การสังเคราะห์พยางค์เสียงหนักและพยางค์เสียงเบาในภาษาไทย

นายนัฐพล พานสมบัติ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมไฟฟ้า ภาควิชาวิศวกรรมไฟฟ้า
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2545

ISBN 974-17-9749-4

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

# SYNTHESIS OF STRESSED AND UNSTRESSED SYLLABLE IN THAI LANGUAGE

Mr.Nattapol Pansombat

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Engineering in Electrical Engineering

Department of Electrical Engineering

Faculty of Engineering

Chulalongkorn University

Academic Year 2002

ISBN 974-17-9749-4

| | |
|---|---|
| หัวข้อวิทยานิพนธ์ | การสังเคราะห์พยางค์เสียงหนักและพยางค์เสียงเบาในภาษาไทย |
| โดย | นายนัฐพล พานสมบัติ |
| สาขาวิชา | วิศวกรรมไฟฟ้า |
| อาจารย์ที่ปรึกษา | รองศาสตราจารย์ ดร.เอกชัย ลีลารัศมี |
| อาจารย์ที่ปรึกษาร่วม | ผู้ช่วยศาสตราจารย์ ดร.สุดาพร ลักษณียนาวิน |

---

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

…….. คณบดีคณะวิศวกรรมศาสตร์
(ศาสตราจารย์ ดร.สมศักดิ์ ปัญญาแก้ว)

คณะกรรมการสอบวิทยานิพนธ์

…… ประธานกรรมการ
(รองศาสตราจารย์ ดร.สมชาย จิตะพันธ์กุล)

…….. อาจารย์ที่ปรึกษา
(รองศาสตราจารย์ ดร.เอกชัย ลีลารัศมี)

…… อาจารย์ที่ปรึกษาร่วม
(ผู้ช่วยศาสตราจารย์ ดร.สุดาพร ลักษณียนาวิน)

…… กรรมการ
(อาจารย์สุวิทย์ นาคพีระยุทธ)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

นัฐพล พานสมบัติ : การสังเคราะห์พยางค์เสียงหนักเบาของพยางค์ในภาษาไทย.
(SYNTHESIS OF STRESSED AND UNSTRESSED SYLLABLES IN THAI LANGUAGE)
อ. ที่ปรึกษา : รศ.ดร.เอกชัย ลีลารัศมี, อ.ที่ปรึกษาร่วม : ผศ.ดร.สุดาพร ลักษณียนาวิน, 90 หน้า.
ISBN 974-17-9749-4.

วิทยานิพนธ์นี้นำเสนอวิธีการสังเคราะห์พยางค์เสียงหนักและเบาในภาษาไทย โดยการเปลี่ยนสัทลักษณะได้แก่ ช่วงเวลา ความถี่มูลฐาน และแอมพลิจูดของสัญญาณเสียง ซึ่งจะช่วยให้เสียงสังเคราะห์ฟังเป็นธรรมชาติมากขึ้น การปรับช่วงเวลาและความถี่มูลฐานของสัญญาณเสียงจะใช้วิธีทีดี-โซลา(Time-Domain Pitch-Synchronous Overlap Add : TD-PSOLA) โดยที่การปรับช่วงเวลาจะใช้การเพิ่มหรือลดสัญญาณช่วงสั้นเพื่อให้ได้ระยะเวลาของพยางค์ตามการลงเสียงหนักเบาและโครงสร้างของหน่วยจังหวะ ในการปรับความถี่มูลฐานจะปรับระยะระห่างระหว่างยอดพิทช์ระบุตามรูปแบบความถี่มูลฐานในฐานข้อมูลความถี่มูลฐานสำหรับพยางค์เสียงเบาที่ได้สร้างไว้ทั้งหมด 14 รูปแบบตามเสียงวรรณยุกต์และโครงสร้างของพยางค์ สำหรับการปรับขนาดแอมพลิจูดของสัญญาณเสียงปรับได้โดยคูณสัญญาณเสียงด้วยอัตราส่วนแอมพลิจูดระหว่างพยางค์เสียงหนักและพยางค์เสียงเบาที่ได้สร้างไว้เป็นฐานข้อมูลตามเสียงสระทั้งหมด 24 หน่วย การประเมินคุณภาพเสียงที่สังเคราะห์ตามวิธีในวิทยานิพนธ์นี้ทำโดยอาสาสมัครจำนวน 10 คน ได้ค่าเอ็มโอเอส (Most Opinion Score : MOS) สำหรับการปรับลักษณะทางสัทศาสตร์ในระดับคำเท่ากับ 3.67 และในระดับประโยค 3.92

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาควิชา วิศวกรรมไฟฟ้า ลายมือชื่อนิสิต .....
สาขาวิชา วิศวกรรมไฟฟ้า ลายมือชื่ออาจารย์ที่ปรึกษา .....
ปีการศึกษา 2545 ลายมือชื่ออาจารย์ที่ปรึกษาร่วม .....

## 4270379721 : MAJOR ELECTRICAL ENGINEERING

KEY WORD : SPEECH SYNTHESIS / STRESSED AND UNSTRESSED SYLLABLE / PITCH-SCALE MODIFICATION / TIME-SCALE MODIFICATION TIME DOMAIN PITCH SYNCHRONOUS OVERLAPPED ADD

NATTAPOL PANSOMBAT : SYNTHESIS OF STRESSED AND UNSTRESSED SYLLABLES IN THAI. THESIS ADVISOR : ASSO. PROF. EKACHAI LEELARASMEE, Ph.D., . THESIS COADVISOR : ASSIST. PROF. SUDAPORN LUKSANEEYANAWIN, Ph.D. 90 pp. ISBN 974-17-9749-4.

This thesis presents the stressed and unstressed syllables synthesis method by modifying acoustic characteristics consisting of duration, fundamental frequency and sound amplitude in order to make the synthesized speech sound more naturally. Time Domain Pitch Synchronous Overlapped Add (TD-PSOLA) is used for modifying duration and fundamental frequency. Duration can be expanded or compressed by creating or eliminating short time signals to derive the desired syllable duration which depends on its rhythmic unit structure. Fundamental frequency(F0) of speech can be modified according to 14 patterns of unstressed syllable fundamental frequency, which are classified by syllable tone and syllable structure. This modification is by manipulating of the duration among the consecutive pitch marks. Amplitude modification is performed by multiplying the speech signal by the amplitude ratio between unstressed and stressed syllables, which are separated by syllable vowels into 24 units. The speech quality of this synthesis method was assessed by 10 volunteers. The results of assessments have MOS (Mean Opinion Score) is 3.67 for acoustic modification in word and 3.92 for acoustic modification in sentience.

Department Electrical Engineering Student's signature .....

Field of study Electrical Engineering Advisor's signature .....

Academic year 2002 Co-Advisor's signature .....

# กิตติกรรมประกาศ

ในโอกาสนี้ข้าพเจ้าใคร่ขอขอบคุณ รองศาสตราจารย์ ดร.เอกชัย ลีลารัศมี อาจารย์ที่ปรึกษาวิทยานิพนธ์ และ ผู้ช่วยศาสตราจารย์ ดร. สุดาพร ลักษณียนาวิน อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม ซึ่งท่านทั้งสองได้ช่วยเหลือให้คำแนะนำ ข้อคิดที่เป็นประโยชน์ อันเป็นส่วนสำคัญให้วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้

ขอขอบคุณทุนสนับสนุนจากโครงการทุนบัณฑิตศึกษาภายในประเทศ สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ ที่ช่วยให้การทำวิทยานิพนธ์นี้เป็นไปอย่างราบรื่น

นอกจากนี้ข้าพเจ้าขอขอบคุณ พี่ๆ เพื่อนๆ และน้องๆห้องปฏิบัติการวิจัยเชิงเลข(DSRL) ซึ่งเป็นชื่อเดิม และสมาชิกห้องปฏิบัติการวิจัยออกแบบและประยุกต์วงจรรวม(IDAR) ซึ่งเป็นชื่อปัจจุบัน ตั้งแต่ปี พ.ศ. 2542 จนถึงปี พ.ศ. 2546 ที่ได้ให้ความร่วมมือ สนับสนุน ช่วยเหลือ จนกระทั่งวิทยานิพนธ์ฉบับนี้สำเร็จได้

ท้ายที่สุดขอขอบคุณ บิดา มารดา น้องสาว ที่ได้มอบ กำลังใจ ความห่วงใย ช่วยเหลือแก่ข้าพเจ้าเสมอมา

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

# สารบัญ


บทคัดย่อภาษาไทย ..... ง
บทคัดย่อภาษาอังกฤษ ..... จ
กิตติกรรมประกาศ ..... ฉ
สารบัญ ..... ช
สารบัญตาราง ..... ญ
สารบัญภาพ ..... ฏ

บทที่ 1 บทนำ ..... 1
1.1 ความเป็นมาและความสำคัญของปัญหา ..... 1
1.2 วัตถุประสงค์ ..... 2
1.3 ขอบเขตของวิทยานิพนธ์ ..... 2
1.4 ประโยชน์ที่คาดว่าจะได้รับ ..... 2
1.5 ขั้นตอนการดำเนินงาน ..... 2
1.6 โครงสร้างวิทยานิพนธ์ ..... 3

บทที่ 2 ทฤษฎีและวรรณกรรมที่เกี่ยวข้อง ..... 4
2.1 ทฤษฎีและวรรณกรรมที่เกี่ยวข้องทางภาษาศาสตร์ ..... 4
2.1.1 ระบบเสียงหนักเบาในภาษาไทย ..... 4
2.1.2 ลักษณะร่วมทางโสตสัทศาสตร์ของพยางค์เสียงหนักและพยางค์เสียงเบา ..... 8
2.1.3 เสียงพูดภาษาไทย ..... 13
2.1.4 พยางค์เป็นพยางค์ตาย ..... 16
2.2 ทฤษฎีในการวิเคราะห์และสังเคราะห์เสียง ..... 18
2.2.1 การตรวจหาพิทช์โดยวิธีการหาค่าอัตสหสัมพันธ์ ..... 18
2.2.2 หน้าต่างแฮนนิ่ง ..... 20
2.2.3 วิธี Time-Domain Pitch Synchronous Overlap Add ..... 21
2.3 หน่วยเสียงอัฒพยางค์ ..... 22

บทที่ 3 การวิเคราะห์สัญญาณเสียงและการสร้างหน่วยเสียง ..... 23
3.1 ฐานข้อมูลความถี่มูลฐานของพยางค์เสียงเบา ..... 23

3.1.1 การสร้างคำทดสอบ 23
3.1.2 การบันทึกเสียง 24
3.1.3 การวิเคราะห์และประมวลผลข้อมูล 25
3.2 ฐานข้อมูลอัตราส่วนแอมพลิจูดระหว่างพยางค์เสียงหนักและพยางค์เสียงเบา 26
3.2.1 การสร้างคำทดสอบ 26
3.2.2 การบันทึกเสียง 26
3.2.3 การวิเคราะห์และประมวลผลข้อมูล 26
3.3 หน่วยเสียง 27
3.3.1 จำนวนหน่วยเสียง 27
3.3.2 การบันทึกเสียง 32
3.3.3 การวิเคราะห์และประมวลผลข้อมูล 32

บทที่ 4 วิธีสังเคราะห์เสียงหนักเบา 33
4.1 ระบบสังเคราะห์เสียงโดยรวม 34
4.2 การเชื่อมต่อหน่วยเสียง 35
4.3 การระบุพิทช์ 35
4.4 การระบุพิทช์เสมือน 39
4.5 การวิเคราะห์ด้วยฟังก์ชันหน้าต่าง 40
4.6 การปรับช่วงเวลา 40
4.6.1 การลดช่วงเวลา 40
4.6.2 การเพิ่มช่วงเวลา 41
4.6.3 ขั้นตอนการเพิ่มและลบสัญญาณตัวอย่างสำหรับการเปลี่ยนช่วงเวลา 43
4.7 การปรับความถี่มูลฐาน 45
4.8 การบวกซ้อน 49
4.9 การปรับแอมพลิจูดของสัญญาณเสียง 50

บทที่ 5 ผลการสังเคราะห์เสียง 52
5.1 รายละเอียดของเสียงสังเคราะห์สำหรับการทดสอบ 52
5.2 การทดสอบการปรับช่วงเวลา 52
5.2.1 ขั้นตอนการทดสอบ 52
5.2.2 ผลการทดสอบ 53

5.2.3 วิเคราะห์ผลการทดสอบการปรับช่วงเวลา 54
5.3 การทดสอบการสังเคราะห์เสียงวรรณยุกต์สำหรับพยางค์เสียงเบา 54
5.3.1 ขั้นตอนการทดสอบ 54
5.3.2 ผลการทดสอบ 65
5.3.3 วิเคราะห์ผลการทดสอบการสังเคราะห์เสียงวรรณยุกต์สำหรับพยางค์เสียงเบา 69
5.4 การทดสอบการสังเคราะห์เสียงเมื่อปรับช่วงเวลาความถี่และความดัง 69
5.4.1 วิธีการการทดสอบ 69
5.4.2 วิเคราะห์ผลการทดสอบการสังเคราะห์เสียงเมื่อปรับช่วงเวลาความถี่และความดัง 69
5.5 เอ็มโอเอส 70
5.2.1 การประเมินผล 70
5.2.2 การสร้างแบบทดสอบ 71
5.2.3 ผลการประเมิน 71
5.2.4 การวิเคราะห์ผลเอ็มโอเอส 72

บทที่ 6 ข้อสรุปและข้อเสนอแนะ 74
6.1 ข้อสรุป 74
6.2 ข้อเสนอแนะ 75

รายการอ้างอิง 76
ภาคผนวก 79
ภาคผนวก ก. การแบ่งหน่วยจังหวะ 80
ภาคผนวก ข. แบบทดสอบสำหรับการประเมินผลด้วยวิธีเอ็มโอเอส 81
ภาคผนวก ค. บทความที่ได้รับการตีพิมพ์ 84
ประวัติผู้เขียนวิทยานิพนธ์ 90

