ์	85
10.2 สรุป	85
10.3 ข้อเสนอแนะสำหรับการพัฒนาในอนาคต	86
บรรณานุกรม	87

สารบัญตาราง

	หน้า
ตารางที่ 2.1 แสดงพยัญชนะในภาษาไทยแบ่งออกเป็นวรรคตามฐานที่เกิดของอักษร	7
ตารางที่ 2.2 แสดงฐานที่เกิดของเสียงพยัญชนะในสมาวรรค	7
ตารางที่ 2.3 แสดงสัญลักษณ์แทนเสียงของหน่วยเสียงพยัญชนะต้น เดี่ยว	9
ตารางที่ 2.4 ลักษณะของเสียงและที่เกิดเสียงของหน่วยเสียงพยัญชนะในภาษาไทย	13
ตารางที่ 2.5 แสดงเสียงพยัญชนะไทยที่สามารถควบกับเสียงอื่นได้	14
ตารางที่ 2.6 แสดงความเกี่ยวเนื่องของเสียงแต่ละเสียงในการออกเสียงสระประสม	16
ตารางที่ 2.7 แสดงสัญลักษณ์แทนเสียงของหน่วยเสียงพยัญชนะต้น เดี่ยว 21 หน่วยเสียง	18
ตารางที่ 2.8 แสดงสัญลักษณ์แทนเสียงของหน่วยเสียงพยัญชนะต้น ควบกล้ำ 12 หน่วยเสียง	19
ตารางที่ 2.9 แสดงสัญลักษณ์แทนเสียงของหน่วยเสียงพยัญชนะต้น ควบกล้ำ ที่มาจากภาษาต่างประเทศ 5 หน่วยเสียง	19
ตารางที่ 2.10 แสดงสัญลักษณ์แทนเสียงของพยัญชนะ ตัวสะกด 8 หน่วยเสียง	19
ตารางที่ 2.11 แสดงสัญลักษณ์แทนเสียงของหน่วยเสียงพยัญชนะ ตัวสะกด ที่มาจากต่างประเทศ 4 หน่วยเสียง	20
ตารางที่ 2.12 แสดงสัญลักษณ์แทนเสียงของหน่วยเสียงสระเดี่ยวเสียงสั้น-ยาว 18 หน่วยเสียง	20
ตารางที่ 2.13 แสดงสัญลักษณ์แทนเสียงของหน่วยเสียงสระผสม 6 หน่วยเสียง	21
ตารางที่ 2.14 แสดงสัญลักษณ์แทนเสียงของหน่วยเสียงวรรณยุกต์ 5 หน่วยเสียง	21
ตารางที่ 3.1 เสียงพยัญชนะต้นเดี่ยวและพยัญชนะควบกล้ำผสมสระเสียงสั้น รวม 444 แฟ้มเสียง	28
ตารางที่ 3.2 สระเสียงสั้น (ไม่มีตัวสะกด) รวม 48 แฟ้มเสียง	30
ตารางที่ 3.3 สระเสียงสั้น (ตัวสะกดเสียงกัก) รวม 72 แฟ้มเสียง	30
ตารางที่ 3.4 สระเสียงสั้น (ตัวสะกดเสียงไม่กัก) รวม 300 แฟ้มเสียง	30
ตารางที่ 3.5 สระเสียงยาว (ไม่มีตัวสะกด) รวม 60 แฟ้มเสียง	31
ตารางที่ 3.6 สระเสียงยาว (ตัวสะกดเสียงกัก) รวม 108 แฟ้มเสียง	32
ตารางที่ 3.7 สระเสียงยาว (ตัวสะกดเสียงไม่กัก) รวม 300 แฟ้มเสียง	32
ตารางที่ 5.1 เปรียบเทียบรูปคลื่นเสียงเชิงเวลา	42
ตารางที่ 7.1 แสดง Field ต่าง ๆ ในพจนานุกรมคำศัพท์	52
ตารางที่ 7.2 แสดงตัวอย่างคำที่เก็บในพจนานุกรมคำศัพท์	53
ตารางที่ 8.1 แสดงการจัดกลุ่มของหน่วยเสียงพยัญชนะต้น	63
ตารางที่ 8.2 แสดงการจัดกลุ่มของหน่วยเสียงพยัญชนะท้าย	64
ตารางที่ 8.3 แสดงการจัดกลุ่มของหน่วยเสียงสระ	64

ตารางที่ 8.4 แสดงการจัดกลุ่มของหน่วยวรรณยุกต์	64
ตารางที่ 8.5 แสดงโครงสร้างฐานข้อมูลที่ใช้เก็บรายละเอียดเกี่ยวกับพยัญชนะต้น	66
ตารางที่ 8.6 แสดงโครงสร้างฐานข้อมูลที่ใช้เก็บรายละเอียดเกี่ยวกับพยัญชนะท้าย	66
ตารางที่ 8.7 แสดงโครงสร้างฐานข้อมูลที่ใช้เก็บรายละเอียดเกี่ยวกับสระ	67
ตารางที่ 8.8 แสดงผลกระบวนการตัดคำโดยวิธี WordBucket	68
ตารางที่ 8.9 แสดงการแปลงพยางค์ตามสมการแทนเสียงเป็นสัญลักษณ์แทนเสียง	70



สารบัญภาพ

	หน้า
รูปที่ 2.1 แสดงความสัมพันธ์ระหว่างส่วนย่อยทั้งสามของวิชาสัตสาศตร์	4
รูปที่ 4.1 การแบ่งเสียงเป็นไบโโฟน	35
รูปที่ 4.2 การแบ่งไฟล์เสียงที่ถูกต้อง	36
รูปที่ 5.1 รูปเสียงคำว่า กิพา	39
รูปที่ 5.2 รูปเสียงคำว่า อิกกา	39
รูปที่ 5.3 รูปเสียงคำว่า กิพา ที่เกิดจากการผสมคำ	39
รูปที่ 5.4 รูปเสียงคำว่า ชำนาญ	40
รูปที่ 5.5 รูปเสียงคำว่า คำเนิน	40
รูปที่ 5.6 รูปเสียงคำว่า ชำนาญ ที่เกิดจากการผสมคำ	40
รูปที่ 5.7 เสียงคำว่า "ควาย" จากการพูดจริง	41
รูปที่ 5.8 เสียงคำว่า "ควาย" จากการสังเคราะห์	41
รูปที่ 5.9 รูปเสียงประโยค "ฉันชอบดื่มนม" ที่พูดปกติ	43
รูปที่ 5.10 รูปเสียงประโยค "ฉันชอบดื่มนม" ที่ได้จากการนำคำมาต่อกัน โดยแทรก Short Pause ไว้ 0.03	43
รูปที่ 5.11 รูปเสียงประโยค "ฉันชอบดื่มนม" ที่ได้จากการนำคำมาต่อกัน โดยแทรก Short Pause ไว้ 0.05 วินาที	43
รูปที่ 6.1 แสดงการบันทึกเสียงเป็นชุดๆ	46
รูปที่ 6.2 แสดงการเลือกเสียงที่ดีที่สุดจากการบันทึกเป็นชุดๆ	46
รูปที่ 6.3 คำว่า "ก้าง" ซึ่งตัดส่วนหน้ายาวไป	47
รูปที่ 6.4 คำว่า "ก้าง" ซึ่งตัดส่วนหน้าสั้น	47
รูปที่ 6.5 คำว่า "ฟัน" ซึ่งยังไม่ได้เพิ่ม แอมพลิจูด ของส่วนที่กักลม	48
รูปที่ 6.6 คำว่า "ฟัน" ที่ทำการเพิ่ม แอมพลิจูดของส่วนที่กักลม	48
รูปที่ 6.7 คำว่า "ฟาม" ซึ่งยังไม่ได้ปรับขนาดเสียง	49
รูปที่ 6.8 คำว่า "ฟาม" ซึ่งได้ตัดส่วนหนึ่งของคลื่นออกเพื่อปรับขนาดของไบโโฟน	49
รูปที่ 6.9 หางเสียงคำว่า "นาม" ซึ่งลากหางเสียงยาวไป	50
รูปที่ 6.10 หางเสียง "นาม" ซึ่งพูดกระชับ	50
รูปที่ 6.11 ประโยค "หลบหน่อยพระเอกมา" โดยไม่มีการแทรก Short Pause	51
รูปที่ 6.12 ประโยค "หลบหน่อยพระเอกมา" โดยมีการแทรก Short Pause	51
รูปที่ 7.1 โครงสร้างของประโยคภาษาไทย	52
รูปที่ 7.2 แสดงวิธีการตัดคำโดยใช้ WordBucket	54
รูปที่ 8.1 แสดงโรงฆ่าการเกิดพยางค์ตามลำดับการเขียนพยัญชนะและสระ	58
รูปที่ 8.2 แสดงขั้นตอนการทำงานของอัลกอริทึม MPT-networks	68

รูปที่ 9.1 โครงสร้างการทำงานของโปรแกรมสังเคราะห์เสียงภาษาไทยด้วยคอมพิวเตอร์	71
รูปที่ 9.2 Diagram การออกแบบระดับที่ 1	72
รูปที่ 9.3 Diagram การออกแบบระดับที่ 2	73
รูปที่ 9.4 Diagram การออกแบบระดับที่ 3	74
รูปที่ 9.5 แสดงโครงสร้างของหน่วยเสียงระดับต่าง ๆ	76
รูปที่ 9.6 แสดง Flowchart ของส่วนรับข้อความจากเพิ่มข้อความภาษาไทย	77
รูปที่ 9.7 แสดง Flowchart ของส่วนรับข้อความจากแป้นพิมพ์	78
รูปที่ 9.8 แสดง Flowchart ของส่วนตัดคำ	79
รูปที่ 9.9 แสดง Flowchart ของส่วนการแยกคำออกเป็นพยางค์	82
รูปที่ 9.10 แสดง Flowchart ของส่วนสร้างสัญลักษณ์แทนเสียง	83
รูปที่ 9.11 แสดง Flowchart ของส่วนประกอบเพิ่มเสียง	85



บทที่ 1

บทนำ

1.1 ความสำคัญและที่มา

ในปัจจุบัน การพัฒนาเทคโนโลยีทางด้านคอมพิวเตอร์เป็นไปอย่างก้าวหน้า และกว้างขวาง มีการนำเอาเทคโนโลยีทางด้านคอมพิวเตอร์มาประยุกต์ใช้กับชีวิตประจำวันอย่างมากมาย ทั้งด้านการอำนวยความสะดวกในการดำรงชีวิต การอำนวยความสะดวกในการปฏิบัติงาน

นอกจากนี้คอมพิวเตอร์ยังสามารถเข้ามาช่วยพัฒนาเทคโนโลยีเพื่อคนพิการได้อีกด้วย ในปัจจุบัน คนพิการได้รับการยอมรับมากขึ้นในสังคม ไม่ว่าจะเป็นด้านการศึกษาร่วมกับคนที่มีร่างกายปกติ คนพิการก็สามารถทำได้ หรือจะเป็นด้านการศึกษา ก็มีกีฬาสำหรับคนพิการ แสดงให้เห็นว่าคนพิการนั้นมีศักยภาพเพียงพอในการทำงานต่าง ๆ หากพวกเขาได้รับการสนับสนุนจากสังคม กระผมจึงมีแนวคิดที่จะพัฒนาโปรแกรมสำหรับผู้พิการทางสายตาและทางเสียงขึ้นเพื่อให้พวกเขามีโอกาสในการใช้ชีวิตในสังคมได้ดีขึ้น และเป็นโอกาสที่จะช่วยให้คนเหล่านั้นได้รับรู้ และเรียนรู้ความรู้และเทคโนโลยีใหม่ ๆ เพื่อให้เขาได้นำไปพัฒนาตนเองให้เป็นบุคคลที่มีคุณภาพในสังคม

โปรแกรมสังเคราะห์เสียงภาษาไทยด้วยคอมพิวเตอร์ เป็นโปรแกรมสร้างเสียงภาษาไทยจากเพิ่มข้อความภาษาไทย เพียงแค่มีเพิ่มข้อความภาษาไทยก็จะสามารถรับทราบข้อความในเพิ่มในด้วยการฟังข้อความผ่านทางคอมพิวเตอร์ จากโปรแกรมสังเคราะห์เสียงนี้ได้

1.2 วัตถุประสงค์ของงานวิจัย

สร้างโปรแกรมสังเคราะห์เสียงภาษาไทยด้วยคอมพิวเตอร์ เพื่อทำการแปลงจากเพิ่มข้อความภาษาไทยให้เป็นเสียงพูดภาษาไทย โดยให้มีความสามารถรับได้ทั้งเพิ่มข้อความและจากทางแป้นพิมพ์ มีความสามารถในการตัดคำได้ถูกต้องตามหลักภาษาไทย และสามารถสังเคราะห์เสียงภาษาไทยออกมาได้ใกล้เคียงกับธรรมชาติมากที่สุด

1.3 ขอบเขตของงานวิจัย

การออกแบบและสร้างโปรแกรมสังเคราะห์เสียงจะสนใจงานในภาพรวมคือ สามารถตัดคำและสร้างเสียงให้ใกล้เคียงกับเสียงพูดธรรมชาติที่สุด สามารถนำไปใช้งานได้จริง วิธีการสร้างเสียงภาษาไทยจะใช้วิธีการสังเคราะห์จากหน่วยเสียงที่เล็กกว่าหน่วยพยางค์ เพราะเป็นวิธีที่ประหยัดเนื้อที่การจัดเก็บ และให้เสียงใกล้เคียงกับธรรมชาติ การตัดคำภาษานั้นจะตัดได้เฉพาะที่มีอยู่ในพจนานุกรมคำศัพท์ของโปรแกรมเท่านั้น แต่สามารถเพิ่มคำศัพท์ที่ไม่มีในพจนานุกรมคำศัพท์ได้

1.4 วิธีการดำเนินงาน

ปริญญานิพนธ์ฉบับนี้จะเริ่มจากการศึกษาทฤษฎีพื้นฐานต่าง ๆ ที่เกี่ยวข้อง ซึ่งมีเรื่องหลัก ๆ คือ การศึกษาระบบเสียงภาษาไทย ,การกำหนดหน่วยเสียงและการวิเคราะห์เสียง ซึ่งอยู่ในบทที่ 2 และ 3 ตามลำดับ จากนั้นนำความรู้ที่ได้ศึกษามาทดลองสังเคราะห์เสียง โดยวิธีบันทึกเสียงและการแบ่งเสียงออกเป็น โฟเน็มจะอยู่ในในบทที่ 4 และบทที่ 5 จะเป็นการทดลองสังเคราะห์เสียงและผลการทดลอง บทที่ 6 จะเป็นการปรับปรุงคุณภาพเสียงที่ได้จากการสังเคราะห์ โดยจะอธิบายวิธีการปรับปรุงคุณภาพเสียง บทที่ 7, บทที่ 8 และบทที่ 9 จะเป็นในส่วนของการเตรียมการเพื่อเขียนโปรแกรมสังเคราะห์เสียง และบทที่ 10 จะเป็นบทวิจารณ์และสรุป



บทที่ 2

ระบบเสียงภาษาไทย

2.1 บทนำ

ระบบเสียงในภาษาไทยเป็นการศึกษาที่เกี่ยวข้องกับการสร้างเสียงภาษาไทย จะเริ่มจากการศึกษาอวัยวะที่ใช้ในการออกเสียงพูด การศึกษาเสียงในภาษาไทยและการศึกษาเสียงของอักษรไทย ซึ่งเสียงของอักษรไทยจะประกอบไปด้วย เสียงพยัญชนะ เสียงสระ และเสียงวรรณยุกต์ และสุดท้ายจะกล่าวถึงการกำหนดสัญลักษณ์แทนเสียงของแต่ละอักษรไทย

2.2 ความหมายของเสียง

พจนานุกรมฉบับราชบัณฑิตยสถานให้ความหมายของเสียงไว้ว่า “สิ่งที่ได้ยินด้วยหู คำที่เปล่งออกมา เช่น เสียงลือ ว่า คำเล่าลือ ความเหิน ความนิยม คะแนนเสียง” และได้ให้ความหมายของคำว่าเสียงเมื่อนำไปประสมกับคำอื่นบางคำเอาไว้ว่า “เสียงเขียว น. แสดงอาการ โกรธจัด เสียงแข็ง น. คำพูดที่กล่าวออกมาโดยไม่ยั้งเกรง เสียงแตก น. เสียงพูดในตอนแตกเนื้อหนุ่ม เสียงทอง น. เสียงเพราะทำให้คำตัวในการร้องเพลงสูงขึ้น เสียงหลง น. เสียงที่แผดดังผิดปกติ หรือผิดระดับเสียงคนตรี เสียงอ่อน น. คำพูดที่เพลาความหมายแข็งแรงลง” จากความหมายดังกล่าวแล้วจะเห็นว่าความหมายของเสียงมีอยู่ 2 ส่วนสำคัญคือ ส่วนแรกหมายถึงสิ่งที่หูอาจหรือสามารถได้ยิน จะเป็นเสียงอะไรก็ได้ เช่น เสียงฝนตก ฟังร้อง เสียงวัตถุกระทบกัน เป็นต้น ส่วนที่สอง หมายถึงเสียงพูดของมนุษย์ ส่วนหลังนี้มีความสำคัญที่เราจำเป็นต้องศึกษา เพราะเป็นเสียงในภาษาที่ใช้สื่อสารในชีวิตประจำวันทุกด้าน การศึกษาเสียงพูดคือการศึกษารายละเอียดของมนุษย์

2.3 การศึกษาเกี่ยวกับเสียง

กาญจนา นาคสกุล¹ ให้ความเห็นว่า “เสียงพูดมีลักษณะที่จะอธิบายได้ด้วยหลักเกณฑ์ทางวิทยาศาสตร์ซึ่งเป็นสากล แม้ว่าภาษาหนึ่ง ๆ จะมีเสียงแตกต่างกันออกไปมากบ้างน้อยบ้าง แต่ละเสียงก็สามารถที่จะนำมาพิจารณาและอธิบายให้รู้ลักษณะการออกเสียง และตำแหน่งการเกิดของเสียงได้ คำอธิบายจะทำให้รู้ และเข้าใจลักษณะของเสียงทุกเสียง ทำให้เห็นความแตกต่าง และคล้ายคลึงกันของเสียงทุกเสียงได้ ซึ่งจะเป็นประโยชน์อย่างมากต่อการเรียนรู้”

เรื่องการศึกษาเกี่ยวกับเสียงนี้ ประสิทธิ์ กาพย์กลอน² ได้สรุปว่ามีวิชาที่ศึกษาเฉพาะอยู่ 2 แขนงใหญ่ ๆ คือ

¹ กาญจนา นาคสกุล ระบบเสียงภาษาไทย(กรุงเทพฯ:จุฬาลงกรณ์มหาวิทยาลัย,2520),หน้า 6-7

² ประสิทธิ์ กาพย์กลอน การศึกษารายละเอียดของเสียงพูด(กรุงเทพฯ:ไทยวัฒนาพานิช,2516)

1. สัทศาสตร์ (Phonetics) คือการศึกษาเสียงที่เราได้ยินว่าเกิดอย่างไร เป็นเสียงชนิดใด และมีธรรมชาติเป็นอย่างไร ในแขนงนี้ จะศึกษาลักษณะของเสียงอย่างละเอียดทุกเสียงที่ได้ยิน แบ่งย่อยออกเป็น 3 ส่วนคือ

1.1 สรีรศาสตร์ (Articulatory Phonetics) จะศึกษาอวัยวะต่าง ๆ ที่ใช้ในการออกเสียง เช่น ใช้ อวัยวะใดทำเสียง เส้นเสียงสั่นสะเทือนหรือไม่ ลมที่พุ่งออกมาแรงหรือไม่ เป็นต้น

1.2 กลศาสตร์ (Acoustic Phonetics) ศึกษาว่าเสียงเปล่งออกไปในอากาศได้อย่างไร การศึกษา ส่วนนี้ต้องอาศัยเครื่องมือวิทยาศาสตร์ประกอบด้วย

1.3 โสตศาสตร์ (Auditory Phonetics) ศึกษาเรื่องการรับฟังเสียงที่เปล่งออกไปว่าผู้ฟังได้ยิน เสียงอย่างไร



รูปที่ 2.1 แสดงความสัมพันธ์ระหว่างส่วนย่อยทั้งสามของวิชาสัทศาสตร์

2. วิชาว่าด้วยหน่วยเสียง (Phonemics) คือการศึกษาเรื่องหน่วยเสียงหรือเสียงสำคัญในภาษา เป็นการนำเอา เสียงที่วิเคราะห์แล้วตามวิชาสัทศาสตร์มาจัดเป็นหมวดหมู่ โดยอาศัยหลักการเฉพาะหน่วยเสียง

2.4 อวัยวะที่ใช้ในการออกเสียงพูด

อวัยวะที่ใช้ในการออกเสียงพูดก็คือ อวัยวะต่าง ๆ ที่มีไว้ทำหน้าที่อื่น ๆ นั้นเอง แต่มนุษย์ใช้อวัยวะเหล่านั้นมาใช้ในการออกเสียงต่าง ๆ ตามต้องการได้ เราสามารถแบ่งอวัยวะที่ใช้ออกเสียงเป็น 3 พวกคือ

1. พวกเริ่มต้น คือพวกที่ก่อให้เกิดกระแสลม กำหนดลมและทำให้ลมเคลื่อนที่ ได้แก่ ปอด กระบังลม และกล้ามเนื้อระหว่างซี่โครง
2. พวกที่ใช้ในการออกเสียง ได้แก่ หลอดลม กล่องเสียงและเส้นเสียง
3. พวกที่ทำให้เสียงเปลี่ยนแปลงไปต่าง ๆ ได้แก่ ช่องปาก ฟัน ปุ่มเหงือก เพดานอ่อน เพดานแข็ง ลิ้น ริมฝีปากรวมทั้งช่องจมูกด้วย

อวัยวะต่าง ๆ เหล่านี้ถ้าจะแบ่งตามหน้าที่สามารถแบ่งออกได้เป็น 2 พวก คือ พวกที่หนึ่ง เรียกว่า “อวัยวะกำลังส่ง” หมายถึงอวัยวะที่ทำให้เกิดกระแสลม กำหนดลม ทำให้ลมวิ่งไปตามหลอดลม ผ่านเส้นเสียงแล้วทำให้เกิดเสียง ได้แก่ ปอด กระบังลม กล้ามเนื้อระหว่างซี่โครง เส้นเสียงและหลอดลม พวกที่สอง เรียกว่า “อวัยวะกล่อมกล่าเสียง” หมายถึงอวัยวะที่ทำให้เสียงเปลี่ยนแปลงไป ได้แก่ อวัยวะต่าง ๆ ภายในปากทั้งหมด ซึ่งแบ่งออกเป็น 2 ชุด คือ ชุดที่หนึ่ง เรียกว่า “ชุดล่าง” เป็นพวกที่เคลื่อนไหวได้ ได้แก่ ลิ้น ริมฝีปากล่าง ฟันล่าง ชุดที่สอง เรียกว่า “ชุดบน” เคลื่อนไหวได้ยาก ได้แก่ ริมฝีปากบน ฟันบน ปุ่ม

เหงือก เพดานอ่อน เพดานแข็ง เป็นต้น อวัยวะกล่อมเกลเสียงทั้ง 2 ชุดนี้ จะทำงานประสานสัมพันธ์อย่างดีและเป็นอัตโนมัติ

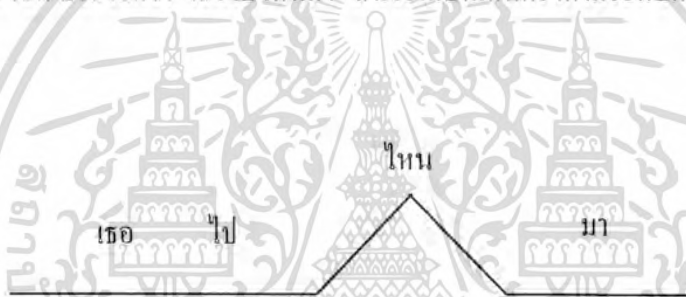
2.5 เสียงในภาษาไทย

เสียงในภาษาไทยจำแนกออกได้เป็น 3 ชนิด คือ เสียงแท้ เสียงแปร และเสียงดนตรี หรือเสียงผัน

2.5.1 เสียงแท้ คือ เสียงที่เปล่งออกมาจากลำคอโดยตรง โดยไม่ถูกอวัยวะต่างๆ ในปากปิดหรือกั้นลมไว้ เป็นเสียงที่เปล่งออกมาได้โดยสะดวก เช่น เสียง อา อิ เอ โอ ออ เป็นต้น สัญลักษณ์หรืออักษรที่ใช้แทนเสียงแท้ เรียกว่า “สระ”

2.5.2 เสียงแปร คือ เสียงที่เปล่งออกมาจากลำคอแล้วถูกอวัยวะต่างๆ ในปากปิดกั้นหรือกั้นลมไว้ ไม่ปล่อยให้ลมแล่นออกมาได้สะดวก ทำให้เสียงแปรไป เช่น เสียง /ป/ เกิดจากริมฝีปากทั้งสองข้างปิดกั้นลม เป็นต้น สัญลักษณ์หรือตัวอักษรที่ใช้แทนเสียงแปร เรียกว่า “พยัญชนะ”

2.5.4 เสียงดนตรี หรือเสียงผัน คือ เสียงแท้หรือเสียงแปรที่เปล่งออกมาแล้วมีระดับเสียงสูง ๆ ต่ำ ๆ คล้ายเสียงดนตรี เช่น ประโยคว่า “เธอไปไหนมา” ถ้าเขียนเป็นเส้นตรงตามระดับความสูง-ต่ำของเสียงจะได้ดังนี้



สัญลักษณ์หรือตัวอักษรที่ใช้แทนเสียงดนตรีเรียกว่า “วรรณยุกต์”

2.6 เสียงอักษรของไทย

เสียงกับตัวอักษรมีความแตกต่างกัน เพราะเสียงในภาษาไทยก็คือเสียงพูดของมนุษย์ ส่วนตัวอักษรคือสัญลักษณ์ที่ใช้แทนเสียงพูดของมนุษย์ ไม่มีภาษาใดในโลกที่มีตัวอักษรถ่ายทอดเสียงพูดได้หมดทุกเสียง ตัวอักษรไทยได้ชื่อว่ามีลักษณะพิเศษ เพราะสามารถถ่ายทอดเสียงพูดในภาษาไทยได้ตรงเกือบหมดทุกเสียง นอกจากนี้ยังสามารถใช้ถ่ายทอดเสียงพูดในภาษาอื่น ๆ ได้มากด้วย แม้ว่าภาษานั้น ๆ จะออกเสียงยากเพียงใดก็ตาม ดังที่เราถ่ายทอดเสียงภาษาอังกฤษเป็นอักษรไทยได้ เป็นต้น

“ชาคริต อนันทราวัน”¹ ได้ตั้งข้อสังเกตว่า ทำไมรูปและเสียงของอักษรไทยจึงมีจำนวนไม่เท่ากัน รวมไปถึงว่า การกลายเสียงสระ เสียงพยัญชนะ และเสียงวรรณยุกต์ ในบางท้องถิ่นของไทยเกิดขึ้นได้อย่างไร ปรากฏการณ์ดังกล่าวนี้มีอิทธิพลต่อกฎเกณฑ์ของภาษาหรือการใช้ภาษาหรือไม่ จากข้อสังเกตนี้ “ชาคริต อนันทราวัน” จึงอธิบายเปรียบเทียบเสียงอักษรไทยตามตำราอักษรวิธีเดิม จากการศึกษาวิเคราะห์เชิง

¹ ชาคริต อนันทราวัน, เอกสารคำสอน ท.111 หลักภาษาไทย (เลข:ภาควิชาภาษาไทยวิทยาลัยครูเดช, 2530), หน้า 19-25

ศาสตร์พบว่าการศึกษาตามอักษรวิธีเดิมนั้นส่วนมากถูกต้องและเป็นจริงเชิงภาษาศาสตร์ ดังจะจำแนกการศึกษาทั้งสองแบบมาสรุปไว้ดังนี้

2.6.1 เสียงอักษรไทยแบบเดิม

2.6.1.1 เสียงพยัญชนะ

พยัญชนะไทยมี 44 รูป 21 เสียง ดังนี้

หน่วยเสียง ก

หน่วยเสียง ข ฦ ค ฦ

หน่วยเสียง ง

หน่วยเสียง จ

หน่วยเสียง ฉ ฦ

หน่วยเสียง ช ฦ ฦ

หน่วยเสียง ฦ ฦ

หน่วยเสียง ฦ ฦ

หน่วยเสียง ฦ ฦ ฦ ฦ

หน่วยเสียง ฦ ฦ

หน่วยเสียง ฦ

หน่วยเสียง ฦ

หน่วยเสียง ฦ ฦ

หน่วยเสียง ฦ ฦ

หน่วยเสียง ฦ

หน่วยเสียง ฦ

หน่วยเสียง ฦ ฦ

หน่วยเสียง ฦ ฦ

หน่วยเสียง ฦ

หน่วยเสียง ฦ ฦ

หน่วยเสียง ฦ

พยัญชนะในภาษาไทยเหล่านี้แบ่งออกเป็นวรรคๆ ตามฐานที่เกิดของอักษรเหมือนภาษาบาลีและภาษาสันสกฤต คือ จัดพยัญชนะที่เกิดจากฐานคอไว้ในวรรคที่ 1, เกิดจากเพดานไว้ในวรรคที่ 2, เกิดจากฐานไม่เหวี่ยงเป็นวรรคที่ 3, เกิดจากฐานฟันเป็นวรรคที่ 4, เกิดจากฐานริมฝีปากเป็นวรรคที่ 5 และพยัญชนะที่เกิดจากฐานต่างๆ กัน ซึ่งรวมกับวรรคทั้ง 5 ข้างต้นไม่ได้ จัดเป็นวรรคที่ 6 ดังนี้

	อโฆมะ		โฆมะ		
	สีถิล	ธนิต	สีถิล	ธนิต	สีถิล
วรรค ก เกิดจากฐานคอ	ก	ข ฃ	ค ฅ	ฌ	ง
วรรค จ เกิดจากฐานเพดาน	จ	ฉ	ช ฌ	จ	ญ
วรรค ฎ เกิดจากปุ่มเหงือก	ฎ ฏ	ฐ	ฑ	ฒ	ณ
วรรค ต เกิดจากฐานฟัน	ด ต .	ถ	ท	ธ	น
วรรค ป เกิดจากฐานริมฝีปาก	บ ป	ผ ฝ	พ ฟ	ภ	ม
อ วรรค หรือเสววรรค เกิดจากฐานต่างๆ กัน ย ร ล ว ส ข ฬ ห พ อ ฮ					

ตารางที่ 2.1 แสดงพยัญชนะในภาษาไทยแบ่งออกเป็นวรรคตามฐานที่เกิดของอักษร

เมื่อจะแสดงให้เห็นว่าเสียงพยัญชนะในเสววรรคมีฐานที่เกิดอย่างไรก็จะได้ทราบดังตาราง

	อโฆมะ		โฆมะ		
	สีถิล	ธนิต	สีถิล	ธนิต	สีถิล
วรรค ก เกิดจากฐานคอ	ก อ*	ข ฃ ห*	ค ฅ	ฌ	ง ฮ*
วรรค จ เกิดจากฐานเพดาน	จ ส*	ฉ	ช ฌ	จ	ญ ย*
วรรค ฎ เกิดจากปุ่มเหงือก	ฎ ฏ บ*	ฐ	ฑ	ฒ	ณ ร* พ*
วรรค ต เกิดจากฐานฟัน	ด ต ส*	ถ	ท	ธ	น ล *
วรรค ป เกิดจากฐานริมฝีปาก	บ ป	ผ ฝ	พ ฟ	ภ	ม ว*
* คือ พยัญชนะเสววรรค					

ตารางที่ 2.2 แสดงฐานที่เกิดของเสียงพยัญชนะในเสววรรค

จากตาราง 2 ตารางข้างต้นมีข้อความที่ควรทราบเพิ่มเติมเกี่ยวกับเสียงพยัญชนะตามอักษรวิธีเดิม ดังนี้
ฐาน คือ ที่เกิดของอักษร

พยัญชนะ 5 วรรคข้างต้น คือ ตั้งแต่ตัว ก ถึงตัว ม มี 33 ตัว รวมเรียกว่า “พยัญชนะวรรค” เพราะจัดเข้าเป็นหมวดหมู่ได้ตามฐานที่เกิด และเรียกอีกอย่างหนึ่งว่า “พยัญชนะมูกระ” หรือ “พยัญชนะใบ้” พยัญชนะพวกนี้จะให้ควบกล้ำในพวกเดียวกันไม่ได้

พยัญชนะกลุ่มสุดท้าย ตั้งแต่ตัว ข ถึงตัว ฮ เรียกว่า “พยัญชนะเสววรรค” หรือ “อ วรรค” เพราะมีเสียงและฐานที่เกิดแตกต่างกัน จึงจัดแยกเป็นวรรคพิเศษ

การจัดเรียงพยัญชนะตามลำดับต้น ตั้งแต่ ก ถึง ฮ นั้น ถือเอาเสียงเป็นสำคัญคือ จัดเรียงลำดับเสียงเบา ไปหาเสียงหนัก และเรียงตามเสียงที่เกิดจากข้างในออกมาข้างนอก คือ จัดเสียงที่เกิดจากฐานคอ ไว้หน้า ต่อมาที่เป็นเพดาน ปุ่มเหงือก ฟัน และริมฝีปากตามลำดับ

อโฆมะ คือ พยัญชนะที่มีเสียงไม่ก้อง

โฆมะ คือ พยัญชนะที่มีเสียงก้อง

สีถิล คือ พยัญชนะที่มีเสียงกระทบฐานเบาๆ

ชนิด คือ พยัญชนะที่มีเสียงกระพือฐานหนัก

อรรธสระ คือ พยัญชนะที่ทำหน้าที่ถึงสระ ทำให้พยัญชนะตัวอื่นที่มาควบด้วยออกเสียงได้ทั้งที่ไม่มีสระได้แก่ ย ร ล ว ตัวอย่างเช่น กัลยา ครู คลา ฯลฯ

2.6.1.2 เสียงสระในภาษาไทย

รูปสระในภาษาไทย มีเพียง 21 รูป แต่มี 32 เสียง ซึ่งแบ่งออกเป็นหมวดหมู่ได้ดังนี้

1. สระพื้นฐานเดี่ยว มี 8 เสียง คือ อะ,อา,อิ,อี,อี,เอ,อุ สระพวกนี้กล่าวกันว่าเป็นเสียงที่เปล่งออกมาจากกระพือฐานใดฐานหนึ่งเพียงฐานเดียว
2. สระแท้ 2 ฐาน มี 10 เสียง คือ เอะ,เอ,แอะ,โอะ,โอ,เออะ,เออ,เอาะ,ออ สระพวกนี้กล่าวกันว่าเป็นสระที่เปล่งออกมาจากกระพือสองฐานพร้อมกัน
3. สระผสม มี 6 เสียง คือ เอียะ,เอีย,เอือะ,เอือ,อัวะ,อัว
4. สระเกิน 8 เสียง คือ อุ,อุา,อุ,อุา,อัว,โอ,โอ,เออ

เราจะสังเกตเห็นว่า เสียงสระบางเสียงต้องใช้รูปสระหลายๆ รูป มาผสมกันให้เป็นเสียงหนึ่ง เช่น เอะ มีรูปไม้หน้า (-) กับ วิสรรชนีย์ (-ะ) มาผสมกันเป็นต้น แต่ในที่นี้เราจะไม่กล่าวถึง ณ ที่นี้

2.6.1.3 เสียงวรรณยุกต์

วรรณยุกต์ มี 4 รูป 5 หน่วยเสียง ซึ่งเราอาจแบ่งออกเป็น 2 ชนิด คือ

1. วรรณยุกต์ที่มีรูป หมายถึง วรรณยุกต์ที่มีเครื่องหมายบอกระดับของเสียงให้เห็นชัดเจนอยู่เบื้องบนของอักษร มี 4 รูป ได้แก่ วรรณยุกต์ เอก โท ตรี จัตวา ตามลำดับ และให้เขียนไว้บนพยัญชนะต้น เช่น ก่, ก้, ก๊, กั ถ้าเป็นอักษรควบหรืออักษรนำให้เขียนไว้บนพยัญชนะตัวที่ 2 เช่น กรุ่น, เกล็ด, กลั่น เป็นต้น

2. วรรณยุกต์ไม่มีรูป ได้แก่ เสียงที่มีทำนองสูง-ต่ำตามหมู่ของอักษร โดยไม่ต้องมีรูปวรรณยุกต์กำกับก็อ่านออกเสียงได้ เหมือนมีรูปวรรณยุกต์กำกับ อยู่ด้วย เช่น

“คุณ” มีเสียงวรรณยุกต์ สามัญ

“จะ” มีเสียงวรรณยุกต์ เอก

“พูด” มีเสียงวรรณยุกต์ โท

“นัก” มีเสียงวรรณยุกต์ ตรี

“หนา” มีเสียงวรรณยุกต์ จัตวา

2.6.2 เสียงอักษรไทยแบบปัจจุบัน

การศึกษาเรื่องเสียงอักษรไทยปัจจุบันได้เจริญก้าวหน้าไปมาก โดยใช้วิธีการทางสัทศาสตร์เข้ามาศึกษาเพื่อหาความจริง จนได้ทราบว่า เสียงและรูปอักษรมีความสัมพันธ์กันบางส่วน แต่บางส่วนก็ไม่มี ความสัมพันธ์กัน

2.6.2.1 เสียงพยัญชนะ

เสียงพยัญชนะเป็นเสียงสำคัญในภาษาพวกหนึ่ง ซึ่งเกิดจากลมที่ผ่านเส้นเสียงแล้วถูกดัดแปลงด้วยอวัยวะออกเสียงส่วนต่างๆ ในปากทำให้เกิดเสียงขึ้น พยัญชนะมีอยู่หลายประเภท มีลักษณะการออกเสียงแตกต่างกันหลายแบบ จะเห็นได้จากพยัญชนะของไทยเรามีถึง 44 รูป แต่มีหน่วยเสียงเพียง 21 หน่วยเสียง ซึ่งความจริงมีเพียง 21 รูปเท่ากับจำนวนหน่วยเสียงที่น่าจะพอ ความข้อนี้ก็ต้องโยงความคิด ไปถึง

ประวัติอักษรไทย ซึ่งมีที่มาและโอกาสให้ใช้ต่างกันออกไป คำที่มีเสียงเหมือนกันแต่มีรูปต่างกันความหมายก็ต่างกันไป

ก. หน่วยเสียงหรือพยัญชนะและรูปอักษรพยัญชนะไทย

สัญลักษณ์แทนเสียง	รูปของพยัญชนะ
/p/	ป
/t/	ต ฏ
/c/	จ
/k/	ก
/ʔ/	อ
/ph/	พ ฝ
/th/	ถ ท ฐ ฑ ฒ
/ch/	ช ฌ ฦ
/kh/	ข ค ฅ ฌ ฦ
/b/	บ
/d/	ด ฎ
/m/	ม หม
/n/	น หน
/ng/	ง หง
/l/	ล ฬ ฬ
/r/	ร ฬ
/i/	ร ฬ
/s/	ส ศ ช ซ
/h/	ห ฮ
/w/	ว ฬ
/j/	ย ญ ฬ ฬ