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

# สารบัญตาราง

ตารางที่ 2.1 ตัวอย่างการลงเสียงหนักในคำสามัญ 2 พยางค์ 5
ตารางที่ 2.2 ตัวอย่างการลงเสียงหนักในคำสามัญ 3 พยางค์ 5
ตารางที่ 2.3 ตัวอย่างการลงเสียงหนักในคำสามัญ 4 พยางค์ 6
ตารางที่ 2.4 สระเดี่ยวในภาษาไทย 13
ตารางที่ 2.5 สระผสมในภาษาไทย 13
ตารางที่ 2.6 เสียงพยัญชนะต้นและรูปเขียนในภาษาไทย 14
ตารางที่ 2.7 เสียงพยัญชนะควบกล้ำ 15
ตารางที่ 2.8 เสียงพยัญชนะสะกดในภาษาไทย 15
ตารางที่ 3.1 โครงสร้างพยางค์และเสียงวรรณยุกต์สำหรับการวิเคราะห์ความถี่มูลฐาน 24
ตารางที่ 3.2 อัตราส่วนแอมพลิจูดระหว่างพยางค์เสียงเบาและพยางค์เสียงหนักสำหรับแต่ละเสียงสระ 27
ตารางที่ 3.3 หน่วยเสียงท้ายของสระเสียงยาวเมื่อยังไม่รวมเสียงวรรณยุกต์ 29
ตารางที่ 3.4 หน่วยเสียงท้ายของสระเสียงสั้นเมื่อยังไม่รวมเสียงวรรณยุกต์ 30
ตารางที่ 3.5 การคำนวณจำนวนหน่วยเสียงท้าย 31
ตารางที่ 5.1 ผลการประเมินด้วยวิธี MOS 72

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

# สารบัญภาพ


รูปที่ 2.1 ความถี่มูลฐานเสียงวรรณยุกต์ในภาษาไทย 16

รูปที่ 2.2 ค่าอัตสหสัมพันธ์ ของสัญญาณเสียงคำว่า “กา” (kaa0) ณ จุดตัวอย่างที่ 2055 (m=2055) 19

รูปที่ 2.3 ลักษณะของ C{s(n)} 19

รูปที่ 2.4 ค่าอัตสหสัมพันธ์ของสัญญาณเสียงคำว่า “กา” (kaa0) ณ จุดตัวอย่างที่ 2055 (m=2055) หลังจากทำ Central Clipping 20

รูปที่ 2.5 หน้าต่างแฮนนิ่ง 20

รูปที่ 2.6 โครงสร้างหน่วยเสียงอัฒพยางค์ 22

รูปที่ 3.1 รูปแบบฐานข้อมูลความถี่มูลฐานสำหรับพยางค์เสียงเบา 25

รูปที่ 4.1 ระบบสังเคราะห์พยางค์เสียงหนักและเบาโดยรวม 34

รูปที่ 4.2 การเลือกยอดพิทช์แรก 37

รูปที่ 4.3 ฟังก์ชันน้ำหนักของค่าอัตสหสัมพันธ์ 38

รูปที่ 4.4 ความผิดพลาดในการหายอดพิทช์โดยการหาค่าอัตสหสัมพันธ์ เมื่อทำซ้ำต่อไปหลายครั้ง 38

รูปที่ 4.5 ยอดพิทช์ล่างและยอดพิทช์เสมือนในสัญญาณเสียง 40

รูปที่ 4.6 การลดช่วงเวลา 41

รูปที่ 4.7 การเพิ่มช่วงเวลา 42

รูปที่ 4.8 แผนผังของการปรับช่วงเวลา 43

รูปที่ 4.9 การเปลี่ยนระดับเสียงสูงต่ำของสัญญาณเสียงโดยใช้วิธี TD-PSOLA 45

รูปที่ 4.10 แผนผังการปรับความถี่มูลฐาน 47

รูปที่ 4.11 การประมาณค่ารูปแบบความถี่มูลฐาน 48

รูปที่ 4.12 การบวกซ้อนสัญญาณเสียง 49

รูปที่ 4.13 แผนผังการปรับแอมพลิจูดของสัญญาณเสียง 50

รูปที่ 4.14 การปรับแอมพลิจูดของสัญญาณเสียงสังเคราะห์คำว่า “นูน” (/nuun0/) 51

รูปที่ 5.1 คลื่นเสียงของเสียงสังเคราะห์ “ปู่ไปเที่ยวฟิลิปปินส์” ที่ปรับช่วงเวลา 53

รูปที่ 5.2 คลื่นเสียงของเสียงสังเคราะห์ “ปู่ไปเที่ยวฟิลิปปินส์” ที่ไม่ปรับช่วงเวลา 54

รูปที่ 5.3 ผลการสังเคราะห์เสียงวรรณยุกต์สามัญสำหรับพยางค์เสียงเบาในพยางค์เป็น 55

รูปที่ 5.4 ผลการสังเคราะห์เสียงวรรณยุกต์เอกสำหรับพยางค์เสียงเบาในพยางค์เป็น 56

รูปที่ 5.5 ผลการสังเคราะห์เสียงวรรณยุกต์โทสำหรับพยางค์เสียงเบาในพยางค์เป็น 57

รูปที่ 5.6 ผลการสังเคราะห์เสียงวรรณยุกต์ตรีสำหรับพยางค์เสียงเบาในพยางค์เป็น ..... 58
รูปที่ 5.7 ผลการสังเคราะห์เสียงวรรณยุกต์จัตวาสำหรับพยางค์เสียงเบาในพยางค์เป็น ..... 59
รูปที่ 5.8 ผลการสังเคราะห์เสียงวรรณยุกต์เอกสำหรับพยางค์เสียงเบาในพยางค์ตายสระเสียงยาว ..... 60
รูปที่ 5.9 ผลการสังเคราะห์เสียงวรรณยุกต์โทสำหรับพยางค์เสียงเบาในพยางค์ตายสระเสียงยาว ..... 61
รูปที่ 5.10 ผลการสังเคราะห์เสียงวรรณยุกต์ตรีสำหรับพยางค์เสียงเบาในพยางค์ตายสระเสียงยาว ..... 62
รูปที่ 5.11 ผลการสังเคราะห์เสียงวรรณยุกต์เอกสำหรับพยางค์เสียงเบาในพยางค์ตายสระเสียงสั้นลงท้ายด้วยเสียงกักที่เส้นเสียง ..... 63
รูปที่ 5.12 ผลการสังเคราะห์เสียงวรรณยุกต์โทสำหรับพยางค์เสียงเบาในพยางค์ตายสระเสียงสั้นลงท้ายด้วยเสียงกักที่เส้นเสียง ..... 64
รูปที่ 5.13 ผลการสังเคราะห์เสียงวรรณยุกต์ตรีสำหรับพยางค์เสียงเบาในพยางค์ตายสระเสียงสั้นลงท้ายด้วยเสียงกักที่เส้นเสียง ..... 65
รูปที่ 5.14 ผลการสังเคราะห์เสียงวรรณยุกต์เอกสำหรับพยางค์เสียงเบาในพยางค์ตายสระเสียงสั้นลงท้ายด้วยเสียงกักที่อื่นๆ ..... 66
รูปที่ 5.15 ผลการสังเคราะห์เสียงวรรณยุกต์โทสำหรับพยางค์เสียงเบาในพยางค์ตายสระเสียงสั้นลงท้ายด้วยเสียงกักที่อื่นๆ ..... 67
รูปที่ 5.16 ผลการสังเคราะห์เสียงวรรณยุกต์ตรีสำหรับพยางค์เสียงเบาในพยางค์ตายสระเสียงสั้นลงท้ายด้วยเสียงกักที่อื่นๆ ..... 68
รูปที่ 5.17 การเปรียบเทียบเสียงสังเคราะห์กับเสียงพูดจริงจากคำว่า “มหาสารคาม” ..... 70

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

# บทที่ 1

# บทนำ

## 1.1 ความเป็นมาและความสำคัญของปัญหา

ในชีวิตประจำวันของเราทุกคน ต้องมีการติดต่อสื่อสารระหว่างกัน เสียงพูดของมนุษย์เป็นตัวกลางสำคัญที่ใช้ในการติดต่อสื่อสาร เป็นรูปแบบการสื่อสารที่ถ่ายทอดได้รวดเร็วและเข้าใจได้ง่าย จึงมีการนำเอาเสียงพูดไปประยุกต์ใช้ในงานต่างๆมากมาย เช่น ระบบสั่งงานอุปกรณ์ด้วยเสียง เสียงเรียกคิวในธนาคาร ดิกชันนารีพูดได้ Screen Reader หลายๆการใช้งานแสดงให้เห็นถึงความต้องการที่จะให้มี อุปกรณ์ เครื่องมือ ที่จะติดต่อ สามารถพูดคุยกับมนุษย์ได้

ระบบสังเคราะห์เสียงพูด (Speech Synthesis System) จึงเริ่มเข้ามามีบทบาท เป็นที่สนใจ ในภาษาหลักหลายภาษา เช่น ภาษาอังกฤษ ญี่ปุ่น จีน ได้มีงานวิจัยทางด้านการสังเคราะห์เป็นจำนวนมาก ได้มีการพัฒนาระบบสังเคราะห์เสียงเพื่อให้ได้เสียงสังเคราะห์ที่มีคุณภาพสูง

สำหรับในภาษาไทย งานวิจัยด้านการสังเคราะห์เสียงมีดังต่อไปนี้ สุดาพร ลักษณียนาวิน (Luksaneeyanawin, 1989,1990,1991 a., 1991 b.) กฤษดา เรเยส (2530) วิเชียร แซ่โล้ว (2539) เอกพล อนุสุเรนทร์ (2541) และ ชัชวาลย์ หาญสกุลบันเทิง (2542) วิธีการสังเคราะห์เสียงส่วนมากจะเป็นการสังเคราะห์เสียงโดยวิธีการต่อคลื่นเสียง (Concatenation) โดยใช้หน่วยหน่วยเสียงอัฒพยางค์ (Demisyllables) มีบางส่วนที่ใช้ หน่วยเสียงอนุภาค (Microphoneme) คู่เส้นสเปกตรัม (Line Spectrum Pairs) งานวิจัยส่วนใหญ่เป็นงานวิจัยที่เน้นการสร้างระบบสังเคราะห์เสียงพูดภาษาไทย เสียงจากการสังเคราะห์ยังมีลักษณะซึ่งแตกต่างจากเสียงพูดปกติมาก การนำความรู้ทางด้านทางสัทศาสตร์มาประยุกต์ใช้ในการสังเคราะห์เสียง จะช่วยให้เสียงสังเคราะห์ที่ได้มีความเป็นธรรมชาติใกล้เคียงเสียงพูดของมนุษย์มากขึ้น

ในภาษาไทยนั้น มีการลงเสียงหนักเบาในคำพูดที่พูดออกมา ดังจะเห็นได้จากงานวิจัยที่ศึกษาลักษณะร่วมทางสัทศาสตร์ของพยางค์เสียงหนักและพยางค์เสียงเบาจำนวนมาก จากงานวิจัยของ สุดาพร ลักษณียนาวิน (Luksaneeyanawin, 1983) สุกัลยา สุรินทร์ไพบูลย์ (2528) และ รุจนา พินิจารมย์ (2532) สามารถสรุปได้ว่า การลงเสียงหนักเกี่ยวพันกับลักษณะร่วมทางสัทศาสตร์ทั้ง 4 คือ ความดัง ความยาว ระดับเสียงสูงต่ำ คุณสมบัติของสระและวรรณยุกต์

การลงเสียงหนักเบานั้น นอกจากจะทำให้คำพูดที่ออกมาดูเป็นธรรมชาติแล้ว ยังเป็นส่วนสำคัญในการเข้าใจความหมายของคำพูดที่พูดออกมาอีกด้วย เช่น คำว่า “เสือหมอบ” ในประโยคว่า “เสือหมอบอยู่ใต้ต้นไม้” ความหมายแรกหมายถึง เสือ ทำกริยา หมอบ อยู่ใต้ต้นไม้ ในความหมายนี้จะลงเสียงหนักทั้งคำว่า ”เสือ” และคำว่า ”หมอบ” ส่วนอีกความหมาย “เสือหมอบ” ซึ่งเป็น

คำผสม หมายถึง จักรยานเสือหมอบ ความหมายนี้จะลงเสียงหนักที่คำว่า "หมอบ" ส่วนคำว่า "เสือ" จะลงเสียงเบา

ดังนั้นวิทยานิพนธ์นี้จึงนำเสนอการสังเคราะห์พยางค์เสียงหนักและพยางค์เสียงเบา จากคำพยางค์เดี่ยวที่ได้จากการสังเคราะห์โดยการปรับลักษณะทางสัทศาสตร์ที่เกี่ยวข้องให้เสียงจากการสังเคราะห์มีเสียงหนักเบา เพื่อให้คำจากการสังเคราะห์รับฟังเป็นธรรมชาติมากขึ้น

## 1.2 วัตถุประสงค์

1. เพื่อสังเคราะห์เสียงหนักเบาในพยางค์ภาษาไทยโดยการปรับ ความดัง ความยาว ระดับเสียงสูงต่ำ
2. เพื่อพัฒนาโปรแกรมสังเคราะห์เสียงพูดภาษาไทยพยางค์เดี่ยวที่สามารถสังเคราะห์เสียงหนักเบาให้กับพยางค์

## 1.3 ขอบเขตของวิทยานิพนธ์

1. สร้างซอฟต์แวร์สังเคราะห์เสียงพูดภาษาไทยโดยใช้การต่อคลื่นเสียงและใช้หน่วยเสียงแบบอัฒพยางค์ (Demisyllables)
2. สร้างหน่วยเสียงโดยบันทึกหน่วยเสียงจากผู้พูดโดยบันทึกเสียงพูดจากคำพูดต่อเนื่อง ซึ่งเป็นคำที่ลงเสียงหนัก เป็นจำนวน 1,536 หน่วยเสียง
3. สามารถสังเคราะห์เสียงพูดเป็นคำที่ลงเสียงหนักเบาได้ โดยการปรับลักษณะร่วมทางสัทศาสตร์ได้แก่ ค่าระยะเวลา ความถี่มูลฐาน แอมพลิจูด โดยใช้วิธี Time-Domain Pitch-Synchronous Overlap-Add

## 1.4 ประโยชน์ที่คาดว่าจะได้รับ

1. เป็นแนวทางพัฒนาเสียงสังเคราะห์ให้มีความเป็นธรรมชาติมากยิ่งขึ้น
2. สามารถนำไปประยุกต์เพื่อสร้างเป็นเครื่องสังเคราะห์เสียงภาษาไทย