ตารางที่ 2.3 แสดงสัญลักษณ์แทนเสียงของหน่วยเสียงพยัญชนะต้น เดี่ยว

จากตารางข้างบนจะพบว่า เสียงบางเสียงมีรูปอักษรหลายรูป แต่มีวิธีการใช้และค้นเสียงวรรณยุกต์ออกไปต่างๆ กัน จากเหตุผลดังกล่าว ในการอธิบายเรื่องของหน่วยเสียงพยัญชนะจะใช้ "สัญลักษณ์แทนหน่วยเสียง" แทน "รูปอักษร" เช่น เมื่อพูดถึงรูปอักษรชุด "จ ท ฐ ฏ ฒ" จะใช้สัญลักษณ์ "/th/" แทน เป็นต้น

ข. ลักษณะของหน่วยเสียงพยัญชนะ ในภาษาไทย

เสียงพยัญชนะในภาษาไทยแบ่งออกเป็น 6 ประเภทคือ เสียงกักหรือเสียงระเบิด เสียงเสียดแทรก เสียงนาสิก เสียงข้างลิ้น เสียงกระทบ และเสียงกึ่งสระหรือเสียงอรรดสระ

1) พยัญชนะระเบิด และพยัญชนะกัก (Plosive & Stop)

“เสียงพยัญชนะระเบิด (Plosive)” คือ เสียงพยัญชนะที่เกิดจากลมที่ปลั่งออกมาถูกกัก ณ ที่หนึ่งใดในช่องปาก แล้วเปิดช่องที่กักนั้นให้ลมพุ่งออกมาโดยแรง ในภาษาไทยมีพยัญชนะเสียงกัก 11 เสียง คือ

- เสียงพยัญชนะระเบิด ไม่ก้อง ไม่มีลม

- เกิดที่ริมฝีปาก คือ /p/

- เกิดที่บริเวณปุ่มเหงือกหรือฟัน คือ /t/

- เกิดที่เพดานแข็ง คือ /c/

- เกิดที่เพดานอ่อน คือ /k/

- เกิดที่ช่องระหว่างเส้นเสียง คือ /ʔ/

การเปล่งเสียง ริมฝีปากอยู่ชิดกันสนิท ลมจึงถูกกักที่ริมฝีปาก เมื่อจะเปล่งเสียงต้องเปิดริมฝีปากปล่อยลมให้ลมระเบิดออกมาทันที เส้นเสียงเปิดปริมาณลมเป็นไปตามปกติหรืออาจกล่าวได้ว่าไม่มีกลุ่มลมตามออกมา คำที่มีเสียงนี้เช่น ปา /pa:/ ปู /pu:/ ปรากฏเป็นเสียงเฉพาะต้นพยางค์เท่านั้น เมื่อเสียงนี้ปรากฏที่ท้ายพยางค์จะถูกเรียกว่าพยัญชนะกัก (Stops) จะไม่มีการเปิดริมฝีปากให้ลมระเบิดออกมาโดยถือว่าพยัญชนะทั้งคู่นี้เป็นเสียงย่อยของหน่วยเสียงเดียวกันกับปรากฏสลับหลักกัน

- เสียงพยัญชนะระเบิด ไม่ก้อง ไม่มีลม

- เกิดที่ริมฝีปาก คือ /ph/

- เกิดที่ฟัน คือ /th/

- เกิดที่เพดานแข็งส่วนหน้า ใกล้ปุ่มเหงือก คือ /ch/

- เกิดที่เพดานอ่อน คือ /kh/

การเปล่งเสียง ริมฝีปากอยู่ชิดกันสนิท ลมถูกกักที่ริมฝีปาก เมื่อจะเปล่งเสียงต้องเปิดริมฝีปากปล่อยลมให้ลมระเบิดออกมาทันที เส้นเสียงเปิด ปริมาณลมออกมามากกว่าปกติ คำที่มีเสียงนี้ เช่น ฟ้า /pha:2/ ภาพ /pha:p2/ พวก /phuak2/ ปรากฏเป็นเสียงต้นพยางค์เท่านั้น เมื่ออยู่ท้ายพยางค์จะเรียกว่าพยัญชนะกัก

- เสียงพยัญชนะระเบิด ก้อง ไม่มีลม

- เกิดที่ริมฝีปาก คือ /b/

- เกิดที่ปุ่มเหงือก คือ /d/

การเปล่งเสียง ริมฝีปากอยู่ชิดกันสนิท ลมถูกกักที่ริมฝีปาก เส้นเสียงอยู่ชิดกันแต่ไม่ชิดกันสนิท ขณะเปล่งเสียงลมจะผ่านเส้นเสียงออกมาทำให้เส้นเสียงสั่นสะเทือน ขณะเดียวกันก็เปิดริมฝีปากปล่อยลมระเบิดออกไปทันที ปริมาณลมเป็นไปตามปกติ คำที่มีเสียงนี้ เช่น บ้าน /ba:n2/ ปรากฏเป็นเสียงต้นพยางค์เท่านั้น เสียงระเบิดจะเกิดเฉพาะกับพยัญชนะต้นเท่านั้นถ้าพยัญชนะพวกนี้ไปอยู่ที่ส่วนท้ายคำหรือตัวสะกดจะทำให้เกิดเสียงอีกอย่างหนึ่ง คือ “เสียงกัก (Stop)” เสียงกักในภาษาไทยมี 4 เสียง คือ

/pʰ/ เสียงพยัญชนะกัก ไม่ก้อง เกิดที่ริมฝีปาก

- /t'/ เสียงพยัญชนะกัก ไม่ก้อง เกิดที่ฟัน
- /k'/ เสียงพยัญชนะกัก ไม่ก้อง เกิดที่เพดานแข็ง
- /'/ เสียงพยัญชนะกัก ไม่ก้อง เกิดที่ช่องระหว่างเส้นเสียง

2) พยัญชนะนาสิก (Nasals)

ในการเปล่งเสียงนาสิก ลมจากปอดขึ้นไปถึงจมูกได้ แต่ผ่านออกทางช่องปากไม่ได้เพราะทางเดินของลมในช่องปากถูกปิดกั้นสนิท ณ ที่ใดที่หนึ่ง เสียงนาสิกในภาษาไทย ปรากฏมี 3 หน่วย มีลักษณะเป็นเสียงก้องทั้งหมด และหน่วยเสียงทั้ง 3 นี้ปรากฏได้ทั้งตำแหน่งพยัญชนะต้นและตัวสะกด คือ

- /m/ พยัญชนะนาสิก ก้อง เกิดที่ริมฝีปาก
- /n/ พยัญชนะนาสิก ก้อง เกิดที่ฟัน
- /ng/ พยัญชนะนาสิก ก้อง เกิดที่เพดานอ่อน

3) พยัญชนะข้างลิ้น (Lateral)

เสียงพยัญชนะข้างลิ้นเป็นเสียงพยัญชนะก้องซึ่งเปล่งโดยใช้ลิ้นปิดบริเวณปุ่มเหงือกและเพดานตรงส่วนกลาง ๆ ไว้ ปล่อยให้ลมออกมาข้าง ๆ ลิ้น มีเพียงหน่วยเสียงเดียวคือ /l/

4) พยัญชนะรวิว (Rolled)

เสียงพยัญชนะรวิวเป็นเสียงพยัญชนะก้อง ในขณะที่เปล่งเสียง ต้องทำปลายลิ้นให้มีลักษณะอ่อนตัวที่สุดจนสามารถกระดกขึ้นไปแตะหลังฟันได้หลายๆ ครั้ง ในภาษาไทยมีเสียงเดียว คือ /r/

5) พยัญชนะเสียงแทรก (Fricatives)

เสียงพยัญชนะเสียงแทรก คือ เสียงพยัญชนะซึ่งเกิดขึ้นเมื่อลมที่ทำให้เกิดเสียงต้องบีบตัวผ่านช่องแคบ ณ ที่ใดที่หนึ่งในช่องปาก และทำให้เกิดเสียงซู่ซ่าขึ้น อวัยวะที่ทำให้เกิดช่องแคบได้ คือ ริมฝีปาก ฟันกับริมฝีปาก ลิ้นกับส่วนต่างๆ ของเพดานปาก และที่ช่องระหว่างเส้นเสียง ในภาษาไทยเสียงเสียดแทรกเป็นเสียงไม่ก้องและปรากฏในตำแหน่งพยัญชนะต้น มี 3 หน่วยเสียง คือ

- /f/ เกิดขึ้นโดยการกดฟันลงบนริมฝีปากล่างเบาๆ แล้วเปล่งเสียงโดยปล่อยลมให้ผ่านฟันออกมา
- /s/ เกิดขึ้นโดยอวัยวะ คือ ฟันทั้งบนและล่างกับลิ้นส่วนปลายทำงานประสานสัมพันธ์กัน
- /h/ เกิดขึ้นตรงช่องระหว่างเส้นเสียง ช่องนี้จะอยู่ในลักษณะที่แคบกว่าในตำแหน่งหายใจยาว ๆ แต่ไม่แคบจนทำให้เส้นเสียงสั่น ถ้าลมที่ถูกบีบผ่านออกมาค่อนข้างแรง

6) เสียงอรรถสระหรือเสียงกึ่งสระ (Semi-vowels)

เสียงพยัญชนะอรรถสระ เป็นเสียงที่เกิดขึ้นระหว่างเสียงสระ 2 เสียง ในการเปล่งเสียงพยัญชนะอรรถสระ อวัยวะที่ใช้ในการออกเสียงจะอยู่ในตำแหน่งของสระใดสระหนึ่งก่อนแล้วเปล่งเสียงออกมา ในขณะที่

เปลี่ยนอวัยวะนั้น ๆ ไปสู่ตำแหน่งของสระอีกตำแหน่งหนึ่ง หน่วยเสียงพยัญชนะอรรธสระในภาษาไทยมี 2 หน่วย คือ /j/ และ /w/ หน่วยเสียงพยัญชนะอรรธสระในภาษาไทยปรากฏทั้งพยัญชนะต้น และพยัญชนะตัวสะกด

ก. ลักษณะเสียงพยัญชนะไทย

นอกจากลักษณะหน่วยเสียงต่างๆ ดังที่กล่าวมาแล้ว พยัญชนะไทยยังมีลักษณะเสียงที่ควรทราบ คือ

1) พยัญชนะ โงมะหรือพยัญชนะเสียงก้อง (Voiced Sounds) คือ เสียงพยัญชนะที่เกิดจากลมผ่านเส้นเสียง ขณะที่เส้นเสียงปิดลงจึงดันเส้นเสียงทำให้เส้นเสียงเกิดอาการสั่นสะบัด มี 8 เสียง คือ /b/, /d/, /m/, /n/, /ng/, /l/, /w/, /j/

2) พยัญชนะ โงมะหรือพยัญชนะเสียงไม่ก้อง (Voiceless Sounds) คือ เสียงพยัญชนะที่เกิดจากลมผ่านเส้นเสียงขณะที่เส้นเสียงเปิด คือ เส้นเสียงห่างจากกันทำให้เกิดช่องระหว่างเส้นเสียง (Glottis) ขึ้น ลมจึงผ่านเส้นเสียงออกมาได้ เส้นเสียงจึงไม่เกิดการสั่นสะบัด มี 12 เสียง คือ /p/, /t/, /k/, /ʔ/, /ph/, /th/, /kh/, /c/, /ch/, /f/, /s/, /h/

3) พยัญชนะติดหรือพยัญชนะเสียงเบา (Unaspirated Sounds) คือ เสียงพยัญชนะที่เปล่งออกมาแล้วไม่มีลมพุ่งตามออกมาด้วย สามารถทดสอบได้โดยการใส่กระดาษบางๆ หรือหลังมือจ่อที่ริมฝีปาก จะไม่มีลมพุ่งตามออกมากระทบกระดาษหรือหลังมือ มี 5 เสียง คือ /p/, /t/, /k/, /ʔ/, /c/

4) พยัญชนะติดหรือพยัญชนะเสียงหนัก (Aspirated Sounds) คือ เสียงพยัญชนะที่เปล่งออกมาแล้วมีลมพุ่งตามออกมาด้วย สามารถทดสอบได้โดยการใส่กระดาษบางๆ หรือหลังมือจ่อที่ริมฝีปาก จะมีลมพุ่งตามออกมากระทบกระดาษหรือหลังมือ มี 4 เสียง คือ /ph/, /th/, /kh/, /ch/ โดยสัญลักษณ์แทนเสียงจะใช้ตัว “h” แทนการที่มีลมพุ่งออกมา

หน่วยเสียงพยัญชนะไทยดังกล่าว มีลักษณะของเสียงและที่เกิดของเสียงดังตารางต่อไปนี้ คือ

ที่เกิดเสียง			ริมฝีปากทั้งสองข้าง	ริมฝีปากกลางและฟันบน	ปุ่มเหงือก	ปุ่มเหงือกต่อเพดานแข็ง	เพดานแข็ง	เพดานอ่อน	เส้นเสียง
ระเบิด			/b/		/d/				
	อโฆมะ	สถิต	/p/		/t/			/k/	/ʔ/
		ชนิต	/ph/		/th/			/kh/	
กึ่งเสียดแทรก อโฆมะ		สถิต				/c/			
		ชนิต				/ch/			
เสียดแทรก อโฆมะ				/f/	/s/				/h/
นาสิก โฆมะ			/m/		/n/			/ng/	
ข้างลิ้น โฆมะ					/l/				
ร้าว โฆมะ					/r/				
อรรถสระ โฆมะ			/w/				/j/		

ตารางที่ 2.4 ลักษณะของเสียงและที่เกิดเสียงของหน่วยเสียงพยัญชนะในภาษาไทย

ง. ตำแหน่งและคุณสมบัติของเสียงพยัญชนะไทย

เสียงพยัญชนะไทยทั้ง 21 เสียงนั้น มีตำแหน่งและคุณสมบัติที่อธิบายได้ดังนี้

1) ตำแหน่งต้นคำ พยัญชนะทุกเสียงในภาษาไทย ตกอยู่ในตำแหน่งต้นคำได้ทั้งหมด เช่น ลา /la/, มี /mi:/, ม้า /ma:/, นะ /na/ เป็นต้น

2) ตำแหน่งท้ายคำ พยัญชนะทุกเสียงในภาษาไทย ตกอยู่ในตำแหน่งท้ายคำได้ มีอยู่ 9 เสียงเท่านั้น คือ ป /p/, ต /t/, ก /k/, ม /m/, ว /w/, น /n/, ย /j/, ง /ng/ และ อ /ʔ/ (ไม่มีตัวสะกด) ตัวอย่างเช่น

ฉาบ /cha:p/ มีเสียง ป /p/ ท้ายคำ แม่กบ

ตาด /ta:t/ มีเสียง ต /t/ ท้ายคำ แม่กด

กัก /kak/ มีเสียง ก /k/ ท้ายคำ แม่กก

โคม /kho:m/ มีเสียง ม /m/ ท้ายคำ แม่กม

นาง /na:ng/ มีเสียง ง /ng/ ท้ายคำ แม่กง

นาย /na:j/ มีเสียง ย /j/ ท้ายคำ แม่เกย

นาว /na:w/ มีเสียง ว /w/ ท้ายคำ แม่เกว

เอ๊ะ /ʔeʔ/ มีเสียง อ /ʔ/ ท้ายคำ ไม่มีตัวสะกด

จากตัวอย่างจะพบว่าในแม่กบ กับแม่กด ไม่ได้เสียง บ /b/ กับเสียง ด /d/ สะกด เพราะว่าทั้งสองคือ บ และ ด เวลาออกเสียงระเบิด(released) คล้ายๆ เสียง b และ d ในภาษาอังกฤษ ดังนั้น คำว่า “ฉาบ” ควรจะออกเสียงเป็น ฉาบ-เบอะ (เบอะ คือ เสียงที่เกิดจากลมออกมาจากริมฝีปาก) และคำว่า “ตาด” ก็จะเป็นตาด-เดอะ

ดังนั้นหน่วยเสียง บ และ ค จะปรากฏท้ายคำในภาษาไทยไม่ได้ ควรเป็นเสียง ป /p/ และ ต /t/ มากกว่า เพราะสองเสียงนี้เฝ้าอยู่ท้ายคำเสียงไม่ระเบิด

3) เสียงพยัญชนะควบกับเสียงอื่นได้ เสียงพยัญชนะไทยที่สามารถควบกับเสียงอื่นได้ หรือสามารถออกเสียงพร้อมกับเสียงพยัญชนะอื่นได้ คือ เสียง ร /r/, ล /l/ และ ว /w/ ตัวอย่างเช่น

เสียงต้น เสียงควบ	ก	ค	ต	ท	ป	พ	บ	ด	ฟ
ร	กร	คร	ตร	ทร	ปร	พร	бр	พร	ฟร
ล	กล	คล			ปล	พล	บล		ฟล
ว	กว	คว	*						

ตารางที่ 2.5 แสดงเสียงพยัญชนะไทยที่สามารถควบกับเสียงอื่นได้

ตัวอย่างของคำที่มีพยัญชนะต้นควบกันได้แก่

กร	เช่น	กราม	กรู	เกรียม
กล	เช่น	กลืน	กลืน	กลับ
กว	เช่น	ไกว	กว้าง	แกว่ง
คร	เช่น	ครบ	ครีบ	คราด
คล	เช่น	คลาน	โคล	คลุก
คว	เช่น	ความ	ควาย	ควั่น
ดร	เช่น	ดรวน	ดรวจ	ดั่ง
ทร	เช่น	จันทร์	นิตรา	ทรง
ปร	เช่น	ปรับ	ปรู	ปรี
ปล	เช่น	ปลง	ปลุก	ปลิง
พร	เช่น	พระ	ไพร	พรา
พล	เช่น	พลัน	พลาย	พลิก
บร*	เช่น	บริดจ์	บรอนซ์	เบรมา
บล*	เช่น	บล็อก	บลู	
ดร*	เช่น	ดรัม	ดรอว์	ดราฟต์
ฟร*	เช่น	ฟรี	ฟรักโทส	
ฟล*	เช่น	ฟลาย	แฟลต	

จากตัวอย่างเราจะเห็นว่าคำภาษาไทยจะไม่ปรากฏเสียงควบกล้ำในตำแหน่งท้ายพยางค์ และคำที่มีเครื่องหมายดอกจัน (*) เป็นคำภาษาต่างประเทศที่ถือเป็นภาษาไทยแล้ว เช่น คำที่ขยืมมาจากภาษาบาลีและสันสกฤต

2.6.2.2 เสียงสระ

เสียงสระเป็นเสียงแท้ หมายถึง เสียงที่เกิดจากลมพุ่งออกมาได้อย่างสะดวกไม่มีอวัยวะส่วนหนึ่งส่วนใดปิดกั้นทางลมไว้ เป็นเสียงที่เกิดจากลมที่ผ่านเส้นเสียงในตำแหน่งที่ปิดเกือบสนิทในขณะที่อวัยวะในปากจะอยู่ในท่าและตำแหน่งต่างๆ ที่ทำให้โพรงปากมีลักษณะต่างกันได้หลายแบบ ลมที่ผ่านออกจึงเกิดเป็นเสียงได้ต่างๆ กัน “ประยุทธ์ จุฑาสคร” ได้สรุปลักษณะสำคัญของหน่วยเสียงไว้เป็นข้อๆ ดังนี้ เป็นเสียงที่ลมแล่นออกมาได้โดยสะดวก เพราะไม่ถูกอวัยวะในปากปิดกั้น

เป็นเสียงก้องทุกเสียง

สามารถออกเสียงได้ยาวนานและชัดเจน

มีทั้งเสียงสั้นและเสียงยาว

ลักษณะของเสียงขึ้นอยู่กับการ ริมฝีปาก ปริมาณของลมที่พุ่งออกมา และระยะเวลาที่ใช้ออกเสียง

ก. ประเภทของเสียงสระ

รูปสระในภาษาไทยมี 21 รูป แต่มีถึง 32 เสียง ในจำนวนทั้งหมด 32 เสียงนั้น สามารถแบ่งประเภทได้ตามแนวการศึกษาทางภาษาศาสตร์ คือ สระเดี่ยว 18 เสียง สระประสม 6 เสียง และสระเกิน 8 เสียง ซึ่งสระเกินเป็นสระที่มีเสียงพยัญชนะรวมอยู่ด้วยจึงไม่ได้นับเป็นเสียงสระที่เรากำลังศึกษากันอยู่

1) สระเดี่ยว (Monophthong) หมายถึง เสียงสระที่มีที่กำเนิดเพียงฐานเดียว คือ เสียงสระที่เกิดจากการบังคับลิ้นให้อยู่ในระดับที่ต้องการเพียงจุดหนึ่งจุดใด แล้วจึงปล่อยเสียงออกมาเกิดเป็นเสียง มี 18 เสียงแบ่งเป็นเสียงสั้น 9 เสียงและเสียงยาวอีก 9 เสียง โดยแบ่งออกเป็นชนิดย่อยได้ 3 ชนิด โดยเรียกตามส่วนของลิ้นที่ยกขึ้นใกล้เพดาน ดังนี้

- สระหน้า เกิดจากการที่ออกเสียงสระโดยปล่อยลมออกจากปอดให้ผ่านขึ้นมาในช่องปาก ในขณะที่ยกลิ้นส่วนหน้าขึ้นใกล้เพดานแข็ง สระหน้าในภาษาไทยมี 3 คู่ คือ อี [i], อือ [i:] ; เอะ [e], เออ [e:] และ แอะ [æ], แอ [æ:] โดยแต่ละคู่คือสระเสียงสั้นและเสียงยาวที่เป็นคู่กัน

- สระกลาง เกิดจากการที่ออกเสียงสระโดยปล่อยลมออกจากปอดให้ผ่านขึ้นมาในช่องปาก ในขณะที่ยกลิ้นระหว่างลิ้นส่วนหน้ากับลิ้นส่วนหลังขึ้นใกล้บริเวณเพดานแข็งต่อกับบริเวณเพดานอ่อน สระกลางในภาษาไทยมี 3 คู่ คือ อะ [a], อา [a:] ; อึ [ʌ], อือ [ʌ:] และ ไออะ [ɪ], ไออ [ɪ:] โดยแต่ละคู่คือเสียงสระเสียงสั้นและสระเสียงยาวที่เป็นคู่กัน

- สระหลัง เกิดจากการที่ออกเสียงสระโดยปล่อยลมออกจากปอดให้ผ่านขึ้นมาในช่องปาก ในขณะที่ยกลิ้นส่วนหลังขึ้นใกล้เพดานอ่อน สระหลังในภาษาไทยมี 3 คู่ คือ อุ [u], อุ [u:] ; โอะ [o], โอ [o:] และ เอาะ [ə], เอา [ə:] โดยแต่ละคู่คือเสียงสระเสียงสั้นและสระเสียงยาวที่เป็นคู่กัน

ในการแบ่งสระอาจใช้วิธีการแบ่งโดยเรียกสระตามระดับความสูง-ต่ำของลิ้น ดังนี้

- สระสูง เป็น

2) สระประสม (Diphthong) หมายถึง เสียงสระที่มีที่กำเนิดตั้งแต่ 2 ฐานขึ้นไป กล่าวคือ เสียงสระที่เกิดจากการใช้ลิ้นออกเสียงสระหนึ่งแล้วลิ้นยังไม่กลับเข้าที่เดิม และลิ้นได้ออกเสียงอีกสระหนึ่ง เป็นลักษณะการเปลี่ยนระดับของลิ้นจากระดับหนึ่งไปอีกระดับหนึ่ง โดยลงเสียงหนักที่สระตัวแรก สระประสมในภาษาไทยมี 6 เสียง คือ

2.6.2.3 เสียงวรรณยุกต์

ภาษาไทยเราเรียกระดับสูง-ต่ำของคำว่า “วรรณยุกต์” ซึ่งเป็นส่วนสำคัญที่ทำให้คำซึ่งมีส่วนประกอบอื่นๆ คือ พยัญชนะ สระ และตัวสะกดเดียวกัน มีความหมายต่างกันได้ ดังนั้น วรรณยุกต์ในภาษาไทยจึงจัดเป็นหน่วยเสียงหน่วยหนึ่ง เรียกว่า “หน่วยเสียงวรรณยุกต์”

ก. รูปและเสียงวรรณยุกต์ในภาษาไทย

ในภาษาไทยที่ใช้กันอยู่ในปัจจุบันนี้ รูปวรรณยุกต์มีอยู่ 4 รูป แต่มี 5 เสียง เรียกตามอักขระวิธีเดิมว่า เสียงสามัญ เสียงเอก เสียงโท เสียงตรี และเสียงจัตวา และใช้รูปวรรณยุกต์ 4 รูปดังกล่าวมาแล้ว ส่วนเสียงสามัญไม่มีรูปวรรณยุกต์ ลักษณะเสียงวรรณยุกต์ 5 หน่วยนี้ เราสามารถจัดได้เป็น 2 พวกใหญ่ๆ คือ วรรณยุกต์ระดับและวรรณยุกต์เปลี่ยนระดับ

1) วรรณยุกต์ระดับ (LEVEL TONE) หมายถึง เสียงวรรณยุกต์ที่มีเสียงคงที่ หรือค่อนข้างคงที่ ตั้งแต่ต้นพยางค์ตลอดไปจนท้ายพยางค์ มี 3 หน่วย คือ

- หน่วยเสียงวรรณยุกต์ระดับต่ำ (LOW TONE) วรรณยุกต์นี้มีต้นเสียงกลางๆ ประมาณ 120 HZ แล้วจะต่ำลงมาถึงประมาณ 110 HZ อย่างรวดเร็ว วรรณยุกต์แบบนี้คือวรรณยุกต์เอก
- หน่วยเสียงวรรณยุกต์ระดับกลาง (MID TONE) วรรณยุกต์นี้มีต้นเสียงกลางๆ ประมาณ 120 HZ แล้วจะลดความถี่ลงมาถึง 110 HZ ที่ท้ายพยางค์ วรรณยุกต์แบบนี้คือวรรณยุกต์สามัญ
- หน่วยเสียงวรรณยุกต์ระดับสูง (HIGH TONE) วรรณยุกต์นี้มีต้นเสียงสูง และสูงขึ้นเรื่อยๆ จากต้นเสียงประมาณ 125 HZ จนถึงประมาณ 135-140 HZ เมื่อสิ้นพยางค์ หรืออาจลดต่ำลงอีกที่ตรงท้ายพยางค์ประมาณ 130 HZ วรรณยุกต์แบบนี้คือวรรณยุกต์ตรี

2) วรรณยุกต์เปลี่ยนระดับ (CONTOUR TONE) มีระดับความถี่ของการออกเสียงเปลี่ยนแปลงไม่คงที่ในช่วงพยางค์หนึ่งๆ เช่น ต้นเสียงมีระดับเสียงสูงแล้วลดต่ำลงอย่างรวดเร็วที่ท้ายพยางค์ หรือต้นพยางค์มีระดับเสียงต่ำแล้วเพิ่มสูงขึ้นที่ท้ายพยางค์ มีอยู่ 2 หน่วยเสียง คือ

- หน่วยเสียงวรรณยุกต์เปลี่ยนตก (FALLING TONE) ระดับเสียงเริ่มต้นจะมีความถี่ประมาณ 140 HZ และลดต่ำลงอย่างรวดเร็วที่ท้ายพยางค์ ประมาณ 100 HZ วรรณยุกต์แบบนี้คือวรรณยุกต์โท
- หน่วยเสียงวรรณยุกต์เปลี่ยนขึ้น (RISING TONE) ระดับเสียงจะเริ่มที่ความถี่ประมาณ 110 HZ แล้วจะลดอีกเล็กน้อย ก่อนจะเปลี่ยนเสียงขึ้นอย่างรวดเร็วไปสู่ความถี่ที่สูงขึ้นกว่าตอนเริ่มต้นที่ประมาณ 140 HZ ตอนท้ายพยางค์ วรรณยุกต์แบบนี้คือวรรณยุกต์จัตวา

การกำหนดสัญลักษณ์แทนเสียงสำหรับพยัญชนะ หรือหน่วยเสียงพยัญชนะ

พยัญชนะไทยมีทั้งหมด 44 รูป 21 เสียง โดยสามารถแบ่งออกเป็น พยัญชนะต้น และพยัญชนะตัวสะกด มีทั้งพยัญชนะเดี่ยวและพยัญชนะควบกล้ำ ดังแสดงในตารางต่อไปนี้

สัญลักษณ์แทนเสียง	รูปพยัญชนะ
/p/	ป
/t/	ต ฏ
/c/	จ
/k/	ก
/ʔ/	อ
/ph/	พ ฝ
/th/	ถ ฑ ฐ ฒ
/ch/	ช ฌ ฦ
/kh/	ข ค ฌ ฦ ค
/b/	บ
/d/	ด ฎ
/m/	ม ฌ
/n/	น ฌ
/ng/	ง ฌ
/l/	ล ฬ ฬ
/r/	ร ฌ
/f/	ฝ ฝ
/s/	ส ศ ษ ษ
/h/	ห ฌ
/w/	ว ฌ
/j/	ย ฌ ฌ ฌ

ตารางที่ 2.7 แสดงสัญลักษณ์แทนเสียงของหน่วยเสียงพยัญชนะต้น ด้วย 21 หน่วยเสียง

สัญลักษณ์แทนเสียง	รูปพยัญชนะ
/pr/	ปร
/phr/	พร
/tr/	ตร
/kr/	กร
/khr/	ขร กร
/pl/	ปล
/phl/	พล ผล
/thr/	ทร
/kl/	กล
/khl/	ขล กล
/kw/	กว
/khw/	ขว กว

ตารางที่ 2.8 แสดงสัญลักษณ์แทนเสียงของหน่วยเสียงพยัญชนะต้น ควบกล้ำ 12 หน่วยเสียง

สัญลักษณ์แทนเสียง	รูปพยัญชนะ
/br/	บร
/bl/	บล
/fr/	ฟร
/fl/	ฟล
/dr/	ดร

ตารางที่ 2.9 แสดงสัญลักษณ์แทนเสียงของหน่วยเสียงพยัญชนะต้น ควบกล้ำ ที่มาจากภาษาต่างประเทศ 5 หน่วยเสียง

สัญลักษณ์แทนเสียง	รูปพยัญชนะ
/p/	/p/
/t/	/t/
/k/	/k/
/n/	/n/
/m/	/m/
/ng/	/ng/
/j/	/j/
/w/	/w/

ตารางที่ 2.10 แสดงสัญลักษณ์แทนเสียงของพยัญชนะ ตัวสะกด 8 หน่วยเสียง

สัญลักษณ์แทนเสียง	รูปพยัญชนะ
/f/	ฟ
/l/	ล
/s/	ส ข ส
/ch/	ช

ตารางที่ 2.11 แสดงสัญลักษณ์แทนเสียงของหน่วยเสียงพยัญชนะ ตัวสะกด ที่มาจากต่างประเทศ 4 หน่วยเสียง

หมายเหตุ สำหรับคำที่มาจากต่างประเทศเราจะยึดการออกเสียงแบบภาษาไทยโดยตัวสะกดภาษาไทย

การกำหนดสัญลักษณ์แทนเสียงสำหรับสระ หรือหน่วยเสียงสระภาษาไทยมีสระทั้งหมด 24 เสียง โดยแบ่งออกเป็นสระเดี่ยว 18 หน่วยเสียง และสระผสม 6 หน่วยเสียง รวมทั้งหมด 24 หน่วยเสียง ดังแสดงในตารางต่อไปนี้

สัญลักษณ์แทนเสียง	รูปสระ
a	อะ
a:	อา
i	อิ
i:	อี
u	อุ
u:	อู
e	เอะ
e:	เอ
x	แอะ
x:	แอ
o	โอะ
o:	โอ
(ə)	เออะ
(ə):	ออ
#	เออะะ
#:	เออ

ตารางที่ 2.12 แสดงสัญลักษณ์แทนเสียงของหน่วยเสียงสระเดี่ยวเสียงสั้น-ยาว 18 หน่วยเสียง

บทที่ 3

การกำหนดหน่วยเสียงเพื่อการสังเคราะห์เสียง

3.1 บทนำ

ในบทนี้จะกล่าวถึงลักษณะของคลื่นเสียง หน่วยเสียงพื้นฐานและการจัดเก็บหน่วยเสียงเพื่อการนำไปใช้ในการสังเคราะห์เสียง ซึ่งมีวิธีแบ่งหน่วยเสียงเพื่อสังเคราะห์เสียงอยู่หลายวิธี

3.2 คลื่นเสียง

ลักษณะพิเศษของคลื่นโดยทั่วไป

คลื่นที่เกิดจากส้อมเสียงจะอยู่ในรูปแบบคลื่นที่ง่ายที่สุดเรียกว่า คลื่นรูปซายน์ (sine wave) เราอาจจะเห็นรูปแบบคลื่นเช่นนี้จากหน้าจอของเครื่องวัดทางอิเล็กทรอนิกส์ ที่เรียกว่าออสซิลอโคป โดยจะเห็นได้ว่าลักษณะที่สำคัญของคลื่น คือ ความถี่ และแอมพลิจูด (amplitude) ค่าที่สูงสุดของการกระพริบได้จากสูงสุดของคลื่น

แต่ในความจริงเราสามารถแยกแยะเสียงหนึ่งว่าแตกต่างจากเสียงอื่นได้โดย คุณสมบัติ 3 ข้อคือ พิตช์ (ความถี่) แอมพลิจูด (ความดัง) และคุณภาพ

พิตช์ (Pitch)

เราบอกว่าคลื่นเสียงถูกสร้างได้ง่ายที่สุดจากการสั่นของบางอย่างในอากาศ เช่น การดีดกีตาร์ เข่า ทัมโบรีน หรือพูด แต่ถ้าอัตราการสั่นของวัตถุได้กำหนดคุณสมบัติอย่างหนึ่งนั่นคือ พิตช์ การที่สายกีตาร์สั่นได้เร็วกว่านั่นคือเราได้ยินเสียงที่มีพิตช์ที่สูงกว่า

พิตช์จะขึ้นอยู่กับความถี่ ซึ่งเป็นลักษณะทางกายภาพโดยทั่วไปของสิ่งอย่างเดียวกัน ความถี่ของเสียงคือ จำนวนของลูกคลื่นที่สมบูรณ์ที่เกิดขึ้นใน 1 วินาที

ความถี่ของเสียงจะถูกวัดในหน่วยที่เรียกว่า เฮิร์ตซ์ เสียงที่เกิดการสั่น 1 ครั้งต่อ 1 วินาที เราจะเรียกว่า 1 เฮิร์ตซ์ ความถี่มักจะถูกเขียนในรูปกิโลเฮิร์ตซ์ หน่วยนี้จะแทน 1,000 เฮิร์ตซ์ คนเราโดยปกติสามารถรับฟังเสียงได้ในช่วง 20 เฮิร์ตซ์ ถึง 20,000 เฮิร์ตซ์ (20 กิโลเฮิร์ตซ์)

และความแตกต่างระหว่างความถี่กับพิตช์คือ ความถี่จะมีความหมายแทนการวัดจำนวนการสั่นสะท้อนต่อวินาที แต่พิตช์จะแทนท้สนะของเราที่จะเปรียบเทียบว่าการสั่นโดยรวมสั้นอย่างไร เช่น เราสามารถพูดได้ว่าพิตช์นี้ต่ำกว่าหรือสูง เช่น กระดิ่งมีพิตช์ที่สูง เบสมีพิตช์ที่ต่ำ

แอมพลิจูด (Amplitude)

การวัดแอมพลิจูดของคลื่นมีความสำคัญเพราะว่ามันจะบอกเราได้ว่าปริมาณของพลังงานในคลื่น ซึ่งจะมาอยู่ในรูปแบบของความดังที่เราได้ยิน แอมพลิจูดก็คือส่วนที่วัดได้ของความสูงของเสียง มันเป็น

การวัดความแรงของแหล่งกำเนิดเสียง เสียงที่ให้แอมพลิจูดที่สูงกว่าก็就会有ความดังมากกว่าเสียงที่แอมพลิจูดที่ต่ำกว่า

การที่เราใช้ลักษณะที่กล่าวมาทั้ง 2 ลักษณะของเสียงคือ ความถี่ และแอมพลิจูด เราก็สามารถวาดรูปแบบเสียงอย่างคร่าว ๆ ได้แล้ว ดังรูปเมื่อเราเขียนกราฟเสียงได้เราก็สามารถที่จะเปรียบเทียบมันได้เราเห็นได้ว่าจะเปรียบเทียบกราฟเสียงได้อย่างไร ถ้าเราเห็นว่าเสียงเสียงหนึ่งมีพิทช์ที่สูงกว่า หรือมีความดังมากกว่า การที่มีคาบเวลาของ 1 ลูกคลื่นที่แคบ ๆ จะหมายถึงความถี่ที่สูงกว่าและมีพิทช์ที่สูงกว่าด้วย และการที่มียอดคลื่นที่สูงกว่า ก็หมายความว่าแอมพลิจูดที่มากกว่า และความดังที่มากกว่า

เดซิเบล คือหน่วยวัดความแรงของสัญญาณ โดยจะถูกใช้เมื่อเปรียบเทียบความดังของ 2 เสียง

แม้ว่าแอมพลิจูดและพิทช์จะทำให้เราสามารถแยกแยะเสียงส่วนใหญ่ได้แต่คุณภาพของเสียงจะมีส่วนช่วยในการตัดสินใจที่ดีกว่า

คลื่นเสียงส่วนมากจะไม่ได้อยู่ในรูปแบบง่าย ๆ บางส่วนของเสียงอาจจะมีความดังหรือมีความถี่สูงมากกว่าส่วนอื่นในเสียงเดียวกัน ซึ่งเสียงส่วนมากก็จะเป็นเช่นนี้ แต่ถ้าเสียงที่ไม่ได้มีพิทช์ที่หลากหลายแต่มีพิทช์เดียว เราเรียกเสียงนั้นว่า โทน และ โทนเดี่ยว (pure tone) จะมีความถี่และแอมพลิจูดคงที่

คุณภาพของเสียง

คุณภาพของเสียงหรือเรียกว่า “สันของเสียง” เป็นคุณสมบัติของเสียงทางดนตรี มันแยกแยะเสียงดนตรีว่ามีโทนจากเครื่องดนตรีชนิดหนึ่งออกจากโทนเครื่องดนตรีชนิดอื่น ที่มีพิทช์และความถี่เดียวกัน เช่นคุณภาพเสียงจะให้เราสามารถแยกเสียงโทนเดียวกันที่เล่นโดย เปียโน กับแซกโซโฟนได้

คุณภาพของโทนเสียงดนตรี คือผลจากการรวมกันของหลากหลายความถี่ เช่น สายเปียโนอาจถูกสั่นทั้งหมด 200 ครั้ง สิ่งนี้เรียกว่า ฟาวเดเมนทอล โทน (fundamental tone) อย่างไรก็ตามฟาวเดเมนทอล โทนนี้จะถูกประกอบไปด้วยโทนอื่นของความถี่ที่สูงกว่า โทนเหล่านี้เกิดจากสายเปียโนสั่นเป็นช่วง แต่ละครึ่งของสาย แต่ละหนึ่งในสามของสาย และไปเรื่อย ๆ และ โทนความถี่ที่สูงกว่าจะเรียกว่า ฮาร์โมนิกส์ (harmonics) หรือ โอเวอร์ โทน (overtone)

ตลอดช่วงระยะเวลาฮาร์โมนิกส์จะมีแอมพลิจูดที่ต่ำกว่าฟาวเดเมนทอล โทน แต่ไม่จริงเสมอไป แต่อย่างไรก็ตามฮาร์โมนิกส์ ยังคงเป็นค่าที่คูณของความถี่ฟาวเดเมนทอล จำนวนของพลังงานของฮาร์โมนิกส์เดี่ยว ๆ จะกำหนดคุณภาพของเสียง