## 1.5 ขั้นตอนการดำเนินงาน

1. ศึกษาลักษณะร่วมทางสัทศาสตร์อันได้แก่ ความยาว ระดับเสียงสูงต่ำ ความดัง ของพยางค์ที่ลงเสียงหนักและพยางค์ที่ลงเสียงเบา
2. ศึกษาวิธีการปรับลักษณะร่วมทางสัทศาสตร์ของพยางค์ที่ลงเสียงหนักและพยางค์ที่ลงเสียงเบา ได้แก่ ระยะเวลา ความถี่มูลฐาน แอมพลิจูด
3. บันทึกเสียงเพื่อสร้างหน่วยเสียง

4. วิเคราะห์เสียงวรรณยุกต์ในพยางค์เสียงเบาเพื่อนำมาสร้างเป็นฐานข้อมูลความถี่มูลฐานสำหรับการปรับความถี่มูลฐาน
5. วิเคราะห์แอมพลิจูดของพยางค์เสียงเบาและเพื่อหาอัตราส่วนแอมพลิจูดระหว่างพยางค์เสียงเบาและพยางค์เสียงหนักเพื่อสร้างเป็นฐานข้อมูลสำหรับการปรับแอมพลิจูด
6. สร้างซอฟต์แวร์สังเคราะห์เสียงพูด
7. ทดสอบและปรับปรุงแก้ไข

### 1.6 **โครงสร้างวิทยานิพนธ์**

ในวิทยานิพนธ์ฉบับนี้จะแบ่งเนื้อหาในการนำเสนอออกเป็น 4 ส่วน คือ ในส่วนของบทที่ 2 จะกล่าวถึง ระบบเสียงหนักเบาในภาษาไทย ลักษณะทางสัทศาสตร์ของพยางค์เสียงเบาและพยางค์เสียงหนัก ลักษณะของเสียงพูดในภาษาไทย และได้กล่าวถึงทฤษฎี เทคนิค สำหรับการวิเคราะห์และสังเคราะห์เสียง

ในบทที่ 3 จะกล่าวถึงการวิเคราะห์เสียงพยางค์เสียงเบาเพื่อสร้างฐานข้อมูลความถี่มูลฐานสำหรับพยางค์เบาและอัตราส่วนแอมพลิจูดระหว่างพยางค์เสียงเบาและพยางค์เสียงหนักสำหรับการปรับความถี่มูลฐานและแอมพลิจูดตามลำดับ ในส่วนท้ายของบทนี้จะกล่าวถึงการสร้างหน่วยเสียงสำหรับวิทยานิพนธ์นี้

ในบทที่ 4 จะกล่าวถึงระบบการสังเคราะห์เสียง วิธีการปรับลักษณะทางสัทศาสตร์อันได้แก่ ระยะเวลา ความถี่มูลฐาน และ แอมพลิจูด

ในบทที่ 5 และบทที่ 6 จะได้กล่าวถึงวิธีการและขั้นตอนในการทดสอบวิธีการสังเคราะห์เสียง นำเสนอผลและวิเคราะห์ผลการทดสอบ และสรุปผลที่ได้จากวิทยานิพนธ์นี้รวมทั้งให้ข้อเสนอแนะในการพัฒนางานวิจัยส่วนการสังเคราะห์เสียงนี้ต่อไป

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

# บทที่ 2

# ทฤษฎีและวรรณกรรมที่เกี่ยวข้อง

ในบทนี้จะได้แบ่งเนื้อหาออกเป็นสามส่วน ส่วนแรกจะได้กล่าวถึงทฤษฎีทางภาษาศาสตร์ที่เกี่ยวข้องกับการทำวิทยานิพนธ์นี้ และในส่วนที่สองจะได้กล่าวถึงทฤษฎีในการวิเคราะห์และสังเคราะห์เสียง และในส่วนสุดท้ายจะได้กล่าวถึงหน่วยเสียงอัฒพยางค์ซึ่งเป็นหน่วยเสียงที่ใช้ในวิทยานิพนธ์นี้

## 2.1 ทฤษฎีและวรรณกรรมที่เกี่ยวข้องทางภาษาศาสตร์

### 2.1.1 ระบบเสียงหนักเบาในภาษาไทย

สุดาพร ลักษณียนาวิน (Luksaneeyanawin, 1983, 1998) ได้อธิบายกฎการกำหนดตำแหน่งของพยางค์เสียงหนักในคำพยางค์เดียว และคำหลายพยางค์สรุปได้ดังนี้

**1. คำสามัญพยางค์เดียว**

ในคำพูดแต่ละคำมีการลงเสียงหนักหรือลงเสียงเบา เกณฑ์กำหนดตำแหน่งของพยางค์เสียงหนักในคำพยางค์เดียว ดังกฎการลงเสียงหนัก ดังนี้

คำหลัก (Content Word) ได้แก่ คำนาม คำกริยา คำวิเศษณ์ และคำคุณศัพท์ จะเน้นเสียงหนักในการพูดปกติ

คำไวยากรณ์ (Grammatical Word) ได้แก่คำประเภทอื่นที่ไม่ใช่คำเนื้อหา จะเน้นเสียงเบาในการพูดปกติ ยกเว้น ในการพูดเพียงลำพังหรือมีการเน้น

ตัวอย่างการลงเสียงหนักของคำพยางค์เดียว ยกตัวอย่างคำว่า"กัน"

| | | | | |
|---|---|---|---|---|
| - | อย่า | กัน | ดี | กว่า |
| | ja:1 | kan0 | di:0 | khwa:1 |
| | (ปฏิเสธ) | (กริยา) | (วิเศษณ์) | |
| | U | S | U | S |
| - | หย่า | กัน | ดี | กว่า |
| | ja:1 | kan0 | di:0 | khwa:1 |
| | (กริยา) | (สรรพนามแสดงการปฏิภาค) | (วิเศษณ์) | |
| | S | U | U | S |

โดย S แทน การลงเสียงหนัก U แทน การลงเสียงเบา

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

2. **คำสามัญหลายพยางค์**

การกำหนดตำแหน่งพยางค์เสียงหนักในคำสามัญหลายพยางค์ ขึ้นอยู่กับ จำนวนพยางค์และโครงสร้างของพยางค์ในคำว่าเป็น พยางค์เชื่อม (L) หรือ พยางค์ไม่เชื่อม (O)[1] โดยที่พยางค์สุดท้ายจะลงเสียงหนักเสมอ คำสามัญหลายพยางค์มีรูปแบบการลงเสียงหนักเบาดังรูปที่ 2.1-2.3 แยกตามจำนวนพยางค์ โดย ◡ แทนการลงเสียงหนัก และ ˈ แทนการลงเสียงเบา

ในคำสามัญสองพยางค์ พยางค์ทั้งสองเป็นได้ทั้งพยางค์เชื่อมและพยางค์ไม่เชื่อมจึงมีโครงสร้างพยางค์ได้ 4 แบบดังตารางที่ 2.1

ตารางที่ 2.1 ตัวอย่างการลงเสียงหนักในคำสามัญ 2 พยางค์

| Accentual Pattern | Example | | Phonetic Realization | |
|---|---|---|---|---|
| | | | Casual | Careful |
| L ‘L | ระยะ | / ra?$^{h}$ ja?$^{h}$/ | ◡ ˈ | ◡ ˈ |
| L ‘O | ฉบับ | /cha?$^{l}$ bap$^{l}$/ | | |
| ‘O 'O | สาหัส | /sa:$^{r}$ hat$^{l}$/ | ◡ ˈ | ˈ ˈ |
| ‘O 'L | วาระ | /wa:$^{m}$ ra?$^{h}$/ | | |

ในคำสามัญสามพยางค์ มีโครงสร้างพยาค์ได้ทั้งหมด 8 แบบ มีรูปแบบการลงเสียงหนักเบา ดังตารางที่ 2.2

ตารางที่ 2.2 ตัวอย่างการลงเสียงหนักในคำสามัญ 3 พยางค์

| Accentual Pattern | Example | | Phonetic Realization | |
|---|---|---|---|---|
| | | | Casual | Careful |
| ‘L L ‘L | สมถะ | /sa?$^{l}$ ma?$^{h}$ tha?$^{l}$/ | ◡ ◡ ˈ | ˈ ◡ ˈ |
| ‘L L ‘O | กะละมัง | /ka?$^{l}$ la?$^{h}$ maŋ$^{m}$/ | | |
| 'O O 'O | กิริยา | /ki?$^{l}$ ri?$^{h}$ ja:$^{m}$/ | ◡ ◡ ˈ | ˈ ◡ ˈ |
| 'O O 'L | สัมภาระ | /sam$^{r}$ pa:$^{m}$ ra?$^{h}$/ | | |
| L 'O 'O | กะทันหัน | /ka?$^{l}$ than$^{m}$ han$^{r}$/ | ◡ ◡ ˈ | ◡ ˈ ˈ |
| L 'O 'L | สภาวะ | /sa?$^{l}$ pha:$^{m}$ wa?$^{h}$/ | | |
| 'O L 'L | ภารตะ | /pha:$^{m}$ ra?$^{h}$ ta?$^{l}$/ | ◡ ◡ ˈ | ˈ ◡ ˈ |
| 'O L 'O | กีรติ | /ki:$^{m}$ ra?$^{h}$ ti?$^{l}$/ | | |

[1] พยางค์เชื่อม คือ พยางค์ที่มีเสียงสระ อะ (/a/) และสะกดด้วยพยัญชนะกักเส้นเสียง (/?/) เช่น “พะ” ในคำว่า “พนัน”, “อะ” ในคำว่า “อะไร” ส่วนพยางค์ไม่เชื่อมคือ พยางค์อื่นๆที่ไม่มีคุณสมบัติดังกล่าว

คำสามัญสี่พยางค์มีโครงสร้างพยางค์ทั้งหมด 16 แบบ โครงสร้างพยางค์ 8 แบบหลัง มีรูปแบบเสียงหนักเบา 2 แบบ ในการพูดทั้งการพูดแบบปกติและการพูดแบบระวังตัวของคำสามัญสี่พยางค์อาจมีการลงเสียงหนักเบามากกว่า 1 แบบ ดังตารางที่ 2.3

ตารางที่ 2.3 ตัวอย่างการลงเสียงหนักในคำสามัญ 4 พยางค์

| Accentual Pattern | Example | | Phonetic Realization | |
|---|---|---|---|---|
| | | | Casual | Careful |
| L 'O O 'O | ปทานุกรม | /pa?$^{l}$ tha:$^{m}$ nu?$^{h}$ krom$^{m}$/ | ˩˧˩˧ or ˩˩˩˧ | ˩˧˩˧ |
| L 'O O 'L | สวัสดิกะ | /sa?$^{l}$ wat$^{l}$ di?$^{l}$ ka?$^{l}$/ | | |
| L 'O L 'O | สถาปนา | /sa?$^{l}$ tha:$^{r}$ pa?$^{l}$ na:$^{m}$/ | | |
| L 'O L 'L | ปกิณณกะ | /pa?$^{l}$ kin$^{m}$ na?$^{h}$ ka?$^{l}$/ | | |
| 'O L O 'O | กัลปังหา | /kan$^{m}$ la?$^{h}$ paŋ$^{m}$ ha:$^{r}$/ | ˧˩˩˧ or ˩˩˩˧ | ˧˩˩˧ |
| 'O L O 'L | กัลยาณะ | / kan$^{m}$ la?$^{h}$ ja:$^{m}$ na?$^{h}$/ | | |
| 'O L L 'O | ฌาปนกิจ | /cha:$^{m}$ pa?$^{l}$ na?$^{h}$ kit$^{l}$/ | | |
| 'O L L 'L | กัณฐกะ | /kan$^{m}$ na?$^{h}$ tha?$^{l}$ ka?$^{l}$/ | | |
| #L ‘L L ‘L or ‘L L L ‘L | ปรมตะ | /pa?$^{l}$ ra?$^{h}$ ma?$^{h}$ ta?$^{l}$/ | #˩˧˩˧ or ˧˩˩˧ or ˩˩˩˧ | #˩˧˩˧ or ˧˩˩˧ |
| | กรหณะ | /kɣa?$^{l}$ ra?$^{h}$ ha?$^{l}$ na?$^{h}$/ | | |
| #L ‘L L ‘O or ‘L L L ‘O | กรกวรรษ | /ka?$^{l}$ ra?$^{h}$ ka?$^{l}$ wat$^{h}$/ | | |
| | กรกฎา | /ka?$^{l}$ ra?$^{h}$ ka?$^{l}$ da:$^{m}$/ | | |
| #L ‘L O ‘O or ‘L L O ‘O | คมนาคม | /kha?$^{h}$ ma?$^{h}$ na:$^{m}$ khom$^{m}$/ | | |
| | ปรมาณู | /pa?$^{l}$ ra?$^{h}$ ma:$^{m}$ nu:$^{m}$/ | | |
| #L ‘L O ‘L or ‘L L O ‘L | ปรเมหะ | /pa?$^{l}$ ra?$^{h}$ me:$^{m}$ ha?$^{l}$/ | | |
| | กรณียะ | /ka?$^{l}$ ra?$^{h}$ ni:$^{m}$ ja?$^{h}$/ | | |
| #O ‘O O ‘O or ‘O O O ‘O | กรรมาธิการ | /kam$^{m}$ ma:$^{m}$ thi?$^{h}$ ka:n$^{m}$/ | | |
| | เปศัสการี | /pe:$^{m}$ sat$^{h}$ ka:$^{m}$ ri:$^{m}$/ | | |
| #O ‘O O ‘L or ‘O O O ‘L | กรรติเกยะ | /kan$^{m}$ ti?$^{l}$ ke:$^{m}$ ja?$^{h}$/ | | |
| | ปาวามัลละ | /pa:$^{m}$ wa:$^{m}$ man$^{m}$ la?$^{h}$/ | | |
| #O ‘O L ‘L or ‘O O L ‘L | ดีไวยกะ | /di:$^{m}$ waj$^{m}$ ja?$^{h}$ ka?$^{l}$/ | | |
| | มาตามหะ | /ma:$^{m}$ ta:$^{m}$ ma?$^{h}$ ha?$^{l}$/ | | |
| #O ‘O L ‘O or ‘O O L ‘O | เนรัญชรา | /ne:$^{m}$ ran$^{m}$ cha?$^{h}$ ra:$^{m}$/ | | |
| | อาราธนา | /?a:$^{m}$ ra:t$^{l}$ tha?$^{h}$ na:$^{m}$/ | | |

3. **คำซ้ำ** เกิดจากมีคำซึ่งเป็นคำฐานเช่นคำว่า “เหลือง” และ มีคำซ้ำคำฐานเช่นคำว่า “เลื้อง” จะเกิดคำซ้ำขึ้นเป็น เลื้องเหลือง

การกำหนดตำแหน่งพยางค์เสียงหนักของคำซ้ำ ขึ้นอยู่กับชนิดของคำซ้ำ ดังนี้

1) คำซ้ำธรรมดา (เกิดจาก คำฐาน O กับคำซ้ำฐาน R ซึ่งมีลักษณะเหมือนคำฐานทุกประการ มาซ้ำข้างหน้าคำซ้ำฐาน) พยางค์สุดท้ายจะลงเสียงหนักเป็นเอก

- คำซ้ำธรรมดา 2 พยางค์ ไม่มีคำลงเสียงหนักเป็นโท / R ‘O / เช่น เด็กๆ
- คำซ้ำธรรมดา 4 พยางค์ จะลงเสียงหนักเป็นโทที่พยางค์ก่อนพยางค์รองสุดท้าย / R ‘R O ‘O / เช่น น้ำตาลๆ

2) คำซ้ำแบบเน้น เป็นคำซ้ำที่มักจะเกิดขึ้นกับ คำกริยา คำวิเศษณ์ จะลงเสียงหนักเน้น (i) ที่พยางค์สุดท้ายของคำซ้ำฐาน และจะลงเสียงหนักเป็นโทที่พยางค์สุดท้าย

- คำซ้ำแบบเน้น 2 พยางค์ / $^{i}$R ‘O / เช่น เลื้องเหลือง
- คำซ้ำแบบเน้น 4 พยางค์ / R $^{i}$R O ‘O / เช่น ทุกวั้น ทุกวัน

3) คำซ้ำบางส่วน คำซ้ำบางส่วนจะมีบางส่วนคล้ายกับคำฐาน อาจมีพยัญชนะต้น วรรณยุกต์ หรือพยัญชนะทั้งหมดกับวรรณยุกต์เหมือนคำฐาน แต่มีสระต่างออกไป

- คำซ้ำบางส่วน 2 พยางค์ / ‘O ‘R / เช่น สวยเสย
  / ‘R ‘O / เช่น มงมือ
- คำซ้ำบางส่วน 4 พยางค์ / O ‘O R ‘R / เช่น ทุเรียนทุเริน
  / R ‘R O ‘O / เช่น ทุรงทุเรียน

4) คำซ้ำพิเศษ คำซ้ำชนิดนี้พิเศษเพราะไม่ได้เกิดจากคำฐานเหมือนกับคำซ้ำชนิดอื่นๆ เป็นคำซ้ำที่เกิดจากพยางค์ที่คู่กันและมีพยัญชนะ วรรณยุกต์ เหมือนๆกัน ความยาวๆเท่ากัน (ด้วยเหตุนี้จึงใช้สัญลักษณ์ R เหมือนกัน) คำซ้ำชนิดนี้จะไม่เกิดตามลำพัง จะเกิดคู่กันเสมอ

- คำซ้ำพิเศษ 2 พยางค์ / ‘R ‘R / เช่น โย้เย้
- คำซ้ำพิเศษ 4 พยางค์ / R ‘R R ‘R / เช่น กระโดกกระเดก

4. **คำผสม**

ตำแหน่งการลงเสียงหนักของคำผสม ขึ้นอยู่กับโครงสร้างของคำที่มาผสมกัน วิธีการผสมจะช่วยบอกตำแหน่งการลงเสียงหนักในในคำผสม คำผสมจะลงเสียงหนักเป็นเอกที่พยางค์สุดท้ายเสมอ จะลงเสียงหนักเป็นโทที่ตำแหน่งลงเสียงหนักเป็นเอกเดิมของคำที่เป็นฐานซึ่งปรากฏอยู่หน้าขอบเขตระหว่างคำที่นำมาผสม ยกเว้นเมื่อพยางค์ที่ปรากฏอยู่หน้าขอบเขตระหว่างคำที่นำมาผสมเป็นพยางค์เชื่อม พยางค์ที่ลงเสียงหนักเป็นโทจะเป็นพยางค์ที่ลงเสียงหนักที่อยู่ก่อนพยางค์เชื่อม ส่วนพยางค์เชื่อมที่เดิมลงเสียงหนักเป็นเอกจะปรากฏเป็นพยางค์ลงเสียงหนักเป็นตรีแทน

ตัวอย่างคำผสม

น้ำส้มคั้น /na:m$^{h}$. ‘som$^{f}$. ‘khan$^{h}$/

คำมูล + คำมูล /‘khan3/

/’na:m$^{h}$/ + /‘som$^{f}$/ /‘khan$^{h}$/

คำผสม + คำมูล

/na:m$^{h}$. ‘som$^{f}$/ + /‘khan$^{h}$/

/na:m$^{h}$. ‘som$^{f}$. ‘khan$^{h}$/

โดยที่ เครื่องหมาย ‘ หน้าพยางค์ใดๆ หมายถึงพยางค์นั้นลงเสียงหนัก

### 2.1.2 ลักษณะร่วมทางโสตสัทศาสตร์ของพยางค์เสียงหนักและพยางค์เสียงเบา

ในเรื่องของลักษณะร่วมทางสัทศาสตร์ซึ่งเป็นลักษณะที่แตกต่างกันของพยางค์ที่ลงเสียงหนักและพยางค์ที่ลงเสียงเบามีนักภาษาศาสตร์หลายท่านได้ศึกษาและกล่าวถึงไว้ดังนี้

ผลการศึกษาของ Fry (1958) ได้แสดงถึงลักษณะซึ่งบ่งชี้ว่าพยางค์ใดลงเสียงหนักหรือลงเสียงเบา คือ ความดังค่อย ระดับเสียงสูงต่ำ ความยาว และคุณสมบัติของสระ

Brosnahan and Malmberg (1970) ได้กล่าวว่า พยางค์ที่ลงเสียงหนัก จะมีความยาวเพิ่มขึ้น ระดับเสียงสูงต่ำสูงขึ้น ความดังค่อยของเสียงดังมากขึ้น

O’Connor (1973) ได้กล่าวถึงลักษณะของพยางค์ที่ลงเสียงหนักและพยางค์ที่ลงเสียงเบา สรุปได้ว่า พยางค์ที่ลงเสียงหนักมีแนวโน้มที่มีความดังมากกว่าพยางค์ที่ลงเสียงเบา แต่ความดังเพียงอย่างเดียวไม่เพียงพอต่อการบ่งชี้ว่าพยางค์ใดลงเสียงหนัก พยางค์ใดลงเสียงเบา ยังต้องประกอบกับ เรื่อง ความยาว ระดับเสียงสูงต่ำ และ คุณสมบัติของสระอีกด้วย

สำหรับในภาษาไทยนั้นมีผู้ศึกษาลักษณะร่วมทางสัทศาสตร์ไว้ ดังนี้

สำอางค์ หิรัญบูรณะ (Hiranburana, 1971) ได้กล่าวถึงลักษณะของพยางค์ที่ลงเสียงเบาไว้ ดังนี้

1. ค่าความยาวสั้นกว่าพยางค์ที่ลงเสียงหนัก
2. ในพยัญชนะควบกล้ำ พยัญชนะตัวที่สองจะหายไป
3. ในพยางค์ที่ลงท้ายด้วยพยัญชนะกึ่งสระ พยัญชนะกึ่งสระจะถูกรวมกับสระ หรือพยัญชนะกึ่งสระจะหายไป
4. พยางค์ที่ลงท้ายด้วยเสียงกัก เสียงกักจะหายไป
5. สระในพยางค์ที่ลงเสียงเบาจะปรากฏเพียง 6 สระโดย / i v u / ยังคงเดิมสระ / x a @ / ขยับขึ้นไปเป็นสระกลางที่สัมพันธ์กับ / e q o / (สัญลักษณ์แทนเสียงสระนี้ใช้ตามระบบการเขียนแบบ LRU ซึ่งกล่าวไว้ใน Schoknecht, 2000)

6. พยางค์ที่มีเสียงวรรณยุกต์ เสียงสูง เสียงขึ้น เสียงตก จะเปลี่ยนระดับเสียงเป็นเสียงสูง ซึ่งแตกต่างจากเสียงวรรณยุกต์เสียงสูง

ธีรพันธ์ เหลืองทองคำ (Luangthongkum, 1977) ได้กล่าวถึงลักษณะร่วมทางสัทศาสตร์ของพยางค์ที่ลงเสียงหนักแตกต่างจากพยางค์ที่ลงเสียงเบา สรุปได้ คือ ความยาว คุณสมบัติของสระและพยัญชนะ และคุณสมบัติของวรรณยุกต์ (ระดับเสียงสูงต่ำ)

จากงานวิจัยของ สุดาพร ลักษณียนาวิน (Luksaneeyanawin, 1993) ได้สรุป การลงเสียงหนักเกี่ยวพันกับลักษณะร่วมทางสัทศาสตร์ 4 อย่าง คือ ระดับเสียงสูงต่ำ ความดังค่อย ความยาว และคุณสมบัติของสระพยัญชนะและวรรณยุกต์

จากงานวิจัยของนักภาษาศาสตร์หลายท่าน ทำให้สรุปได้ว่า การลงเสียงหนักเกี่ยวพันกับลักษณะร่วมทางสัทศาสตร์ทั้ง 4 คือ ความดังค่อย ความสั้นยาว ระดับเสียงสูงต่ำ และคุณสมบัติของพยัญชนะสระและวรรณยุกต์ โดยมีเนื้อหารายละเอียดดังจะกล่าวในหัวข้อต่อไป

1. **ระดับเสียงสูงต่ำ** (Pitch)

สุดาพร ลักษณียนาวิน (Luksaneeyanawin, 1983, 1998) สุกัลยา สุรินทร์ไพบูลย์ (2528) และ รุจนา พินิจารมย์ (2533) ได้ศึกษาการเปลี่ยนแปลงของระดับเสียงสูงต่ำในพยางค์เสียงเบาในพยางค์ 2 ประเภท สรุปได้ดังนี้

**ประเภทพยางค์เชื่อม**

เสียงสูง (เสียงตรี) เมื่อไม่ลงเสียงหนักส่วนมากจะปรากฏเป็นเสียงกลางระดับ แต่บางครั้งปรากฏเป็นเสียงสูงระดับแต่ต่ำลงกว่าเดิม

เสียงต่ำ (เสียงเอก) เมื่อไม่ลงเสียงหนักส่วนมากจะปรากฏเป็นเสียงกลางระดับ แต่บางครั้งปรากฏเป็นเสียงต่ำระดับ

**ประเภทพยางค์ไม่เชื่อม**

เสียงตก (เสียงโท) เมื่อไม่ลงเสียงหนักจะปรากฏเป็นได้ทั้งเสียงสูงระดับและเสียงสูงตกที่มีช่วงการตกสั้นลง

เสียงขึ้น (เสียงจัตวา) เมื่อไม่ลงเสียงหนักจะปรากฏเป็นเสียงขึ้น

เสียงสูง (เสียงตรี) เมื่อไม่ลงเสียงหนักส่วนมากจะปรากฏเป็นเสียงสูงระดับ แต่บางส่วนปรากฏเป็นเสียงกลางระดับ

เสียงกลาง (เสียงสามัญ) เมื่อไม่ลงเสียงหนักจะปรากฏเป็นเสียงกลางระดับ

เสียงต่ำ (เสียงเอก) เมื่อไม่ลงเสียงหนักจะปรากฏเป็นเสียงต่ำระดับ แต่บางส่วนปรากฏเป็นเสียงกลางระดับ

2. **ความยาว** (Length)

พยางค์ที่ลงเสียงหนักจะมีความยาวของช่วงเวลามากกว่าพยางค์ที่ลงเสียงเบา

ผลการวิจัยของสุกัลยา สุรินทร์ไพบูลย์ (2528) สอดคล้องกับการวิจัยของธีรพันธ์ เหลืองทองคำ ซึ่งความสั้นยาวขึ้นอยู่กับโครงสร้างจำนวนพยางค์ในหน่วยจังหวะ

มีนักภาษาศาสตร์หลายท่านกล่าวถึงลักษณะของความยาวของช่วงเวลาไว้ สำอาง หิรัญบูรณะ (Hiranburana, 1971) ธีรพันธ์ เหลืองทองคำ (Luangthongkum, 1977) สุกัลยา สุรินทร์ไพบูลย์ (2528) ผลงานวิจัยของแต่ละท่านชี้ให้เห็นว่า พยางค์ที่ลงเสียงหนักจะมีความยาวมากกว่าพยางค์ที่ลงเสียงเบา

ธีรพันธ์ เหลืองทองคำ (Luangthongkum, 1977) ได้ศึกษาความยาวของช่วงเวลาโดยการวัดค่าระยะเวลาของพยางค์ในหน่วยจังหวะ (รายละเอียดของการแบ่งหน่วยจังหวะได้กล่าวไว้ในภาคผนวก ก.) ได้ทำการศึกษาจากโครงสร้างทางจังหวะ แบ่งหน่วยเวลาตามจำนวนพยางค์ในหน่วยจังหวะ ตั้งแต่หน่วยจังหวะที่ประกอบด้วย 1 พยางค์หนัก (| S |) จนถึงหน่วยจังหวะที่ประกอบด้วย 1 พยางค์หนัก 3 พยางค์เบา (| S U U U| ) ดังนี้

1. $|S| = 3$
2. $|S\ U| = 2:1$
3. $|S\ U\ U| = 1\frac{1}{2}:\frac{3}{4}:\frac{3}{4}$
4. $|S\ U\ U\ U| = 1:\frac{2}{3}:\frac{2}{3}:\frac{2}{3}$

เมื่อ | | แทน ขอบเขตของหน่วยจังหวะ
S แทน พยางค์ที่ลงเสียงหนัก
U แทน พยางค์ที่ลงเสียงเบา

ช่วงเวลาของหน่วยจังหวะที่จำนวนพยางค์ไม่เท่ากัน โดยมีค่าเฉลี่ยเวลาของโครงสร้างหน่วยจังหวะแต่ละแบบ และค่าเฉลี่ยช่วงเวลาของพยางค์ในโครงสร้างหน่วยจังหวะแต่ละแบบแสดงดังนี้