โดยรวมการออกแบบเครื่องดนตรีกำหนดธรรมชาติของฮาร์โมนิกส์ที่ถูกรวมเข้ากับฟาวเดเมนทอล โทนของมัน จากตัวกีตาร์ถ้ามีลักษณะแตกต่างกัน เช่น มีช่วงตัวที่อุ่นกว่าหรือเล็กกว่า จะให้คุณภาพเสียงที่โทนเดียวกันแตกต่างกัน

คอมพิวเตอร์และเสียง

เราจะเห็นว่าสามารถกำหนดลักษณะของเสียงจาก 3 สิ่ง คือ พิตช์ (ความถี่) แอมพลิจูด (ความดัง) และคุณภาพของเสียง และเพราะเห็นว่าคอมพิวเตอร์ทำการประมวลผลข้อมูล (ให้เสียงเป็นชนิดของข้อมูล)

ก่อนที่คอมพิวเตอร์ของเราจะทำการบันทึก ถ่ายเท และเล่นกลับข้อมูลเสียง เสียงจะต้องทำให้อยู่ในรูปที่คอมพิวเตอร์เข้าใจได้ คือจากเสียงที่เป็นอนาล็อกมาเป็น ข้อมูลดิจิทัลโดยการกระทำที่เรียกว่า การแปลงอนาล็อกเป็นดิจิทัล (Analog-to-Digital Conversion , ADC) โดยทำให้เสียงๆหนึ่งมาอยู่ในรูปแบบข้อมูลเป็นไบนารีทำให้เราสามารถนำพลังงานจากคอมพิวเตอร์มาประมวลผลแต่ละไบนารีของข้อมูลเสียงได้ เช่น เราสามารถ ใส่เสียงก้องลงไปได้ คอมพิวเตอร์สามารถรับรู้เสียงได้ว่าเป็นคำ "ใช่" หรือ "ไม่ใช่" สามารถ รวมเสียงจากหลายไฟล์รวมกัน ปรับเปลี่ยนพิตช์ของเสียงได้

และการแปลงดิจิทัลเป็นอนาล็อก (Digital-to-Analog Conversion , DAC) จะกระทำการเปลี่ยนข้อมูลเสียงที่เป็นไบนารีกลับไปเป็นสัญญาณอิเล็กทรอนิกส์ที่เป็นอนาล็อกและส่งออกไปยังลำโพงให้ได้เสียงออกมา

การสุ่มข้อมูล (Sampling)

จะเริ่มจากการที่เราจับเสียงที่พูดให้ได้ลงไมโครโฟน เมื่อคลื่นเสียงกระทบกับไมโครโฟนจะมีกลไกจับความเคลื่อนไหวคลื่นเสียงให้มาอยู่ในรูปสัญญาณไฟฟ้า สัญญาณนี้จะถูกเรียกว่า สัญญาณอนาล็อก เพราะว่าเป็นสัญญาณที่มีการเปลี่ยนแปลงอย่างต่อเนื่อง

การแปลงอนาล็อกเป็นดิจิทัล

จะทำการเปลี่ยนสัญญาณที่ต่อเนื่องมาเป็นรูปของลำดับค่าที่ต่อเนื่องทางดิจิทัล โดยการสุ่มข้อมูล ทำได้โดยการวัดค่าที่คงที่ของ แอมพลิจูด ของสัญญาณทุกช่วงเวลา โดยใช้อัตราในการสุ่มข้อมูลคงที่ ถ้าการวัดได้กระทำไปอย่างมีประสิทธิภาพคือใช้อัตราการสุ่มข้อมูลที่สูง ทำให้รูปแบบของสัญญาณเสียงไม่เปลี่ยนแปลงไปมากในระหว่างแต่ละข้อมูล (sample) แสดงว่าการแปลงอนาล็อกเป็นดิจิทัลจะกระทำการประมาณค่าจากอนาล็อกไปเป็นดิจิทัลได้อย่างดี

การทำการแปลงอนาล็อกเป็นดิจิทัลทำงานอย่างไร สังเกตว่าคลื่นเสียงที่ต่อเนื่องจะถูกประมาณค่าโดยค่าทางดิจิทัลที่ใช้ 8 บิต ค่าที่ได้จะเปลี่ยนจาก +128 ถึง -127 ซึ่งจะถูกรวบรวมในไฟล์เป็นไบนารีและโดยการให้ข้อมูลที่มากกว่า โดยใช้ข้อมูลขนาด 16 บิต เราจะสามารถสร้างเสียงที่มีคุณภาพได้ดีกว่า

อัตราการสุ่มข้อมูลและขนาดของข้อมูล

การที่จะให้ได้เสียงที่ผ่านการเปลี่ยนเป็นดิจิทัลโดยการแปลงดิจิทัลเป็นอนาล็อก และเปลี่ยนกลับเป็นอนาล็อก โดยการแปลงดิจิทัลเป็นอนาล็อกอีกครั้งหนึ่งให้มีความเสมือนจริงกับแหล่งต้นกำเนิดมากที่สุด (เราจะเรียกว่า ไฮไฟ) จะขึ้นอยู่กับทางเลือกใช้อัตราการสุ่มข้อมูล และขนาดของข้อมูลที่ถูกต้อง

อัตราการสุ่มข้อมูล

อัตราการสุ่มข้อมูลจะต้องมากเพียงพอสำหรับเสียงที่มีความถี่สูง เช่น เสียงกระทบของแก้วคริสตัล หรือเสียงสีไวโอลิน เสียงความถี่สูงนี้จะต้องโดยการจับมาอย่างแน่นยำ ตามทฤษฎีของไนควิสต์ว่า "เราสามารถจำลองรูปแบบคลื่นให้ได้ความน่าเชื่อถือได้ ถ้าอัตราการสุ่มข้อมูลจะต้องเป็นอย่างน้อย 2 เท่าของ ความถี่ที่สูงที่สุด" ความถี่ที่สูงที่สุดนั้นมนุษย์เราสามารถรับรู้ได้ไม่เกินกว่า 20 กิโลเฮิร์ตซ์เล็กน้อย ดังนั้นการวัดเสียง 16 บิต โดยทั่วไปจะมีอัตราการสุ่มข้อมูลสูงสุดอยู่ที่ 44.1 กิโลเฮิร์ตซ์ และด้วยอัตราเดียวกันนี้ในปัจจุบันก็ถูกใช้กับเครื่องเล่นซีดีด้วย

เราสามารถเลือกอัตราการสุ่มข้อมูลในการบันทึกไฟล์เสียง โดยอัตราการสุ่มข้อมูลที่สูงกว่าจะให้คุณภาพที่ดีกว่า เช่นการใช้อัตราการสุ่มข้อมูล 6,000 เฮิร์ตซ์ (คือมี 6,000 ข้อมูล ต่อ 1 วินาที) จะดีเพียงพอ กับเสียงของผู้ชาย แต่ถ้าเป็นเสียงของผู้หญิง ที่จะมีส่วนของความถี่ที่สูงกว่า เราควรจะใช้อัตราการสุ่มข้อมูล 8,000 เฮิร์ตซ์ จะดีกว่าและให้คุณภาพที่ดีกว่าด้วย

3.3 หน่วยเสียงพื้นฐาน (Basic Phoneme)

หน่วยเสียง (phoneme) เป็นลักษณะเฉพาะแทนเสียงที่ไม่ซ้ำกัน หน่วยเสียงพื้นฐานในภาษาไทย ได้แก่ หน่วยเสียงพยัญชนะ หน่วยเสียงสระ และหน่วยเสียงวรรณยุกต์ ซึ่งเป็นเสียงของพยัญชนะ สระ และวรรณยุกต์ที่ประกอบเป็น โครงสร้างพยางค์ไทย การเก็บบันทึกโอมารถทำได้หลายวิธี แต่ประเด็นสำคัญจะต้องนำหน่วยเสียงเหล่านี้ไปใช้ให้เกิดความชัดเจน และเมื่อรวมกันแล้วต้องสามารถให้เสียงที่ถูกต้องเหมือนเสียงพูดตามธรรมชาติ รูปแบบการเก็บบันทึกหน่วยเสียงพอสรุปได้ดังต่อไปนี้

แบบที่ 1

1. บันทึกพยัญชนะต้นเดี่ยว 21 หน่วยเสียง และพยัญชนะควบกล้ำ 17 หน่วยเสียง รวมเป็น 38 หน่วยเสียง

2. บันทึกเสียงสระเสียงเดี่ยว 9 หน่วยเสียง และสระสั้นผสม 3 หน่วยเสียง ผันเสียงวรรณยุกต์ 5 เสียง ได้ 60 หน่วยเสียง

3. บันทึกเสียงสระเสียงยาวเดี่ยว 9 หน่วยเสียง และสระยาวผสม 3 หน่วยเสียง ผันเสียงวรรณยุกต์ 5 เสียง ได้ 60 หน่วยเสียง

หน่วยเสียงรวมทั้งสิ้น $38 + 60 + 60 = 158$ หน่วยเสียง

ตัวอย่างการผสมเสียง

รัก	จะได้	ร+อะ+ก
เหนียว	จะได้	น+อิ+อำ+ว
ปลด	จะได้	ปล+อ่อ+ด

เป็นต้น

การบันทึกเสียงในแบบที่ 1 จะประหยัดเนื้อที่ในหน่วยความจำมากที่สุด แต่เสียงที่ผสมจะไม่ใช่เป็นธรรมชาติ และบางคำจะฟังไม่รู้เรื่อง

แบบที่ 2

1. บันทึกพยัญชนะต้น (เดี่ยว+ควบกล้ำ) 38 หน่วยเสียง ผสมเสียงสระ อะ ได้ทั้งหมด 38 หน่วยเสียง
2. บันทึกสระเสียงสั้น 12 หน่วยเสียง ผันเสียงวรรณยุกต์ 3 หน่วยเสียง คือ เสียงเอก เสียงโท และเสียงตรี รวมทั้งหมด 36 เฟรมเสียง
3. บันทึกสระเสียงสั้นที่มีตัวสะกดที่เป็นเสียงกัก ได้แก่ แม่กก , แม่กด และแม่กข ผันเสียงวรรณยุกต์ 3 หน่วยเสียง หน่วยเสียงวรรณยุกต์เอก เสียงโท และเสียงตรี รวมทั้งสิ้น 108 หน่วยเสียง
4. บันทึกเสียงสระสั้นผสมตัวสะกด แม่กง แม่กน แม่กบ แม่เกย แม่เกอว ผันเสียงวรรณยุกต์ 5 หน่วยเสียง รวมทั้งสิ้น 300 เฟรมเสียง
5. บันทึกสระเสียงยาว 12 หน่วยเสียง ผันเสียงวรรณยุกต์ 5 หน่วยเสียง คือเสียงสามัญ เสียงเอก เสียงโท เสียงตรี และเสียงจัตวา รวมทั้งหมด 60 เฟรมเสียง
6. บันทึกสระเสียงยาวที่มีตัวสะกดที่เป็นเสียงกัก ได้แก่ แม่กก แม่กด และแม่กข ผันเสียงวรรณยุกต์ 3 หน่วยเสียง หน่วยเสียงวรรณยุกต์เอก เสียงโท และเสียงตรี รวมทั้งสิ้น 108 หน่วยเสียง
7. บันทึกเสียงสระเสียงยาวผสมตัวสะกด แม่กง แม่กน แม่กบ แม่เกย แม่เกอว ผันเสียงวรรณยุกต์ 5 หน่วยเสียง รวมทั้งสิ้น 300 เฟรมเสียง

หน่วยเสียงรวมทั้งสิ้น $38+36+108+300+60+108+300 = 950$ หน่วยเสียง

ตัวอย่างการผสมเสียง

บ้าน	จะได้	ปะ+อ้าน
เชือด	จะได้	ชะ+เอือด
เหนียว	จะได้	นะ+เอี้ยว

เป็นต้น

แบบที่ 3

1. บันทึกเสียงพยัญชนะต้น (เดี่ยว+ควบกล้ำ) ผสมเสียงสระเสียงสั้น ได้ทั้งหมด 456 หน่วยเสียง
2. บันทึกสระเสียงสั้น 12 หน่วยเสียง ผันเสียงวรรณยุกต์ 4 หน่วยเสียง คือ เสียงเอก เสียงโท เสียงตรี และเสียงจัตวา รวมทั้งหมด 48 เฟรมเสียง
3. บันทึกสระเสียงสั้นที่มีตัวสะกดที่เป็นเสียงกัก ได้แก่ แม่กก แม่กด และแม่กข ผันเสียงวรรณยุกต์ 4 หน่วยเสียง หน่วยเสียงวรรณยุกต์เอก เสียงโท เสียงตรี และเสียงจัตวา รวมทั้งสิ้น 144 หน่วยเสียง
4. บันทึกสระเสียงสั้นผสมตัวสะกด แม่กง แม่กน แม่กบ แม่เกย แม่เกอว ผันเสียงวรรณยุกต์ 5 หน่วยเสียง รวมทั้งสิ้น 300 เฟรมเสียง

5. บันทึกสระเสียงยาว 12 หน่วยเสียง ผันเสียงวรรณยุกต์ 5 หน่วยเสียง คือ เสียงสามัญ เสียงเอก เสียงโท เสียงตรี และเสียงจัตวา รวมทั้งหมด 60 เฟ้มเสียง

6. บันทึกสระเสียงยาวที่มีตัวสะกดที่เป็นเสียงกัก ได้แก่ แม่กก แม่กด และแม่กบ ผันเสียงวรรณยุกต์ 4 หน่วยเสียง หน่วยเสียงวรรณยุกต์เอก เสียงโท เสียงตรี และเสียงจัตวา รวมทั้งสิ้น 144 หน่วยเสียง

7. บันทึกเสียงสระเสียงยาวผสมตัวสะกด แม่กง แม่กน แม่กบ แม่เกย แม่เกอว ผันเสียงวรรณยุกต์ 5 หน่วยเสียง รวมทั้งสิ้น 300 เฟ้มเสียง

รวมหน่วยเสียงทั้งสิ้น $456+48+144+300+60+144+300 = 1452$ หน่วยเสียง

ตัวอย่างการผสมเสียง

บ้าน	จะได้	ปะ+อ้าน
รัก	จะได้	ระ+อัก
เหนียว	จะได้	เนยะ+เอี้ยว

เป็นต้น

แบบที่ 4

1. บันทึกพยัญชนะต้น (เดี่ยว+ควมกล้า) 38 หน่วยเสียง ผสมสระ (สั้น+ยาว) ทั้งหมด 24 หน่วยเสียง ผันเสียงวรรณยุกต์ 5 เสียง รวม 4560 หน่วยเสียง

2. บันทึกเสียงพยัญชนะต้น (เดี่ยว+ควมกล้า) ผสมสระ และตัวสะกด ผันเสียงวรรณยุกต์ 5 เสียง รวม 36,000 หน่วยเสียง

รวมหน่วยเสียงทั้งสิ้น $4560+36000 = 40560$ หน่วยเสียง

แบบนี้คล้ายเป็นการอ่าน และเป็นคำที่ครอบคลุมทุกเสียงในภาษาไทยทั้งหมด

ตัวอย่างการออกเสียง

ประชากร	จะได้	ประ+ชา+กร
สลอน	จะได้	สะ+หลอน
บุรุษ	จะได้	บุ+หฺรุด
ชำเลื่อง	จะได้	ชำ+เลื่อง
ไปรษณีย์	จะได้	ไปร+สะ+นี

เป็นต้น

แบบที่ 5

เป็นการบันทึกเสียงเป็นคำ ๆ ซึ่งวิธีนี้จะได้เสียงที่เป็นธรรมชาติที่สุด แต่ก็จะต้องเปลืองหน่วยความจำที่สุด

จากข้อมูลการจัดเก็บทั้ง 5 รูปแบบนั้น รูปแบบที่ 3 เป็นวิธีที่เหมาะสมที่สุด แม้ว่าจะให้คุณภาพของเสียงออกมามีไม่เท่ากับแบบที่ 4 และแบบที่ 5 ไม่ได้แต่จะมีข้อเด่นกว่าตรงที่ไม่เปลืองการจัดเก็บข้อมูลในระบบฐานข้อมูลเสียง

3.4 การจัดเก็บหน่วยเสียง

การบันทึกหน่วยเสียงแบบไบโฟน (biphone) นั้นต้องทำการจัดเก็บแฟ้มเสียงเป็นจำนวนดังนี้ เสียงพยัญชนะต้นเดี่ยวและพยัญชนะควบกล้ำ ผสมสระเสียงสั้น ได้ 456 แฟ้มเสียง

สระเสียงสั้น 12 หน่วย ผันวรรณยุกต์ 4 หน่วยเสียง คือ เอก โท ตรี และจัตวา รวม 48 แฟ้มเสียง

สระเสียงสั้นที่มีตัวสะกดเสียงกัก ได้แก่ แม่ก ก แม่ก ด แม่ก บ ผันวรรณยุกต์ 2 หน่วยเสียง คือ เอก และตรี รวม 72 แฟ้มเสียง

สระเสียงสั้นที่มีตัวสะกดเสียงไม่กัก ได้แก่ แม่ก น แม่ก ง แม่ก ม แม่ก ย แม่ก ว ผันวรรณยุกต์ 5 หน่วยเสียง คือ สามัญ เอก โท ตรี และจัตวา รวม 300 แฟ้มเสียง

สระเสียงยาว 12 หน่วย ผันวรรณยุกต์ 5 หน่วยเสียง คือ สามัญ เอก โท ตรี และจัตวา รวม 60 แฟ้มเสียง ***

สระเสียงยาวที่มีตัวสะกดเสียงกัก ได้แก่ แม่ก ก แม่ก ด แม่ก บ ผันวรรณยุกต์ 3 หน่วยเสียง คือ เอก โท และตรี รวม 108 แฟ้มเสียง ***

สระเสียงยาวที่มีตัวสะกดเสียงไม่กัก ได้แก่ แม่ก น แม่ก ง แม่ก ม แม่ก ย แม่ก ว ผันวรรณยุกต์ 5 หน่วยเสียง คือ สามัญ เอก โท ตรี และจัตวา รวม 300 แฟ้มเสียง ***

รวมต้องมีแฟ้มเสียงทั้งหมด $444 + 48 + 72 + 300 + 60 + 108 + 300 = 1332$ แฟ้มเสียง ***

การเอาเสียงสระสั้นมาผสมกันให้เป็นเสียงสระยาวนั้นต้องคำนึงถึงสิ่งต่อไปนี้ด้วย เสียงวรรณยุกต์ของเสียงสระสั้นที่เอามาผสมนั้นต้องตรงกับเสียงสระยาวที่ต้องการ ตัวสะกดของเสียงสระสั้นที่เอามาประกอบกันนั้นต้องเป็นตัวสะกดที่เป็นแม่เดียวกัน

	อะ	อิ	อุ	อี	เอะ	โอะ	เอาะ	เออะ	แอะ	ัวะ	เือะ	เือะ
	(a)	(i)	(u)	(v)	(e)	(o)	(@)	(#)	(x)	(ua)	(va)	(ia)
ก (k)	ka	ki	ku	kv	ke	ko	k@	k#	kae	kua	kva	kia
ค,ข,ช (kh)	kha	khi	khu	khv	khe	kho	kh@	kh#	khae	khua	khva	khia
ง (ng)	nga	ng i	ngu	ngv	nge	ngo	ng@	ng#	ngae	ngua	ngva	ngia
จ,ช,ฉ (sh)	sha	shi	shu	shv	she	sho	sh@	sh#	shae	shua	shva	shia
จ (ch)	cha	chi	chu	chv	che	cho	ch@	ch#	chae	chua	chva	chia
ซ,ส,ษ,ศ (s)	sa	si	su	sv	se	so	s@	s#	sae	sua	sva	sia

	อะ (a)	อิ (i)	อุ (u)	อี (v)	เอะ (e)	โอะ (o)	เอาะ (@)	เอดะ (#)	แอะ (x)	ัวะ (ua)	เื้อะ (va)	เือะ (ia)
จ, ช (j)	ja	ji	ju	jv	je	jo	j@	j#	jae	jua	jva	jia
ด, ต (d)	da	di	du	dv	de	do	d@	d#	dae	dua	dva	dia
ต, ถ (t)	ta	ti	tu	tv	te	to	t@	t#	tae	tua	tva	tia
ฐ, ถ, ท, ฑ, ท (th)	tha	thi	thu	thv	the	tho	th@	th#	thae	thua	thva	thia
น, ณ (n)	na	ni	nu	nv	ne	no	n@	n#	nae	nua	nva	nia
บ (b)	ba	bi	bu	bv	be	bo	b@	b#	bae	bua	bva	bia
ป (p)	pa	pi	pu	pv	pe	po	p@	p#	paе	pua	pva	pia
พ, ผ (ph)	pha	phi	phu	phv	phe	pho	ph@	ph#	phae	phua	phva	phia
ฝ, ฟ (f)	fa	fi	fu	fv	fe	fo	f@	f#	fae	fua	fva	fia
ม (m)	ma	mi	mu	mv	me	mo	m@	m#	mae	mua	mva	mia
ร (r)	ra	ri	ru	rv	re	ro	r@	r#	rae	rua	rva	ria
ล (l)	la	li	lu	lv	le	lo	l@	l#	lae	lua	lva	lia
ว (w)	wa	wi	wu	wv	we	wo	w@	w#	wae	wua	wva	wia
อ (?)	?a	?i	?u	?v	?e	?o	?@	?#	?ae	?ua	?va	?ia
ฮ, ห (h)	ha	hi	hu	hv	he	ho	h@	h#	hae	hua	hva	hia
ก, ข (kw)	kwa	kwi	kwu	kwv	kwe	kwo	kw@	kw#	kwae	kwua	kwva	kwia
กร (kr)	kra	kri	kru	krv	kre	kro	kr@	kr#	krae	krua	krva	kria
กล (kl)	kla	kli	klu	klv	kle	klo	kl@	kl#	klae	klua	klva	klia
ค, ฆ (khr)	khra	khri	khru	khv	khre	khro	kh@	kh#	khrae	khrua	khvua	khria
ค, ฆ (khl)	khla	khli	khlu	khv	khle	khlo	kh@	kh#	khlae	khlua	khvua	khlia
พ, ฐ (phr)	phra	phri	phru	phv	phre	phro	ph@	ph#	phrae	phrua	phvua	phria
พ, ฐ (phl)	phla	phli	phlu	phv	phle	phlo	ph@	ph#	phlae	phlua	phvua	phlia
ป, ฐ (pr)	pra	pri	pru	prv	pre	pro	pr@	pr#	prae	prua	prva	pria
ป, ฐ (pl)	pla	pli	plu	plv	ple	plo	pl@	pl#	plae	plua	plva	plia
ต, ฐ (tr)	tra	tri	tru	trv	tre	tro	tr@	tr#	trae	trua	trva	tria
ท, ฐ (thr)	thra	thri	thru	thv	thre	thro	thr@	thr#	thrae	thrua	thvua	thria
ด, ฐ (dr)	dra	dri	dru	drv	dre	dro	dr@	dr#	drae	drua	drva	dria
ฟ, ฐ (fr)	fra	fri	fru	frv	fre	fro	fr@	fr#	frae	frua	frva	fria
ฟ, ฐ (fl)	fla	fli	flu	flv	fle	flo	fl@	fl#	flae	flua	flva	flia

	อะ (a)	อิ (i)	อุ (u)	อี (v)	เอะ (e)	โอะ (o)	เอาะ (@)	เอะ (#)	แอะ (x)	อัวะ (ua)	เอือะ (va)	เอียะ (ia)
บร (br)	bra	bri	bru	brv	bre	bro	br@	br#	brae	brua	brva	bria
บล (bl)	bla	bli	blu	blv	ble	blo	bl@	bl#	blae	blua	blva	blia

ตารางที่ 3.1 เสียงพยัญชนะต้นเดี่ยวและพยัญชนะควบกล้ำผสมสระเสียงสั้น รวม 444 แฟ้มเสียง

	อะ (a)	อิ (i)	อุ (u)	อี (v)	เอะ (e)	โอะ (o)	เอาะ (@)	เอะ (#)	แอะ (ae)	อัวะ (ua)	เอือะ (va)	เอียะ (ia)
เอก	a1	i1	u1	v1	e1	o1	@1	#1	ae1	ua1	va1	ia1
โท	a2	i2	u2	v2	e2	o2	@2	#2	ae2	ua2	va2	ia2
ตรี	a3	i3	u3	v3	e3	o3	@3	#3	ae3	ua3	va3	ia3
จัตวา	a4	i4	u4	v4	e4	o4	@4	#4	ae4	ua4	va4	ia4

ตารางที่ 3.2 สระเสียงสั้น (ไม่มีตัวสะกด) รวม 48 แฟ้มเสียง

	อะ (a)	อิ (i)	อุ (u)	อี (v)	เอะ (e)	โอะ (o)	เอาะ (@)	เอะ (#)	แอะ (ae)	อัวะ (ua)	เอือะ (va)	เอียะ (ia)
แม้มก เอก	ak1	ik1	uk1	vk1	ek1	ok1	@k1	#k1	ae1	uak1	vak1	iak1
แม้มก ตรี	ak3	ik3	uk3	vk3	ek3	ok3	@k3	#k3	ae3	uak3	vak3	iak3
แม้มก เอก	at1	it1	ut1	vt1	et1	ot1	@t1	#t1	act1	uat1	vat1	iat1
แม้มก ตรี	at3	it3	ut3	vt3	et3	ot3	@t3	#t3	act3	uat3	vat3	iat3
แม้มกบ เอก	ap1	ip1	up1	vp1	ep1	op1	@p1	#p1	aep1	uap1	vap1	iap1
แม้มกบ ตรี	ap3	ip3	up3	vp3	ep3	op3	@p3	#p3	aep3	uap3	vap3	iap3

ตารางที่ 3.3 สระเสียงสั้น (ตัวสะกดเสียงกัก) รวม 72 แฟ้มเสียง

	อะ (a)	อิ (i)	อุ (u)	อี (v)	เอะ (e)	โอะ (o)	เอาะ (@)	เอะ (#)	แอะ (ae)	อัวะ (ua)	เอือะ (va)	เอียะ (ia)
แม้มกน สามัญ	an1	in1	un1	vn1	en1	on1	@n1	#n1	aen1	uan1	van1	ian1
เอก	an1	in1	un1	vn1	en1	on1	@n1	#n1	aen1	uan1	van1	ian1
โท	an2	in2	un2	vn2	en2	on2	@n2	#n2	aen2	uan2	van2	ian2

	อะ (a)	อิ (i)	อุ (u)	อี (v)	เอะ (e)	โอะ (o)	เอาะ (@)	เอยะ (#)	แอะ (ae)	อัวะ (ua)	เอือะ (va)	เอียะ (ia)
ตรี	an3	in3	un3	vn3	en3	on3	@n3	#n3	aen3	uan3	van3	ian3
จัตวา	an4	in4	un4	vn4	en4	on4	@n4	#n4	aen4	uan4	van4	ian4
แม่กง สามัญ	ang1	ing1	ung1	vng1	eng1	Eng1	@ng1	#ng1	aeng1	uang1	vang1	iang1
เอก	ang1	ing1	ung1	vng1	eng1	Eng1	@ng1	#ng1	aeng1	uang1	vang1	iang1
โท	ang2	ing2	ung2	vng2	eng2	Eng2	@ng2	#ng2	aeng2	uang2	vang2	iang2
ตรี	ang3	ing3	ung3	vng3	eng3	Eng3	@ng3	#ng3	aeng3	uang3	vang3	iang3
จัตวา	ang4	ing4	ung4	vng4	eng4	Eng4	@ng4	#ng4	aeng4	uang4	vang4	iang4
แม่กม สามัญ	am1	im1	um1	vm1	em1	Om1	@m1	#m1	aem1	uam1	vam1	iam1
เอก	am1	im1	um1	vm1	em1	Om1	@m1	#m1	aem1	uam1	vam1	iam1
โท	am2	im2	um2	vm2	em2	Om2	@m2	#m2	aem2	uam2	vam2	iam2
ตรี	am3	im3	um3	vm3	em3	Om3	@m3	#m3	aem3	uam3	vam3	iam3
จัตวา	am4	im4	um4	vm4	em4	Om4	@m4	#m4	aem4	uam4	vam4	iam4
แม่เกย สามัญ	aj1	ij1	uj1	vj1	ej1	oj1	@j1	#j1	aej1	uaj1	vaj1	iaj1
เอก	aj1	ij1	uj1	vj1	ej1	oj1	@j1	#j1	aej1	uaj1	vaj1	iaj1
โท	aj2	ij2	uj2	vj2	ej2	oj2	@j2	#j2	aej2	uaj2	vaj2	iaj2
ตรี	aj3	ij3	uj3	vj3	ej3	oj3	@j3	#j3	aej3	uaj3	vaj3	iaj3
จัตวา	aj4	ij4	uj4	vj4	ej4	oj4	@j4	#j4	aej4	uaj4	vaj4	iaj4
แม่ เกอว สามัญ	aw1	iw1	uw1	vw1	ew1	ow1	@w1	#w1	aw1	uaw1	vaw1	iaw1
เอก	aw1	iw1	uw1	vw1	ew1	ow1	@w1	#w1	aw1	uaw1	vaw1	iaw1
โท	aw2	iw2	uw2	vw2	ew2	ow2	@w2	#w2	aw2	uaw2	vaw2	iaw2
ตรี	aw3	iw3	uw3	vw3	ew3	ow3	@w3	#w3	aw3	uaw3	vaw3	iaw3
จัตวา	aw4	iw4	uw4	vw4	ew4	ow4	@w4	#w4	aw4	uaw4	vaw4	iaw4

ตารางที่ 3.4 สระเสียงสั้น (ตัวสะกดเสียงไม่กัก) รวม 300 แฟ้มเสียง

	อา (a:)	อี (i:)	อุ (u:)	อึ (v:)	เอ (e:)	โอ (o:)	ออ (@:)	เออ (#:)	แอ (x:)	อัว (u:a)	เอือ (v:a)	เอีย (i:a)
เอก	a:1	i:1	u:1	v:1	e:1	o:1	@:1	#:1	x:1	u:a1	v:a1	i:a1
โท	a:2	i:2	u:2	v:2	e:2	o:2	@:2	#:2	x:2	u:a2	v:a2	i:a2
ตรี	a:3	i:3	u:3	v:3	e:3	o:3	@:3	#:3	x:3	u:a3	v:a3	i:a3
จัตวา	a:4	i:4	u:4	v:4	e:4	o:4	@:4	#:4	x:4	u:a4	v:a4	i:a4

ตารางที่ 3.5 สระเสียงยาว (ไม่มีตัวสะกด) รวม 60 แฝ้มเสียง

	อา (a:)	อี (i:)	อุ (u:)	อึ (v:)	เอ (e:)	โอ (o:)	ออ (@:)	เออ (#:)	แอ (x:)	อัว (u:a)	เอือ (v:a)	เอีย (i:a)
แมกก เอก	a:k1	i:k1	u:k1	v:k1	e:k1	o:k1	@:k1	#:k1	x:k1	u:ak1	v:ak1	i:ak1
แมกก โท	a:k2	i:k2	u:k2	v:k2	e:k2	o:k2	@:k2	#:k2	x:k2	u:ak2	v:ak2	i:ak2
แมกก ตรี	a:k3	i:k3	u:k3	v:k3	e:k3	o:k3	@:k3	#:k3	x:k3	u:ak3	v:ak3	i:ak3
แมกค เอก	a:t1	i:t1	u:t1	v:t1	e:t1	o:t1	@:t1	#:t1	x:t1	u:at1	v:at1	i:at1
แมกค โท	a:t2	i:t2	u:t2	v:t2	e:t2	o:t2	@:t2	#:t2	x:t2	u:at2	v:at2	i:at2
แมกค ตรี	a:t3	i:t3	u:t3	v:t3	e:t3	o:t3	@:t3	#:t3	x:t3	u:at3	v:at3	i:at3
แมกบ เอก	a:p1	i:p1	u:p1	v:p1	e:p1	o:p1	@:p1	#:p1	x:p1	u:ap1	v:ap1	i:ap1
แมกบ โท	a:p2	i:p2	u:p2	v:p2	e:p2	o:p2	@:p2	#:p2	x:p2	u:ap2	v:ap2	i:ap2
แมกบ ตรี	a:p3	i:p3	u:p3	v:p3	e:p3	o:p3	@:p3	#:p3	x:p3	u:ap3	v:ap3	i:ap3

ตารางที่ 3.6 สระเสียงยาว (ตัวสะกดเสียงกัก) รวม 108 แฝ้มเสียง

แมกน สามัญ	a:n1	i:n1	u:n1	v:p1	e:n1	o:n1	@:n1	#:n1	x:n1	u:an1	v:an1	i:an1
เอก	a:n1	i:n1	u:n1	v:p1	e:n1	o:n1	@:n1	#:n1	x:n1	u:an1	v:an1	i:an1
โท	a:n2	i:n2	u:n2	v:n2	e:n2	o:n2	@:n2	#:n2	x:n2	u:an2	v:an2	i:an2
ตรี	a:n3	i:n3	u:n3	v:n3	e:n3	o:n3	@:n3	#:n3	x:n3	u:an3	v:an3	i:an3
จัตวา	a:n4	i:n4	u:n4	v:n4	e:n4	o:n4	@:n4	#:n4	x:n4	u:an4	v:an4	i:an4

แม่กง สามัญ	a:ng1	i:ng1	u:ng1	v:ng1	e:ng1	e:ng1	@:ng1	#:ng1	x:ng1	u:ang1	v:ang1	i:ang1
เอก	a:ng1	i:ng1	u:ng1	v:ng1	e:ng1	e:ng1	@:ng1	#:ng1	x:ng1	u:ang1	v:ang1	i:ang1
โท	a:ng2	i:ng2	u:ng2	v:ng2	e:ng2	e:ng2	@:ng2	#:ng2	x:ng2	u:ang2	v:ang2	i:ang2
ตรี	a:ng3	i:ng3	u:ng3	v:ng3	e:ng3	e:ng3	@:ng3	#:ng3	x:ng3	u:ang3	v:ang3	i:ang3
จัตวา	a:ng4	i:ng4	u:ng4	v:ng4	e:ng4	e:ng4	@:ng4	#:ng4	x:ng4	u:ang4	v:ang4	i:ang4
แม่กม สามัญ	a:m1	i:m1	u:m1	v:m1	e:m1	o:m1	@:m1	#:m1	x:m1	u:am1	v:am1	i:am1
เอก	a:m1	i:m1	u:m1	v:m1	e:m1	o:m1	@:m1	#:m1	x:m1	u:am1	v:am1	i:am1
โท	a:m2	i:m2	u:m2	v:m2	e:m2	o:m2	@:m2	#:m2	x:m2	u:am2	v:am2	i:am2
ตรี	a:m3	i:m3	u:m3	v:m3	e:m3	o:m3	@:m3	#:m3	x:m3	u:am3	v:am3	i:am3
จัตวา	a:m4	i:m4	u:m4	v:m4	e:m4	o:m4	@:m4	#:m4	x:m4	u:am4	v:am4	i:am4
แม่เกย สามัญ	a:j1	i:j1	u:j1	v:j1	e:j1	o:j1	@:j1	#:j1	x:j1	u:aj1	v:aj1	i:aj1
เอก	a:j1	i:j1	u:j1	v:j1	e:j1	o:j1	@:j1	#:j1	x:j1	u:aj1	v:aj1	i:aj1
โท	a:j2	i:j2	u:j2	v:j2	e:j2	o:j2	@:j2	#:j2	x:j2	u:aj2	v:aj2	i:aj2
ตรี	a:j3	i:j3	u:j3	v:j3	e:j3	o:j3	@:j3	#:j3	x:j3	u:aj3	v:aj3	i:aj3
จัตวา	a:j4	i:j4	u:j4	v:j4	e:j4	o:j4	@:j4	#:j4	x:j4	u:aj4	v:aj4	i:aj4
แม่ เกว สามัญ	a:w1	i:w1	u:w1	v:w1	e:w1	o:w1	@:w1	#:w1	x:w1	u:aw1	v:aw1	i:aw1
เอก	a:w1	i:w1	u:w1	v:w1	e:w1	o:w1	@:w1	#:w1	x:w1	u:aw1	v:aw1	i:aw1
โท	a:w2	i:w2	u:w2	v:w2	e:w2	o:w2	@:w2	#:w2	x:w2	u:aw2	v:aw2	i:aw2
ตรี	a:w3	i:w3	u:w3	v:w3	e:w3	o:w3	@:w3	#:w3	x:w3	u:aw3	v:aw3	i:aw3
จัตวา	a:w4	i:w4	u:w4	v:w4	e:w4	o:w4	@:w4	#:w4	x:w4	u:aw4	v:aw4	i:aw4

ตารางที่ 3.7 สระเสียงยาว (ตัวสะกดเสียงไม่กัก) รวม 300 แห่งเสียง

บทที่ 4

การบันทึกเสียงและการแบ่งเสียงออกเป็นไบโโฟน

4.1 บทนำ

ในบทนี้จะกล่าวถึงวิธีการบันทึกเสียงเข้าสู่เครื่องคอมพิวเตอร์ และการแบ่งเสียงที่บันทึกนั้นออกเป็นไบโโฟน

4.2 รูปแบบของไฟล์

ในการทดลองนี้จะทำการบันทึกเสียงลงในเครื่องคอมพิวเตอร์ เพื่อใช้ในการทดลองสังเคราะห์เป็นคำใหม่ โดยสัญญาณเสียงที่ทำการบันทึกจะอยู่ในรูปแบบของไฟล์ .wav ที่ Sampling Rate เท่ากับ 44 kHz ในแบบ 16-bit Mono

4.3 การบันทึกเสียง

การบันทึกเสียงนี้เป็นขั้นตอนแรกในการทำงาน อุปกรณ์ที่ใช้ในการบันทึกเสียงก็คือ ไมโครโฟนสำหรับคอมพิวเตอร์ และใช้โปรแกรม Cool Edit 2000 ในการบันทึกเสียง

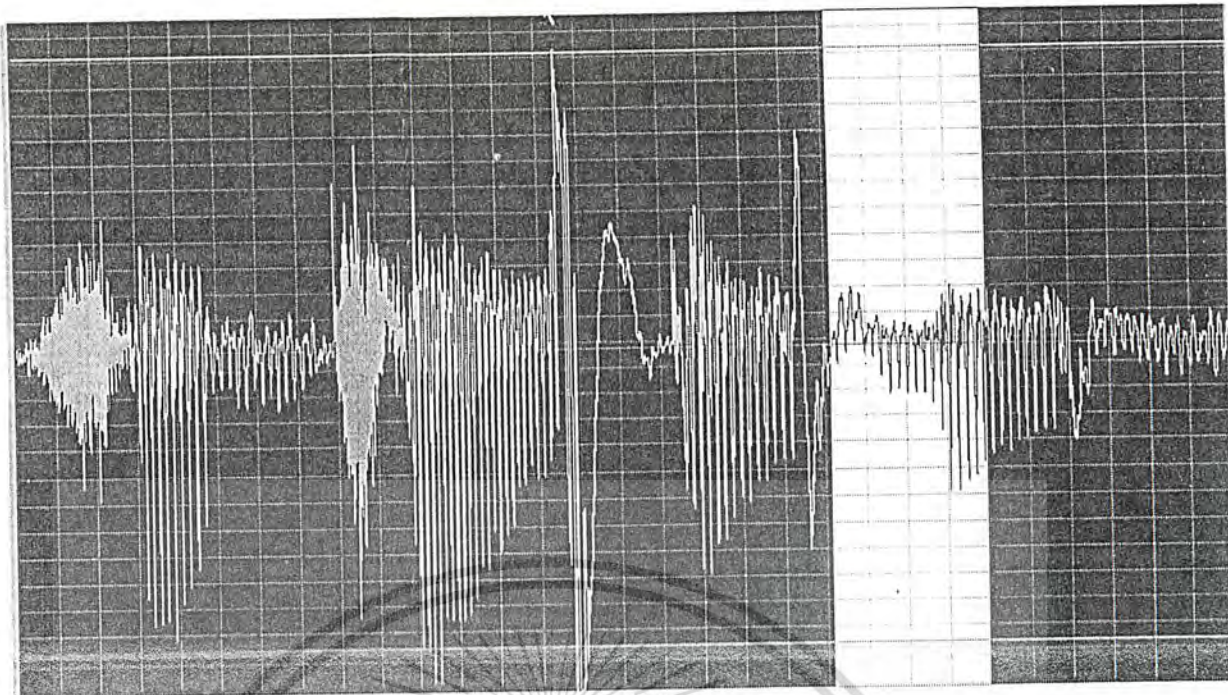
เสียงที่ทำการบันทึกนี้อาจอยู่ในลักษณะเป็นคำคำเดียวหรือเป็นประโยคก็ได้ แต่จากการทดลองได้พบว่าการบันทึกเป็นประโยคจะให้ผลลัพธ์ออกมาดีกว่าการบันทึกเป็นคำคำเดียว เพราะว่าการบันทึกเป็นคำคำเดียวนั้น โดยธรรมชาติของมนุษย์แล้วจะลากเสียงออกไปมากกว่าปกติ ไม่เหมือนที่พูดคุยกันปกติ ซึ่งเสียงจะออกมาสั้นกว่าการพูดคำเดียว ๆ