$|S| = 0.30$ วินาที

$|S\ U| = 0.40$ วินาที หรือ $|0.25:0.15|$ วินาที

$|S\ U\ U| = 0.49$ วินาที หรือ $|0.23:0.13:0.13|$ วินาที

$|S\ U\ U\ U| = 0.59$ วินาที หรือ $|0.21:0.12:0.13:0.13|$ วินาที

อัตราส่วนของช่วงเวลาเท่ากับ 3 : 4 : 5 : 6 สำหรับ 1 พยางค์, 2 พยางค์, 3 พยางค์ และ 4 พยางค์ ตามลำดับ

นอกจากการศึกษาความยาวของพยางค์ ธีรพันธ์ยังพบว่าจังหวะมีผลต่อความยาวของแต่ละพยางค์ ไม่เฉพาะความยาวของสระเท่านั้นที่ได้รับผลจากจังหวะนี้ แต่ความยาวของพยัญชนะต้นและพยัญชนะท้ายก็เปลี่ยนไปตามจังหวะด้วย

สุกัลยา สุรินทร์ไพบูลย์ (2528) พบว่า พยางค์ที่ลงเสียงหนัก ส่วนใหญ่จะยาวกว่าพยางค์ที่ไม่ลงเสียงหนัก ในคำสามัญคำเดียวกันพบพยางค์ลงเสียงหนักที่มีความยาวเท่ากันหรือน้อยกว่าพยางค์เสียงเบาเพียงร้อยละ 2.15 ของคำทั้งหมดเท่านั้น และได้ทำการคำนวณค่าเฉลี่ยของพยางค์ที่ลงเสียงหนักและไม่ลงเสียงหนักในหน่วยจังหวะพยางค์เดียวที่ปรากฏหน้าการหยุดเว้นระยะ หน่วยจังหวะพยางค์เดียว จนถึงหน่วยจังหวะ 3 พยางค์ ได้ผล ดังนี้

$|S|\#$ = 0.38 วินาที

$|S|$ = 0.24 วินาที

$|S\ U|$ = 0.35 วินาที หรือ $|0.22 : 0.13|$ วินาที

$|S\ U\ U|$ = 0.51 วินาที หรือ $|0.23 : 0.15 : 0.13|$ วินาที

เมื่อ |S|# แทน หน่วยเสียงพยางค์เดียวที่ปรากฏหน้าการหยุดเว้นระยะ

3. **คุณสมบัติของสระและพยัญชนะ** (Vowel and Consonant Quality)

จากการศึกษาของ สำอาง หิรัญบูรณะ (Hiranburana, 1971) ธีรพันธ์ เหลืองทองคำ (Luangthongkum, 1977) และสุกัลยา สุรินทร์ไพบูลย์ (2528) สรุปได้ ดังนี้

**คุณสมบัติของสระ**

1. สระจะเปลี่ยนไปเป็นสระกลางมากขึ้น
2. ความแตกต่างระหว่างสระเสียงสั้นและสระเสียงยาวหายไป
3. สระผสมจะกลายเป็นสระเดี่ยว

**คุณสมบัติของพยัญชนะ**

1. เสียงกักเส้นเสียง /?/ จะหายไป
2. เสียง ร และ ล ในพยัญชนะควบกล้ำหายไป เช่น
3. ความแตกต่างระหว่าง ร และ ล ซึ่งมีนัยสำคัญทางภาษาศาสตร์หายไป
4. เสียงพยัญชนะท้าย /t/ ของพยางค์ที่ไม่ลงเสียงหนักอาจหายไปได้
5. เสียงกึ่งเสียดแทรก /ch/ กลายเป็น เสียงเสียดแทรก /ɕ/

6. เสียงพยัญชนะท้ายของพยัญชนะที่ไม่ลงเสียงหนักจะกลมกลืนกับพยัญชนะต้นของพยางค์ที่ลงเสียงหนักตัวที่ติดกันตัวถัด
7. เสียงพยัญชนะท้ายของพยัญชนะที่ไม่ลงเสียงหนักจะหายไป ถ้าพยัญชนะท้ายกับพยัญชนะต้นตัวที่ตามมาเป็นเสียงเดียวกัน (เช่น pp, tt, kk, mm, nn, ngng)

4. **ความดังค่อย** (Loudness)

สำอาง หิรัญบูรณะ (Hiranburana, 1971) พบว่า พยางค์ที่ลงเสียงหนักจะเปล่งด้วยเสียงที่ดังมากขึ้น

จากการศึกษาของ สุกัลยา สุรินทร์ไพบูลย์ (2528) พบว่า พยางค์ที่ลงเสียงหนักจะรับฟังได้ว่าดังกว่าพยางค์ที่ลงเสียงเบา ยกเว้น กรณีที่พยางค์สองพยางค์ซึ่งอยู่ชิดกันมีความสั้นยาวของพยางค์แตกต่างกันมาก โดยที่พยางค์ที่สั้นกว่า ดังกว่า เล็กน้อย พยางค์ที่ยาวกว่าแต่มีความดังน้อยกว่าเป็นพยางค์ที่ลงเสียงหนัก

ลักษณะของความดังค่อยสรุปได้ว่า พยางค์ที่ลงเสียงหนักมีแนวโน้มที่จะมีความดังมากกว่าพยางค์ที่ลงเสียงเบา

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

### 2.1.3 เสียงพูดภาษาไทย

พยางค์ในภาษาไทยมีโครงสร้างเป็น C(C)V(V)C คือมีเสียงพยัญชนะต้น เสียงพยัญชนะตาม เสียงสระ เสียงพยัญชนะสะกด และเสียงวรรณยุกต์ โดยแต่ละเสียงมีรายละเอียด ดังนี้

1. เสียงสระ (Vowel Sounds) เสียงสระในภาษาไทยประกอบด้วยสระเดี่ยวและสระผสม ดังนี้

1.1 สระเดี่ยว (Monophthongs) ในภาษาไทยมีทั้งหมด 18 เสียง แบ่งเป็นสระเสียงยาว 9 เสียง สระเสียงสั้น 9 เสียง จับกันเป็นคู่ แต่ละคู่มีคุณลักษณะของเสียงใกล้เคียงกัน แต่มีความยาวต่างกัน (สุดาพร ลักษณียนาวิน, 1993) ดังตารางที่ 2.4

ตารางที่ 2.4 สระเดี่ยวในภาษาไทย (Luksaneeyanawin, 1993)

| ความสูงลิ้น (Tongue Height) | ตำแหน่ง (Tongue Advancement) | | |
|---|---|---|---|
| | หน้า (Front) | กลาง (Central) | หลัง (Back) |
| สูง (High) | i, ii<br>อิ, อี | v, vv<br>อึ, อื | u, uu<br>อุ, อู |
| กลาง (Mid) | e, ee<br>เอะ, เอ | q, qq<br>เออะ, เออ | o, oo<br>โอะ, โอ |
| ต่ำ (Low) | x, xx<br>แอะ, แอ | a, aa<br>อะ, อา | @, @@<br>เอาะ, ออ |

1.2 สระผสม (Diphthongs) คือสระที่เกิดจากการเรียงต่อกันเป็นลำดับของสระเดี่ยวสองตัว ในภาษาไทยมีสระผสมเพียงแค่ 6 เสียง เป็นสระเสียงสั้นและสระเสียงยาวอย่างละ 3 เสียง จับกันเป็นคู่ ดังตารางที่ 2.5

ตารางที่ 2.5 สระผสมในภาษาไทย (Luksaneeyanawin, 1993)

| | ตำแหน่งลิ้นของเสียงสระตัวหน้า | | |
|---|---|---|---|
| | หน้า (Front) | กลาง (Central) | หลัง (Back) |
| สระผสม (Diphthongs) | ia, iia<br>เอียะ, เอีย | va, vva<br>เออะ, เออ | ua, uua<br>อัวะ, อัว |

2. เสียงพยัญชนะ (Consonant Sounds) แบ่งได้เป็น 2 กลุ่มคือ

1) เสียงพยัญชนะต้น ในภาษาไทยมีทั้งหมด 21 เสียง แบ่งเป็นพยัญชนะเสียงกัก 11 เสียง พยัญชนะเสียงไม่กัก 10 เสียง ดังตารางที่ 2.6 ซึ่งแสดงเสียงพยัญชนะต้นพร้อมรูปเขียน

ตารางที่ 2.6 เสียงพยัญชนะต้นและรูปเขียนในภาษาไทย (Luksaneeyanawin, 1993)

| ลักษณะการออกเสียง (Manner of Articulation) | | ตำแหน่งการออกเสียง (Place of Articulation) | | | | |
|---|---|---|---|---|---|---|
| | | ริมฝีปาก (Labial) | ปุ่มเหงือก (Alveolar) | เพดานบน (Palatal) | เพดานอ่อน (Velar) | เส้นเสียง (Glottal) |
| เสียงกัก (Stop) | ไม่ก้อง ไม่พ่นลม (Voiceless Unaspirated) | p<br>ป | t<br>ต, ฏ | c<br>จ | k<br>ก | ?<br>อ |
| | ไม่ก้อง พ่นลม (Voiceless Aspirated) | ph<br>พ, ภ, ผ | th<br>ท, ธ, ฒ, ฑ, ถ, ฐ | ch<br>ช, ฌ, ฉ | kh<br>ค | |
| | ก้อง (Voiced) | b<br>บ | d<br>ด, ฎ, ฑ | | | |
| เสียงไม่กัก (Non-Stop) | นาสิก (Nasal) | m<br>ม | n<br>น, ณ | | ng<br>ง | |
| | เสียดแทรก (Fricative) | f<br>ฟ, ฝ | s<br>ซ, ศ, ษ, ส | | | h<br>ฮ, ห |
| | รัว (Trill) | | r<br>ร, ฤ | | | |
| | ข้างลิ้น (Lateral) | | l<br>ล, ฤ | | | |
| | อัฒสระ (Approximant) | w<br>ว | | j<br>ย, ญ | | |

2) เสียงพยัญชนะควบกล้ำ (Double Consonant Sounds) ในภาษาไทยมีทั้งหมด 12 เสียง ดังตารางที่ 2.7 ซึ่งแสดงเสียงพยัญชนะควบกล้ำพร้อมรูปเขียน

ตารางที่ 2.7 เสียงพยัญชนะควบกล้ำ (Luksaneeyanawin, 1993)

| | | | |
|---|---|---|---|
| กลุ่มกักไม่พ่นลม (Unaspirated Stop Set) | pr<br>ปร, ปฤ | tr<br>ตร, ตฤ | kr<br>กร, กฤ |
| | pl<br>ปล | | kl<br>กล |
| | | | kw<br>กว |
| กลุ่มกักไม่พ่นลม (Aspirated Stop Set) | phr<br>พร, พฤ, ภร ,ภฤ | thr<br>ทร, ทฤ | khr<br>คร, ขร |
| | phl<br>พล, ผล | | khl<br>คล, ขล |
| | | | khw<br>คว, ขว |

3) เสียงตัวสะกดในภาษาไทยมีเสียงพยัญชนะเพียง 8 เสียงที่สามารถเกิดที่ตำแหน่งท้ายพยางค์ ทำหน้าที่เป็นตัวสะกด ดังตารางที่ 2.8 ซึ่งแสดงเสียงพยัญชนะท้ายและรูปเขียน จะสังเกตได้ว่ารูปเขียนที่ใช้แทนเสียงพยัญชนะท้ายแต่ละหน่วยมีจำนวนค่อนข้างมากโดยเฉพาะในกลุ่มพยัญชนะปุ่มเหงือก

ตารางที่ 2.8 เสียงพยัญชนะสะกดในภาษาไทย (Luksaneeyanawin, 1993)

| ลักษณะการออกเสียง (Manner of Articulation) | ตำแหน่งการออกเสียง (Place of Articulation) | | | |
|---|---|---|---|---|
| | ริมฝีปาก (Labial) | ปุ่มเหงือก (Alveolar) | เพดานบน (Palatal) | เพดานอ่อน (Velar) |
| กัก (Stops) | p<br>บ, ป, พ, ภ, ฟ | t<br>ด, ฎ, ต, ฏ,ท, ธ, ฒ, ฑ, ถ, ฐ, จ, ช,ซ, ศ, ษ, ส | | k<br>ก, ต, ฏ, ข |
| นาสิก (Nasals) | m<br>ม | n<br>น, ณ, ร, ล, ฬ, ญ | | ng<br>ง |
| อัฒสระ (Approximants) | w<br>ว | | j<br>ย | |

3. เสียงวรรณยุกต์ (Tones) (Luksaneeyanawin, 1993) ในภาษาไทยมี 5 เสียง มีลักษณะดังรูปที่ 2.1 แบ่งเป็น 2 กลุ่ม ได้ ดังนี้

วรรณยุกต์คงระดับ (Static Tone) ประกอบด้วย 3 เสียงคือ

1) เสียงสูง (High) หรือเสียงวรรณยุกต์ตรี

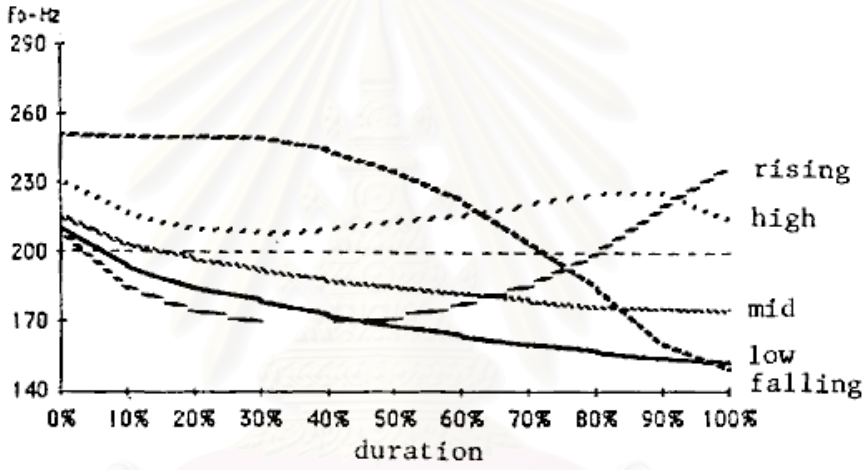
2) เสียงกลาง (Mid) หรือเสียงวรรณยุกต์สามัญ