แต่การบันทึกเสียงในลักษณะเป็นประโยคก็มีข้อเสียเช่นกัน คือเสียงอาจจะเพี้ยนได้ง่าย เพราะในการฟังเสียงที่บันทึกเป็นประโยคนั้น เราสามารถรับรู้ได้ว่าคำแต่ละคำนั้นควรจะได้ยินว่าอะไร เนื่องจากความหมายของประโยคจะทำให้เราสามารถเข้าใจคำนั้น ๆ ได้ แต่หากเลือกฟังเฉพาะบางคำจากประโยคจะพบว่าคำคำนั้นออกเสียงเพี้ยนไป จึงต้องระวังในการบันทึกเสียง

4.4 การแบ่งเสียงออกเป็นไบโโฟน

เมื่อทำการบันทึกเสียงแล้ว ขั้นตอนนี้เป็นการแบ่งเสียงที่ได้บันทึกไว้ออกเป็นไบโโฟนตามที่เราต้องการ โดยใช้โปรแกรม Cool Edit 2000 เลือกช่วงที่เป็นไบโโฟนที่ต้องการ แล้วลองฟังดูว่าไบโโฟนที่ต้องการหรือไม่ ถ้าไม่ใช่ก็ลองเลื่อนไปเลือกช่วงอื่น ๆ จนกว่าจะได้ไบโโฟนที่ต้องการจริง ๆ ดังรูปที่ 4.3

ตำแหน่งเสียงที่แบ่งเป็นไบโโฟนนั้น รูปคลื่นเสียงด้านหน้าจะเป็นส่วนพยัญชนะต้นกับสระกึ่งหนึ่ง และส่วนหลังของ รูปคลื่นเสียงนั้นคือ สระกึ่งที่เหลือและตัวสะกดรวมวรรณยุกต์



รูปที่ 4.1 การแบ่งเสียงเป็นไบโโฟน

4.5 สรุปหลักในการบันทึกเสียง

1. เลือกอุปกรณ์ (ไมโครโฟน) ที่มีคุณภาพดีที่สุด เพราะไมโครโฟนทุกตัวจะให้คุณภาพเสียงไม่เท่ากัน และเมื่อเสียงเข้ากับพอร์ตพร้อมที่จะทำงานแล้ว จะมีเสียงรบกวนในตัวของมันเองในระดับหนึ่ง เราจะต้องเลือกให้มีเสียงรบกวนน้อยที่สุด

2. ทำการบันทึกเสียงในห้องที่เงียบที่สุด อย่าให้มีเสียงรบกวนจากภายนอก

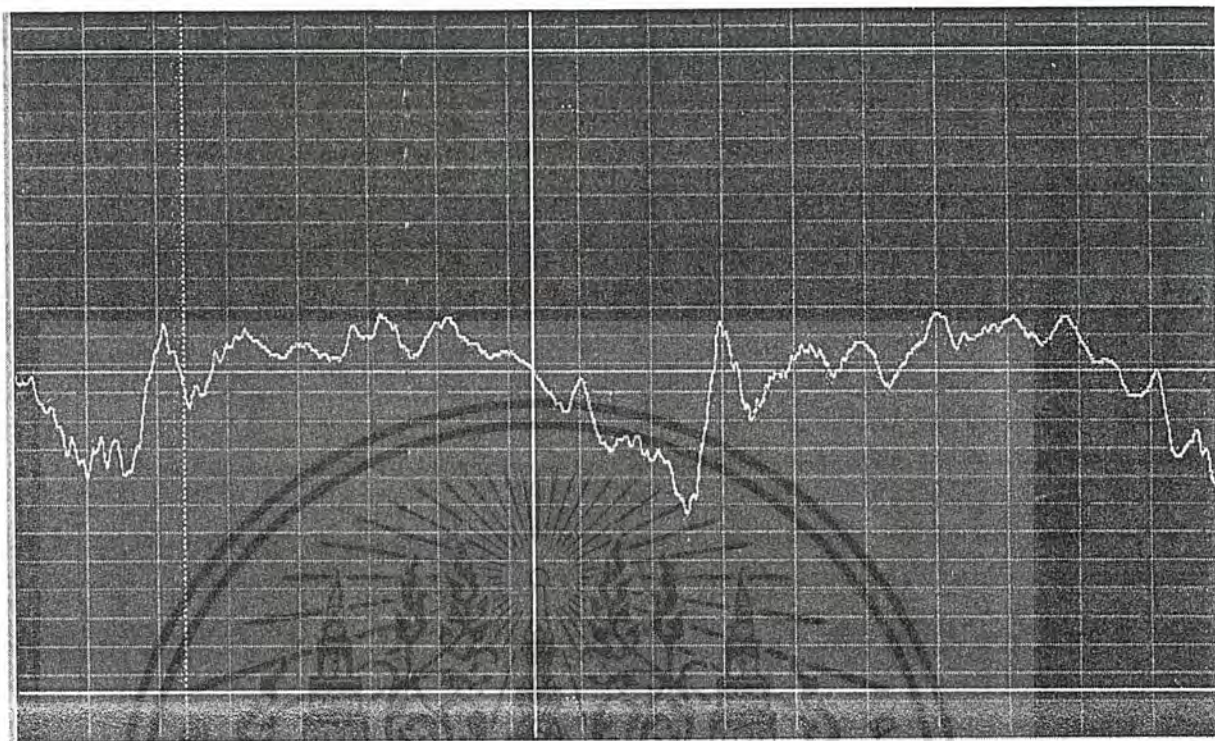
3. การบันทึกเสียงแต่ละครั้งต้องพยายามออกเสียงให้ชัดถ้อยชัดคำมากที่สุด อย่าให้เพี้ยนทั้งเสียงพยัญชนะต้น โดยเฉพาะควบกล้ำหรือเสียงสระที่ผันวรรณยุกต์แล้ว และโดยเฉพาะเสียงพยัญชนะสะกดที่ผันวรรณยุกต์ก็จะมีเพี้ยนได้ง่ายโดยไม่รู้ตัว เพราะเมื่อฟังเป็นคำดั้งเดิม (ที่บันทึกไว้) ก็จะฟังออกดี แต่พอแบ่งเป็นส่วนๆ แล้วนำไปผสมกับส่วนอื่นที่มาจากคำอื่น จึงจะรู้ว่าเสียงนั้นเพี้ยนออกไปเพราะภาษาไทยคำเสียงวรรณยุกต์เปลี่ยน ไปก็ทำให้ความหมายของคำนั้นเปลี่ยนไปด้วย เช่นเรามีเสียงพยัญชนะต้น /ล/ มีเสียงสระอาที่ผันวรรณยุกต์ตรีและเราตัดเสียงพยัญชนะสะกด /น/ ที่ผันเสียงวรรณยุกต์ตรีมาจากคำว่า ลีน แล้วนำมาผสมกัน เสียงที่ควรจะได้คือ ลีน แต่บางทีอาจออกเสียงเป็น หลาน ได้ ต้องระวังให้ดี

4. การบันทึกเสียงแต่ละครั้งต้องพยายามให้ระดับเสียงหรือความดังของเสียง (Amplitude) ใกล้เคียงกันมากที่สุด เพราะถ้าต่างกันมากเกินไปอาจทำให้เพี้ยนได้ ซึ่งตรงนี้อาจแก้ไขได้โดยการใช้คำสั่งในโปรแกรมที่ใช้บันทึกเสียงในการลดหรือเพิ่มแอมพลิจูดให้อยู่ในระดับที่ต้องการ

4.6 สรุปหลักในการแบ่งไฟล์เสียง

1. จะต้องแบ่งให้ตรงตามเสียงที่ต้องการ ไม่มีส่วนที่ไม่ต้องการติดมาด้วย เช่น ถ้าต้องการเสียงพยัญชนะต้น /ล/ ก็อย่าให้มีเสียงสระหรือวรรณยุกต์ติดมา

2. เนื่องจากไฟล์เสียงที่เราบันทึกไว้จะมีลักษณะเป็นคลื่นรูปไซน์ (sine wave) เวลาแบ่งจะต้องแบ่งตรงจุดที่ครบลูกคลื่นพอดี (crossing zero) ในทุกไฟล์ ดังรูปที่ 4.4



รูปที่ 4.2 การแบ่งไฟล์เสียงที่ถูกต้อง

3. ในบางครั้งแม้ว่าเราจะแบ่งตรงจุดที่ครบลูกคลื่นพอดีแล้วก็ตาม เสียงที่ได้ก็อาจจะยังไม่ดีพอ เราจะต้องสังเกตให้ดี แล้วลองเลื่อนไปหรือเลื่อนกลับไปยังลูกคลื่นถัดไปแทน

มาถึงตอนนี้เราก็ได้ไฟล์เสียงที่เป็นโบ โฟน ที่จะนำมาใช้ในการสังเคราะห์เสียงภาษาไทยได้ ไฟล์เสียงเหล่านี้จะถูกเก็บไว้เป็นฐานข้อมูลในการให้โปรแกรมสังเคราะห์เสียงภาษาไทยนำมารวมกันเป็นคำ และเป็นประโยคต่อไป

บทที่ 5

การสังเคราะห์เสียง

5.1 บทนำ

ในบทนี้จะกล่าวถึงการทดลองสังเคราะห์เสียงและผลการทดลอง เพื่อหาวิธีการสังเคราะห์เสียงที่ทำให้ได้เสียงที่มีความเป็นธรรมชาติมากที่สุด

5.2 วิธีการผสมคำในการทำการทดลอง

ในการทดลองที่ได้ทำนี้ เราได้นำคำที่ได้ตัดแบ่งไว้ มาทำการต่อกันเพื่อให้เป็นคำใหม่ดังนี้



โดยทำการแบ่งคำหนึ่งคำออกเป็น สองส่วน โดยในส่วนแรกนั้น จะเป็นเสียงพยัญชนะต้น เช่น
ช
ก
กร
ปร

โดยเสียงที่ยกตัวอย่างมานี้จะเป็นเสียงพยัญชนะต้น แต่ว่าเสียงของสระมาด้วยกึ่งหนึ่ง ถ้าเรานำแต่เฉพาะเสียงพยัญชนะต้นมานั้น ในการผสมคำ จะได้เสียงที่ไม่ค่อยเป็นธรรมชาติมาก เสียงจะเหนือ ๆ บางคำไม่สามารถฟังเข้าใจได้

ดังนั้น เสียงที่เรายกตัวอย่างมานั้น จึงกล่าวได้ดังนี้

- ช ถ้าเราจะนำไปรวมเป็น คำว่า ช้า เราต้องตัดเป็น ชะ
- ก ถ้าเราจะนำไปรวมเป็น คำว่า กี่ เราต้องตัดเป็น กิ
- กร ถ้าเราจะนำไปรวมเป็น คำว่า กราบ เราต้องตัดเป็น กระ

ดังตัวอย่างจะเห็นได้ว่าในส่วนของพยัญชนะต้นจะต้องมีการเพิ่มในส่วนเสียงของสระเข้าไปด้วย จึงจะได้เสียงที่เป็นธรรมชาติที่สุด

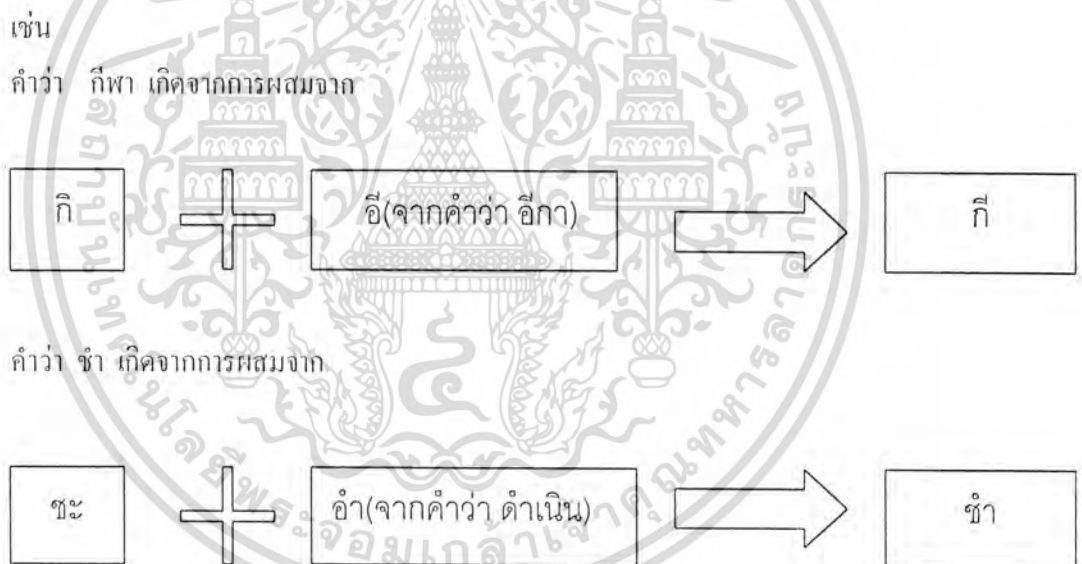
ในส่วนของสระและตัวสะกดนั้น ก็เช่นเดียวกัน เราต้องตัดเสียงสระมาครึ่งหนึ่งในครั้งแรก
เช่น

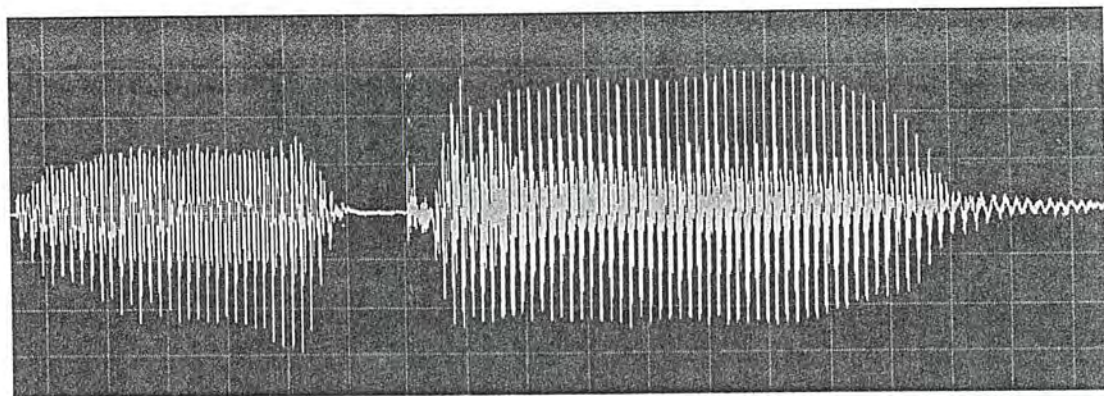
- คำว่า การ เกิดจากการผสมจาก ก + ร = กระ + อาน กลายเป็นคำว่า การ
- คำว่า สวย เกิดจากการผสมจาก ส + ว = สัวะ + อัว กลายเป็นคำว่า สวย

จากผลการทดลองนั้นคำที่เกิดจากการผสมระหว่าง พยัญชนะต้นที่แตกต่างกันแต่ออกเสียงสระเดียวกัน กับ สระ ตัวสะกด วรรณยุกต์ที่ต่างกัน

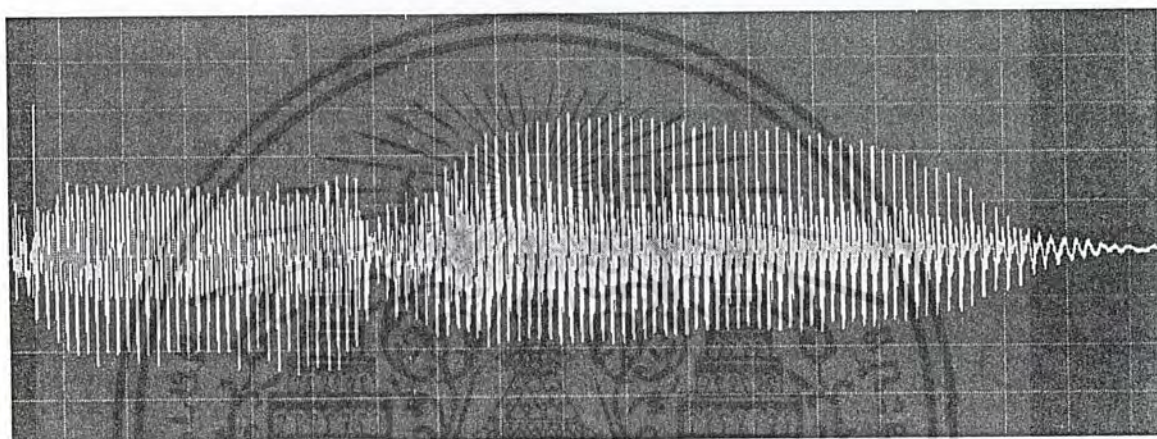
เช่น ชะ + อำ = ชำ
 ชะ + อำ = ฉำ
 ชะ + อำ = ฉำ
 กิ + อี = กิ
 กิ + อี = กิ
 กิ + อี = กิ เป็นต้น

จากคำที่ได้ยกตัวอย่างให้ดูนี้และจากการผสมคำที่มีพยัญชนะต้นต่างกัน ผลที่ได้จากการฟังนั้น เสียงที่ออกมาไม่ได้แตกต่างจากเสียงที่ไม่ได้เกิดจากการผสมเลย เสียงใกล้เคียงกันมาก แทบไม่แตกต่างกันเลย

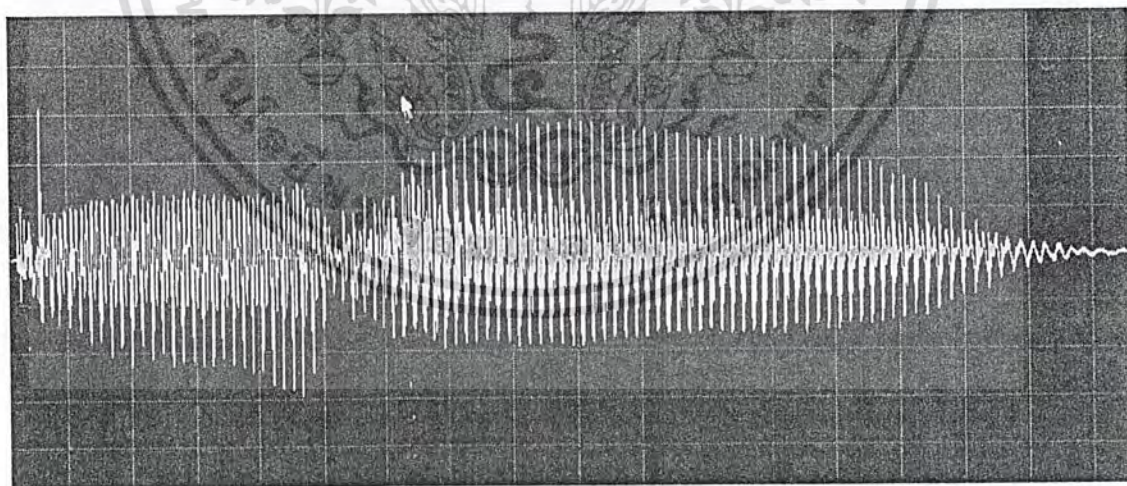




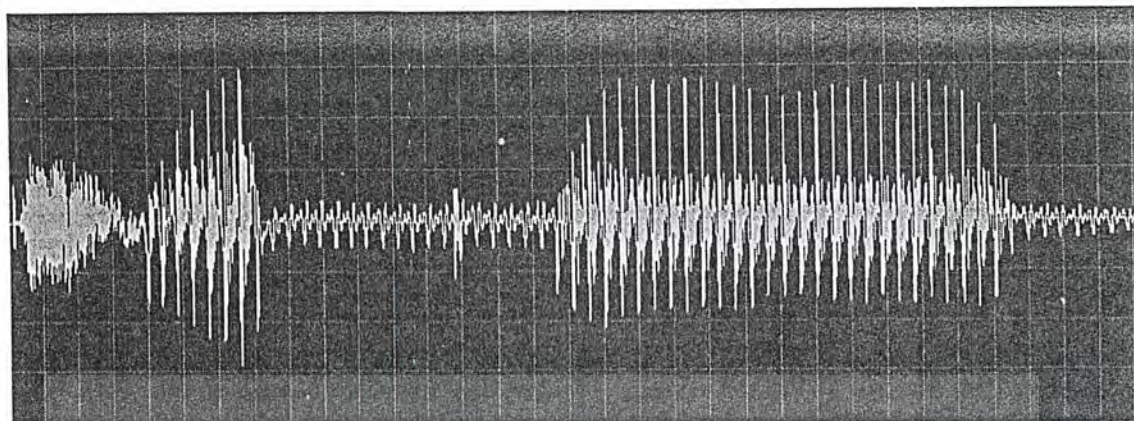
รูปที่ 5.1 รูปเสียงคำว่า กีฬา



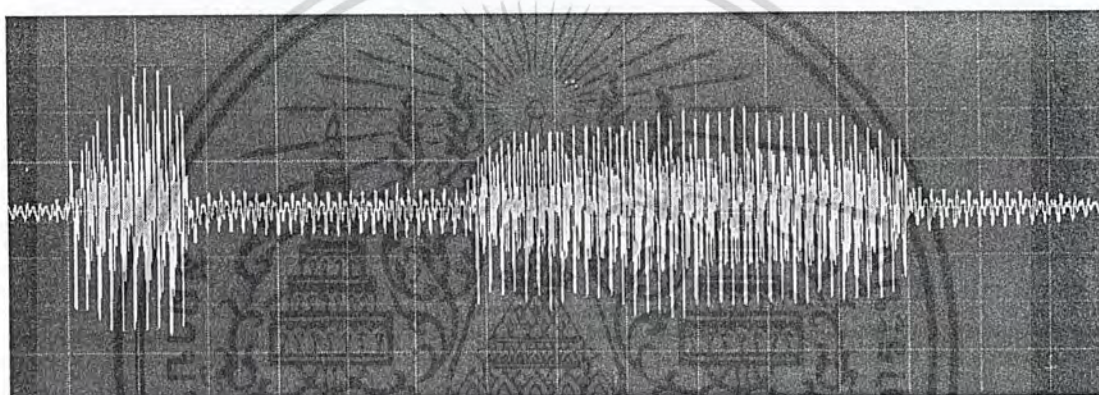
รูปที่ 5.2 รูปเสียงคำว่า อีคา



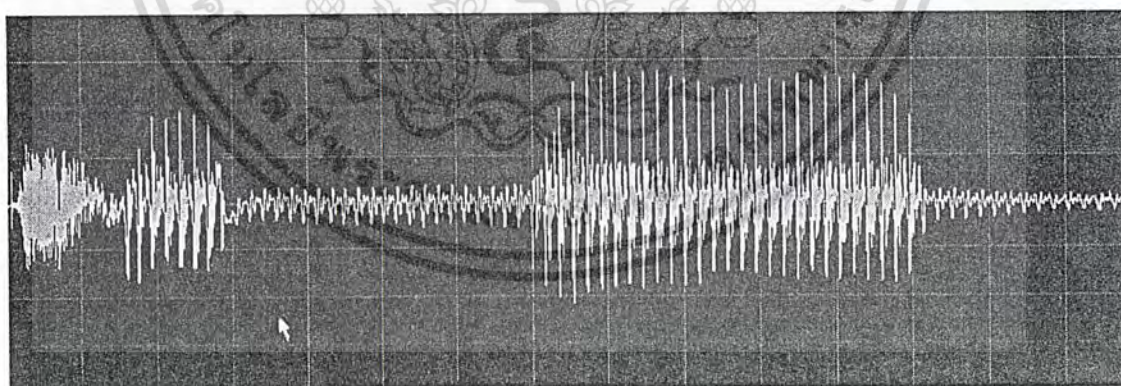
รูปที่ 5.3 รูปเสียงคำว่า กีฬา ที่เกิดจากการผสมคำ



รูปที่ 5.4 รูปเสียงคำว่า ชำนาญ



รูปที่ 5.5 รูปเสียงคำว่า คำเนิน

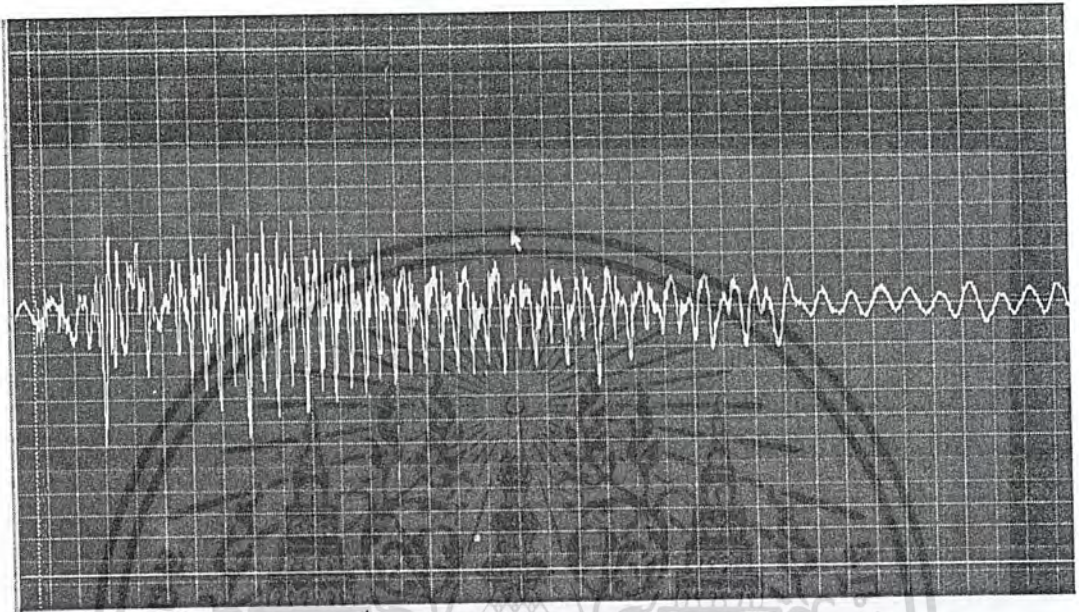


รูปที่ 5.6 รูปเสียงคำว่า ชำนาญ ที่เกิดจากการผสมคำ

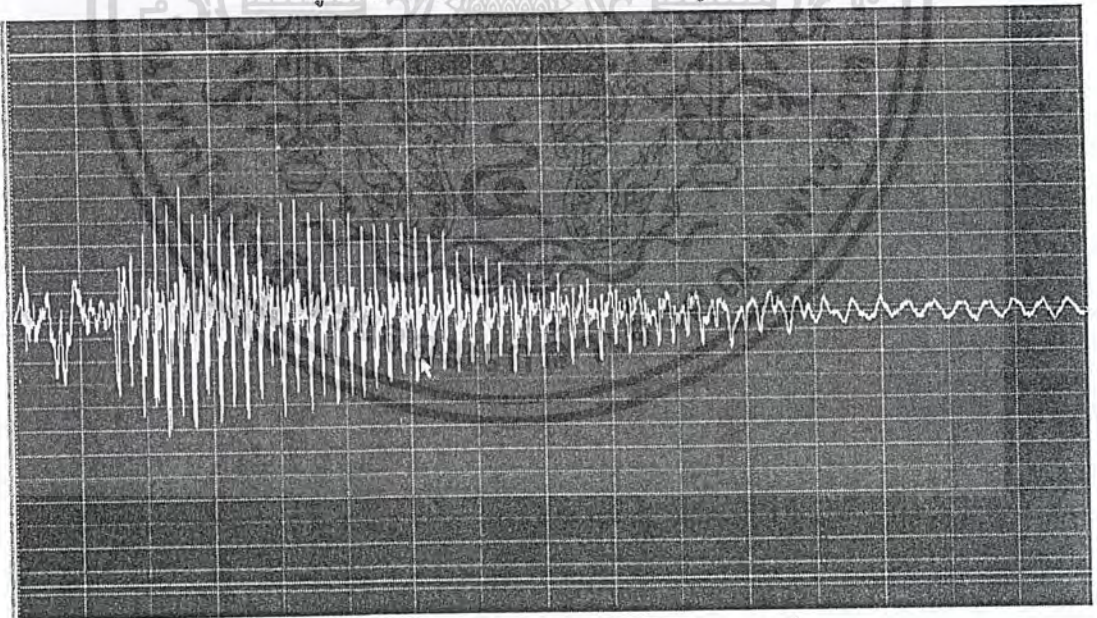
จากที่ได้ยกตัวอย่างมานี้เสียงที่ได้ฟังจากการผสมนั้น เสียงจะคล้ายกับต้นฉบับมากแต่รูปร่างของคลื่นเสียงที่เกิดขึ้นนั้นไม่เหมือนต้นฉบับซะทีเดียวอาจเนื่องมาจากในขณะที่อัดเสียงนั้น มีความไม่แน่นอนในขณะอัดเสียง จึงทำให้ รูปร่างของคลื่นเสียงที่เกิดขึ้น และขนาดแอมพลิจูดไม่เท่ากัน

5.3 การเปรียบเทียบระหว่างเสียงที่ได้จากการสังเคราะห์กับเสียงธรรมชาติ

การทดลองนี้เป็นการทดลองเพื่อเปรียบเทียบความแตกต่างของเสียงที่ได้จากการสังเคราะห์ กับเสียงธรรมชาติ โดยนำไบโโฟนที่ได้จากบทที่แล้ว มาผสมเป็นคำ แล้วบันทึกเสียงคำเดียวกันนั้นด้วยการพูดจริง ซึ่งแล้วลองนำรูปคลื่นเสียงมาเปรียบเทียบกัน

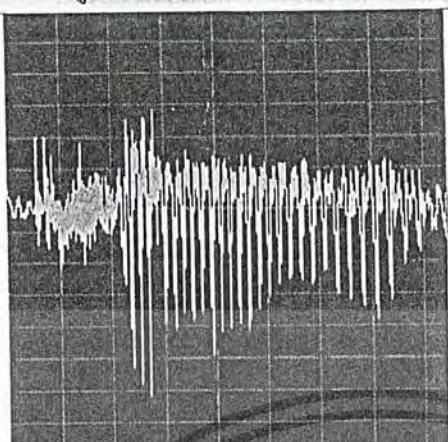
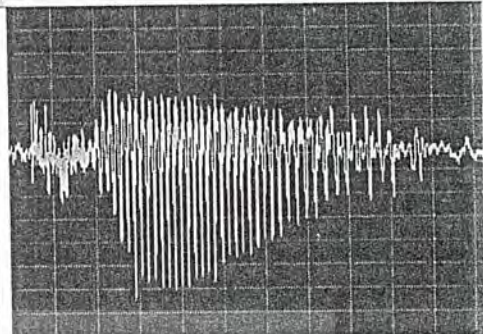
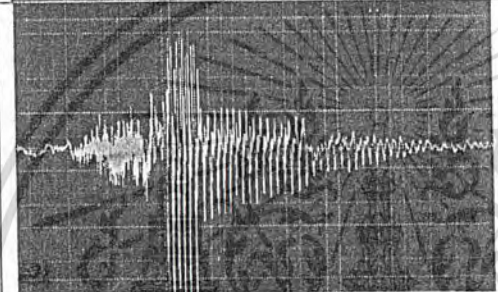
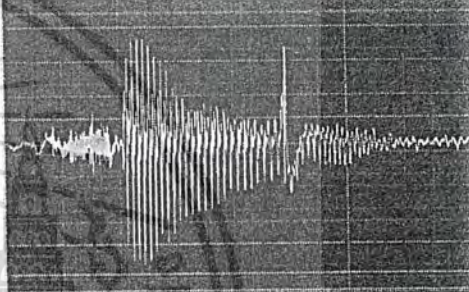
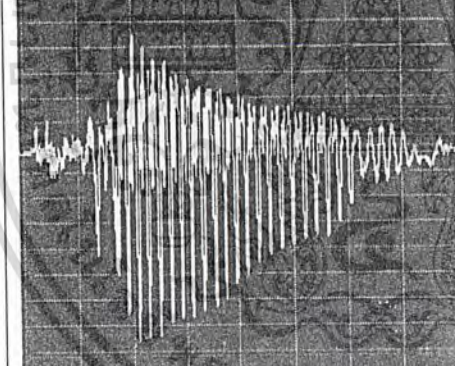
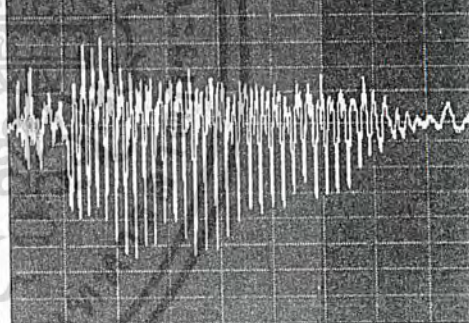


รูปที่ 5.7 เสียงคำว่า "ควาย" จากการพูดจริง



รูปที่ 5.8 เสียงคำว่า "ควาย" จากการสังเคราะห์

ตารางเปรียบเทียบรูปคลื่นเสียงเชิงเวลา

คำ	รูปคลื่นเสียงที่ได้จากการสังเคราะห์	รูปคลื่นเสียงที่ได้จากการบันทึกเสียงพูดจริง
ข้าว		
สาม		
เธอ		

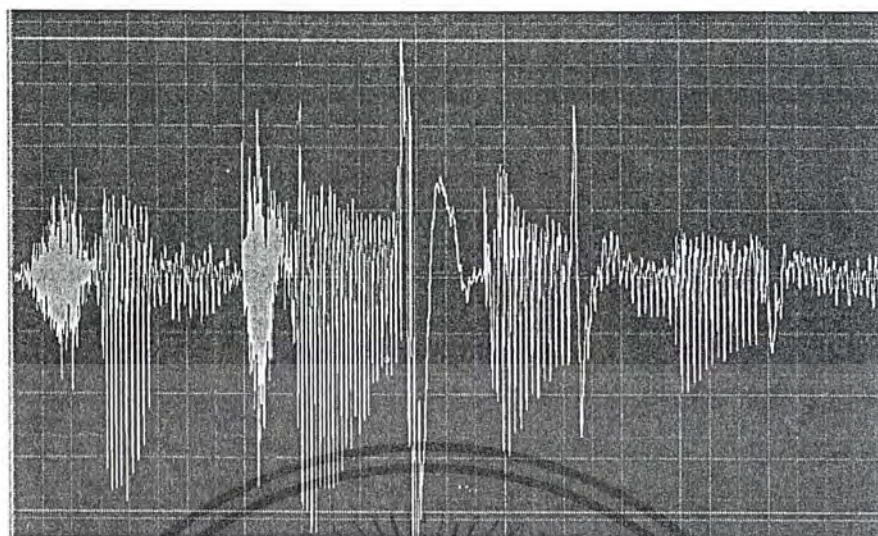
ตารางที่ 5.1 เปรียบเทียบรูปคลื่นเสียงเชิงเวลา

จากตารางการเปรียบเทียบรูปคลื่นเสียงเชิงเวลาของคำจะสังเกตได้ว่าลักษณะของรูปคลื่นที่ได้จากการสังเคราะห์มีความคล้ายกันกับลักษณะของรูปคลื่นที่ได้จากการบันทึกเสียงพูดจริง

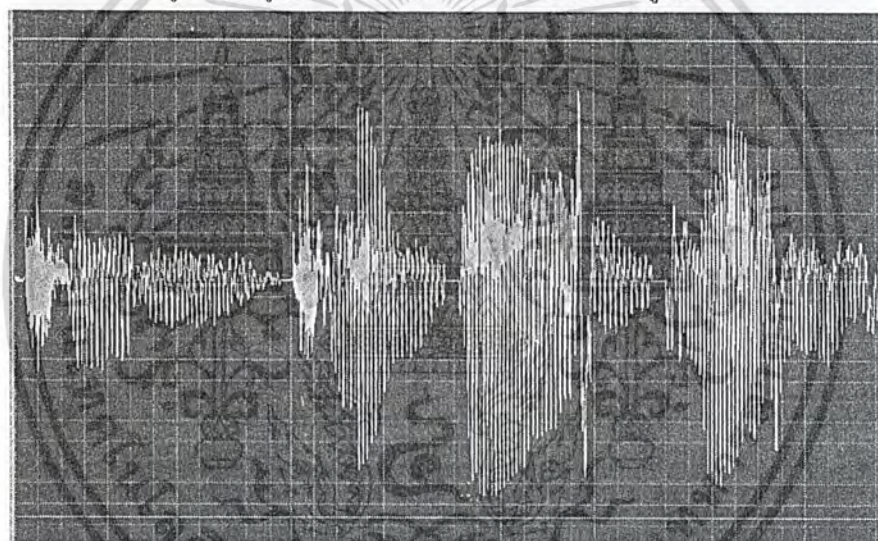
5.4 การผสมคำเป็นประโยค

ในการผสมคำให้เป็นประโยคนั้น ระหว่างคำแต่ละคำในประโยคนั้นอาจมีการเว้นเป็นช่วงสั้นๆ หรือระหว่างคำนั้นมีการเว้นไว้ตามธรรมชาติของการพูด ดังนั้นในการสังเคราะห์คำก็เช่นเดียวกัน เราอาจต้องมีการแทรก ช่องว่างเหล่านั้นด้วยเพื่อให้ฟังเป็นธรรมชาติ

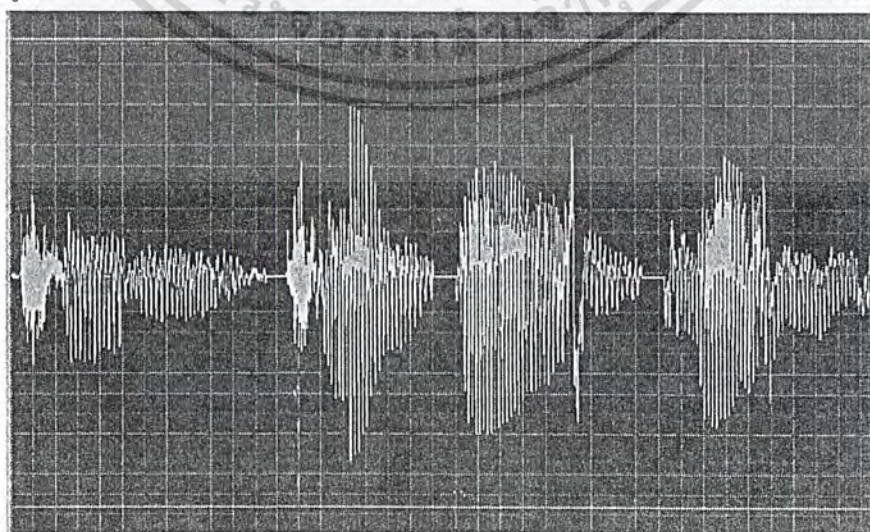
การทดลองนี้จึงเป็นการทดลองเพื่อหาระยะเวลาที่เหมาะสมของ Short Pause ในการแทรกเข้าไป เพื่อให้ประโยคเป็นธรรมชาติมากขึ้น โดยทดลองแทรก Short Pause หลาย ๆ ขนาดเข้าไป ฟังเทียบกับของจริงว่าคล้ายกันหรือไม่