3) เสียงต่ำ (Low) หรือเสียงวรรณยุกต์เอก

วรรณยุกต์เปลี่ยนระดับ (Dynamic Tone) ประกอบด้วย 2 เสียงคือ

1) เสียงขึ้น (Rising) หรือเสียงวรรณยุกต์จัตวา

2) เสียงตก (Falling) หรือเสียงวรรณยุกต์โท




รูปที่ 2.1 ความถี่มูลฐานของเสียงวรรณยุกต์ในภาษาไทย (ปิยฉัตร ปานโรจน์, 2534)

### 2.1.4 พยางค์เป็นพยางค์ตาย

พยางค์ในภาษาไทยมี 2 ชนิด คือ พยางค์เป็น (Live Syllables หรือ Sonorant Ending Syllables) เป็นพยางค์ที่ลงท้ายด้วยเสียงก้อง และ พยางค์ตาย (Dead Syllable หรือ Obstruent Ending Syllables) เป็นพยางค์ที่ลงท้ายด้วยเสียงกัก

พยางค์เป็นมี 3 ประเภท

1) พยางค์ที่ลงท้ายด้วยสระหรือพยางค์เปิด (CV:) เช่น กา

2) พยางค์สระเสียงสั้นที่ลงท้ายด้วยเสียงนาสิก (CVN) หรือเสียงกึ่งสระ (CVW) เช่น กัน กลัว

3) พยางค์สระเสียงยาวที่ลงท้ายด้วยเสียงนาสิก (CV:N) หรือเสียงกึ่งสระ (CV:W) เช่น กาน กาว

พยางค์ตายมี 3 ประเภท

1) พยางค์สระเสียงสั้นที่ลงท้ายด้วยเสียงกักที่เส้นเสียง (CV?) เช่น กะ
2) พยางค์สระเสียงสั้นที่ลงท้ายด้วยเสียงกักอื่นๆ (CVS) เช่น กัก
3) พยางค์สระเสียงยาวที่ลงท้ายด้วยเสียงกักอื่นๆ (CV:S) เช่น กาก

โดยที่ C แทน พยัญชนะต้น
V แทน สระเสียงสั้น
V: แทน สระเสียงยาว
N แทน ตัวสะกดเสียงนาสิก หรือตัวสะกดที่มีหน่วยเสียงเป็น /m/, /n/, /ng/
W แทน ตัวสะกดกึ่งสระ หรือตัวสะกดที่มีหน่วยเสียงเป็น /w/, /j/
? แทน เสียงกักที่เส้นเสียง
S แทน ตัวสะกดเสียงกักที่ไม่ใช่เสียงกักที่เส้นเสียง หรือตัวสะกดที่มีหน่วยเสียงเป็น /k/, /t/, /p/

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

## 2.2 ทฤษฎีในการวิเคราะห์และสังเคราะห์เสียง

### 2.2.1 การตรวจหาพิทช์โดยวิธีการหาค่าอัตสหสัมพันธ์

มีผู้ใช้การหาค่าอัตสหสัมพันธ์ในการตรวจหาพิทช์ (Pitch Detection) กันอย่างแพร่หลาย ดังจะเห็นได้จาก Sondhi (1968); Dubnowski, Schafer and Rabiner (1976); Rabiner (1977)

อัตสหสัมพันธ์ (Autocorrelation) เป็นเครื่องมือวัดความสม่ำเสมอของสัญญาณ สำหรับสัญญาณเสียงนั้น สัญญาณเสียงซึ่งมีค่าอัตสหสัมพันธ์ที่มากและมีลักษณะสม่ำเสมอเป็นลักษณะของสัญญาณเสียงก้อง และในทางกลับกันสัญญาณเสียงซึ่งมีค่าอัตสหสัมพันธ์ที่น้อยและไม่สม่ำเสมอเป็นลักษณะของสัญญาณเสียงไม่ก้อง ค่าอัตสหสัมพันธ์สามารถหาได้ ดังสมการที่ 2.1

$$r_x(\eta;m) = \frac{1}{N}\sum_{n=-\infty}^{\infty} s(n)\,w(m-n)\,s(n+|\eta|)\,w(m-n+|\eta|) \qquad (2.1)$$

เมื่อ $r_x(\eta;m)$ แทน ค่าอัตสหสัมพันธ์ที่ตำแหน่ง m เมื่อหน่วงเวลาไปเป็นระยะ $\eta$

s(n) แทน สัญญาณเสียงที่ตำแหน่ง n

w(n) แทน ฟังก์ชันหน้าต่าง

N แทน ความกว้างของฟังก์ชันหน้าต่าง

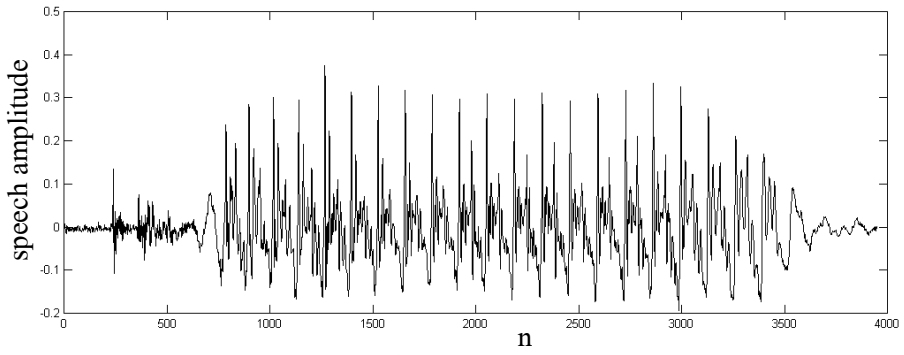
ในสัญญาณที่เป็นสัญญาณคาบที่มีคาบเท่ากับ P จะทำให้ความสัมพันธ์ตามสมการที่ 2.1 มีค่ามากเมื่อ $\eta = 0, \pm P, \pm 2P, \ldots$

สำหรับเสียงพูดที่เป็นเสียงก้อง ค่าอัตสหสัมพันธ์จะมีค่ามากทุกๆตำแหน่งที่เป็นจำนวนเท่าของคาบพิทช์ ดังรูปที่ 2.2

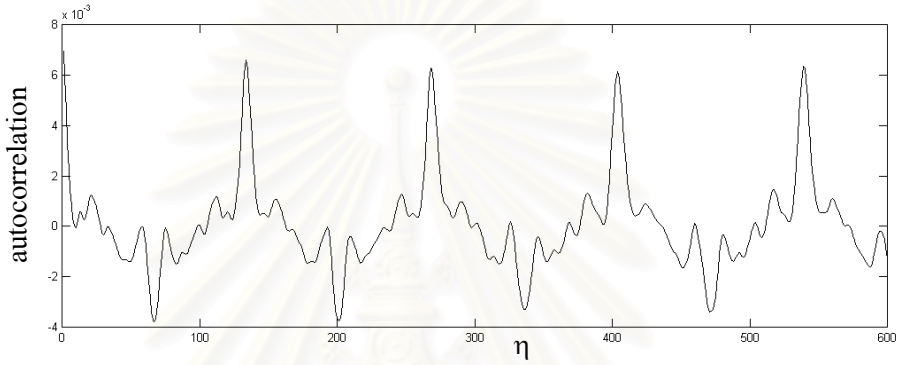
จากรูปที่ 2.2 ก. แสดงสัญญาณเสียงคำว่า “กา” (kaa0) รูปที่ 2.2 ข. ได้แสดงค่าอัตสหสัมพันธ์ของสัญญาณเสียง ณ จุดตัวอย่างที่ 2055 ค่าอัตสหสัมพันธ์จะมีลักษณะซ้ำเป็นคาบ และระยะห่างระหว่างค่าสูงสุดจะเท่ากัน ซึ่งเท่ากับคาบของสัญญาณเสียง

การตรวจหาพิทช์โดยวิธีการใช้ค่าอัตสหสัมพันธ์สามารถลดความผิดพลาดได้โดยการใช้วิธีขลิบกลาง (Central Clipping) ร่วมด้วย โดยวิธีการ Central Clipping มีฟังก์ชันดังสมการที่ 2.2

$$C\{s(n)\} = \begin{cases} s(n) - C^+, & s(n) > C^+ \\ 0, & C^- \le s(n) \le C^+ \\ s(n) - C^-, & s(n) < C^- \end{cases} \qquad (2.2)$$



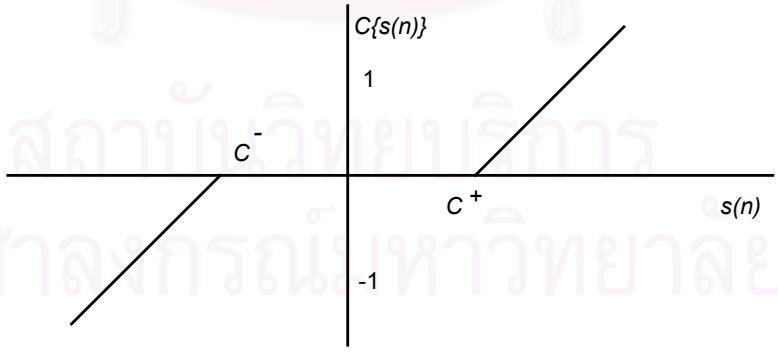

(ก)




(ข)

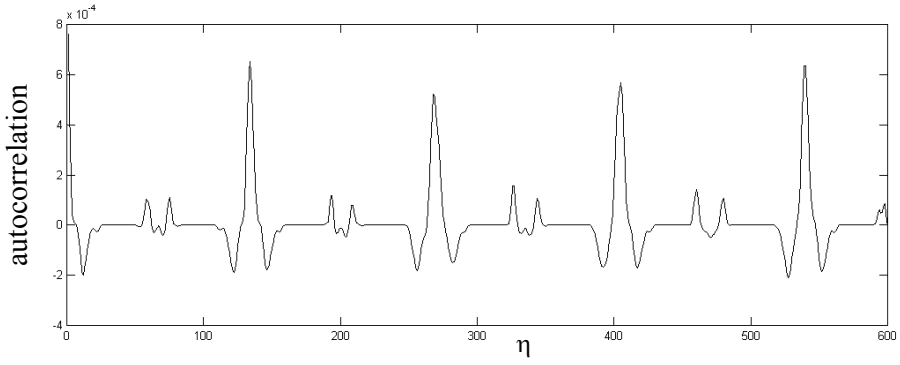
รูปที่ 2.2 ค่าอัตสหสัมพันธ์ ของสัญญาณเสียงคำว่า “กา” (kaa0)
ณ จุดตัวอย่างที่ 2055 (m=2055) (ก) สัญญาณเสียง (ข) ค่าอัตสหสัมพันธ์

โดยทั่วไป $C^+$, $C^-$ จะมีค่าประมาณ 30% ของค่าสูงสุดของค่าสัมบูรณ์ของสัญญาณ (Deller et al., 1993) C{s(n)} มีลักษณะดังแสดงในรูปที่ 2.3




รูปที่ 2.3 ลักษณะของ C{s(n)}

เมื่อนำสัญญาณเสียงมาทำ Central Clipping แล้วนำมาหาค่าอัตสหสัมพันธ์จะให้ค่าสูงสุดที่เด่นชัดขึ้นดังรูปที่ 2.4




รูปที่ 2.4 ค่าอัตสหสัมพันธ์ ของสัญญาณเสียงคำว่า “กา” (kaa0)
ณ จุดตัวอย่างที่ 2055 (m=2055) หลังจากทำ Central Clipping

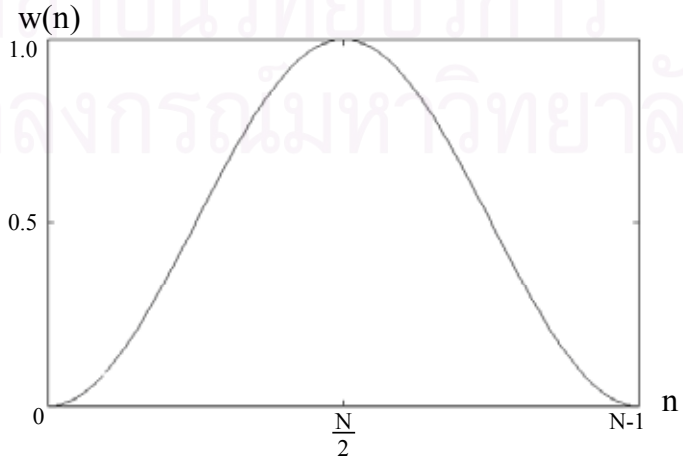
ลักษณะของค่าอัตสหสัมพันธ์หลังจากการทำ Central Clipping จะเห็นว่าจุดสูงสุดชัดเจนขึ้น ช่วยให้หาจุดสูงสุดได้แม่นยำขึ้น

### 2.2.2 **หน้าต่างแฮนนิง** (Hanning Window)

ฟังก์ชันหน้าต่างแฮนนิง w(n) มีค่าดังสมการที่ 2.3

$$w(n) = \begin{cases} 0.5 \times \left[1 - \cos\left(\frac{2\pi n}{N-1}\right)\right] & , \quad n = 0, 1, 2, \ldots, N-1 \\ 0 & , \quad n \quad \text{ค่าอื่นๆ} \end{cases} \qquad (2.3)$$

หน้าต่างแฮนนิงเป็นหน้าต่างที่ใช้งานได้หลากหลาย หน้าต่างแฮนนิงมีลักษณะในเชิงเวลาดังรูปที่ 2.5 หน้าต่างแฮนนิงเป็นหน้าต่างที่นำมาใช้วิเคราะห์ด้วยฟังก์ชันหน้าต่างสำหรับการทำ Time-Domain Pitch-Synchronous Overlap Add




รูปที่ 2.5 หน้าต่างแฮนนิ่ง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

โดยทั่วไปความกว้างของหน้าต่างสำหรับการทำ Time-Domain Pitch-Synchronous Overlap Add มีขนาดแปรผันตามคาบเวลาของสัญญาณเสียง ซึ่งมีขนาดตั้งแต่ 2 ถึง 4 เท่า ของคาบเวลาของสัญญาณเสียง (Charpentier and Moulines, 1989)