รูปที่ 5.9 รูปเสียงประโยค “ฉันชอบคัมภีร์” ที่พูดปกติ



รูปที่ 5.10 รูปเสียงประโยค “ฉันชอบคัมภีร์” ที่ได้จากการนำคำมาต่อกัน โดยแทรก Short Pause ไว้ 0.03



รูปที่ 5.11 รูปเสียงประโยค “ฉันชอบคัมภีร์” ที่ได้จากการนำคำมาต่อกัน โดยแทรก Short Pause ไว้ 0.05 วินาที

จากการทดลองข้างต้น การแทรก Short Pause ขนาด 0.03 และ 0.05 นี้ฟังแล้วใกล้เคียงกับการพูดปกติ และนอกจากการแทรก Short Pause ระหว่างคำในประโยคแล้วยังต้องแทรก Short Pause ระหว่างประโยค ซึ่งต้องเป็นช่วงยาวกว่าระหว่างคำอีกด้วย

5.5 สรุปผลที่ได้จากการทดลอง

จากการทดลองทั้งหมดนี้ ทำให้หาวิธีที่ทำให้การสังเคราะห์เสียงนี้เป็นธรรมชาติมากขึ้น และทำให้ทราบโครงสร้างของรูปคลื่นเสียงได้ดีขึ้นอีกด้วย



บทที่ 6

การปรับปรุงคุณภาพเสียงสังเคราะห์

6.1 บทนำ

ในบทนี้จะเป็นเรื่องของการปรับปรุงคุณภาพของเสียงสังเคราะห์ เพราะโดยปกติแล้วคุณภาพเสียงที่ได้จากการสังเคราะห์นี้ จะแตกต่างกับเสียงพูดธรรมชาติมาก ดังนั้นจึงต้องมีการหาวิธีที่จะปรับปรุงคุณภาพเสียงเพื่อให้เสียงที่ได้ฟังแล้วเป็นธรรมชาติมากที่สุด

6.2 วิธีการบันทึกเสียง

จากการทดลองที่ได้กล่าวในบทก่อน ทำให้เราตัดสินใจเลือกลักษณะการบันทึกเสียงดังนี้

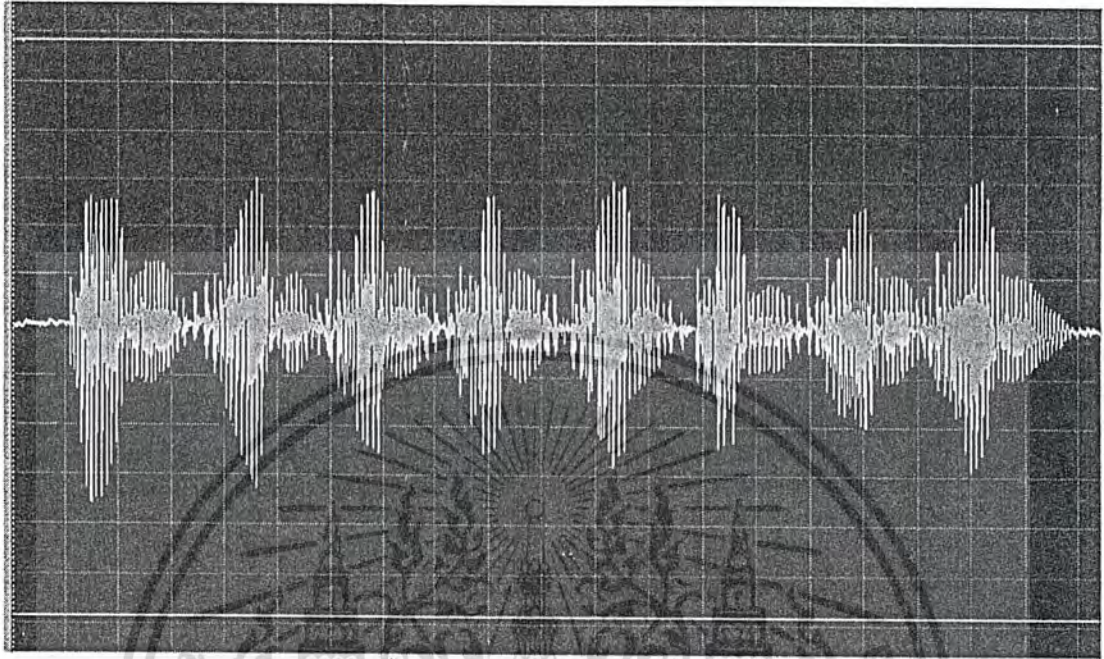
1. บันทึกเสียงพยางค์สั้น (เดี่ยว+ควบกล้ำ) ผสมเสียงสระเสียงสั้น ได้ทั้งหมด 456 หน่วยเสียง
2. บันทึกสระเสียงสั้น 12 หน่วยเสียง ผันเสียงวรรณยุกต์ 4 หน่วยเสียง คือ เสียงเอก เสียงโท เสียงตรี และเสียงจัตวา รวมทั้งหมด 48 แฟ้มเสียง
3. บันทึกสระเสียงสั้นที่มีตัวสะกดที่เป็นเสียงกัก ได้แก่ แม่กก แม่กด และแม่กบ ผันเสียงวรรณยุกต์ 4 หน่วยเสียง หน่วยเสียงวรรณยุกต์เอก เสียงโท เสียงตรี และเสียงจัตวา รวมทั้งสิ้น 144 หน่วยเสียง
4. บันทึกสระเสียงสั้นผสมตัวสะกด แม่กง แม่กน แม่กบ แม่เกย แม่เกอว ผันเสียงวรรณยุกต์ 5 หน่วยเสียง รวมทั้งสิ้น 300 แฟ้มเสียง
5. บันทึกสระเสียงยาว 12 หน่วยเสียง ผันเสียงวรรณยุกต์ 5 หน่วยเสียง คือ เสียงสามัญ เสียงเอก เสียงโท เสียงตรี และเสียงจัตวา รวมทั้งหมด 60 แฟ้มเสียง
6. บันทึกสระเสียงยาวที่มีตัวสะกดที่เป็นเสียงกัก ได้แก่ แม่กก แม่กด และแม่กบ ผันเสียงวรรณยุกต์ 4 หน่วยเสียง หน่วยเสียงวรรณยุกต์เอก เสียงโท เสียงตรี และเสียงจัตวา รวมทั้งสิ้น 144 หน่วยเสียง
7. บันทึกเสียงสระเสียงยาวผสมตัวสะกด แม่กง แม่กน แม่กบ แม่เกย แม่เกอว ผันเสียงวรรณยุกต์ 5 หน่วยเสียง รวมทั้งสิ้น 300 แฟ้มเสียง

รวมหน่วยเสียงทั้งสิ้น $456+48+144+300+60+144+300 = 1452$ หน่วยเสียง

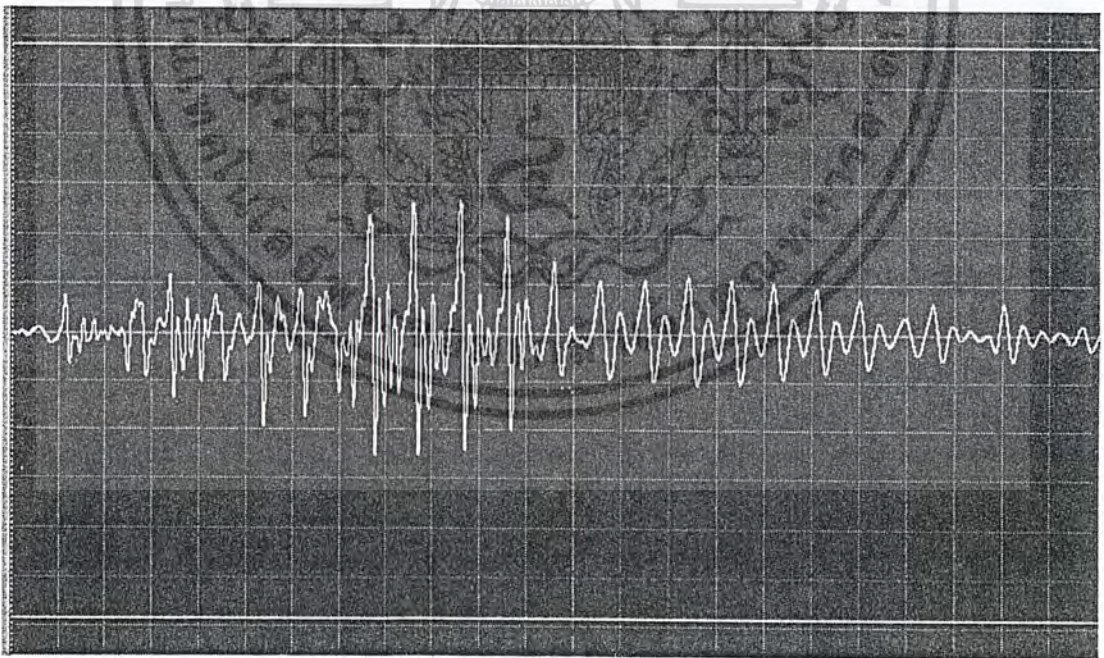
ตัวอย่างการผสมเสียง

บ้าน	จะได้	บะ+อ้าน	
รัก	จะได้	ระ+อັก	
เหนียว	จะได้	เนืะ+เอื้อว	เป็นต้น

ซึ่งลักษณะการบันทึกเสียงแบบนี้เป็นวิธีที่ดีที่สุดและประหยัดเนื้อที่ เมื่อเทียบกับวิธีอื่น ๆ โดยเมื่อเริ่มบันทึกเสียง จะใช้วิธีการบันทึกเป็นชุด ๆ ก็คือจะพูดคำที่ต้องการบันทึกติด ๆ กันหลาย ๆ ครั้งเพื่อนำมาดูแลักษณะคลื่นเสียง แล้วเลือกอันที่ดีที่สุดเพื่อนำมาตัดเป็นไบโโฟนต่อไป



รูปที่ 6.1 แสดงการบันทึกเสียงเป็นชุด ๆ



รูปที่ 6.2 แสดงการเลือกเสียงที่ดีที่สุดจากการบันทึกเป็นชุด ๆ

เมื่อเราได้คำแล้วเราก็นำมาทำการแบ่งเป็นไบโโฟน แต่ถ้าเราไม่สามารถแบ่งคำเป็นไบโโฟนได้ เราก็สามารถเลือกคำอื่นมาทำการตัดแบ่งเป็นไบโโฟน แล้วทำการคัดคำงานเลือกได้ส่วนของไบโโฟนที่ดีที่สุด

การที่เราทำการอัดเสียงแบบชุดเพราะเห็นว่าเป็นวิธีที่เหมาะสมกว่าวิธีอื่น เช่นอัดเป็นคำ หรือว่าอัดเป็นประโยค ในการอัดเป็นประโยคนั้น จากการทดลองจริงนั้นจะพบว่าไบโโฟนที่ได้จะมีความสั้นๆ

ยาวๆ ไม่เท่ากันนำมาตัดเป็นไบโโฟนที่ได้ความยาวที่ต้องการได้ยาก เนื่องจากการพูดเป็นประโยคนั้น การพูดจะสั้นๆยาวเพราะขึ้นอยู่กับลักษณะของประโยค แต่ในการอัดเป็นค่านั้นเราจะได้คำที่ไม่เป็นธรรมชาติ เพราะจะมีการลากเสียงในการอัดเป็นคำ

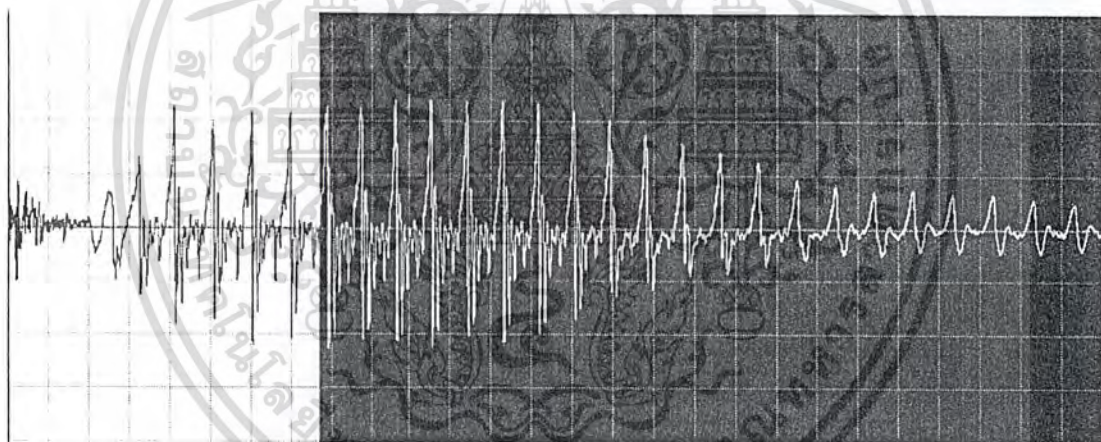
6.3 การปรับปรุงคุณภาพและการปรับลักษณะไบโโฟน

การตัดแบ่งคำที่บันทึกออกเป็นไบโโฟนนี้ จะมี 2 ประเภทคือ

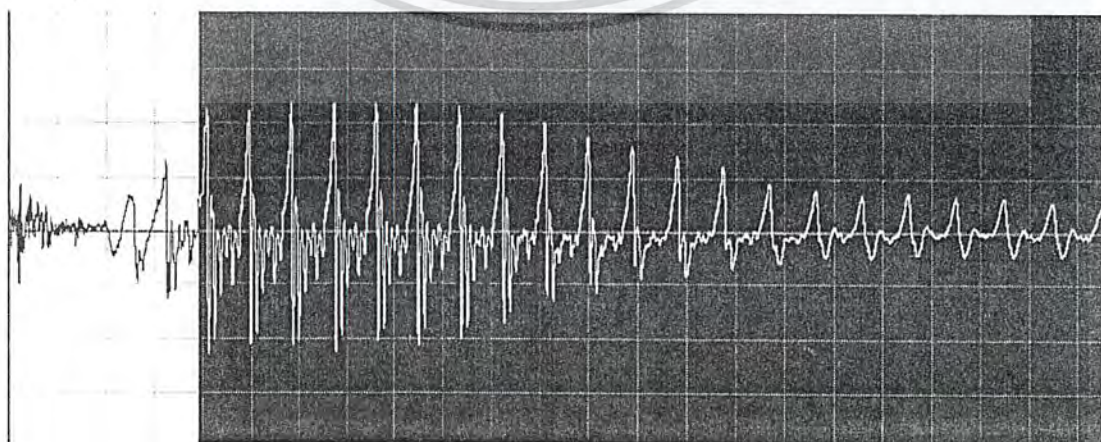
1. ไบโโฟนส่วนที่เป็นพยัญชนะต้น
2. ไบโโฟนส่วนที่เป็นสระ, ตัวสะกด และวรรณยุกต์

6.3.1 การตัดเลือกไบโโฟนส่วนหน้า

ไบโโฟนที่เป็นพยัญชนะต้นนี้จะได้จากการตัดจากคำที่ได้บันทึกในหัวข้อที่ผ่านมา โดยจะตัดเฉพาะส่วนหน้าของคำเพื่อเอาแค่เสียงพยัญชนะต้น และจะต้องพยายามตัดให้เสียงออกมาสั้นที่สุด เพราะถ้าเราตัดไม่สั้นแล้วเมื่อนำมาซ้อนกันเป็นคำแล้วจะทำให้คุณภาพของเสียงที่ออกมาไม่ดี จะมีการกระโดดของเสียงในคำ



รูปที่ 6.3 คำว่า “ก้าง” ซึ่งตัดส่วนหน้ายาวไป



รูปที่ 6.4 คำว่า “ก้าง” ซึ่งตัดส่วนหน้าสั้น

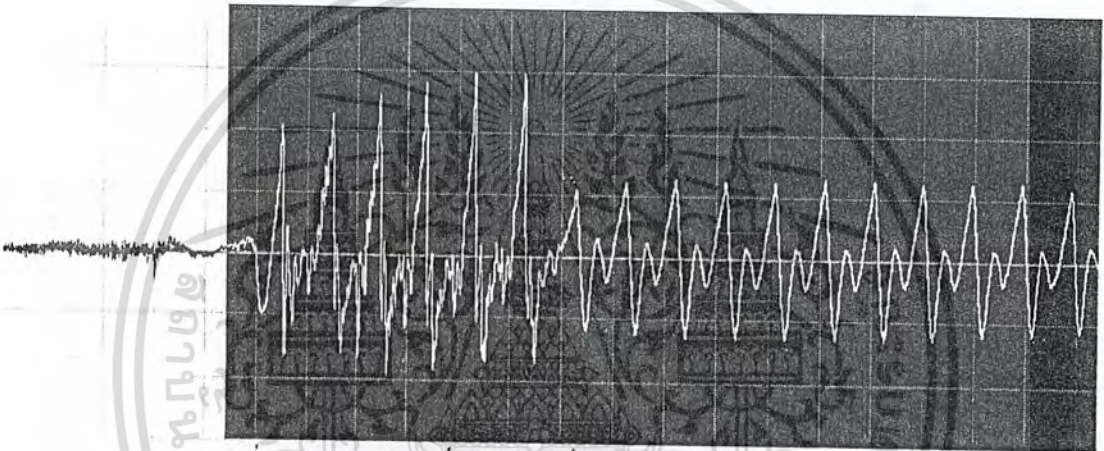
จากรูปที่ 6.3 นั้น จะมีลักษณะเสียงที่กระโดดเพราะว่าไบโโฟนส่วนหน้าที่เรากำลังมานั้นหายไป แต่ในรูปที่ 6.4 เราตัดไบโโฟนมาส่วนที่ติดจึงได้เสียงที่เรียบฟังเป็นคำ

แถบสีขาวที่เห็นในส่วนหน้าของรูปนั้นคือ ส่วนของไบโโฟนส่วนหน้า

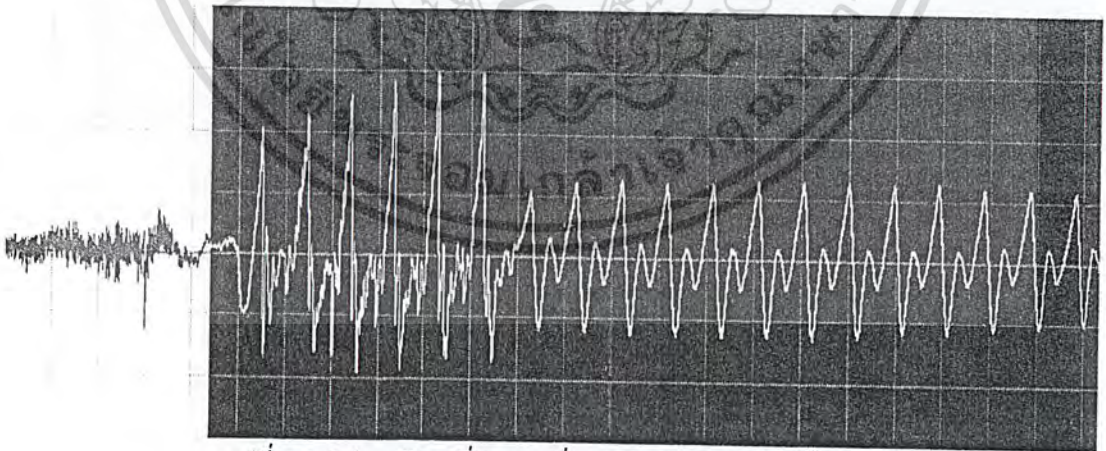
การตัดไบโโฟนส่วนของพยัญชนะต้นนั้น ต้องตัดให้มีส่วนของเสียงสระติดมาบางส่วนกับส่วนของพยัญชนะเล็กน้อย เพราะจะทำให้เสียงที่สังเคราะห์ออกมามีความใกล้เคียงกับเสียงที่พูดจริงมากขึ้น

6.3.2 การปรับปรุงความชัดเจนไบโโฟนส่วนหน้าของเสียง ข,ช,ฟ ที่มีการกักของลมก่อนออกเสียง

การปรับปรุงเสียงที่มีการกักลมก่อนออกเสียงของอักษรตัว ข,ช,ฟ นั้น ในรูปคลื่นนั้นจะเห็นส่วนของที่มีการกักเสียงนั้นชัดเจนซึ่งเราสามารถเพิ่มความชัดเจนได้โดยการเพิ่ม แอมพลิจูด ให้กับส่วนกักของลมนั้น ทำให้เสียงที่ออกมาชัดเจนมากขึ้น



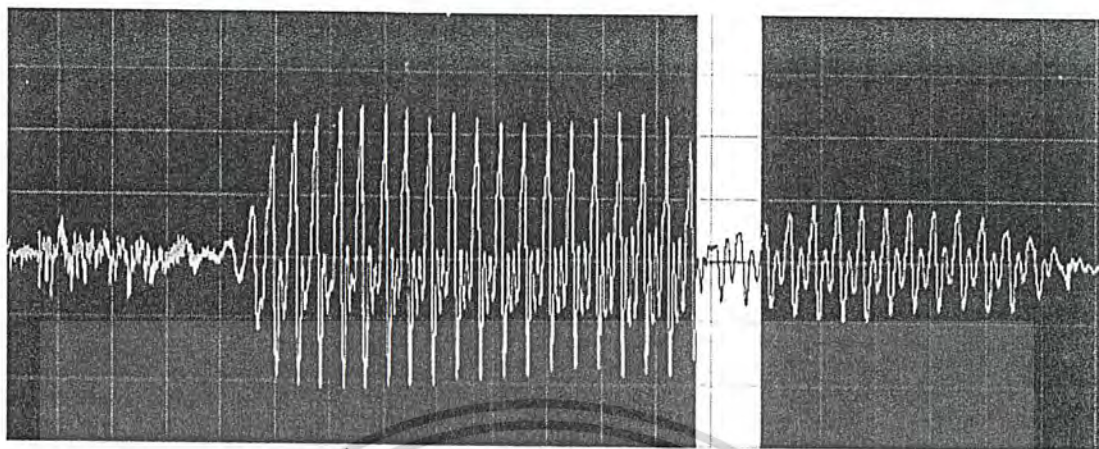
รูปที่ 6.5 คำว่า "ฟัน" ซึ่งยังไม่ได้เพิ่ม แอมพลิจูด ของส่วนที่กักลม



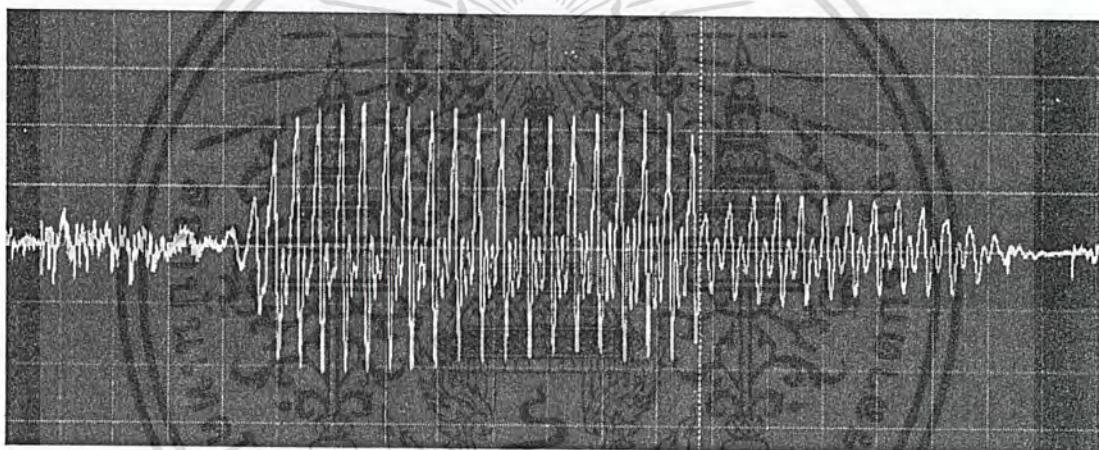
รูปที่ 6.6 คำว่า "ฟัน" ที่ทำการเพิ่ม แอมพลิจูดของส่วนที่กักลม

จากรูป จะแสดงถึงคำว่า "ฟัน" ซึ่งได้ทำการปรับปรุงโดยการเพิ่ม แอมพลิจูด ให้มากขึ้น ซึ่งทำให้เสียง "ฟ" ในส่วนหน้า นั้นชัดเจนมากขึ้น โดยไม่ทำให้เสียงที่ได้นั้นเพี้ยนไป

6.3.3 การปรับขนาดของไบโโฟนส่วนท้ายให้สั้นลงโดยไม่กระทบต่อคุณภาพเสียงที่ได้



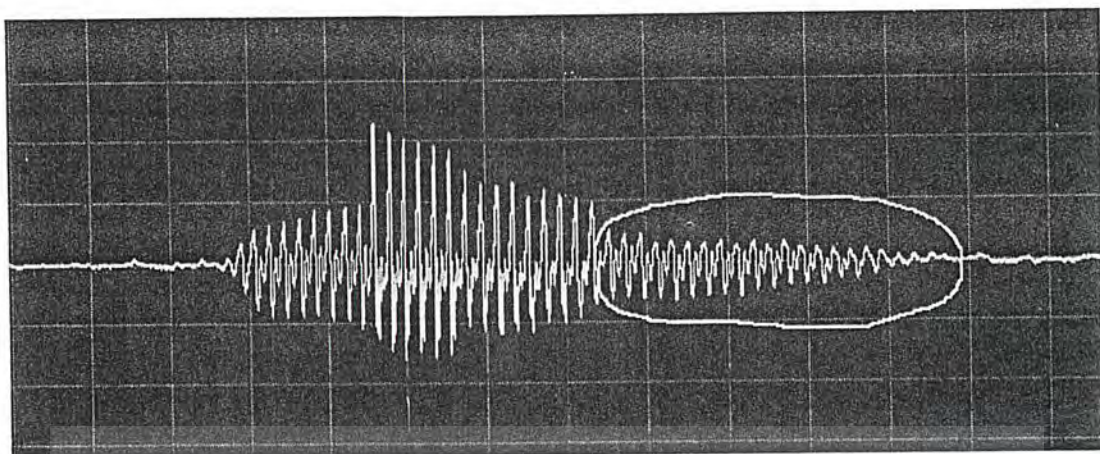
รูปที่ 6.7 คำว่า “ฟาม” ซึ่งยังไม่ได้ปรับขนาดเสียง



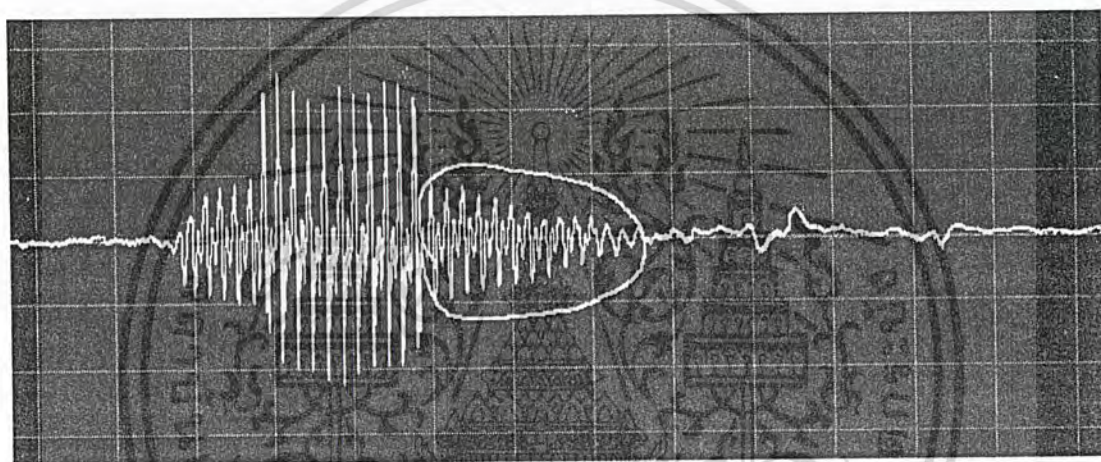
รูปที่ 6.8 คำว่า “ฟาม” ซึ่งได้ตัดส่วนหนึ่งของคลื่นออกเพื่อปรับขนาดของไบโโฟน

จากรูปที่ 6.7 นั้น เราจะเอาส่วนของไบโโฟนส่วนหลังไปใช้ นั่นก็คือเสียง “อาม” ถ้าหากว่าไบโโฟนที่ได้นั้นยาวไป ซึ่งจะพบมากในการอัดเสียง เราสามารถตัดส่วนที่เป็นแถบสีขาวในรูปที่ 6.7 ออกไปได้โดยไบโโฟนที่ได้นั้น จะไม่มีผลต่อเสียงที่ออกมา ทำให้เราสามารถปรับขนาดให้สั้นลงได้โดยไม่ทำให้เสียงนั้นเพี้ยนไป

และยังมีอีกวิธีหนึ่งซึ่งไม่ต้องตัดส่วนใดส่วนหนึ่งของไบโโฟน แต่อยู่ที่ขั้นตอนการอัดเสียงเพื่อให้ได้เสียงที่กระชับและหางเสียงที่สั้นไม่ขัด ซึ่งไม่แนะนำให้ใช้ ฟังก์ชัน “Stretch” ในโปรแกรม “Cool Edit” เพราะจะทำให้เสียงที่ได้สั้นเพี้ยนจากของเดิมไป เพราะเป็นการ ย่นเวลาให้สั้นลง โปรแกรมจะคำนวณออกมาให้รูปคลื่นที่เปลี่ยนไป ดังนั้นเสียงก็จะเพี้ยนตามไปด้วย โดยวิธีที่จะทำให้ได้ไบโโฟนส่วนหลังให้กระชับหางเสียงไม่ยาวเกินไป ต้องอยู่ในกระบวนการอัดเสียง คือพูดให้หางเสียงกระชับ ไม่ลากหางเสียง โดยหางเสียงที่ได้นั้น จะมีความยาวมากและมีผลมากในการต่อเป็นประโยคถ้าหางเสียงยาวไป ประโยคที่ได้ฟัง จะไม่เป็นจังหวะที่เหมาะสม



รูปที่ 6.9 หางเสียงคำว่า “นาม” ซึ่งลากหางเสียงยาวไป



รูปที่ 6.10 หางเสียง “นาม” ซึ่งพูดกระชับ

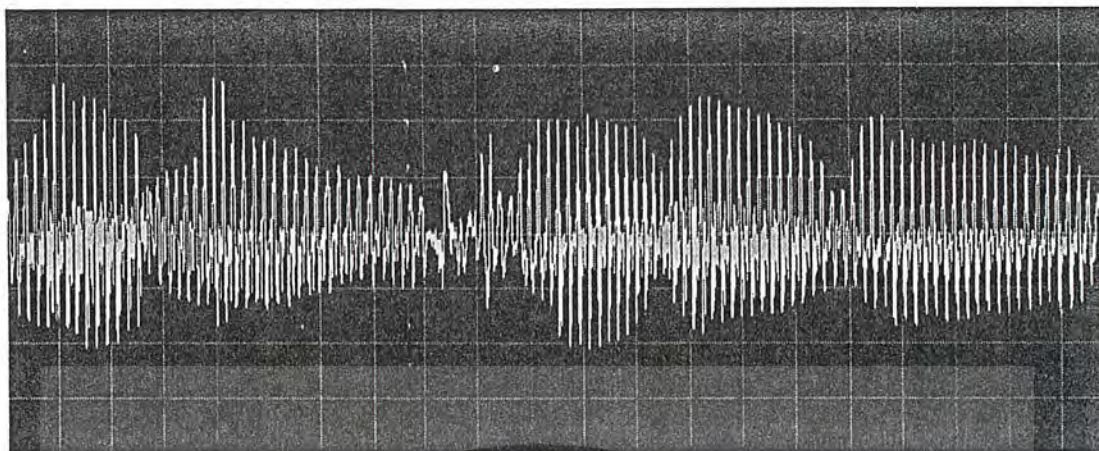
จากรูปที่ 6.9 นั้นหางเสียง “ม” ที่ได้จะบีด ไม่กระชับ ซึ่งอาจทำให้ได้ ไบโพนที่ส่วนหลังยาวเกินขนาดที่ต้องการ จึงต้องทำการพูดให้กระชับเพื่อให้ได้เสียง “ม” ที่สั้นกระชับ

และหากบันทึกสระเสียงสั้นที่มีตัวสะกดที่เป็นเสียงกัก ได้แก่ แม่ก ก แม่กด และแม่กบ นั้นจะต้องมีการเพิ่ม Short Pause ไว้ท้ายเสียงเล็กน้อยด้วย เพื่อให้เสียงสังเคราะห์ เมื่อนำมารวมเป็นประโยคแล้วจะฟังเป็นคำมากขึ้น ไม่เกิดการเชื่อมเสียงกับคำถัดไป เพราะหากมีการเชื่อมเสียงกับคำถัดไป จะทำให้ประโยคที่ได้ฟังไม่เข้าใจ

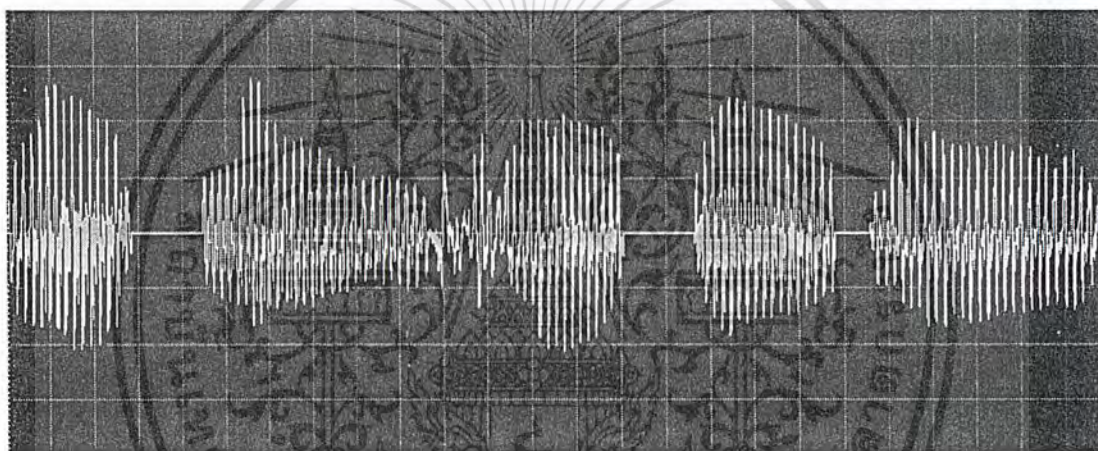
ไบโพนที่ตัดได้จะถูกนำไปเก็บในฐานะข้อมูลสำหรับโปรแกรมสังเคราะห์เสียง เพื่อให้โปรแกรมสังเคราะห์เสียงเรียกใช้ต่อไป

6.4 การแทรก Short Pause

เพื่อให้เสียงที่ได้จากการสังเคราะห์เป็นธรรมชาติมากที่สุดนั้น จำเป็นต้องมีการแทรก Short Pause ระหว่างคำ และ ประโยคด้วย เพราะการพูดของมนุษย์นั้นมีจังหวะ มีการเว้นวรรคระหว่างประโยค เพื่อให้ผู้ฟังเข้าใจ



รูปที่ 6.11 ประโยค “หลบหน่อยพระเอกมา” โดยไม่มีการแทรก Short Pause



รูปที่ 6.12 ประโยค “หลบหน่อยพระเอกมา” โดยมีการแทรก Short Pause

จากรูปที่ 6.11 นั้นไม่มีการแทรก Short Pause ทำให้ประโยคที่ได้ฟังนั้นไม่สามารถฟังได้รู้เรื่องเลย แต่เมื่อทำการแทรก Short Pause แล้วทำให้ฟังได้รู้เรื่อง

โดย Short Pause ที่จะแทรกไประหว่างคำนั้น เราจะทำการแทรก ในใบโฟนส่วนหลังที่เป็นเสียงกัก หรือ คำตายโดย ถ้าเป็นเสียงสระสั้นแล้วสะกดด้วยเสียงกัก จะทำการแทรก Short Pause 80 มิลลิวินาที

ถ้าเป็นเสียงสระยาวแล้วสะกดด้วยเสียงกัก จะทำการแทรก Short Pause 40 มิลลิวินาที

6.5 สรุป

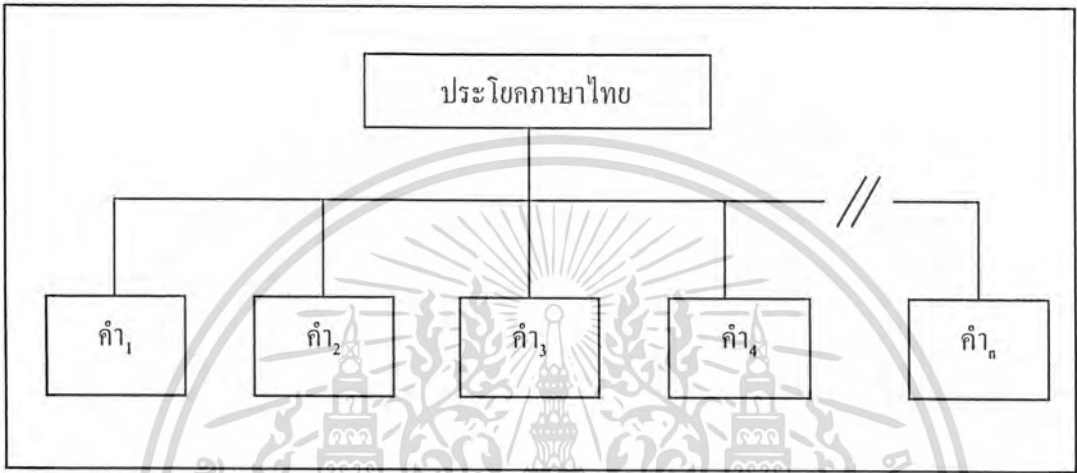
เราสามารถปรับปรุงคุณภาพเสียงสังเคราะห์ให้เป็นธรรมชาติมากที่สุด โดยต้องเริ่มตั้งแต่การเลือกลักษณะวิธีบันทึกเสียง วิธีการตัดใบโฟนโดยต้องมีการตรวจสอบใบโฟนที่ตัดด้วยว่าสามารถนำไปผสมเป็นคำได้จริงหรือไม่ นอกจากนั้นต้องมีการแทรก Short Pause ระหว่างประโยคเพื่อให้เสียงสังเคราะห์ที่ได้ฟังแล้วใกล้เคียงกับการพูดของมนุษย์มากที่สุด

บทที่ 7

การตัดคำ

7.1 บทนำ

การตัดคำเป็นการแยกคำภาษาไทยออกจากประโยคภาษาไทย เพื่อจะนำคำที่ได้ไปใช้ในการสร้างเสียง วิธีที่ใช้ในการตัดคำในครั้งนี้ คือ การตัดคำโดยใช้วิธี WordBucket



รูปที่ 7.1 โครงสร้างของประโยคภาษาไทย

7.2 พจนานุกรมคำศัพท์

พจนานุกรมคำศัพท์ประกอบด้วย field ต่าง ๆ ดังต่อไปนี้

No	Field Name	Type	Size	Description
1	คำ	Text	25	เก็บคำศัพท์ไว้ในพจนานุกรม
2	จำนวนพยางค์	Integer	4	จำนวนพยางค์ของคำนี้
3	พยางค์ที่ 1	Text	10	คำอ่านของพยางค์ที่ 1
4	พยางค์ที่ 2	Text	10	คำอ่านของพยางค์ที่ 2
5	พยางค์ที่ 3	Text	10	คำอ่านของพยางค์ที่ 3
6	พยางค์ที่ 4	Text	10	คำอ่านของพยางค์ที่ 4
7	พยางค์ที่ 5	Text	10	คำอ่านของพยางค์ที่ 5
8	พยางค์ที่ 6	Text	10	คำอ่านของพยางค์ที่ 6
9	พยางค์ที่ 7	Text	10	คำอ่านของพยางค์ที่ 7
10	พยางค์ที่ 8	Text	10	คำอ่านของพยางค์ที่ 8
11	พยางค์ที่ 9	Text	10	คำอ่านของพยางค์ที่ 9
12	พยางค์ที่ 10	Text	10	คำอ่านของพยางค์ที่ 10

ตารางที่ 7.1 แสดง Field ต่าง ๆ ในพจนานุกรมคำศัพท์

ตัวอย่างคำที่เก็บในพจนานุกรมคำศัพท์

คำ	จำนวนพยางค์	พยางค์ที่ 1	พยางค์ที่ 2	พยางค์ที่ 3	พยางค์ที่ 4	พยางค์ที่ 5	พยางค์ที่ 6	พยางค์ที่ 7	พยางค์ที่ 8	พยางค์ที่ 9	พยางค์ที่ 10
กรุง	1	กรุง									
กรุงเทพ	2	กรุง	เทพ								
กรุงเทพมหานคร	3	กรุง	เทพ	มะ	หา	นะ	คอน				
ของ	1	ของ									
คน	1	คน									
ช้าง	1	ช้าง									
เทพ	1	เทพ									
นคร	2	นะ	คอน								
ประ	1	ประ									
ประเดิน	2	ประ	เดิน								
เป็น	1	เป็น									
มหา	2	มะ	หา								
เรื่อง	1	เรื่อง									
ใหญ่	1	ใหญ่									
อพยพ	3	อพ	พะ	ยพ							

ตารางที่ 7.2 แสดงตัวอย่างคำที่เก็บในพจนานุกรมคำศัพท์

7.3 วิธีการตัดคำโดยใช้ WordBucket

การตัดคำโดยใช้ WordBucket คือ การตัดคำโดยวิธีเก็บ “คำที่เป็นไปได้ในการตัดคำ” แต่ละครั้งไว้ใน WordBucket แล้วเลือกคำที่ยาวที่สุดใน WordBucket เป็นคำที่ตัดได้ “คำที่เป็นไปได้ในการตัดคำ” หมายถึง คำที่มีอยู่ในพจนานุกรมคำศัพท์

ตัวอย่างเช่น

ข้อมูลใน WordBucket ในการตัดคำว่า “กรุงเทพมหานคร” คือ

WordBucket (1) กรุง

WordBucket (2) กรุงเทพ

WordBucket (3) กรุงเทพมหานคร

คำที่เป็นไปได้ในการตัด คือ กรุง, กรุงเทพ , กรุงเทพมหานคร

คำที่เป็นไปได้ในการตัดที่ยาวที่สุด คือ กรุงเทพมหานคร

ตัวอย่างการตัดคำออกจากประโยค

ประโยคที่จะตัด คือ

“เรื่องของช้างอพยพเป็นประเด็นใหญ่ของคนกรุงเทพมหานคร”

หลังการตัดคำโดยใช้ WordBucket จะได้ดังต่อไปนี้

เรื่อง-ของ-ช้าง-อพยพ-เป็น-ประเด็น-ใหญ่-ของ-คน-กรุงเทพมหานคร

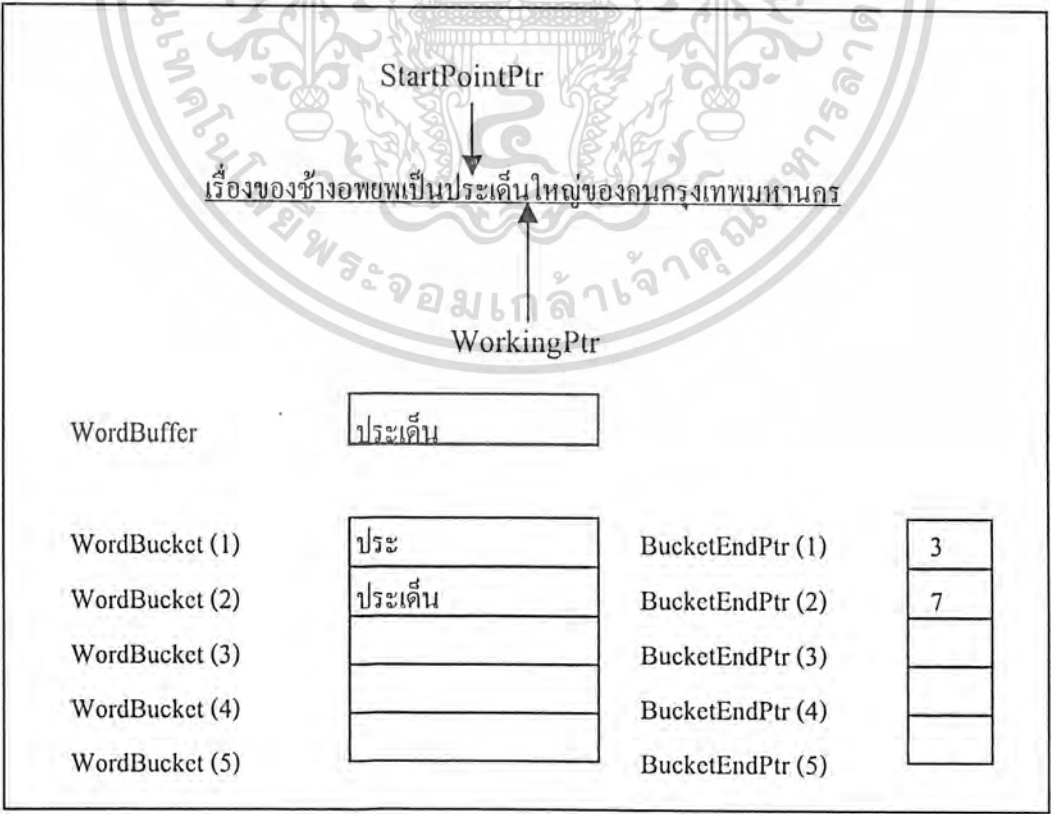
ตัวแปรที่ใช้ในการตัดคำ

ตัวแปรที่ใช้ข้อมูล

- 1. StartPointPtr เป็นตัวแปรที่ใช้ชี้ตำแหน่งตัวอักษรเริ่มต้นของคำที่กำลังตัด
- 2. WorkingPtr เป็นตัวแปรที่ใช้ชี้ตำแหน่งตัวอักษรตัวต่อไปที่จะอ่านเข้ามายัง WordBuffer

ตัวแปรที่ใช้ในการเก็บข้อมูล

- 1. WordBuffer เป็นตัวแปรที่ใช้เก็บกลุ่มของตัวอักษรที่จะนำไปเปรียบเทียบกับพจนานุกรมคำศัพท์
- 2. WordBucket เป็นตัวแปรที่ใช้เก็บคำที่เป็นไปได้ในการตัดคำทั้งหมด เพื่อทำการเลือกคำที่ยาวที่สุดต่อไป
- 3. BucketEndPtr เป็นตัวแปรที่ใช้เก็บจำนวนตัวอักษรของคำที่อยู่ใน WordBucket เพื่อใช้สำหรับเลื่อน StartPointPtr ไปตำแหน่งเริ่มต้นของคำต่อไป



รูปที่ 7.2 แสดงวิธีการตัดคำโดยใช้ WordBucket

อธิบายการทำงานของการทำงานการตัดคำ

1. กำหนดค่าเริ่มต้นให้กับ StartPointPtr และ WorkingPtr ให้เท่ากับ 1 เพื่อชี้ไปยังตัวอักษรตัวแรกของข้อความที่ต้องการตัดคำ
2. นำตัวอักษรที่ชี้โดย WorkingPtr มาเก็บไว้ใน WordBuffer ซึ่งเป็น Buffer ที่ใช้เปรียบเทียบกับพจนานุกรมคำศัพท์ และเพิ่มค่า WorkingPtr อีกหนึ่ง เพื่อชี้ไปยังตัวอักษรตัวต่อไป
3. เปรียบเทียบกลุ่มของตัวอักษรที่อยู่ใน WordBuffer กับพจนานุกรมคำศัพท์ ว่ามีค่าที่อยู่ในพจนานุกรมคำศัพท์ที่ขึ้นต้นด้วยกลุ่มของตัวอักษรที่อยู่ใน WordBuffer หรือไม่ ถ้ามีแสดงว่ายังคงมีค่าที่เป็นไปได้ในการตัดที่ยาวกว่านี้อยู่ให้ไปขั้นตอนต่อไป ถ้าไม่มีแสดงว่าไม่มีค่าที่เป็นไปได้ที่ยาวกว่านี้แล้วให้ไปขั้นตอนที่ 7
4. เปรียบเทียบกลุ่มของตัวอักษรที่อยู่ใน WordBuffer กับพจนานุกรมคำศัพท์ ว่ามีค่าที่อยู่ในพจนานุกรมคำศัพท์ที่ตรงกับค่าที่อยู่ใน WordBuffer นี้หรือไม่ ถ้ามีแสดงว่าพบค่าที่สามารถตัดได้ในพจนานุกรมคำศัพท์ให้ไปขั้นตอนต่อไป ถ้าไม่มีให้กลับไปขั้นตอนที่ 2 เพื่อหาค่าที่สามารถตัดได้ต่อไป
5. นำค่าที่ได้จากขั้นตอนที่ 4 เก็บลง WordBucket ซึ่งเป็น Buffer ขนาด 5 ช่อง สำหรับเก็บค่าทั้งหมดที่สามารถตัดได้
6. ตรวจสอบว่า WordBucket เต็มหรือไม่ ถ้าเต็มไปขั้นตอนต่อไป ถ้าไม่เต็มกลับไปขั้นตอนที่ 2 เพื่อหาค่าที่สามารถตัดได้ยาวที่สุดต่อไป
7. เลือกค่าใน WordBucket ที่ยาวที่สุด โดยปกติจะเป็นค่าสุดท้ายที่ตัดได้ แต่บางครั้งค่าที่ยาวที่สุดที่ตัดได้ อาจทำให้คำต่อไปอ่านไม่ได้ ดังนั้นเราจึงต้องกลับมาเลือกค่าที่ยาวที่สุดลำดับถัดมาที่ทำให้คำต่อไปอ่านได้ โดยการเขียนโปรแกรมแบบ Recursive
8. ตรวจสอบว่าค่าที่เลือกเป็นค่าที่เข้ามาใน WordBucket เป็นค่าสุดท้ายหรือไม่ ถ้าใช่ไม่ต้องทำการเปลี่ยนแปลงค่า Workingptr เพราะจะชี้ไปยังตัวอักษรตัวแรกของคำต่อไปอยู่แล้ว ให้ไปขั้นตอนที่ 10 แต่ถ้าไม่ใช่ให้ไปขั้นตอนต่อไป
9. ถ้าค่าที่เลือกไม่ได้เป็นค่าที่เข้ามาใน WordBucket เป็นค่าสุดท้าย แสดงว่าหลังจากเลือกค่านี้ Workingptr ไม่ได้ชี้ไปยังตัวอักษรตัวแรกของคำถัดไป ต้องทำการย้าย Workingptr ไปยังตำแหน่งที่ถูกต้องดังนี้

$$\text{Workingptr} = \text{StartPointptr} + \text{BucketEndPtr}$$

โดย StartPointptr เป็นตัวแปร Pointer ที่ชี้ตำแหน่งเริ่มต้นของคำปัจจุบัน

BucketEndPtr เป็นตัวแปรที่เก็บจำนวนตัวอักษรของคำ

10. ให้ StartPointptr = Workingptr เพื่อให้ StartPointptr ชี้ไปยังตัวอักษรตัวแรกของคำถัดไป ซึ่งก่อนหน้านี้เราจะไปทำการตัดคำต่อไป ต้องนำค่าที่ตัดได้ไปสร้างเสียงก่อน

7.4 สรุป

การตัดคำโดยวิธี WordBucket เป็นการตัดคำที่เลือกค่าที่ยาวที่สุดใน WordBucket เป็นค่าที่ตัดได้ เมื่อตัดคำได้แล้วก็จะส่งค่าที่ได้ไปสร้างเป็นเสียงต่อไป

บทที่ 8

การแปลงหน่วยคำเป็นสัญลักษณ์แทนเสียง โดยวิธี MPT-networks

8.1 บทนำ

วิธีการแปลงหน่วยคำในประโยคภาษาไทยให้เป็นสัญลักษณ์แทนเสียงโดยวิธี MPT-networks เป็นกระบวนการแปลงหน่วยคำที่อาศัยเงื่อนไขทางโครงสร้างองค์ประกอบพยางค์ตามไวยากรณ์ภาษาไทย เป็นหลักในการพิจารณา เริ่มต้นจากการวิเคราะห์โครงสร้างของพยางค์ของเสียงอ่าน เพื่อเปลี่ยนรูปพยางค์ ที่เขียนได้หลากหลายรูปแบบให้อยู่ในรูปแบบกลางสำหรับใช้หาเสียงวรรณยุกต์เป้าหมาย โดยการสร้าง กฎในรูปของโครงข่ายความสัมพันธ์ระหว่างรูปแบบการเขียนและการออกเสียง จากนั้นจึงเปลี่ยนรูปแบบ การเขียนเป็นสมการเสียงพยางค์ และแปลงเป็นสัญลักษณ์แทนเสียงต่อไป

การแปลงหน่วยคำในภาษาไทยให้เป็นสัญลักษณ์แทนเสียง โดยวิธี MPT-networks ประกอบด้วย ขั้นตอนหลัก 3 ขั้นตอน คือ

1. การวิเคราะห์องค์ประกอบโครงสร้างพยางค์ เพื่อหารูปแบบกลาง
2. การวิเคราะห์รูปแบบกลางเพื่อหาเสียงเป้าหมาย
3. การแปลงเป็นสัญลักษณ์แทนเสียง

8.2 การวิเคราะห์องค์ประกอบโครงสร้างพยางค์ เพื่อหารูปแบบกลาง

จากความรู้ในเรื่องของระบบเสียงในภาษาไทยและลักษณะโครงสร้างของพยางค์ไทยพบว่า พยางค์ประกอบด้วย พยัญชนะ สระ และ วรรณยุกต์ ซึ่งสามารถเขียนได้หลายรูปแบบ รวมทั้งเสียงของ พยางค์ไม่ขึ้นอยู่กับรูปวรรณยุกต์ที่ปรากฏ การที่จะแปลงเป็นสัญลักษณ์แทนเสียงจะต้องทำการวิเคราะห์ ความสัมพันธ์ระหว่างเสียงกับรูปแบบการเขียนของคำอ่านที่เป็นไปได้ทั้งหมด เพราะพยางค์สามารถเขียน ได้หลายรูปแบบ เนื่องจากสาเหตุใหญ่ ๆ 6 ประการคือ

- | | |
|-------------|---|
| ประการที่ 1 | รูปแบบของไวยากรณ์ที่ขอมให้อักษรตัวแรกของพยางค์เป็นได้ทั้งสระหรือ พยัญชนะก็ได้ |
| ประการที่ 2 | ในกลุ่มพยัญชนะควบกล้ำ อักษรตัวที่ต่อจากพยัญชนะตัวแรกไม่จำเป็นต้อง เป็นสระเสมอไป |
| ประการที่ 3 | แต่ละพยางค์อาจจะมีหรือไม่มีตัวสะกด (ซึ่งอยู่ในรูปของพยัญชนะ) |
| ประการที่ 4 | สระบางรูป ที่มีรูปสระมากกว่า 1 ตัว จะเขียนไม่ติดต่อกัน เช่น สระ เอะ (ใช้ รูปของสระ “เ” , “า” และ “ะ”) |
| ประการที่ 5 | สระบางรูป มีการใช้รูปของพยัญชนะเป็นส่วนหนึ่งของสระ เช่น สระ เออะ (มี พยัญชนะ “อ” รวมอยู่ในสระ) |
| ประการที่ 6 | มีการเปลี่ยนรูปพยางค์ใหม่จากการเปลี่ยนหรือลดรูปสระ เช่น คำว่า “คน” มี การลดรูปสระ โะ เป็นตัน |

ดังนั้นการที่จะสร้างรูปแบบกลาง เพื่อใช้แทนรูปแบบการเขียนและใช้หาเสียงเป้าหมายของ พยางค์ ต้องทำการวิเคราะห์องค์ประกอบโครงสร้างพยางค์จากรูปแบบที่เป็นไปได้ทั้งหมด จากสาเหตุดังกล่าวข้างต้น ถ้าพิจารณาการเกิดของรูปแบบตามลำดับการเขียนเฉพาะพยัญชนะและสระ จะได้อักษร เรียงตามกันดังนี้

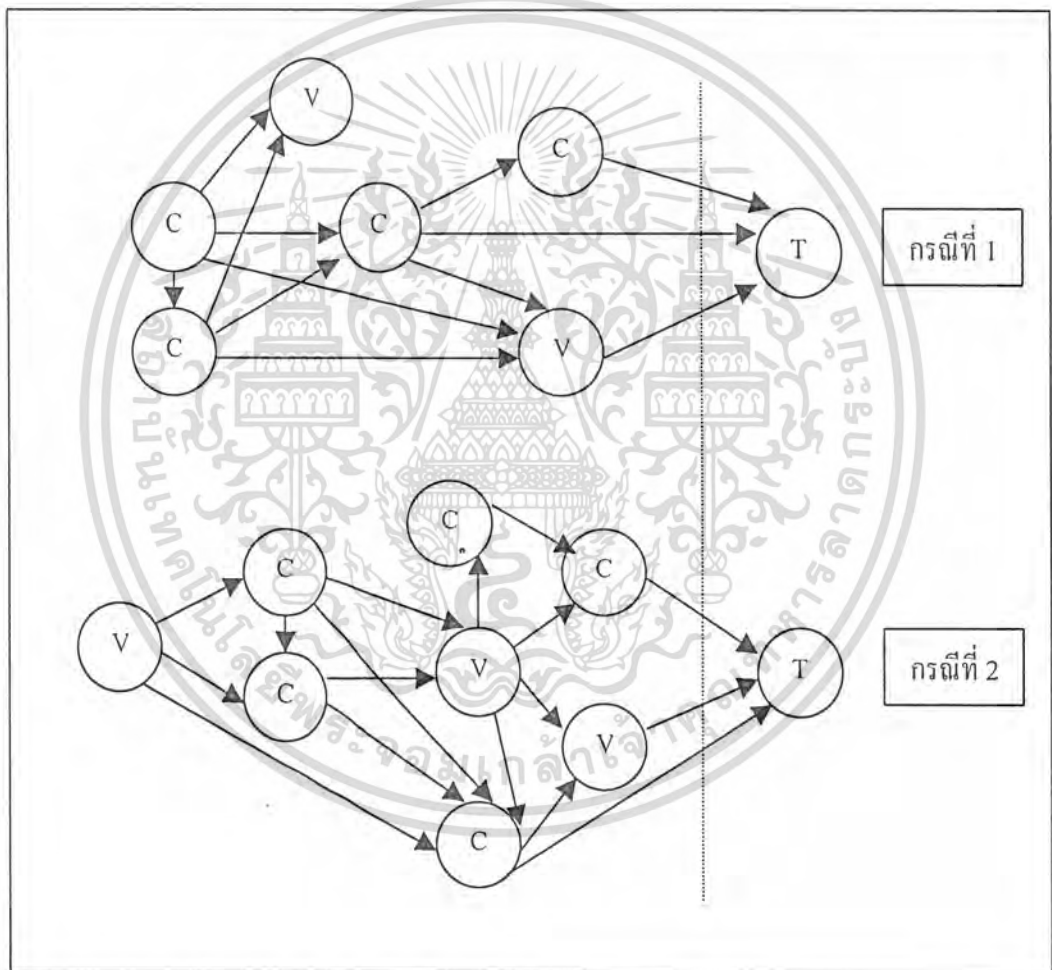
รูปแบบการเขียน	กรณีที่ 1 พยางค์ที่อักษรแรกเป็นพยัญชนะ	ตัวอย่าง
รูปแบบที่		
1 พยัญชนะ+สระ		สา
2 พยัญชนะ+สระ+พยัญชนะ		กาบ
3 พยัญชนะ+สระ+พยัญชนะ+สระ		พัะ
4 พยัญชนะ+พยัญชนะ		กน
5 พยัญชนะ+พยัญชนะ+พยัญชนะ		สวบ
6* พยัญชนะ+พยัญชนะ+สระ		พลี
7* พยัญชนะ+พยัญชนะ+สระ+พยัญชนะ		กราบ
8* พยัญชนะ+พยัญชนะ+สระ+พยัญชนะ+สระ		พลัวะ
9* พยัญชนะ+พยัญชนะ+พยัญชนะ		กลบ
10* พยัญชนะ+พยัญชนะ+พยัญชนะ+สระ		กลอน
หมายเหตุ	รูปแบบที่ 6*-10* เป็นรูปของพยัญชนะควบกล้ำ	

รูปแบบการเขียน	กรณีที่ 2 พยางค์ที่อักษรแรกเป็นสระ	ตัวอย่าง
รูปแบบที่		
1 สระ+พยัญชนะ		โต
2 สระ+พยัญชนะ+พยัญชนะ		โตด
3 สระ+พยัญชนะ+สระ		เขา
4 สระ+พยัญชนะ+พยัญชนะ+สระ		เถอะ
5 สระ+พยัญชนะ+สระ+สระ		เงาะ
6 สระ+พยัญชนะ+สระ+พยัญชนะ		เกิด
7 สระ+พยัญชนะ+สระ+พยัญชนะ+พยัญชนะ		เสียง
8 สระ+พยัญชนะ+สระ+พยัญชนะ+สระ		เพียะ
9* สระ+พยัญชนะ+พยัญชนะ		แคว
10* สระ+พยัญชนะ+พยัญชนะ+พยัญชนะ		โกรง
11* สระ+พยัญชนะ+พยัญชนะ+สระ		เพลา
12* สระ+พยัญชนะ+พยัญชนะ+พยัญชนะ+สระ		เขลอะ
13* สระ+พยัญชนะ+พยัญชนะ+สระ+สระ		เพราะ

- | | | |
|-----|---|---------|
| 14* | สระ+พยัญชนะ+พยัญชนะ+สระ+พยัญชนะ | เสริฐ |
| 15* | สระ+พยัญชนะ+พยัญชนะ+สระ+พยัญชนะ+พยัญชนะ | เพ็ลียง |
| 16* | สระ+พยัญชนะ+พยัญชนะ+สระ+พยัญชนะ+สระ | เพ็ลียะ |

หมายเหตุ รูปแบบที่ 9*-16* เป็นรูปของพยัญชนะควบกล้ำ

จากรูปแบบการเขียนของพยางค์ที่เป็นไปได้ทั้งหมด ถ้าแทนพยัญชนะด้วยสัญลักษณ์ C (Consonant) แทนสระด้วยสัญลักษณ์ V (Vowel) และกำหนดให้วรรณยุกต์เป็นส่วนที่พิจารณาเป็นส่วนตัวที่สุด โดยแทนด้วยสัญลักษณ์ T (Tonality) สามารถเขียนโครงข่ายแสดงการเกิดรูปแบบพยางค์ตามลำดับการเขียนพยัญชนะและสระ ได้ดังรูป



รูปที่ 8.1 แสดงโครงข่ายการเกิดพยางค์ตามลำดับการเขียนพยัญชนะและสระ

จากโครงข่ายการเกิดพยางค์ที่เป็นไปได้ทั้งหมด ใช้วิเคราะห์หาองค์ประกอบของโครงสร้างพยางค์ในรูปของพยัญชนะต้น สระ ตัวสะกด และวรรณยุกต์ เพื่อกำหนดรูปแบบกลางที่จะใช้หาเสียงของวรรณยุกต์ โดยกำหนดให้

$$S = \{C_1 V | C_2\} T \quad \dots(5.1)$$

โดยที่	S	เป็นพยางค์ใด ๆ
	C_1	เป็นพยัญชนะต้น
	V	เป็นสระ
	C_2	เป็นตัวสะกด (อาจจะมีหรือไม่มีก็ได้)
	T	เป็นวรรณยุกต์

สร้างเป็นกฎความสัมพันธ์ของลำดับองค์ประกอบพยางค์กับรูปแบบกลาง ดังต่อไปนี้

กฎที่	รูปแบบพยางค์ตามโครงข่าย		
1	$S = \{CV\}T$		
	$C_1 \rightarrow C$		ในอักขระตัวที่ 1
	$V \rightarrow V$		ในอักขระตัวที่ 2
2	$S = \{CVC\}T$		
	$C_1 \rightarrow C$		ในอักขระตัวที่ 1
	$V \rightarrow V$		ในอักขระตัวที่ 2
	$C_2 \rightarrow C$		ในอักขระตัวที่ 3
3	$S = \{CVCV\}T$		
	$C_1 \rightarrow C$		ในอักขระตัวที่ 1
	$V \rightarrow V$		ในอักขระตัวที่ 2,4
	$C_2 \rightarrow C$		ในอักขระตัวที่ 3
4	$S = \{CC\}T$		
	$C_1 \rightarrow C$		ในอักขระตัวที่ 1
	$V \rightarrow V$		แทนด้วยสระ โ-ะ
	$C_2 \rightarrow C$		ในอักขระตัวที่ 2
5	$S = \{CCC\}T$		
	$C_1 \rightarrow C$		ในอักขระตัวที่ 1
	$V \rightarrow C$		ในอักขระตัวที่ 2 (-ว, -อ)
	$C_2 \rightarrow C$		ในอักขระตัวที่ 3
6	$S = \{CCV\}T$		เมื่ออักขระตัวที่ 2 เป็น [ร, ล, ว]
	$C_1 \rightarrow C+C$		ในอักขระตัวที่ 1,2
	$V \rightarrow V$		ในอักขระตัวที่ 3

- 7 $C_2 \rightarrow$ ไม่มี
 $S = \{CCVC\}T$ เมื่ออักษรตัวที่ 2 เป็น [ร, ล, ว]
 $C_1 \rightarrow C+C$ ในอักษรตัวที่ 1,2
 $V \rightarrow V$ ในอักษรตัวที่ 3
 $C_2 \rightarrow C$ ในอักษรตัวที่ 4
- 8 $S = \{CCVCV\}T$ เมื่ออักษรตัวที่ 2 เป็น [ร, ล, ว]
 $C_1 \rightarrow C+C$ ในอักษรตัวที่ 1,2
 $V \rightarrow V+V$ ในอักษรตัวที่ 3,5
 $C_2 \rightarrow C$ ในอักษรตัวที่ 4
- 9 $S = \{CCC\}T$ เมื่ออักษรตัวที่ 2 เป็น [ร, ล]
 $C_1 \rightarrow C+C$ ในอักษรตัวที่ 1,2
 $V \rightarrow V$ แทนด้วยสระ โ-ะ
 $C_2 \rightarrow C$ ในอักษรตัวที่ 3
- 10 $S = \{CCCC\}T$ เมื่ออักษรตัวที่ 2 เป็น [ร, ล, ว]
 $C_1 \rightarrow C+C$ ในอักษรตัวที่ 1,2
 $V \rightarrow C$ ในอักษรตัวที่ 3 (-ว, -อ)
 $C_2 \rightarrow C$ ในอักษรตัวที่ 4
- 11 $S = \{VC\}T$
 $C_1 \rightarrow C$ ในอักษรตัวที่ 2
 $V \rightarrow V$ ในอักษรตัวที่ 1
 $C_2 \rightarrow$ ไม่มี
- 12 $S = \{VCC\}T$
 $C_1 \rightarrow C$ ในอักษรตัวที่ 2
 $V \rightarrow V$ ในอักษรตัวที่ 1
 $C_2 \rightarrow C$ ในอักษรตัวที่ 3
- 13 $S = \{VCV\}T$
 $C_1 \rightarrow C$ ในอักษรตัวที่ 2
 $V \rightarrow V+V$ ในอักษรตัวที่ 1,3
 $C_2 \rightarrow$ ไม่มี
- 14 $S = \{VCCV\}T$
 $C_1 \rightarrow C$ ในอักษรตัวที่ 2
 $V \rightarrow V+C+V$ ในอักษรตัวที่ 1,3,4
 $C_2 \rightarrow$ ไม่มี

- 15 $S = \{VCVV\}T$
 $C_1 \rightarrow C$ ในอักขระตัวที่ 2
 $V \rightarrow V+V+V$ ในอักขระตัวที่ 1,3,4
 $C_2 \rightarrow$ ไม่มี
- 16 $S = \{VCVC\}T$
 $C_1 \rightarrow C$ ในอักขระตัวที่ 2
 $V \rightarrow V+V$ ในอักขระตัวที่ 1,3 (สระเอ็)
 $C_2 \rightarrow C$ ในอักขระตัวที่ 4
- 17 $S = \{VCVCC\}T$
 $C_1 \rightarrow C$ ในอักขระตัวที่ 2
 $V \rightarrow V+V+C$ ในอักขระตัวที่ 1,3,4
 $C_2 \rightarrow C$ ในอักขระตัวที่ 5
- 18 $S = \{VCVVCV\}T$
 $C_1 \rightarrow C$ ในอักขระตัวที่ 2
 $V \rightarrow V+V+C+V$ ในอักขระตัวที่ 1,3,4,5
 $C_2 \rightarrow$ ไม่มี
- 19 $S = \{VCCC\}T$ เมื่ออักขระตัวที่ 3 เป็น [ร, ล, ว]
 $C_1 \rightarrow C+C$ ในอักขระตัวที่ 2,3
 $V \rightarrow V$ ในอักขระตัวที่ 1
 $C_2 \rightarrow$ ไม่มี
- 20 $S = \{VCCCC\}T$ เมื่ออักขระตัวที่ 3 เป็น [ร, ล, ว]
 $C_1 \rightarrow C+C$ ในอักขระตัวที่ 2,3
 $V \rightarrow V$ ในอักขระตัวที่ 1
 $C_2 \rightarrow C$ ในอักขระตัวที่ 4
- 21 $S = \{VCCCV\}T$ เมื่ออักขระตัวที่ 3 เป็น [ร, ล, ว]
 $C_1 \rightarrow C+C$ ในอักขระตัวที่ 2,3
 $V \rightarrow V+V$ ในอักขระตัวที่ 1,4
 $C_2 \rightarrow$ ไม่มี
- 22 $S = \{VCCCCV\}T$ เมื่ออักขระตัวที่ 3 เป็น [ร, ล, ว]
 $C_1 \rightarrow C+C$ ในอักขระตัวที่ 2,3
 $V \rightarrow V+C+V$ ในอักขระตัวที่ 1,4,5
 $C_2 \rightarrow$ ไม่มี

- 23 $S = \{VCCVV\}T$ เมื่ออักขระตัวที่ 3 เป็น [ร, ล, ว]
 $C_1 \rightarrow C+C$ ในอักขระตัวที่ 2,3
 $V \rightarrow V+V+V$ ในอักขระตัวที่ 1,4,5
 $C_2 \rightarrow$ ไม่มี
- 24 $S = \{VCCVC\}T$ เมื่ออักขระตัวที่ 3 เป็น [ร, ล, ว]
 $C_1 \rightarrow C+C$ ในอักขระตัวที่ 2,3
 $V \rightarrow V+V$ ในอักขระตัวที่ 1,4
 $C_2 \rightarrow C$ ในอักขระตัวที่ 5
- 25 $S = \{VCCVCC\}T$ เมื่ออักขระตัวที่ 3 เป็น [ร, ล, ว]
 $C_1 \rightarrow C+C$ ในอักขระตัวที่ 2,3
 $V \rightarrow V+V+C$ ในอักขระตัวที่ 1,4,5
 $C_2 \rightarrow C$ ในอักขระตัวที่ 6
- 26 $S = \{VCCVCV\}T$ เมื่ออักขระตัวที่ 3 เป็น [ร, ล, ว]
 $C_1 \rightarrow C+C$ ในอักขระตัวที่ 2,3
 $V \rightarrow V+V+C+V$ ในอักขระตัวที่ 1,4,5,6
 $C_2 \rightarrow$ ไม่มี

กฎความสัมพันธ์ของลำดับองค์ประกอบพยางค์กับรูปแบบกลางทั้งหมด 26 กฎ เราใช้หารูปแบบกลางที่ใช้อธิบายโครงสร้างของพยางค์ตามหน้าที่ แต่ยังไม่สามารถอธิบายเสียง เนื่องจากรูปวรรณยุกต์ (T) ในสมการดังกล่าวจะถูกกำหนดเสียงได้หลายเสียง แปรผันตามหน่วยพื้นฐานที่นำมาประกอบเป็นพยางค์ ดังนั้นการที่จะหาเสียงพยางค์ จะต้องทำการวิเคราะห์หาเสียงวรรณยุกต์ที่ปรากฏรูปในรูปแบบกลาง ภายใต้ข้อกำหนดของไวยากรณ์ทางโครงสร้างพยางค์กับระบบเสียงหรือความสัมพันธ์ของหน่วยเสียง (Phoneme) ประเภทต่าง ๆ ที่ประกอบกันเป็นเสียงของพยางค์นั้น ซึ่งจะได้อธิบายในรายละเอียดต่อไป

8.3 การวิเคราะห์รูปแบบกลางเพื่อหาเสียงเป้าหมาย

จากฐานความรู้เรื่องระบบเสียงภาษาไทย พบว่าเสียงพยางค์ เกิดจากการรวมกันของหน่วยเสียง อย่างน้อยที่สุดต้องมี 3 หน่วยเสียงพื้นฐาน คือ เสียงพยัญชนะต้น เสียงสระ และเสียงวรรณยุกต์ โดยมีเสียงสระเป็นแกนของเสียงพยางค์ การสร้างเสียงพยางค์ใด ๆ ได้จากการนำหน่วยเสียงพื้นฐานที่มีอยู่ทั้งหมดมา รวมกันเป็นโครงสร้างพยางค์ตามหลักไวยากรณ์ ดังที่ได้กล่าวมาแล้ว

การหาสมการเพื่ออธิบายเสียงพยางค์ ต้องมีความสอดคล้องกับสมการรูปแบบกลาง คือ สามารถ แทนหน่วยเสียงในตำแหน่งของพยัญชนะต้น สระ และตัวสะกด เข้าไปในรูปแบบกลางได้เลย แต่สำหรับ ระดับเสียงหรือเสียงวรรณยุกต์ จะต้องทำการวิเคราะห์ความสัมพันธ์ตามรูปวรรณยุกต์ที่ปรากฏในรูปแบบ กลาง เราไม่สามารถใช้รูปแบบกลางแทนสมการของเสียงพยางค์โดยตรง เนื่องจากรูปของวรรณยุกต์กับ เสียงวรรณยุกต์ไม่จำเป็นต้องตรงกันเสมอไป ซึ่งต่างจากหน่วยเสียงพยัญชนะหรือหน่วยเสียงสระที่แสดง หรืออธิบายตามรูปที่ปรากฏ แต่เสียงของวรรณยุกต์ขึ้นอยู่กับโครงสร้างพยางค์กับการออกเสียงตามหลัก ไวยากรณ์ภาษาไทย ระบบเสียงในภาษาไทยประกอบด้วยหน่วยเสียงพื้นฐานที่สำคัญ ๆ คือ หน่วยเสียง พยัญชนะ (ทำหน้าที่เป็นเสียงพยัญชนะต้นและพยัญชนะท้ายหรือตัวสะกด) หน่วยเสียงสระ (ทำหน้าที่เป็น แกนของเสียงพยางค์) และหน่วยเสียงวรรณยุกต์ (ทำหน้าที่กำหนดระดับเสียง) โดยกำหนดสัญลักษณ์ให้ เสียงของวรรณยุกต์เป็น 1, 2, 3, 4, 5 แทนระดับเสียง สามัญ, เอก, โท, ศรี, จัตวา ตามลำดับ จะแสดงความ สัมพันธ์ระหว่างระดับเสียงกับรูปวรรณยุกต์ที่ปรากฏในสมการกลาง

เราสามารถแบ่งกลุ่มระดับเสียงที่สัมพันธ์กับรูปวรรณยุกต์ในทิศทางเดียวกันได้ 3 กลุ่ม ตามเสียง ของพยัญชนะต้น และจัดกลุ่มหน่วยเสียงพื้นฐานได้ดังตารางที่ 7.1 - 7.4

หน่วยเสียงพยัญชนะต้น แทนด้วยสัญลักษณ์ C_i เมื่อ $i = 1, 2, 3$

กลุ่มที่ (i)	เดี่ยว	กวางกล้า	ต่างประเทศ
1 อักษรกลาง (บางส่วน)	/k/-ก, /c/-จ, /d/-ดฎ, /t/-ตฎ, /b/-บ, /p/-ป, /?/-อ	/kr/-กร, /kl/-กล, /kw/-กว, /pr/-ปร, /pl/-ปล, /tr/-ตร	/dr/-ดร, /br/-บร, /bl/-บล
2 อักษรสูง (บางส่วน)	/kh/-ข, /ch/-ฉ, /s/-สศษ, /th/-ถฐ, /ph/-ผ, /f/-ฝ, /h/-ฮ	/khr/-ขร, /khl/-ชล, /khw/-ขว, /phr/-พร, /phl/-พล	หร, หล, หว, หน, หม, หย, หญ
3 อักษรต่ำ (บางส่วน)	/kh/-คฌ, /ng/-ง, /ch/-ชฌ, /s/-ซ, /y/-ยญ, /th/-ทธฒา, /n/-นณ, /ph/-พภ, /f/-ฟ, /m/-ม, /r/-ร, /l/-ล, /w/-ว, /h/-ฮ	/khr/-คร, /khl/-คล, /khw/-คว, /phr/-พร, /phl/-พล	/fr/-ฟร, /n/-ฟล, /thr/-ทร

ตารางที่ 8.1 แสดงการจัดกลุ่มของหน่วยเสียงพยัญชนะต้น

หน่วยเสียงพยัญชนะท้าย (ตัวสะกด) แทนด้วยสัญลักษณ์ C_{k2} เมื่อ $k = 1, 2$

กลุ่มที่ (k)	เสียงสะกด
1 (ทำให้พยางค์มีเสียงกักตอนท้าย)	/-k/-ก-ข-ค, /-p/-ก-ค-จ-ศ-ษ-ฏ-ฎ-ท-ช-ฐ-ฒ-ณ-ธ, /-ʔ/-ะ, /-p/-บ-ป-พ-ฟ-ภ
2 (ทำให้พยางค์มีเสียงก้องตอนท้าย)	/-ng/-ง, /-n/-น-ณ-ร-ล, /-m/-ม, /-y/-ย-ญ, /-w/-ว