### 2.2.3 วิธี Time-Domain Pitch-Synchronous Overlap Add (TD-PSOLA)

วิธี Time-Domain Pitch-Synchronous Overlap-Add (Charpentier and Stella, 1986; Charpentier and Moulines, 1989) เป็นวิธีที่ใช้กันอย่างแพร่หลายในการสังเคราะห์เสียงแบบต่อกัน (Concatenative synthesis) เพราะให้คุณภาพการสังเคราะห์ดีและคำนวณไม่มากนัก (Thierry Dutoit, 1997) วิธี Time-Domain Pitch-Synchronous Overlap-Add ใช้ในการปรับปรุงคุณภาพของเสียงให้เป็นธรรมชาติมากขึ้น โดยการปรับระยะระหว่างยอดพิทช์ของสัญญาณเสียงให้มีระดับเสียงสูงต่ำตามต้องการและใช้ในการปรับระยะเวลาของสัญญาณเสียงโดยการเพิ่มหรือลดสัญญาณช่วงสั้นให้มีช่วงเวลาของเสียงตามต้องการได้

วิธี Time-Domain Pitch-Synchronous Overlap-Add เริ่มจากการวิเคราะห์สัญญาณเสียงออกเป็นสัญญาณช่วงสั้น โดยการวิเคราะห์สัญญาณเสียงด้วยฟังก์ชันหน้าต่างที่ตำแหน่งยอดพิทช์ของเสียงก้องและกำหนดคาบเวลาของเสียงไม่ก้องจากค่าคาบเวลาที่วิเคราะห์ได้จากเสียงก้อง ในการปรับระยะเวลาการเพิ่มช่วงเวลาจะทำได้โดยเพิ่มและการลดระยะเวลาจะทำได้โดยลดสัญญาณเสียงออก ในการปรับระดับเสียงสูงต่ำทำได้โดยปรับระยะเวลาระหว่างยอดพิทช์ให้กว้างขึ้นเป็นการลดค่าความถี่มูลฐานและการปรับระยะเวลาระหว่างยอดพิทช์ให้แคบลงเป็นการเพิ่มค่าความถี่มูลฐาน ซึ่งเป็นค่าที่กำหนดเสียงสูงต่ำ การปรับระดับเสียงสูงต่ำนั้นจะทำเฉพาะในส่วนที่เป็นเสียงก้องเท่านั้น ในการสังเคราะห์ซ้ำ (Resynthesis) จะใช้การบวกซ้อน (Overlap - Add) ในการควบคุมเวลาและความถี่มูลฐานของเสียงตามที่ต้องการได้

รายละเอียดของการลดเพิ่มช่วงเวลาและการเปลี่ยนระดับเสียง จะได้กล่าวเพิ่มเติมในบทต่อๆไป

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

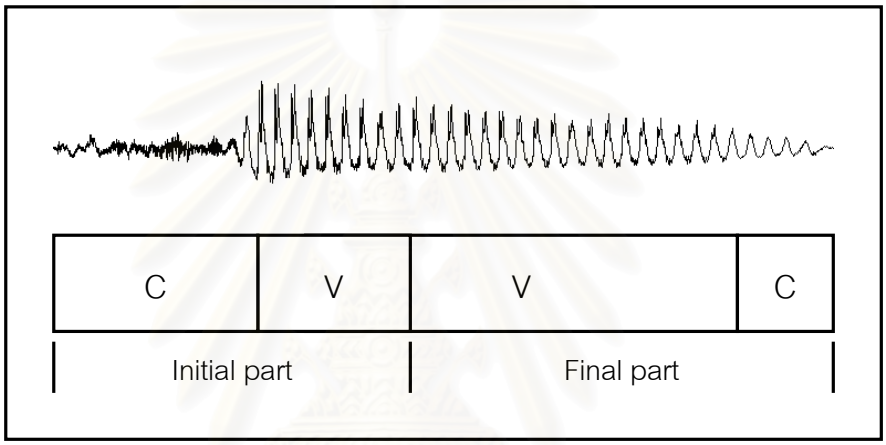
**2.3 หน่วยเสียงอัฒพยางค์** (Demisyllable)

หน่วยเสียงอัฒพยางค์เป็นหน่วยเสียงที่เกิดจากการแบ่งสัญญาณเสียงหนึ่งพยางค์ออกเป็นหน่วยเสียงเริ่ม (Initial) และหน่วยเสียงท้าย (Final)

หน่วยเสียงเริ่ม (Initial) จะเก็บเสียงของพยัญชนะต้น พยัญชนะควบกล้ำ จนถึงช่วงของเสียงสระคงที่ตรงกลางพยางค์

หน่วยเสียงท้าย (Final) จะเก็บเสียงของสระและเสียงอื่นๆที่เหลือจนหมด โดยจะเริ่มที่ช่วงของเสียงสระคงที่เป็นต้นไป

การแบ่งหน่วยเสียงเป็นหน่วยเสียงเริ่มและหน่วยเสียงท้ายแบ่งได้ดังรูปที่ 2.6




รูปที่ 2.6 โครงสร้างหน่วยเสียงอัฒพยางค์

พยางค์สามารถสร้างได้โดยนำหน่วยเสียงส่วนหน้าและหน่วยเสียงส่วนท้ายมาต่อเชื่อมกัน (Concatenate) เช่น

- คำว่า “กาก (kaak1)” สร้างจาก หน่วยเสียงเริ่ม “ka” และหน่วยเสียงท้าย “aak1”

  ka + aak1 ➔ kaak1

- คำว่า “เกวียน (kwiian0)” สร้างจาก หน่วยเสียงเริ่ม “kwia” และหน่วยเสียงท้าย “iian0”

  kwia + iian0 ➔ kwiian0

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

# บทที่ 3

# การวิเคราะห์สัญญาณเสียงและการสร้างหน่วยเสียง

ในบทนี้จะแบ่งเนื้อหาออกเป็น 3 ส่วน ในส่วนแรกได้กล่าวถึงการวิเคราะห์ความถี่มูลฐานของพยางค์เสียงเบาเพื่อสร้างเป็นข้อมูลสำหรับการปรับความถี่มูลฐาน ส่วนที่สองได้กล่าวถึงการวิเคราะห์อัตราส่วนแอมพลิจูดระหว่างพยางค์เสียงหนักและพยางค์เสียงเบาเพื่อสร้างเป็นฐานข้อมูลสำหรับการปรับความดังค่อย และในส่วนสุดท้ายได้กล่าวถึงการสร้างหน่วยเสียงสำหรับการสังเคราะห์เสียงในวิทยานิพนธ์นี้

## 3.1 ฐานข้อมูลความถี่มูลฐานของพยางค์เสียงเบา

ระดับเสียงของเสียงวรรณยุกต์ต่างๆของพยางค์เสียงเบาจะแตกต่างจากระดับเสียงวรรณยุกต์ของพยางค์เสียงหนัก ดังงานวิจัยของ สุดาพร ลักษณียนาวิน (Luksaneeyanawin, 1983, 1998 ) สุกัลยา สุรินทร์ไพบูลย์ (2528 ) รุจนา พินิจารมณ์ (2534 ) ผณินทรา ธีรานนท์ (2543) สำหรับหน่วยเสียงซึ่งสร้างมาจากพยางค์เสียงหนักเมื่อนำมาสังเคราะห์เป็นพยางค์เสียงเบาจะต้องแปลงความถี่มูลฐานตามฐานข้อมูลความถี่มูลฐานของพยางค์เสียงเบา ดังนั้นจึงต้องมีการวิเคราะห์ความถี่มูลฐานของพยางค์เสียงเบาเพื่อนำมาสร้างเป็นฐานข้อมูลความถี่มูลฐานของพยางค์เสียงเบา

ความถี่มูลฐานของแต่ละเสียงวรรณยุกต์จะมีลักษณะแตกต่างกันไปตามโครงสร้างพยางค์ ดังงานวิจัยของ ผณินทรา ธีรานนท์ (2543) ดังนั้นในการสร้างฐานข้อมูลความถี่มูลฐานสำหรับพยางค์เสียงเบาจะวิเคราะห์ความถี่มูลฐานตามโครงสร้างพยางค์สำหรับแต่ละวรรณยุกต์ ดังนี้

1) พยางค์เป็น สำหรับเสียงวรรณยุกต์ สามัญ เอก โท ตรี และจัตวา
2) พยางค์ตายสระเสียงยาว สำหรับเสียงวรรณยุกต์ เอก โท และตรี
3) พยางค์ตายสระเสียงสั้นลงท้ายด้วยเสียงกักที่เส้นเสียง สำหรับเสียงวรรณยุกต์ เอก โท และตรี
4) พยางค์ตายสระเสียงสั้นลงท้ายด้วยเสียงกักอื่นๆ สำหรับเสียงวรรณยุกต์ เอก โท และตรี

### 3.1.1 การสร้างคำทดสอบ

ในการสร้างฐานข้อมูลความถี่มูลฐานสำหรับพยางค์เสียงเบาจะสร้างฐานข้อมูลความถี่แบ่งตามโครงสร้างพยางค์ข้างต้น

ฉะนั้นการสร้างฐานข้อมูลความถี่มูลฐานของพยางค์เสียงเบาจะแบ่งตามโครงสร้างพยางค์และเสียงวรรณยุกต์ 14 เสียง ดังในตารางที่ 3.1

ตารางที่ 3.1 โครงสร้างพยางค์และเสียงวรรณยุกต์สำหรับการวิเคราะห์ความถี่มูลฐาน

| ชนิด | เสียงวรรณยุกต์ | | | | |
|---|---|---|---|---|---|
| | สามัญ | เอก | โท | ตรี | จัตวา |
| พยางค์เป็น | √ | √ | √ | √ | √ |
| พยางค์ตายสระเสียงยาว | | √ | √ | √ | |
| พยางค์ตายสระเสียงสั้นลงท้ายด้วยเสียงกักที่เส้นเสียง | | √ | √ | √ | |
| พยางค์ตายสระเสียงสั้นลงท้ายด้วยเสียงกักอื่นๆ | | √ | √ | √ | |

จากโครงสร้างพยางค์ข้างต้น เลือกพยางค์ทดสอบสำหรับแต่ละโครงสร้างพยางค์และเสียงวรรณยุกต์ทั้ง 14 แบบ แบบละ 3 พยางค์ รวมทั้งหมด 42 พยางค์ทดสอบ

สร้างคำทดสอบโดยที่พยางค์ที่เลือกไว้ทดสอบลงเสียงเบา รวมทั้งหมด 42 คำทดสอบ

### 3.1.2 การบันทึกเสียง

ในการเก็บข้อมูลเสียงจะใช้เสียงที่บันทึกจากเสียงพูดจริงโดยมีรายละเอียด ดังนี้

1) ผู้บอกภาษาเป็นเพศชาย
2) คนออกเสียงใช้ภาษาไทยกรุงเทพฯเป็นหลักจำนวน 1 คน
3) ในการบันทึกเสียง ผู้บอกภาษาออกเสียงเป็นธรรมชาติ
4) บันทึกเสียงเป็นไฟล์โดยไมโครโฟนผ่านการ์ดเสียงเข้าคอมพิวเตอร์โดยตรง

อุปกรณ์และซอฟต์แวร์ที่ใช้ในบันทึกเสียง

1) การ์ดเสียง Sound Blaster AWE 64 Value ของบริษัท Creative Technology
2) ไมโครโฟน NPE รุ่น Pro -18 PROLOGUE
3) โปรแกรมบันทึกเสียงและแต่งเสียง Goldwave ของ Chris Craig

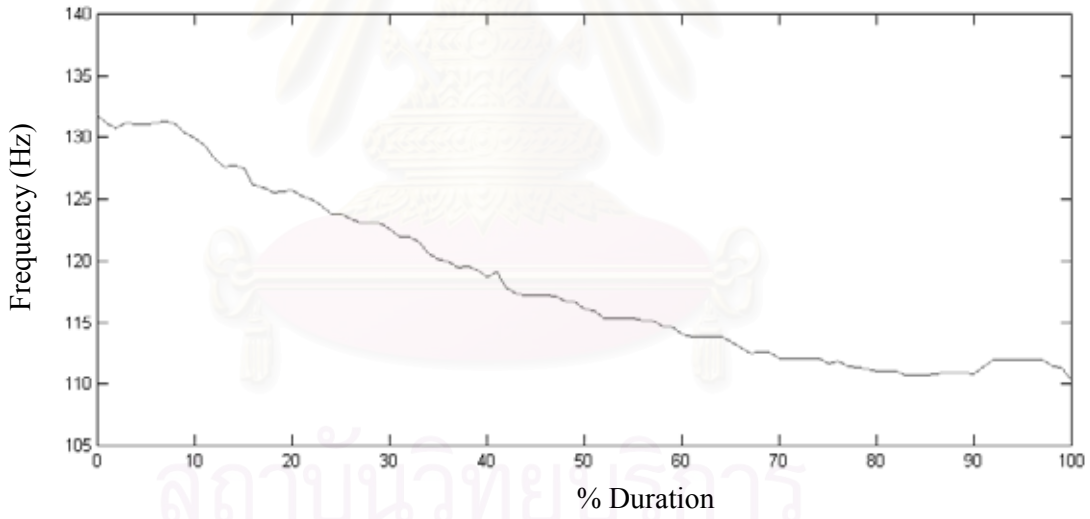
รายละเอียดของเสียงที่บันทึก

1) อัตราสุ่ม 16,000 เฮิรตซ์
2) ความละเอียดของสัญญาณเสียง 16 บิต

### 3.1.3 การวิเคราะห์และประมวลผลข้อมูล

หลังจากการบันทึกเสียง เสียงจะถูกนำมาวิเคราะห์และประมวลผลเพื่อเก็บเป็นฐานข้อมูลความถี่มูลฐานต่อไป โดยการวิเคราะห์และประมวลผลข้อมูลเสียงมีขั้นตอน ดังนี้

1) นำข้อมูลเสียงที่บันทึกไว้มาแบ่งพยางค์ โดยเลือกตัดพยางค์ทดสอบที่เลือกไว้จากการสร้างพยางค์ทดสอบ
2) นำพยางค์ที่ตัดไว้มาวิเคราะห์ความถี่มูลฐานด้วยโปรแกรมวิเคราะห์เสียง Praat ของ Pual Boersma
3) ในการสร้างฐานข้อมูลความถี่มูลฐานของแต่ละโครงสร้างพยางค์ต้องนำความถี่มูลฐานมาหาค่าเฉลี่ย เนื่องจากพยางค์ทดสอบแต่ละพยางค์มีความสั้นยาวของพยางค์ไม่เท่ากัน จึงต้องนำข้อมูลความถี่มูลฐานที่ได้มาประมวลผลแบ่งเป็น 0 – 100 เปอร์เซ็นต์ของช่วงเวลา
4) หาค่าเฉลี่ยของข้อมูลความถี่มูลฐานสำหรับโครงสร้างพยางค์และเสียงวรรณยุกต์ทั้ง 14 แบบ จัดเก็บเป็นฐานข้อมูลความถี่มูลฐานที่มีลักษณะดังรูปที่ 3.1