ตารางที่ 8.2 แสดงการจัดกลุ่มของหน่วยเสียงพยัญชนะท้าย

หน่วยเสียงสระ แทนด้วยสัญลักษณ์ V_j เมื่อ $j = 1, 2$

กลุ่มที่ (j)	สระเดี่ยว	สระผสม
1 (สระเสียงสั้น)	/a/-อะ, /i/-อิ, /u/-อุ, /uc/-อี, /c/-เอะ, /o/-โอะ, /oa/-เอาะ, /cr/-เออะ, /ac/-แออะ	/ua/-อัวะ, /uca/-เอือะ, /ia/-เอือะ
2 (สระเสียงยาว)	/a_/-อา, /i_/-อี, /u_/-อุ, /uc_/-อีอ, /c_/-เออ, /o_/-โออ, /oa_/-เอออ, /cr_/-เอออ, /ac_/-แออ	/ua_/-อัว, /uca_/-เอืออ, /ia_/-เอืออ

ตารางที่ 8.3 แสดงการจัดกลุ่มของหน่วยเสียงสระ

หน่วยเสียงวรรณยุกต์ แทนด้วยสัญลักษณ์ T_i เมื่อ $i = 1, 2, 3, 4, 5$

กลุ่มที่ (i)	เสียงวรรณยุกต์
1 (รูปวรรณยุกต์ - สามัญ)	/1/ เสียงสามัญ
2 (รูปวรรณยุกต์ - เอก)	/2/ เสียงเอก
3 (รูปวรรณยุกต์ - โท)	/3/ เสียงโท
4 (รูปวรรณยุกต์ - ตรี)	/4/ เสียงตรี
5 (รูปวรรณยุกต์ - จัตวา)	/5/ เสียงจัตวา

ตารางที่ 8.4 แสดงการจัดกลุ่มของหน่วยวรรณยุกต์

เราสามารถหาเสียงเป้าหมายของพยางค์ และสร้างสมการแทนเสียงพยางค์ ดังนี้

$$S \rightarrow \{C_{ii}V_j[C_{2k}]\}T_{1 \rightarrow m} \quad \dots(5.2)$$

โดยที่ S เป็นเสียงพยางค์

C_{ii} เป็นเสียงพยัญชนะต้นในกลุ่มที่ $i = 1, 2, 3$

V_j เป็นเสียงสระในกลุ่มที่ $j = 1, 2$

C_{2k} เป็นเสียงตัวสะกดในกลุ่มที่ $k = 1, 2$

$T_{1 \rightarrow m}$ เป็นเสียงวรรณยุกต์เสียงที่ m ตามรูปวรรณยุกต์ที่ $1 = 1, 2, 3, 4, 5$

จากความสัมพันธ์ตามรูปสมการเสียง $\{C_{ii}V_j[C_{2k}]\}T_{1 \rightarrow m}$ ทำให้เราทราบจำนวนเสียงพยางค์ที่มีในภาษาไทย จะมีจำนวนจำกัด สามารถแสดงได้ดังนี้

เสียงพยัญชนะต้น C_{11}, C_{12}, C_{13} หรือ C_1 38 เสียง

เสียงสระ V_1 12 เสียง

V_2 12 เสียง

เสียงตัวสะกด C_{21} 4 เสียง

C_{22} 5 เสียง

เสียงวรรณยุกต์ พื้นได้ 5 เสียง

หาเสียงของพยางค์ ตามเงื่อนไขทางไวยากรณ์ จะได้

C_1V_1 พื้นวรรณยุกต์ได้ 4 เสียง $= 38 \times 12 \times 4$ 1824 เสียง

C_1V_2 พื้นวรรณยุกต์ได้ 5 เสียง $= 38 \times 12 \times 5$ 2280 เสียง

$C_1V_1C_{21}$ พื้นวรรณยุกต์ได้ 4 เสียง $= 38 \times 12 \times 4 \times 4$ 7296 เสียง

$C_1V_2C_{21}$ พื้นวรรณยุกต์ได้ 4 เสียง $= 38 \times 12 \times 4 \times 4$ 7296 เสียง

$C_1V_1C_{22}$ พื้นวรรณยุกต์ได้ 5 เสียง $= 38 \times 12 \times 5 \times 5$ 11400 เสียง

$C_1V_2C_{22}$ พื้นวรรณยุกต์ได้ 5 เสียง $= 38 \times 12 \times 5 \times 5$ 11400 เสียง

รวมเสียงของพยางค์ที่มีในภาษาทั้งสิ้น 41,496 เสียง

8.4 การวิเคราะห์สมการเสียงเพื่อกำหนดสัญลักษณ์แทนเสียง

สมการเสียงพยางค์ ใช้อธิบายเสียงของพยางค์แต่ละพยางค์ สามารถแปลงเป็นสัญลักษณ์แทนเสียง ซึ่งประกอบด้วยวิธีการ 2 ขั้นตอน คือ

การค้นหาลัญลักษณ์แทนเสียงในฐานข้อมูล

จากสมการรูปแบบกลางและสมการเสียงพยางค์ เราทราบว่าโครงสร้างพยางค์ที่ประกอบด้วย พยัญชนะต้น สระ ตัวสะกด อยู่ในกลุ่มใดบ้าง ซึ่งการจัดเก็บในฐานข้อมูล จะมีการแยกเป็นกลุ่มตามประเภทของอักขระ ดังนี้

พยัญชนะต้น เก็บบันทึกอักขระทั้งหมดที่สามารถเกิดเป็นพยัญชนะต้นได้ ทั้งประเภทเดี่ยวและควบกล้ำ (รวมทั้งพยัญชนะต้นที่ใช้อักษรนำ หรือที่ใช้กับภาษาต่างประเทศ เช่น หร, หล, หลย หรือ ฟร, ฟล เป็นต้น) ฐานข้อมูลจะมีโครงสร้างและการจัดเก็บ ดังนี้

อักขระ	กลุ่มที่	สัญลักษณ์แทนเสียง
ก	1	/k/
ข	2	/kh/
ค	3	/kh/
:	:	:
:	:	:
อ	1	/ʔ/
ฮ	3	/h/

ตารางที่ 8.5 แสดง โครงสร้างฐานข้อมูลที่ใช้เก็บรายละเอียดเกี่ยวกับพยัญชนะต้น

พยัญชนะท้ายหรือตัวสะกด เก็บบันทึกอักขระที่สามารถเกิดเป็นตัวสะกดได้ ฐานข้อมูลจะมีโครงสร้างและการจัดเก็บ ดังนี้

อักขระ	กลุ่มที่	สัญลักษณ์แทนเสียง
ก	1	/k/
ค	1	/t/
บ	1	/b/
.	.	.
:	:	:
:	:	:
ว	2	/w/

ตารางที่ 8.6 แสดง โครงสร้างฐานข้อมูลที่ใช้เก็บรายละเอียดเกี่ยวกับพยัญชนะท้าย

สระ เก็บบันทึกอักขระที่สามารถเกิดเป็นสระได้ทั้งหมด ทั้งประเภทสระเดี่ยวและสระผสม ฐานข้อมูลจะมีโครงสร้างและการจัดเก็บ ดังนี้

อักขระ	กลุ่มที่	สัญลักษณ์แทนเสียง	กึ่งหนึ่งของเสียงสระ	สัญลักษณ์แทนเสียงกึ่งหนึ่ง
-ะ	1	/a/	-ะ	/a/
-า	2	/a_ /	-ะ	/a/
-เ	2	/e_ /	-ะ	/e/
:	:	:	:	:
:	:	:	:	:
-อ	2	/oa_ /	-เ-ะ	/oa/

ตารางที่ 8.7 แสดงโครงสร้างฐานข้อมูลที่ใช้เก็บรายละเอียดเกี่ยวกับสระ

สำหรับเสียงและสัญลักษณ์แทนเสียงวรรณยุกต์ เราแทนด้วยตัวเลขที่บอกระดับเสียงตามสมการเสียง เช่น

$$S \rightarrow [C_1 V_1 C_2] T_{1-2} \quad \text{สัญลักษณ์แทนเสียงวรรณยุกต์} = 2$$

การกำหนดสัญลักษณ์แทนเสียงให้กับพยางค์

เมื่อเราได้สมการแทนเสียงวรรณยุกต์ของพยางค์เสียงแล้ว ก็นำสมการเสียงนั้นมาแทนด้วยสัญลักษณ์แทนเสียง ตามกฎการแปลงเสียง โดยกฎนี้จะขึ้นอยู่กับการจัดเก็บบันทึกหน่วยเสียงในหน่วยจัดเก็บข้อมูล ดังนั้นการสังเคราะห์เสียงโดยวิธีนี้เป็นวิธีการสังเคราะห์เสียงโดยใช้กฎการแปลงเสียง (Synthesis by Rules) นั้นเอง ซึ่งกฎการแปลงเสียงจะขึ้นอยู่กับการจัดเก็บบันทึกหน่วยทางเสียง ที่มีอยู่หลายวิธี ดังที่ได้กล่าวถึงในบทที่ 3

8.5 ตัวอย่างการทำงานของระบบการแปลงหน่วยคำเป็นสัญลักษณ์แทนเสียงโดยวิธี MPT-networks

วิธี MPT-networks เป็นอัลกอริทึมที่ใช้ในการแปลงพยางค์ของคำ ที่ได้จากการแยกคำออกจากข้อความ ตามวิธีการ Wordbucket ให้อยู่ในรูปสัญลักษณ์แทนเสียงของหน่วยเสียงพื้นฐาน โดยมีลำดับขั้นตอนการทำงาน ดังนี้

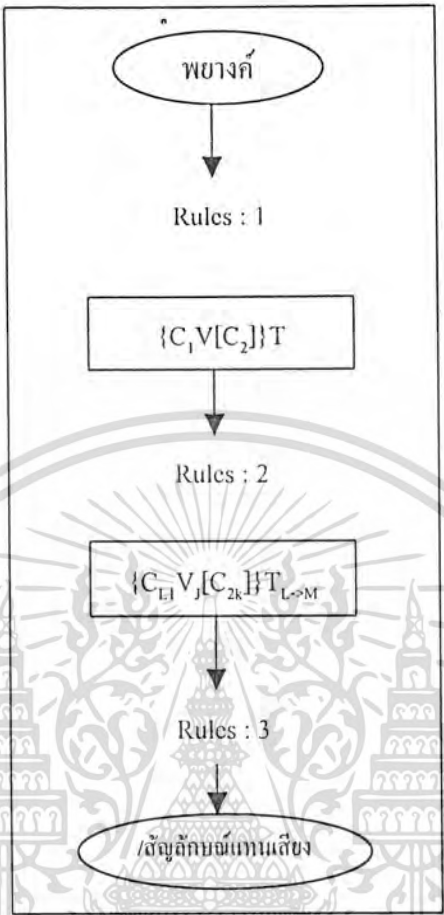
ขั้นตอนที่ 1 นำแต่ละพยางค์มาวิเคราะห์หาองค์ประกอบภายในพยางค์ เพื่อหารูปแบบการเขียนของพยางค์นั้นโดยแยกวรรณยุกต์ไว้ในตำแหน่งท้ายสุด

ขั้นตอนที่ 2 นำรูปแบบการจัดเรียงที่ได้มาวิเคราะห์ความสัมพันธ์ตามโครงข่ายการเกิดพยางค์ที่เป็นไปได้ทั้งหมด เพื่อกำหนดรูปแบบกลาง

ขั้นตอนที่ 3 นำรูปแบบกลางมาวิเคราะห์หาเสียงของวรรณยุกต์ตามโครงข่ายความสัมพันธ์ระหว่างโครงข่ายพยางค์กับการกำหนดเสียงวรรณยุกต์ เพื่อกำหนดเป็นสมการแทนเสียงพยางค์

ขั้นตอนที่ 4 ทำการแปลงสมการแทนเสียงพยางค์ให้เป็นสัญลักษณ์แทนเสียง ที่ตรงกับลักษณะของหน่วยเสียงพื้นฐานที่ทำการจัดเก็บ ซึ่งในปริยญาณิพนธ์ฉบับนี้ใช้การจัดเก็บหน่วยเสียงตามแบบที่ 3 ของหน่วยเสียงพื้นฐานในบทที่ 4 แล้วทำขั้นตอนตั้งแต่ 1-4 กับพยางค์ที่เหลือทั้งหมด

จากขั้นตอนการทำงานเราสามารถแสดงดังรูป



รูปที่ 8.2 แสดงขั้นตอนการทำงานของอัลกอริทึม MPT-networks

จากข้อความ “คณะวิศวกรรมศาสตร์ พระจอมเกล้าลาดกระบัง” เมื่อผ่านกระบวนการตัดคำ โดยวิธี WordBucket จะได้หน่วยคำและเสียงอ่านดังนี้

คณะ	วิศวกรรมศาสตร์	พระจอมเกล้า	ลาดกระบัง
กะ-นะ	วิด-สะ-วะ-กำ-มะ-สาด	พระ-จอม-เกล้า	ลาด-กระ-บัง

ตารางที่ 8.8 แสดงผลกระบวนการตัดคำโดยวิธี WordBucket

ในที่นี้จะใช้ตัวอย่างจากหน่วยคำ “วิศวกรรมศาสตร์” ที่ประกอบด้วยพยางค์ 6 พยางค์ คือ “วิ-ส-วะ-กะ-กำ-มะ-สาธ” มาทำการวิเคราะห์เพื่อแปลงเป็นสัญลักษณ์แทนเสียงโดยแสดงเป็นลำดับขั้นตอนดังนี้

ขั้นตอนที่ 1 ทำการแยกแต่ละพยางค์ออกจากกันจะได้

พยางค์ที่ 1	“วิ”
พยางค์ที่ 2	“ส”
พยางค์ที่ 3	“วะ”
พยางค์ที่ 4	“กำ”
พยางค์ที่ 5	“มะ”
พยางค์ที่ 6	“สาธ”

นำแต่ละพยางค์มาวิเคราะห์หาองค์ประกอบภายในพยางค์ เพื่อหารูปแบบการเขียนของพยางค์นั้น โดยแยกดูพยางค์ที่อยู่ในตำแหน่งท้ายสุด โดยที่พยางค์แรกได้แก่พยางค์ “วิ” ทำการแยกอักขระแต่ละตัว แล้วนำไปเทียบหาในฐานข้อมูลเพื่อตรวจสอบชนิดของอักขระ ได้ดังนี้

อักขระ “วะ”	เป็นพยัญชนะ	แทนด้วย	สัญลักษณ์	“C”
“อิ”	เป็นสระ	แทนด้วย	สัญลักษณ์	“V”
“ค”	เป็นพยัญชนะ	แทนด้วย	สัญลักษณ์	“C”
ไม่มีรูปวรรณยุกต์สามัญ		แทนด้วย	สัญลักษณ์	“T”

ดังนั้นพยางค์ “วิ” จัดเรียงลำดับในสได้ คือ $S = \{CVC\}T$

ขั้นตอนที่ 2 นำรูปแบบการจัดวางที่ได้ คือ $S = \{CVC\}T$ มาวิเคราะห์ความสัมพันธ์ตามโครงข่ายการเกิดพยางค์ที่เป็นไปได้ทั้งหมด เพื่อใช้กำหนดรูปแบบกลาง พบว่าตรงกับกำหนดข้อที่ 2 ดังนั้นเราได้สมการรูปแบบกลาง

$$S = \{C_1VC_2\}T$$

โดยที่

$C_1 \rightarrow C$	ในอักขระตัวที่ 1	หรือ	$C_1 \rightarrow \text{“ว”}$
$V \rightarrow V$	ในอักขระตัวที่ 2	หรือ	$V \rightarrow \text{“อิ”}$
$C_2 \rightarrow C$	ในอักขระตัวที่ 3	หรือ	$C_2 \rightarrow \text{“ค”}$

ขั้นตอนที่ 3 นำสมการรูปแบบกลาง $S = \{C_1VC_2\}T$ มาวิเคราะห์หาเสียงของวรรณยุกต์ โดยการพิจารณาหากลุ่มของแต่ละองค์ประกอบที่อยู่ในสมการรูปแบบกลาง ดังนี้

“วะ”	เป็นพยัญชนะต้น	ในกลุ่มที่	แทนด้วย C_{13}
“อิ”	เป็นสระ	ในกลุ่มที่	แทนด้วย V
“ค”	เป็นพยัญชนะท้าย/สะกด	ในกลุ่มที่	แทนด้วย C_{21}
และ	รูปวรรณยุกต์สามัญ		แทนด้วย $T = 1$

จากโครงข่ายความสัมพันธ์ระหว่างโครงสร้างพยางค์กับการกำหนดเสียงวรรณยุกต์ เราจะได้เสียงวรรณยุกต์ คือ เสียงตรี (4) หรือ T_{1-4} และได้สมการเสียงเป็น

$$S = \{C_{13}V_1C_{21}\}T_{1-4}$$

ขั้นตอนที่ 4 ทำการแปลงพยางค์ตามสมการแทนเสียงเป็นสัญลักษณ์แทนเสียง

ในรูปหน่วยพยางค์ วิด = /wit4/

ในรูปหน่วยอັตพยางค์ วิด = /wi-it4/

แล้วทำขั้นตอนตั้งแต่ 1-4 กับพยางค์ที่เหลือทั้งหมด จะได้สัญลักษณ์แทนเสียงแต่ละพยางค์ ดังนี้

พยางค์	หน่วยพยางค์	หน่วยอັตพยางค์
สะ	/sa2/	/sa-a2/
วะ	/wa4/	/wa-a4/
กำ	/kam1/	/ka-am1/
มะ	/ma4/	/ma-a4/
สาต	/sa_t2/	/sa-a_t2/

ตารางที่ 8.9 แสดงการแปลงพยางค์ตามสมการแทนเสียงเป็นสัญลักษณ์แทนเสียง

ท้ายที่สุด เมื่อนำสัญลักษณ์แทนเสียงของแต่ละพยางค์มารวมเข้าด้วยกัน เราจะได้สัญลักษณ์แทนเสียง /wi-it4, sa-a2, sa-a2, ka-am1, ka-am1, sa-a_t2/ เป็น ตัวแทนเสียงในแต่ละพยางค์ของคำว่า “วิศวกรรมศาสตร์”

8.6 สรุป

โดยวิธีการของ MPT-networks เป็นอัลกอริทึมที่ใช้แปลงรูปแบบการเขียนให้อยู่ในรูปแบบกลาง ที่อธิบายหน้าที่ของแต่ละองค์ประกอบภายในพยางค์จากนั้นแบ่งกลุ่มหรือประเภทของแต่ละองค์ประกอบเพื่อหาเสียงเป้าหมายตามกฎความสัมพันธ์โครงสร้างพยางค์กับเสียงวรรณยุกต์ได้เป็นสมการเสียงพยางค์แล้วจึงนำไปกำหนดสัญลักษณ์แทนเสียงของแต่ละหน่วยอັตพยางค์ตามกฎการแปลงพยางค์เป็นสัญลักษณ์แทนเสียง

บทที่ 9

การออกแบบและการสร้างโปรแกรม

9.1 บทนำ

ในบทนี้จะกล่าวถึงการออกแบบโปรแกรมสังเคราะห์เสียงตั้งแต่ระดับล่างจนถึงระดับบน และกล่าวถึงการสร้างโปรแกรมสังเคราะห์เสียงซึ่งได้จากการออกแบบในระดับสุดท้ายของการออกแบบโปรแกรม

9.2 การออกแบบโปรแกรม

การออกแบบโปรแกรมสังเคราะห์เสียงภาษาไทยด้วยคอมพิวเตอร์ จะออกแบบตามโครงสร้างการทำงานดังต่อไปนี้

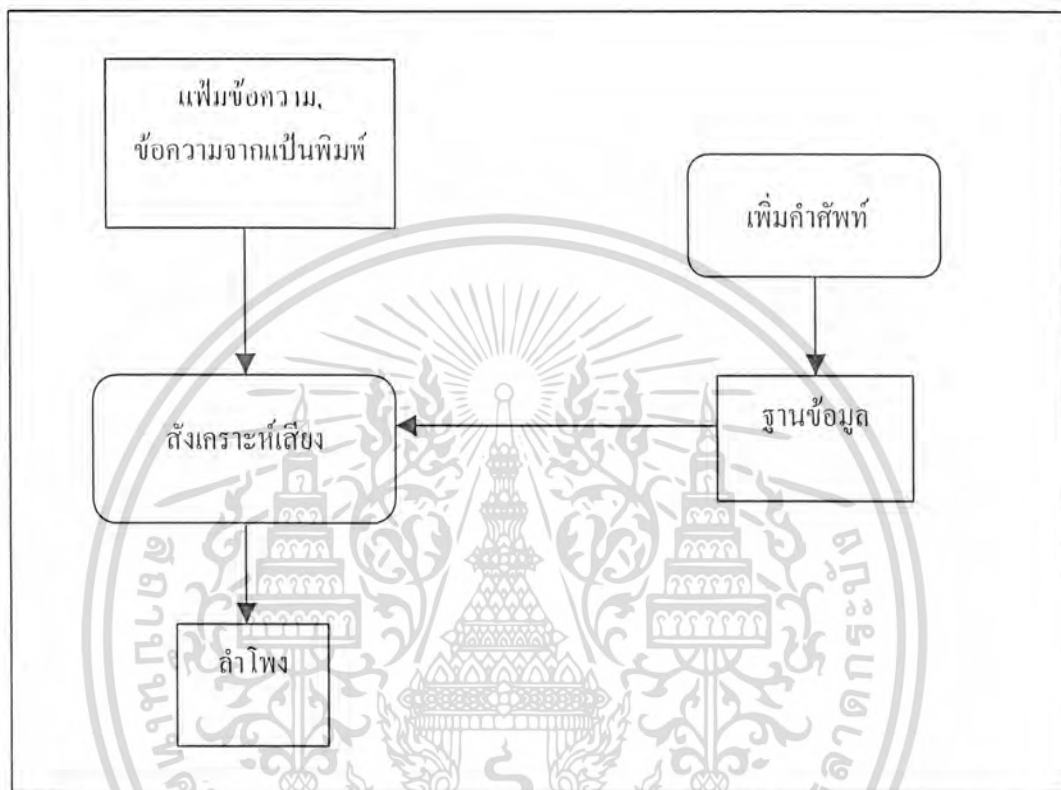


รูปที่ 9.1 โครงสร้างการทำงานของโปรแกรมสังเคราะห์เสียงภาษาไทยด้วยคอมพิวเตอร์

Diagram การออกแบบ

Diagram การออกแบบ จะแสดงการออกแบบโปรแกรมในแต่ละระดับตั้งแต่ระดับล่างที่มีรายละเอียดน้อยไปจนถึงระดับบนที่มีรายละเอียดมาก

Diagram การออกแบบระดับที่ 1

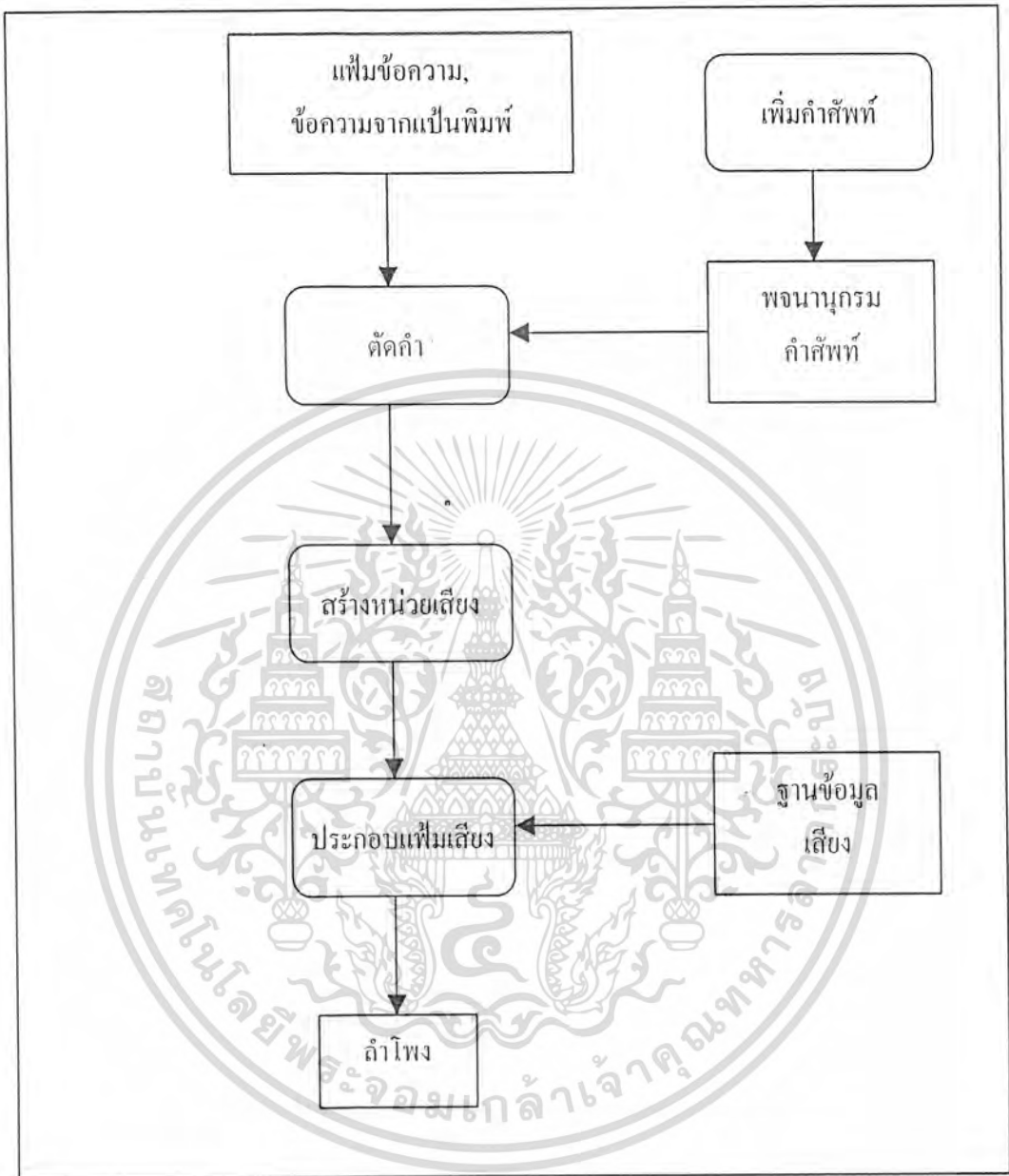


รูปที่ 9.2 Diagram การออกแบบระดับที่ 1

อธิบายการทำงานของ Diagram การออกแบบระดับที่ 1

Diagram การออกแบบระดับที่ 1 แสดงการทำงานของโปรแกรมสังเคราะห์เสียง ซึ่งทำการรับข้อมูลได้ 2 ทาง คือ จากเพิ่มข้อความ หรือ แป้นพิมพ์ เมื่อรับข้อความแล้วจะแปลงข้อความให้เป็นแฟ้มเสียงในกระบวนการสังเคราะห์เสียง และส่งแฟ้มเสียงแล้วส่งผ่านทางลำโพงต่อไป ส่วนของการเพิ่มคำศัพท์ใหม่จะเป็นการเพิ่มคำศัพท์ที่ยังไม่มีในฐานข้อมูลลงไปในฐานข้อมูลตามที่ใช้ต้องการ

Diagram การออกแบบระดับที่ 2



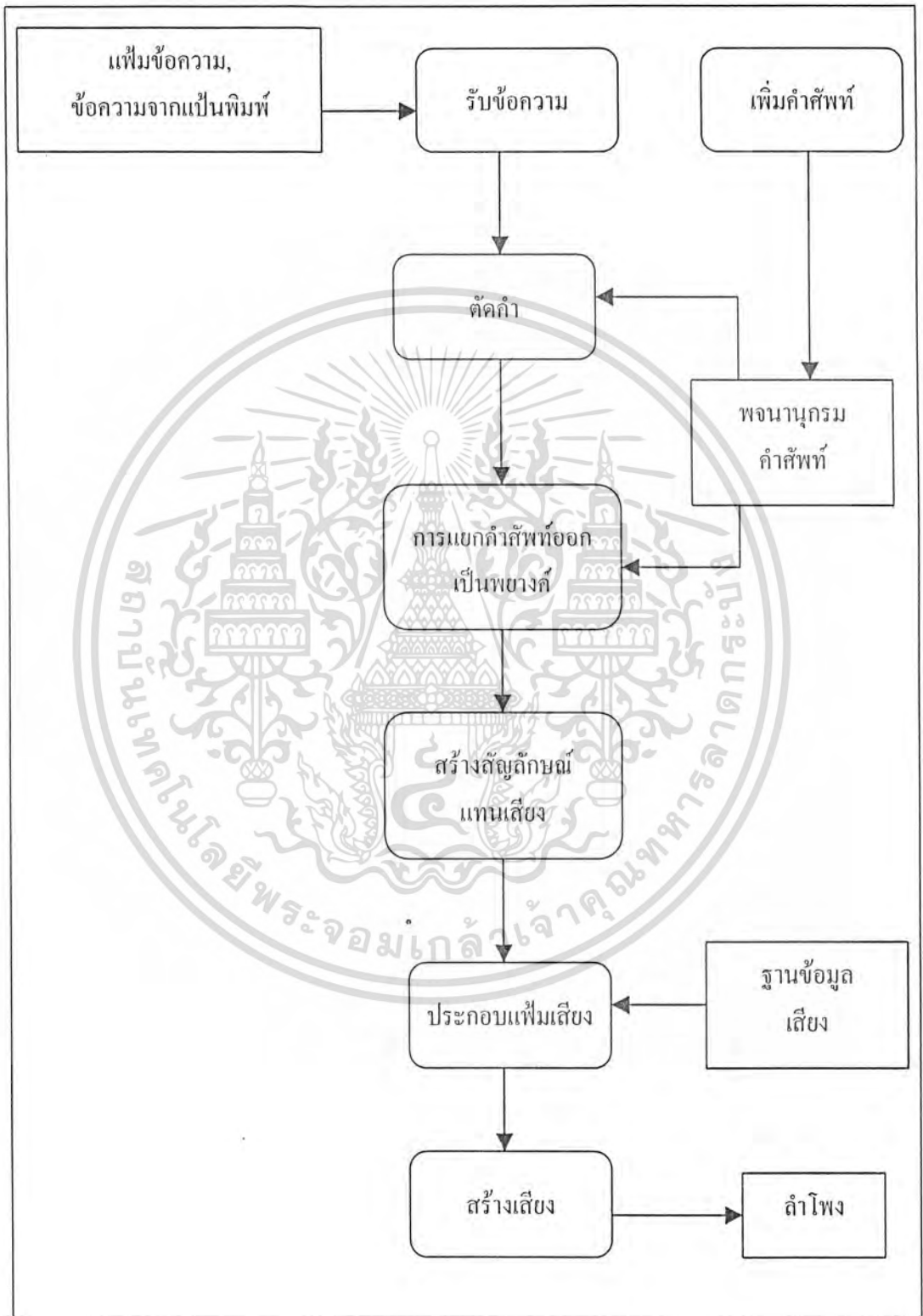
รูปที่ 9.3 Diagram การออกแบบระดับที่ 2

อธิบายการทำงานของ Diagram การออกแบบระดับที่ 2

Diagram การออกแบบระดับที่ 2 จะแบ่งส่วนของการสังเคราะห์เสียงใน Diagram ระดับที่ 1 ออกเป็น 2 ส่วนคือ

1. “ส่วนของการสร้างหน่วยเสียง” เป็นการหาหน่วยเสียงย่อยๆ ที่จะมาประกอบเป็นเสียง 1 เสียง
2. “ส่วนของการประกอบเพิ่มเสียง” เป็นการรวมเพิ่มเสียงของแต่ละหน่วยเสียงเข้าด้วยกัน เพื่อสำหรับการออกเสียง 1 เสียง

Diagram การออกแบบระดับที่ 3

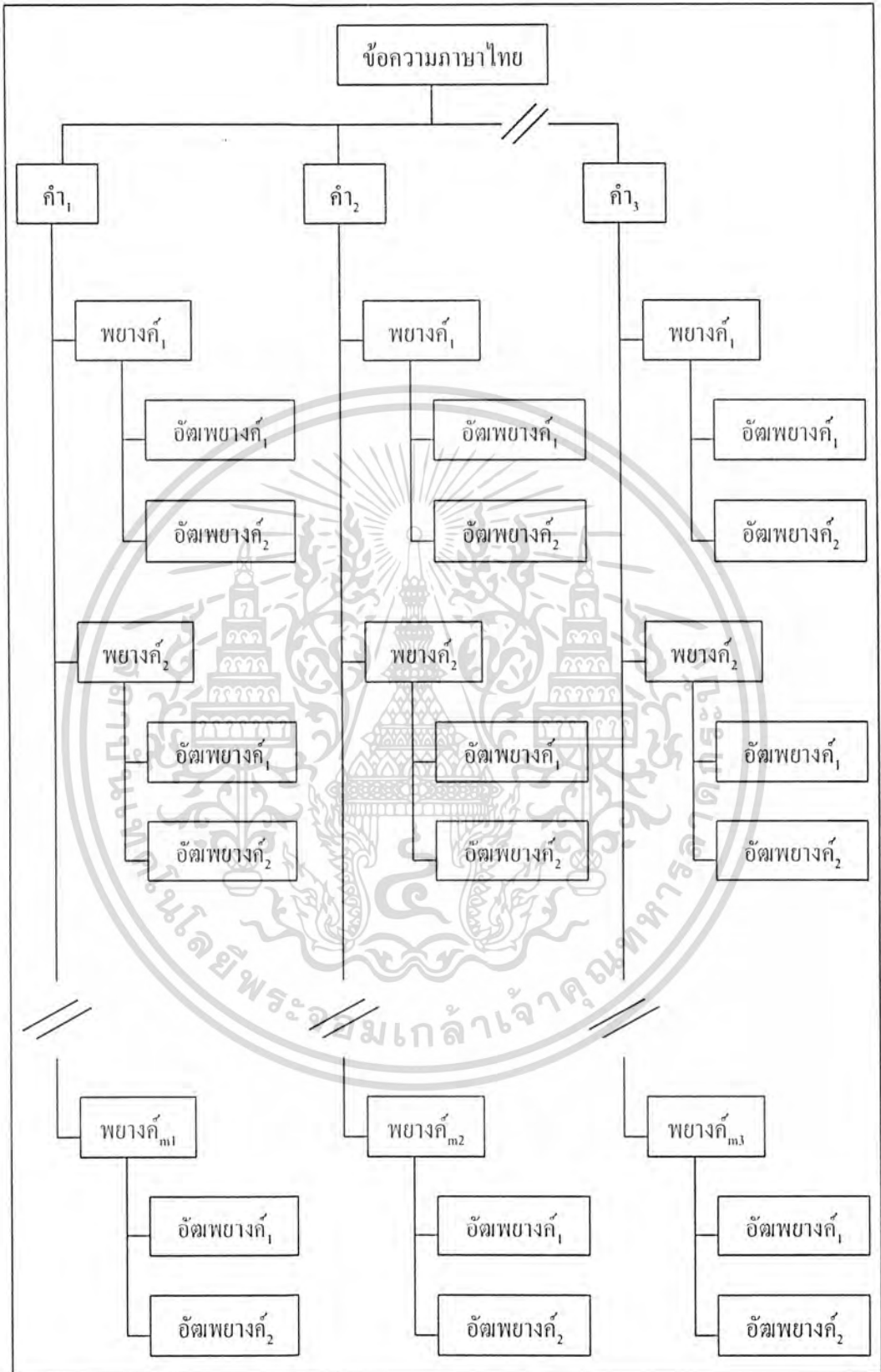


รูปที่ 9.4 Diagram การออกแบบระดับที่ 3

อธิบายการทำงานของ Diagram การออกแบบระดับที่ 3

Diagram ระดับที่ 3 เป็น Diagram ที่แสดงองค์ประกอบทั้งหมดของโปรแกรม ซึ่ง Diagram ระดับนี้ จะเป็น Diagram ที่จะนำไปพัฒนาเป็นโปรแกรมสำหรับใช้งานจริง Diagram ระดับที่ 3 ประกอบไปด้วย ส่วนต่าง ๆ ดังต่อไปนี้

1. “ส่วนรับข้อความ” เป็นส่วนที่ทำหน้าที่รับข้อความจาเพิ่มข้อความหรือจากแป้นพิมพ์ที่ได้จากผู้ใส่
2. “ส่วนคิดค่า” เป็นส่วนที่แยกข้อความออกเป็นคำภาษาไทย
3. “ส่วนแยกคำออกเป็นพยางค์” เป็นการเปลี่ยนคำภาษาไทยออกเป็นพยางค์ย่อยในคำภาษาไทยนั้น ๆ
4. “ส่วนสร้างสัญลักษณ์แทนเสียง” เป็นการแยกพยางค์ออกเป็นสัญลักษณ์แทนเสียงที่แทนหน่วยพยางค์ย่อยที่เรียกว่า อัฒพยางค์ ตามวิธี MPT Network โดย 1 พยางค์จะประกอบไปด้วยสัญลักษณ์แทนเสียงอัฒพยางค์ 2 อัฒพยางค์
5. “ส่วนของการประกอบเพิ่มเสียง” คือ ส่วนที่รวบรวมเพิ่มเสียงของแต่ละอัฒพยางค์ให้เป็นเพิ่มเสียงสำหรับการออกเสียง 1 พยางค์ ซึ่ง 1 พยางค์คือ 1 เสียง สำหรับการสร้างเสียง 1 เสียง จะประกอบด้วยการรวมอัฒพยางค์ 2 อัฒพยางค์นั่นเอง
6. “ส่วนของการสร้างเสียง” เป็นส่วนที่นำเพิ่มเสียงไปแปลงเป็นเสียงส่งออกมาทางลำโพงให้เราได้ยิน



รูปที่ 9.5 แสดงโครงสร้างของหน่วยเสียงระดับต่าง ๆ

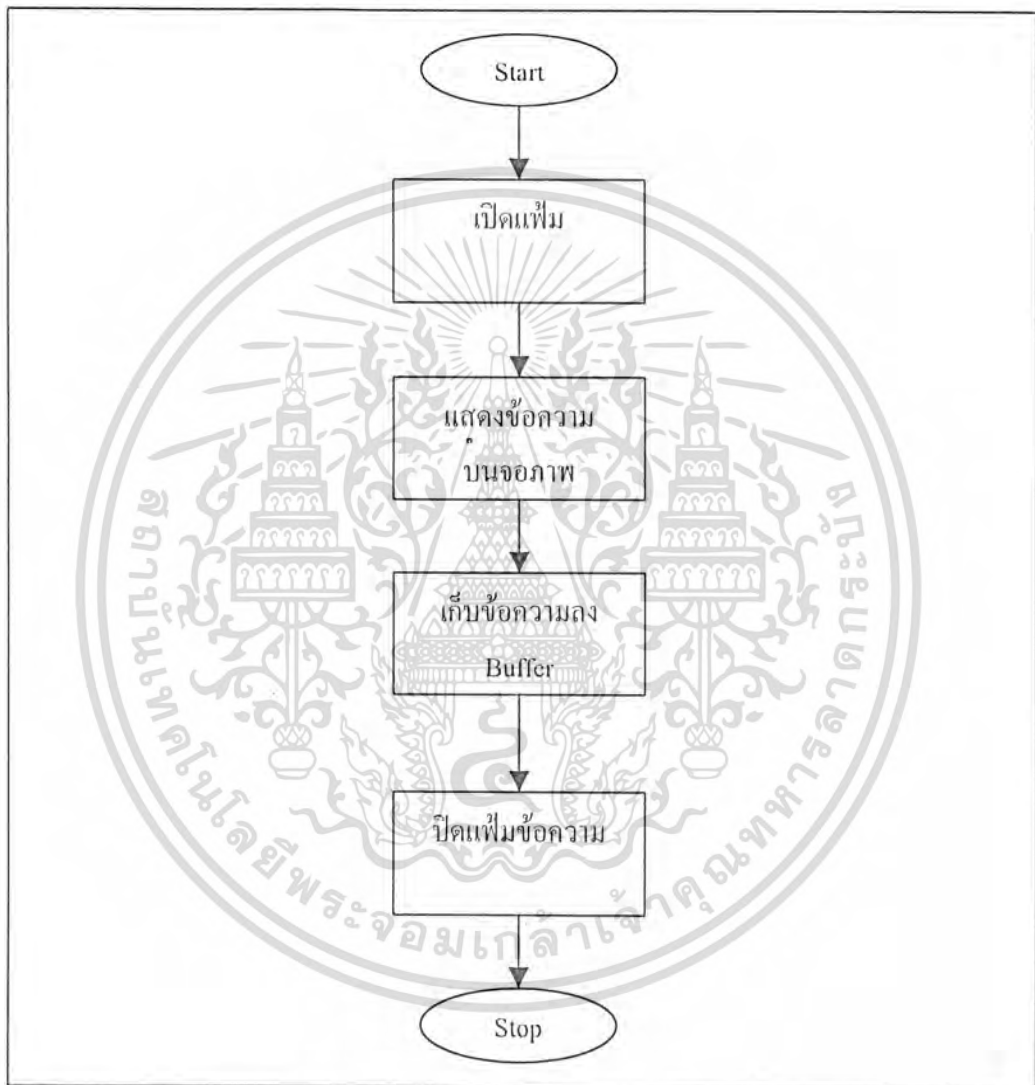
9.3 การสร้างโปรแกรม

การสร้างโปรแกรมเป็นการนำเอา Diagram การออกแบบมาทำการพัฒนาเป็น โปรแกรม ซึ่งจะแสดงการทำงานของแต่ละ โปรแกรมที่ทำการพัฒนาได้ดังต่อไปนี้

1. ส่วนรับข้อความ

ส่วนรับข้อความจะสามารถรับข้อความได้ 2 ทางคือ

1.1 ส่วนรับข้อความจากแฟ้มข้อความ

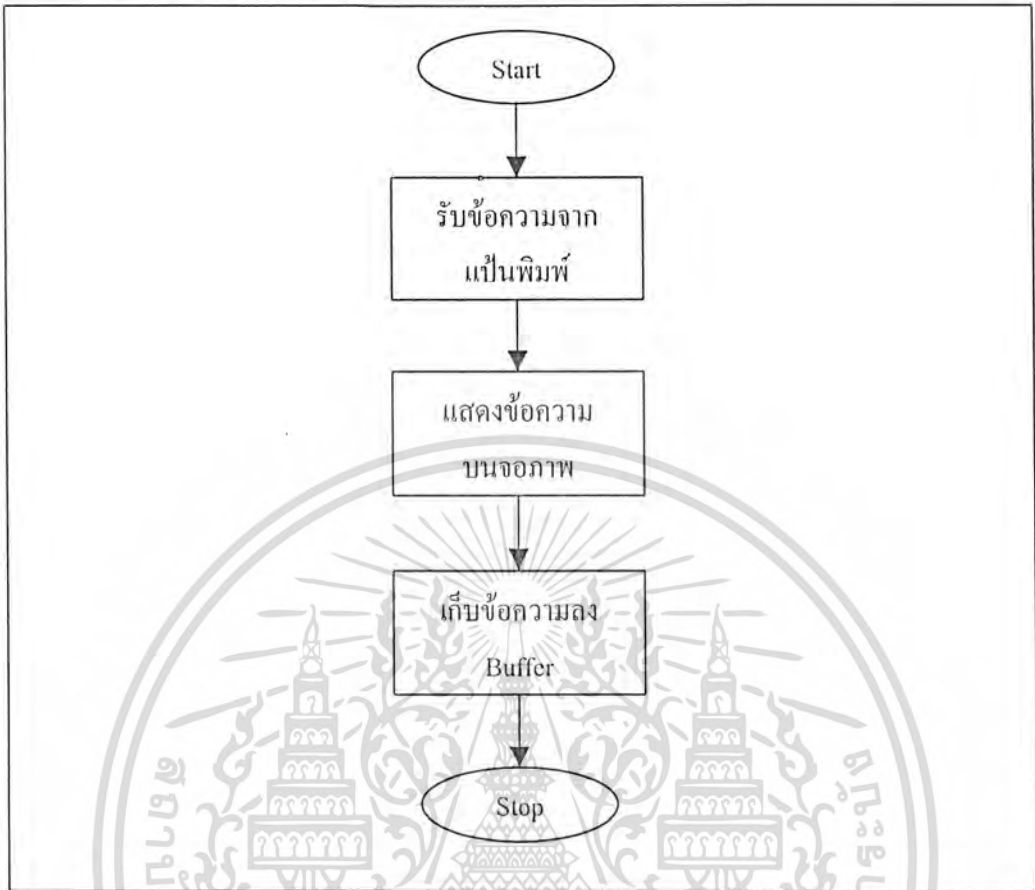


รูปที่ 9.6 แสดง Flowchart ของส่วนรับข้อความจากแฟ้มข้อความภาษาไทย

อธิบายการทำงานของส่วนรับข้อความจากแฟ้มข้อความ

1. ทำการเปิดแฟ้มข้อความ เพื่อทำการอ่านข้อความมาใช้ในกระบวนการสังเคราะห์เสียง
2. อ่านข้อความจากแฟ้มข้อความมาแสดงผลยังส่วนของ โปรแกรม
3. เก็บข้อความทั้งหมดลง Buffer เพื่อนำไปใช้ในขั้นตอนต่อไป
4. ปิดแฟ้มข้อความ

1.2 ส่วนรับข้อความจากแป้นพิมพ์

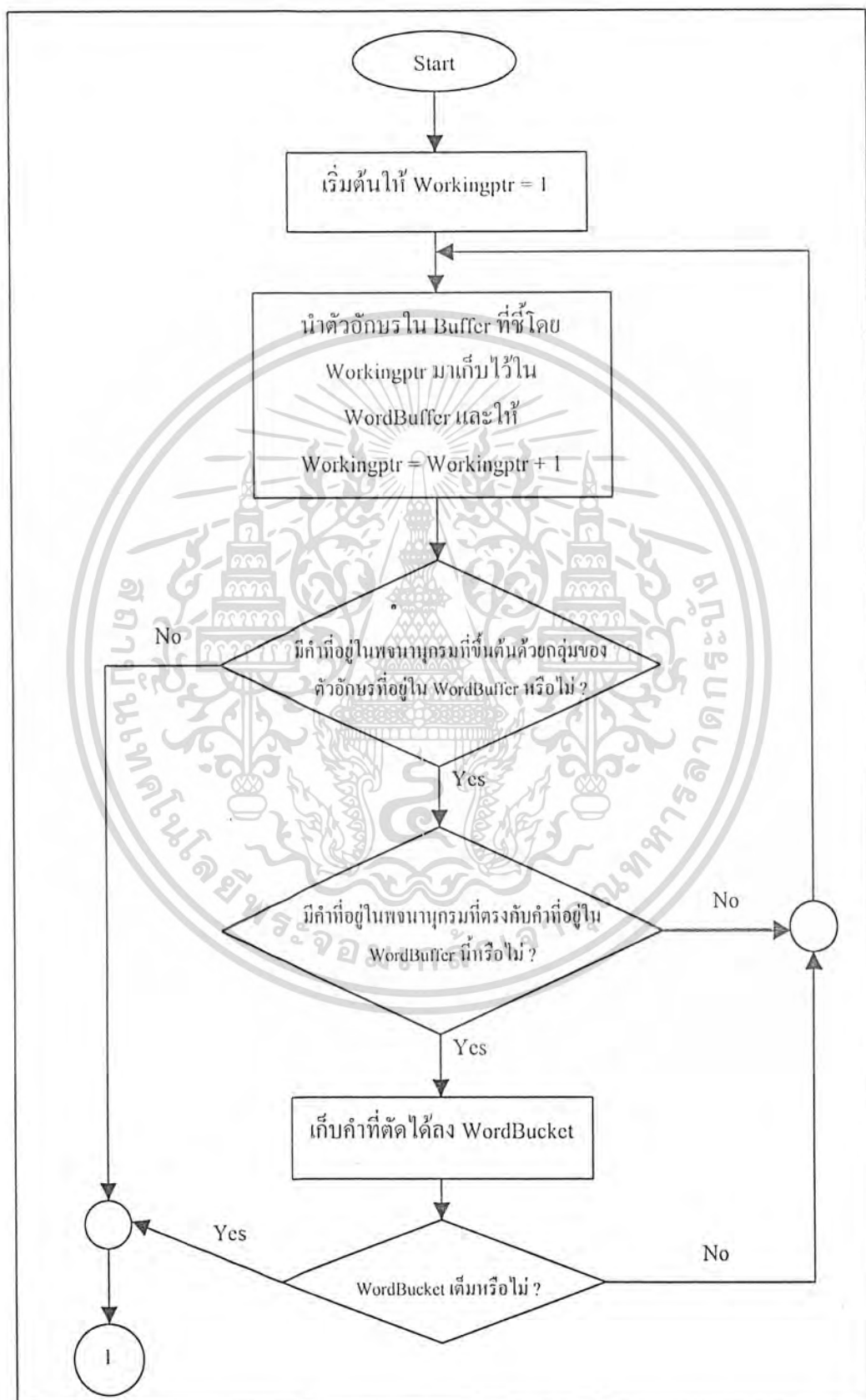


รูปที่ 9.7 แสดง Flowchart ของส่วนรับข้อความจากแป้นพิมพ์

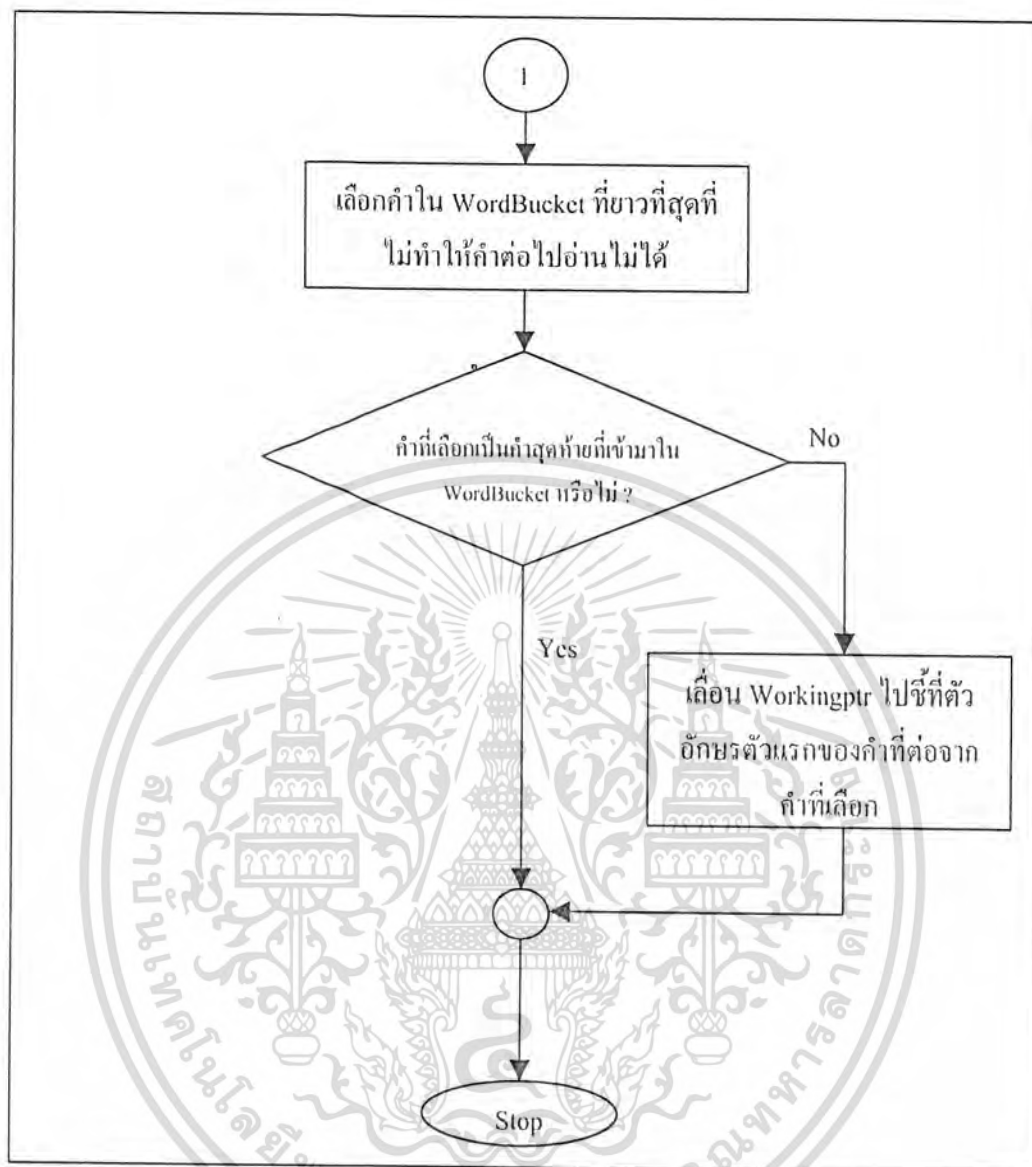
อธิบายการทำงานของส่วนรับข้อความจากแป้นพิมพ์

1. ทำการรับข้อความจากแป้นพิมพ์ เพื่อนำมาใช้ในกระบวนการสังเคราะห์เสียง
2. แสดงข้อความที่ได้จากแป้นพิมพ์ยังส่วนแสดงผลของโปรแกรม
3. เก็บข้อความที่ได้จากแป้นพิมพ์ลง Buffer เพื่อนำไปใช้ในขั้นตอนต่อไป

2. ส่วนตัดคำ



รูปที่ 9.8 แสดง Flowchart ของส่วนตัดคำ



รูปที่ 9.8 (ต่อ) แสดง Flowchart ของส่วนคัดคำ

อธิบายการทำงานของการทำงาน

1. กำหนดค่าเริ่มต้นให้กับ StartPointPtr และ WorkingPtr ให้เท่ากับ 1 เพื่อชี้ไปยังตัวอักษรตัวแรกของข้อความที่ต้องการตัดคำ
2. นำตัวอักษรที่ชี้โดย WorkingPtr มาเก็บไว้ใน WordBuffer ซึ่งเป็น Buffer ที่ใช้เปรียบเทียบกับพจนานุกรมคำศัพท์ และเพิ่มค่า WorkingPtr อีกหนึ่ง เพื่อชี้ไปยังตัวอักษรตัวต่อไป
3. เปรียบเทียบกลุ่มของตัวอักษรที่อยู่ใน WordBuffer กับพจนานุกรมคำศัพท์ ว่ามีคำที่อยู่ในพจนานุกรมคำศัพท์ที่ขึ้นต้นด้วยกลุ่มของตัวอักษรที่อยู่ใน WordBuffer หรือไม่ ถ้ามีแสดงว่ายังคงมีคำที่เป็นไปได้ในการตัดที่ยาวกว่านี้อยู่ให้ไปขั้นตอนต่อไป ถ้าไม่มีแสดงว่าไม่มีคำที่เป็นไปได้ที่ยาวกว่านี้แล้วให้ไปขั้นตอนที่ 7

4. เปรียบเทียบกลุ่มของตัวอักษรที่อยู่ใน WordBuffer กับพจนานุกรมคำศัพท์ ว่ามีคำที่อยู่ในพจนานุกรมคำศัพท์ที่ตรงกับคำที่อยู่ใน WordBuffer นี้หรือไม่ ถ้ามีแสดงว่าพบคำที่สามารถตัดได้ในพจนานุกรมคำศัพท์ให้ไปขั้นตอนต่อไป ถ้าไม่มีให้กลับไปขั้นตอนที่ 2 เพื่อหาคำที่สามารถตัดได้ต่อไป

5. นำคำที่ได้จากขั้นตอนที่ 4 เก็บลง WordBucket ซึ่งเป็น Buffer ขนาด 5 ช่อง สำหรับเก็บคำทั้งหมดที่สามารถตัดได้

6. ตรวจสอบว่า WordBucket เต็มหรือไม่ ถ้าเต็มไปขั้นตอนต่อไป ถ้าไม่เต็มกลับไปขั้นตอนที่ 2 เพื่อหาคำที่สามารถตัดได้ยาวที่สุดต่อไป

7. เลือกคำใน WordBucket ที่ยาวที่สุด โดยปกติจะเป็นคำสุดท้ายที่ตัดได้ แต่บางครั้งคำที่ยาวที่สุดที่ตัดได้ อาจทำให้คำต่อไปอ่านไม่ได้ ดังนั้นเราจึงต้องกลับมาเลือกคำที่ยาวที่สุดลำดับถัดมาที่ทำให้คำต่อไปอ่านได้ โดยการเขียนโปรแกรมแบบ Recursive

8. ตรวจสอบว่าคำที่เลือกเป็นคำที่เข้ามาใน WordBucket เป็นคำสุดท้ายหรือไม่ ถ้าใช่ไม่ต้องทำการเปลี่ยนแปลงค่า Workingptr เพราะจะชี้ไปยังตัวอักษรตัวแรกของคำต่อไปอยู่แล้ว ให้ไปขั้นตอนที่ 10 แต่ถ้าไม่ใช่ให้ไปขั้นตอนต่อไป

9. ถ้าคำที่เลือกไม่ได้เป็นคำที่เข้ามาใน WordBucket เป็นคำสุดท้าย แสดงว่าหลังจากเลือกคำนี้ Workingptr ไม่ได้ชี้ไปยังตัวอักษรตัวแรกของคำถัดไป ต้องทำการย้าย Workingptr ไปยังตำแหน่งที่ถูกต้องดังนี้

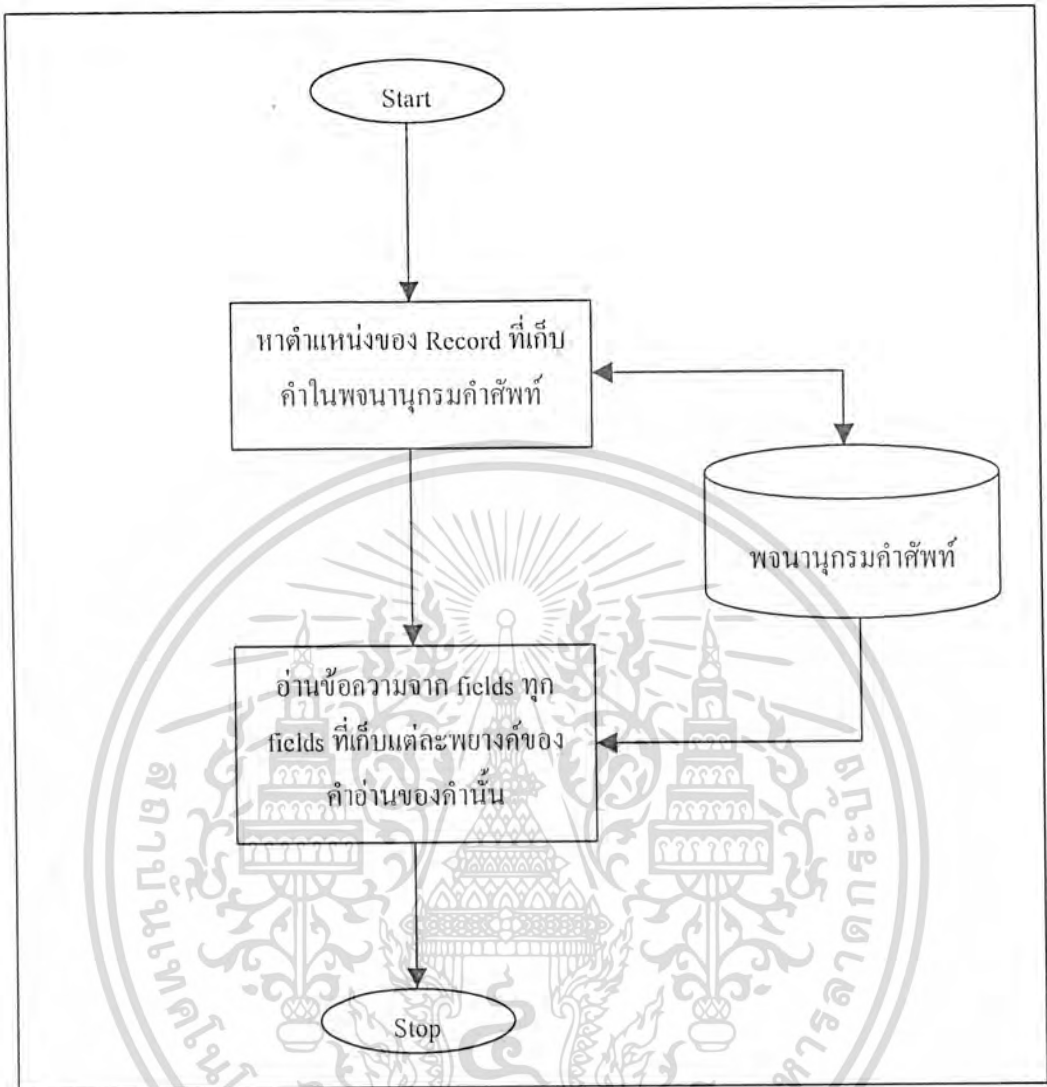
$$\text{Workingptr} = \text{StartPointptr} + \text{BucketEndPtr}$$

โดย StartPointptr เป็นตัวแปร Pointer ที่ชี้ตำแหน่งเริ่มต้นของคำปัจจุบัน

BucketEndPtr เป็นตัวแปรที่เก็บจำนวนตัวอักษรของคำ

10. ให้ StartPointptr = Workingptr เพื่อให้ StartPointptr ชี้ไปยังตัวอักษรตัวแรกของคำถัดไป ซึ่งก่อนหน้าที่เราจะไปทำการตัดคำต่อไป ต้องนำคำที่ตัดได้ไปสร้างเสียงก่อน

3. ส่วนการแยกคำออกเป็นพยางค์



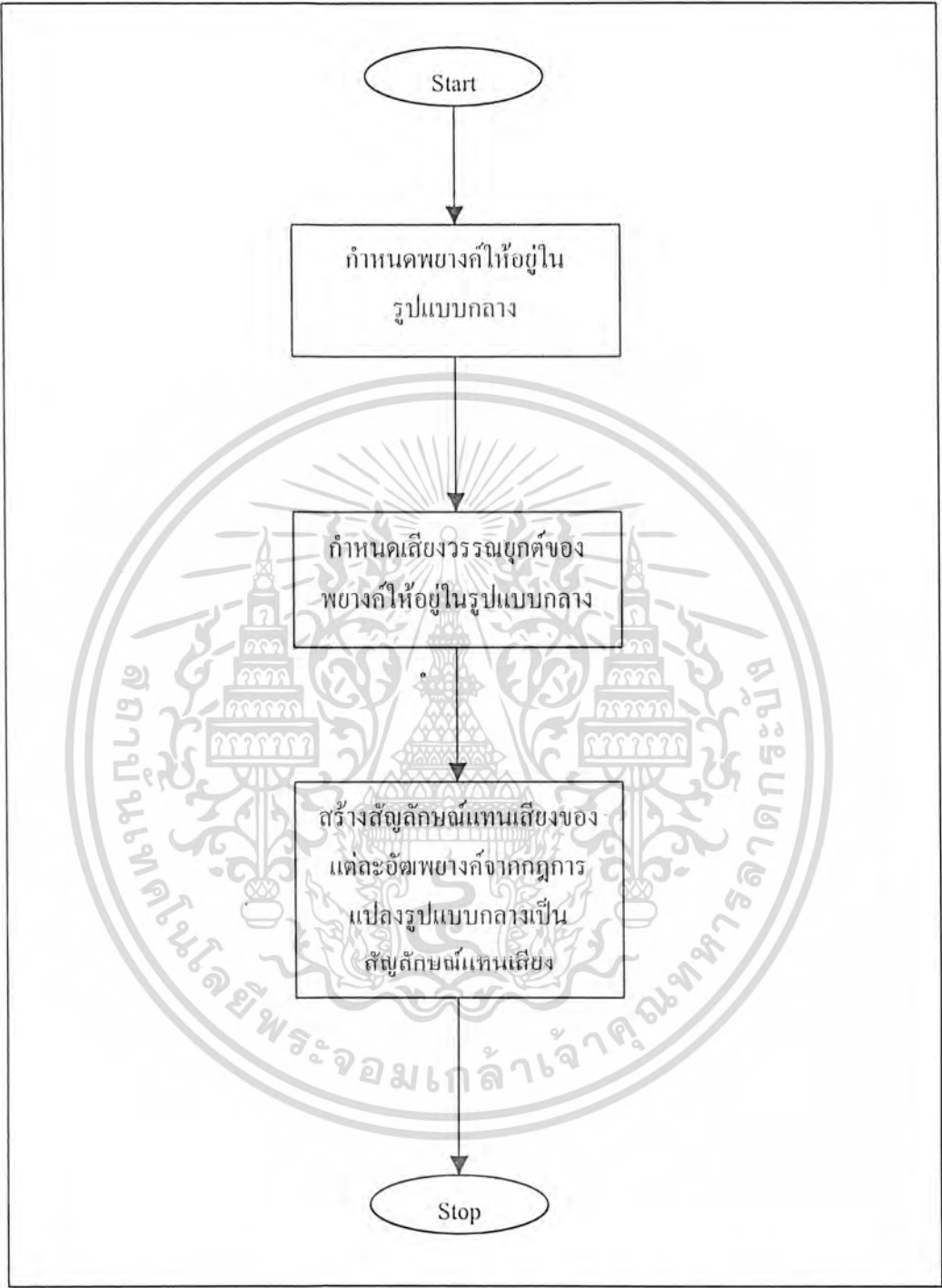
รูปที่ 9.9 แสดง Flowchart ของส่วนการแยกคำออกเป็นพยางค์

อธิบายการทำงานของส่วนแยกคำออกเป็นพยางค์

1. ขั้นตอนแรกในการแยกคำออกเป็นพยางค์ต้องทำการค้นหา Record ที่เก็บคำนั้นในพจนานุกรมคำศัพท์ก่อน ซึ่งแต่ละ Record ที่เก็บคำจะประกอบด้วย field ที่เก็บคำอ่านแต่ละพยางค์ของคำ
2. ขั้นตอนที่สองคืออ่านคำอ่านทุกพยางค์ที่เก็บใน Record ของคำนั้นมาเพื่อทำงานในขั้นตอนต่อไป

หมายเหตุ ตัวอย่างคำอ่านแต่ละพยางค์ของคำ คือ วิศวกรรมศาสตร์ อ่านว่า วิด-สะ-วะ-กำ-มะ-สาด

4. ส่วนสร้างสัญลักษณ์แทนเสียง



รูปที่ 9.10 แสดง Flowchart ของส่วนสร้างสัญลักษณ์แทนเสียง

อธิบายการทำงานของส่วนสร้างสัญลักษณ์แทนเสียง

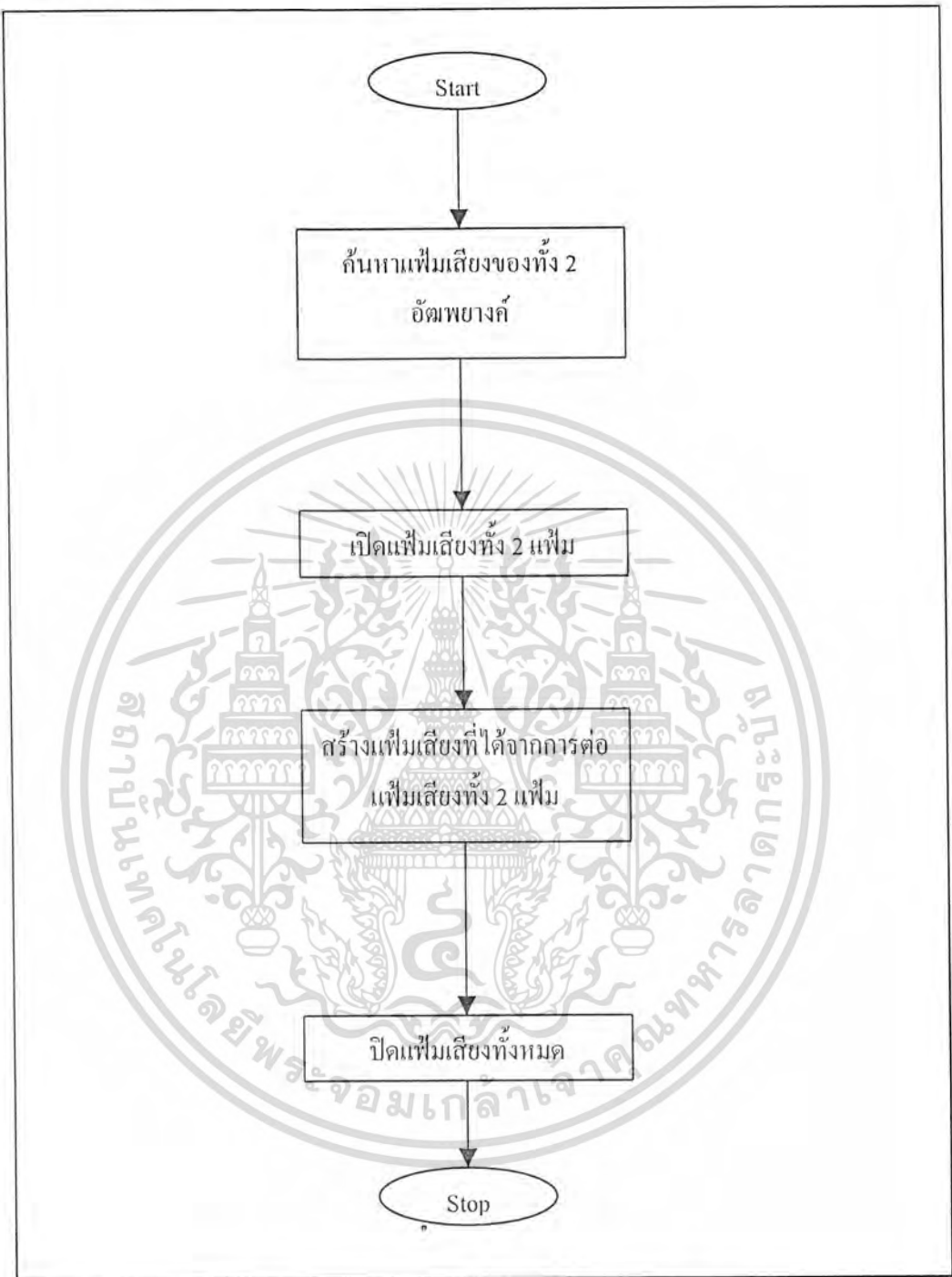
1. ขั้นตอนแรกในการสร้างสัญลักษณ์แทนเสียงของแต่ละอวัยวะที่ตั้งทำการกำหนดรูปแบบกลางของพยางค์ก่อน โดยการใช้โครงข่ายการเกิดพยางค์ ของวิธี MPT Network จะได้พยางค์ที่อยู่ในรูปแบบกลางดังนี้ $(C_1V[C_2])T$

2. เมื่อได้รูปแบบกลางของพยางค์แล้วจะทำการกำหนดเสียงวรรณยุกต์ของพยางค์ให้อยู่ในรูปแบบกลาง โดยใช้โครงข่ายความสัมพันธ์ระหว่างโครงสร้างพยางค์กับการกำหนดเสียงวรรณยุกต์ ซึ่งจะได้รูปแบบกลางอยู่ในรูปแบบดังนี้ $\{C_1V[C_2]T\}_{1..M}$

3. หลังจากได้รูปแบบกลางที่กำหนดวรรณยุกต์เรียบร้อยแล้ว จะทำการแปลงรูปแบบกลางนี้ให้อยู่ในรูปสัญลักษณ์แทนเสียงของแต่ละอวัยวะโดยใช้กฎดังต่อไปนี้



5. ส่วนประกอบเพิ่มเสียง



รูปที่ 9.11 แสดง Flowchart ของส่วนประกอบเพิ่มเสียง

อธิบายการทำงานของส่วนประกอบเพิ่มข้อมูล

1. ขั้นตอนแรกคือทำการค้นหาเพิ่มข้อมูลเสียงของทั้ง 2 อัดพยางค์จากฐานข้อมูลเสียง
2. ทำการเปิดเพิ่มข้อมูลเสียงเพื่อทำการต่อเสียง
3. สร้างเพิ่มข้อมูลเสียงที่ได้จากการนำข้อมูลเสียงทั้ง 2 อัดพยางค์มาต่อกัน
4. ปิดเพิ่มข้อมูลเสียงที่ทำการเปิดมาใช้งานทุกเพิ่ม

6. ส่วนการสร้างเสียง

ส่วนของการสร้างเสียง จะเป็นการ Run ไฟล์เสียง เพื่อให้เกิดเป็นเสียงส่งผ่านออกทางลำโพง ไปยังผู้ฟัง หลังจากทำการสร้างเสียงของพยางค์จนครบ 1 คำแล้ว ก็จะกลับไปยังส่วนของการตัดคำ เพื่อทำงานตามขั้นตอนต่าง ๆ ตั้งแต่ขั้นตอนการตัดคำจนถึงขั้นตอนสร้างเสียงใหม่ เพื่อทำการสร้างเสียงของคำต่อไป ซึ่งจะทำตามขั้นตอนเหล่านี้จนกระทั่งอ่านคำจนหมดเพิ่มข้อความ หรือจนกระทั่งอ่านคำที่รับเข้ามาจากแป้นพิมพ์จนหมด

9.4 สรุป

จากการออกแบบและการสร้างโปรแกรมสังเคราะห์เสียง ทำให้เราได้โปรแกรมที่จะนำไปใช้ในการปฏิบัติงานจริง



บทที่ 10

บทวิจารณ์และสรุป

10.1 บทวิจารณ์

จากการทดลองใช้งานทำให้เราได้พบข้อจำกัดต่างๆของโปรแกรมดังต่อไปนี้

1. แฟ้มที่จะนำมาทำการอ่านจะต้องมีแต่ตัวหนังสือเท่านั้น ห้ามมีรูปภาพอยู่ด้วย เพราะโปรแกรมไม่สามารถวิเคราะห์ได้
2. ไม่สามารถกลับไปอ่านคำก่อนหน้าได้ถ้าฟังไม่ทัน นอกจากจะสั่งให้หยุดโปรแกรมแล้วเริ่มต้นอ่านใหม่ตั้งแต่แรก
3. ไม่สามารถอ่านคำที่ไม่มีอยู่ในพจนานุกรมได้ เพราะการตัดคำของเราต้องทำการเปรียบเทียบกับคำศัพท์ที่มีอยู่ในพจนานุกรมคำศัพท์ แต่สามารถแก้ไขได้โดยการเพิ่มคำศัพท์นั้นลงไปพจนานุกรมคำศัพท์
4. ไม่สามารถอ่านตัวเลขได้แบบเป็นธรรมชาติ เช่น “ฉันมีเงิน 234 บาท” ควรจะอ่านว่า “ฉันมีเงินสองร้อยสามสิบสี่บาท” แต่โปรแกรมจะอ่านว่า “ฉันมีเงินสองสามสี่บาท”
6. ไม่สามารถอ่านคำคำควบให้ถูกต้องตามสถานการณ์ได้ เช่น “ตากลม” อ่านได้ 2 วิธีคือ “ตา+ลม” หรือ “ตาก+ลม” ในปริญญานิพนธ์นี้จะสามารถอ่านได้เพียงวิธีเดียวคือ “ตาก+ลม” เพราะว่าการตัดคำของเราจะใช้วิธีเลือกคำที่ยาวที่สุดที่มีอยู่ในพจนานุกรมคำศัพท์ คำว่า “ตา” และ “ตาก” ต่างก็เป็นคำที่มีอยู่ในพจนานุกรมคำศัพท์ แต่คำว่า “ตาก” เป็นคำที่ยาวที่สุด โปรแกรมจึงอ่านได้ว่า “ตาก+ลม” แต่เราอาจสามารถแก้ไขได้โดยใช้วิธีการวิเคราะห์ประโยคเข้ามาช่วย เพื่อที่ว่าควรจะเลือกตัดคำที่ตำแหน่งใด ซึ่งเป็นวิธีทางด้านปัญญาประดิษฐ์

10.2 สรุป

โปรแกรมสังเคราะห์เสียงภาษาไทยด้วยคอมพิวเตอร์เป็นโปรแกรมที่มีประโยชน์ต่อผู้ที่พิการทางสายตา ทำให้สามารถรับฟังเสียงภาษาไทยที่ได้จากการสังเคราะห์เสียงจากข้อความภาษาไทยหรือสำหรับผู้พิการทางการพูดให้สามารถใช้คอมพิวเตอร์พูดแทนตนเอง ซึ่งมีการทำงานดังต่อไปนี้

1. รับข้อความที่ต้องการสังเคราะห์เสียงจากแฟ้มข้อความภาษาไทย(Text file) เพื่อนำข้อความเหล่านั้นมาสังเคราะห์เสียงออกทางลำโพง นอกจากสามารถรับข้อความจากแฟ้มข้อความแล้ว ยังสามารถรับข้อความจากทางแป้นพิมพ์ได้อีกด้วย
2. เมื่อได้ข้อความมาแล้ว เราก็นำข้อความนั้นมาทำการแยกออกเป็นคำๆ โดยเลือกคำที่ยาวที่สุดเท่าที่จะเป็นไปได้ โดยการเปรียบเทียบกับพจนานุกรมคำศัพท์
เมื่อตัดคำเสร็จเรียบร้อยแล้ว ก็จะนำคำนั้นมาทำการหาคำอ่านโดยการค้นหาคำอ่านของแต่ละคำได้จากพจนานุกรมคำศัพท์
3. เมื่อได้คำอ่านแล้ว ก็จะนำคำอ่านนั้นมาหาสัญลักษณ์แทนเสียง

4. เมื่อได้สัญลักษณ์แทนเสียงแล้ว ก็นำเพิ่มเสียงของแต่ละเสียงที่แทนด้วยสัญลักษณ์แทนเสียงนั้นมารวมกันเป็นเพิ่มเสียงทั้งหมด
5. ขั้นตอนสุดท้ายคือการนำเพิ่มเสียงไปแปลงเป็นเสียงส่งออกทางลำโพงให้เราได้ยินต่อไป

10.3 แนวทางการพัฒนาต่อไป

นำวิธีการทางด้านปัญญาประดิษฐ์มาใช้ร่วมกับโปรแกรม เพื่อให้โปรแกรมสามารถอ่านคำได้ทันที



บรรณานุกรม

- [1] รุ่งจิตต์พงษ์ พุทธเจริญและอัสวิน อุดมศิริ การสังเคราะห์เสียงภาษาไทยด้วยคอมพิวเตอร์, ปรินญา นิพนธ์วิศวกรรมศาสตรบัณฑิต, กรุงเทพฯ : ภาควิชา วิศวกรรมคอมพิวเตอร์ คณะ วิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
- [2] เสริมศักดิ์ กิตติชนนพรและอิทธิพล ดิลกรัตนพิจิตร การสังเคราะห์เสียงภาษาไทยโดยใช้ HTK, ปรินญา นิพนธ์วิศวกรรมศาสตรบัณฑิต, กรุงเทพฯ : ภาควิชา วิศวกรรมคอมพิวเตอร์ คณะ วิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