รูปที่ 3.1 รูปแบบฐานข้อมูลความถี่มูลฐานสำหรับพยางค์เสียงเบา

## 3.2 ฐานข้อมูลอัตราส่วนแอมพลิจูดระหว่างพยางค์เสียงเบาและพยางค์เสียงหนัก

### 3.2.1 การสร้างคำทดสอบ

สำหรับการสร้างฐานข้อมูลอัตราส่วนแอมพลิจูดระหว่างพยางค์เสียงเบาและพยางค์เสียงหนัก จะสร้างฐานข้อมูลแบ่งตามสระในภาษาไทย โดยในภาษาไทยมีเสียงสระทั้งสระเสียงสั้นและสระเสียงยาวรวมกันทั้งหมด 24 เสียง

เลือกพยางค์ทดสอบสำหรับแต่ละเสียงสระทั้ง 24 เสียง เสียงละ 3 พยางค์ รวมทั้งหมด 72 พยางค์ทดสอบ

### 3.2.2 การบันทึกเสียง

การบันทึกเสียงสำหรับการสร้างฐานข้อมูลอัตราส่วนแอมพลิจูดระหว่างพยางค์เสียงเบาและพยางค์เสียงหนักมีรายละเอียดเช่นเดียวกับการบันทึกเสียงในหัวข้อ 3.1.2

### 3.2.3 การวิเคราะห์และประมวลผลข้อมูล

หลังจากการบันทึกเสียง เสียงจะถูกนำมาวิเคราะห์และประมวลผลเพื่อเก็บเป็นฐานข้อมูลความถี่มูลฐานต่อไป โดยการวิเคราะห์และประมวลผลข้อมูลเสียงมีขั้นตอน ดังนี้

1) นำข้อมูลเสียงที่บันทึกเป็นประโยคมาแบ่งพยางค์ โดยเลือกตัดพยางค์ทดสอบที่เลือกไว้จากการสร้างพยางค์ทดสอบ
2) นำพยางค์ที่ตัดไว้มาวิเคราะห์หาค่าแอมพลิจูดเฉลี่ยโดยการหารากเฉลี่ยกำลังสอง (Root mean square) ของสัญญาณเสียงทั้งในพยางค์เสียงหนักและพยางค์เสียงเบา
3) นำข้อมูลค่าแอมพลิจูดเฉลี่ยที่ได้มาหาค่าอัตราส่วนแอมพลิจูดระหว่างพยางค์เสียงเบาและพยางค์เสียงหนัก ดังสมการที่ 3.1

$$\text{อัตราส่วนแอมพลิจูดระหว่างพยางค์เสียงหนักและพยางค์เสียงเบา} = \frac{\text{แอมพลิจูดเฉลี่ยของพยางค์เสียงเบา}}{\text{แอมพลิจูดเฉลี่ยของพยางค์เสียงหนัก}} \quad (3.1)$$

4) นำค่าอัตราส่วนแอมพลิจูดระหว่างพยางค์เสียงเบาและพยางค์เสียงหนักดังตารางที่ 3.2 จัดเก็บเป็นฐานข้อมูลสำหรับการปรับความดัง

ตารางที่ 3.2 อัตราส่วนแอมพลิจูดระหว่างพยางค์เสียงเบาและพยางค์เสียงหนักสำหรับแต่ละเสียงสระ

| เสียงสระ | อัตราส่วนแอมพลิจูดระหว่างพยางค์เสียงหนักและพยางค์เสียงเบา | เสียงสระ | อัตราส่วนแอมพลิจูดระหว่างพยางค์เสียงหนักและพยางค์เสียงเบา |
|---|---|---|---|
| ii | 0.79 | i | 0.77 |
| vv | 0.69 | v | 0.73 |
| uu | 0.80 | u | 0.63 |
| ee | 0.68 | e | 0.66 |
| qq | 0.71 | q | 0.78 |
| oo | 0.66 | o | 0.76 |
| xx | 0.67 | x | 0.77 |
| aa | 0.67 | a | 0.69 |
| @@ | 0.94 | @ | 0.75 |
| iia | 0.94 | ia | 0.88 |
| vva | 0.68 | va | 0.69 |
| uua | 0.95 | ua | 0.67 |

### 3.3 **หน่วยเสียง**

#### 3.3.1 **จำนวนหน่วยเสียง**

1) หน่วยเสียงเริ่ม จะพิจารณาจากเสียงพยัญชนะต้น พยัญชนะควบและเสียงสระ

ในภาษาไทยมีเสียงพยัญชนะต้น 21 เสียง พยัญชนะควบ 12 เสียง รวมมีเสียงพยัญชนะสำหรับหน่วยเสียงเริ่ม 33 เสียง

สำหรับสระในหน่วยเสียงเริ่ม สระเสียงสั้นหรือสระเสียงยาว พิจารณาเป็นเสียงสระเดียวกัน ดังนั้นเสียงสระสำหรับหน่วยเสียงเริ่มมี 12 เสียง

จำนวนรูปแบบหน่วยเสียงเริ่มพิจารณาได้จากผลคูณของจำนวนเสียงพยัญชนะกับจำนวนเสียงสระ ดังนี้

จำนวนหน่วยเสียงเริ่ม = 33 x 12 = 396 หน่วยเสียง

2) หน่วยเสียงท้าย จะพิจารณาจากเสียงสระ เสียงตัวสะกดและเสียงวรรณยุกต์

ในภาษาไทย พยางค์ที่มีเสียงตัวสะกดเป็นคำเป็นจะมีเสียงวรรณยุกต์ได้ทั้ง 5 เสียง แต่ในพยางค์ที่มีเสียงตัวสะกดตายคำเป็นจะมีเสียงวรรณยุกต์ได้ 3 เสียง คือ เสียงเอก เสียงโท และ เสียงตรี

เสียงสระบางเสียงจะไม่มีเสียงตัวสะกดบางตัวต่อท้าย เช่น เสียงสระ อี (/ii/) ไม่มีเสียงตัวสะกด /ng/, /w/ และ /j/ ต่อท้าย

ดังนั้นในการพิจารณาจำนวนหน่วยเสียงท้ายต้องพิจารณาแยกตามชนิดของตัวะกด และ เสียงสระ ดังนี้

- สระเสียงยาว

  สามารถลงท้ายด้วยเสียงตัวสะกดทั้ง 8 เสียง และลงท้ายด้วยพยางค์เปิด แยกเป็น พยางค์เป็น 6 เสียง คือ /m/, /n/, /ng/, /w/, /j/ และ พยางค์เปิด

  พยางค์ตาย 3 เสียง คือ /p/, /t/ และ /k/

- สระเสียงสั้น

  สามารถลงท้ายด้วยเสียงตัวสะกดทั้ง 8 เสียง และลงท้ายด้วยเสียงกักที่เส้นเสียง (/?/) แยกเป็น

  พยางค์เป็น 5 เสียง คือ /m/, /n/, /ng/, /w/ และ /j/

  พยางค์ตาย 4 เสียง คือ /p/, /t/, /k/ และ เสียงกักที่เส้นเสียง (/?/)

  โดยที่แต่ละสระมีตัวสะกดที่สะกดได้ดังตารางที่ 3.3 และ 3.4

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 3.3 หน่วยเสียงท้ายของสระเสียงยาวเมื่อยังไม่รวมเสียงวรรณยุกต์

| สระ | เสียงตัวสะกด | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | พยางค์เป็น | | | | | | พยางค์ตาย | | |
| | m | n | ng | w | j | พยางค์เปิด | p | t | k |
| ii | iim | iin | | | | ii | iip | iit | iik |
| vv | vvm | vvn | | | | vv | vvp | vvt | vvk |
| uu | uum | uun | uung | | uuj | uu | uup | uut | uuk |
| ee | eem | een | eeng | eew | | ee | eep | eet | eek |
| qq | qqm | qqn | qqng | qqw | qqj | qq | qqp | qqt | qqk |
| oo | oom | oon | oong | | ooj | oo | oop | oot | ook |
| xx | xxm | xxn | xxng | xxw | | xx | xxp | xxt | xxk |
| aa | aam | aan | aang | aaw | aaj | aa | aap | aat | aak |
| @@ | @@m | @@n | @@ng | | @@j | @@ | @@p | @@t | @@k |
| iia | iiam | iian | iiang | iiaw | | iia | iiap | iiat | iiak |
| vva | vvam | vvan | vvang | | vvaj | vva | vvap | vvat | vvak |
| uua | uuam | uuan | uuang | | uuaj | uua | uuap | uuat | uuak |

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 3.4 หน่วยเสียงท้ายของสระเสียงสั้นเมื่อยังไม่รวมเสียงวรรณยุกต์

| สระ | เสียงตัวสะกด | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | พยางค์เป็น | | | | | พยางค์ตาย | | | |
| | m | n | ng | w | j | ? | p | t | k |
| i | im | in | ing | iw | | i | ip | it | ik |
| v | vm | vn | vng | | vj | v | vp | vt | vk |
| u | um | un | ung | | uj | u | up | ut | uk |
| e | em | en | eng | ew | | e | ep | et | ek |
| q | qm | qn | qng | | qj | q | qp | qt | qk |
| o | om | on | ong | | | o | op | ot | ok |
| x | xm | xn | xng | xw | | x | xp | xt | xk |
| a | am | an | ang | aw | aj | a | ap | at | ak |
| @ | @m | @n | @ng | | @j | @ | @p | @t | @k |
| ia | iam | ian | iang | iaw | | ia | iap | iat | iak |
| va | vam | van | vang | | vaj | va | vap | vat | vak |
| ua | uam | uan | uang | | uaj | ua | uap | uat | uak |

จากตารางที่ 3.3 และ 3.4 จำนวนหน่วยเสียงท้ายสามารถคำนวณได้ดังตารางที่ 3.5 และเมื่อนำจำนวนหน่วยเสียงเริ่มและหน่วยเสียงท้ายมารวมกัน จะได้จำนวนหน่วยเสียงทั้งหมดดังนี้

จำนวนหน่วยเสียงทั้งหมด = จำนวนหน่วยเสียงเริ่ม+จำนวนหน่วยเสียงท้าย

= 396+782 = 1,178 หน่วยเสียง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 3.5 การคำนวณจำนวนหน่วยเสียงท้าย (ตัวเลขในวงเล็บคือจำนวนเสียงตัวสะกดของสระนั้นๆ คูณกับ จำนวนวรรณยุกต์ที่ผันได้ ซึ่งได้ผลคูณเท่ากับตัวเลขหน้าวงเล็บ)

| สระ | หน่วยเสียงท้าย | |
|---|---|---|
| | พยางค์เป็น | พยางค์ตาย |
| | เสียงตัวสะกด x จำนวนวรรณยุกต์ | เสียงตัวสะกด x จำนวนวรรณยุกต์ |
| ii | 15 (3x5) | 9 (3x3) |
| vv | 15 (3x5) | 9 (3x3) |
| uu | 25 (5x5) | 9 (3x3) |
| ee | 25 (5x5) | 9 (3x3) |
| qq | 30 (6x5) | 9 (3x3) |
| oo | 25 (5x5) | 9 (3x3) |
| xx | 25 (5x5) | 9 (3x3) |
| aa | 30 (6x5) | 9 (3x3) |
| @@ | 25 (5x5) | 9 (3x3) |
| iia | 25 (5x5) | 9 (3x3) |
| vva | 25 (5x5) | 9 (3x3) |
| uua | 25 (5x5) | 9 (3x3) |
| i | 20 (4x5) | 12 (4x3) |
| v | 20 (4x5) | 12 (4x3) |
| u | 20 (4x5) | 12 (4x3) |
| e | 20 (4x5) | 12 (4x3) |
| q | 20 (4x5) | 12 (4x3) |
| o | 15 (3x5) | 12 (4x3) |
| x | 20 (4x5) | 12 (4x3) |
| a | 25 (5x5) | 12 (4x3) |
| @ | 20 (4x5) | 12 (4x3) |
| ia | 20 (4x5) | 12 (4x3) |
| va | 20 (4x5) | 12 (4x3) |
| ua | 20 (4x5) | 12 (4x3) |
| รวม | 530 | 252 |
| รวมทั้งหมด | 782 | |

### 3.3.2 การบันทึกเสียง

รายละเอียดของการบันทึกเสียงสำหรับการสร้างหน่วยเสียงมีรายละเอียดเช่นเดียวกับการบันทึกเสียงในหัวข้อ 3.1.2

### 3.3.3 การวิเคราะห์และประมวลผลข้อมูล

หลังจากการบันทึกเสียง เสียงที่บันทึกจะถูกนำมาประมวลผลเพื่อสร้างเป็นหน่วยเสียงต่อไป โดยการประมวลผลข้อมูลเสียงมีขั้นตอน ดังนี้

1) นำข้อมูลเสียงที่บันทึกเข้าโปรแกรม Praat ซึ่งสามารถวิเคราะห์สัญญาณเสียงและช่วยในการแบ่งส่วน (segment) สัญญาณเสียงได้ง่าย
2) แบ่งส่วนสัญญาณเสียงแบ่งหน่วยเสียงเริ่มและหน่วยเสียงท้าย
3) ตัดเสียงออกเป็นหน่วยเสียงตามตำแหน่งที่ได้แบ่งส่วนไว้ และจัดเก็บหน่วยเสียง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

# บทที่ 4

# วิธีการสังเคราะห์เสียงหนักเบา

ในบทที่ 2 ได้กล่าวถึงลักษณะร่วมทางสัทศาสตร์ของพยางค์เสียงหนักและเบาว่าประกอบด้วย ความยาว ความถี่พิทช์ ความดัง คุณสมบัติของสระและวรรณยุกต์

ฉะนั้นการเปลี่ยนลักษณะทางสัทศาสตร์ของสัญญาณเสียง ทำให้สามารถสังเคราะห์เสียงหนักหรือเบาให้แก่พยางค์ได้ วิทยานิพนธ์นี้จะกล่าวถึงการเปลี่ยนลักษณะทางสัทศาสตร์เพียง 3 ประการ ได้แก่ การปรับช่วงเวลา การปรับความถี่มูลฐาน การปรับแอมพลิจูด ซึ่งสัมพันธ์กับ ความยาวของเสียง ความถี่พิทช์และความดังตามลำดับ

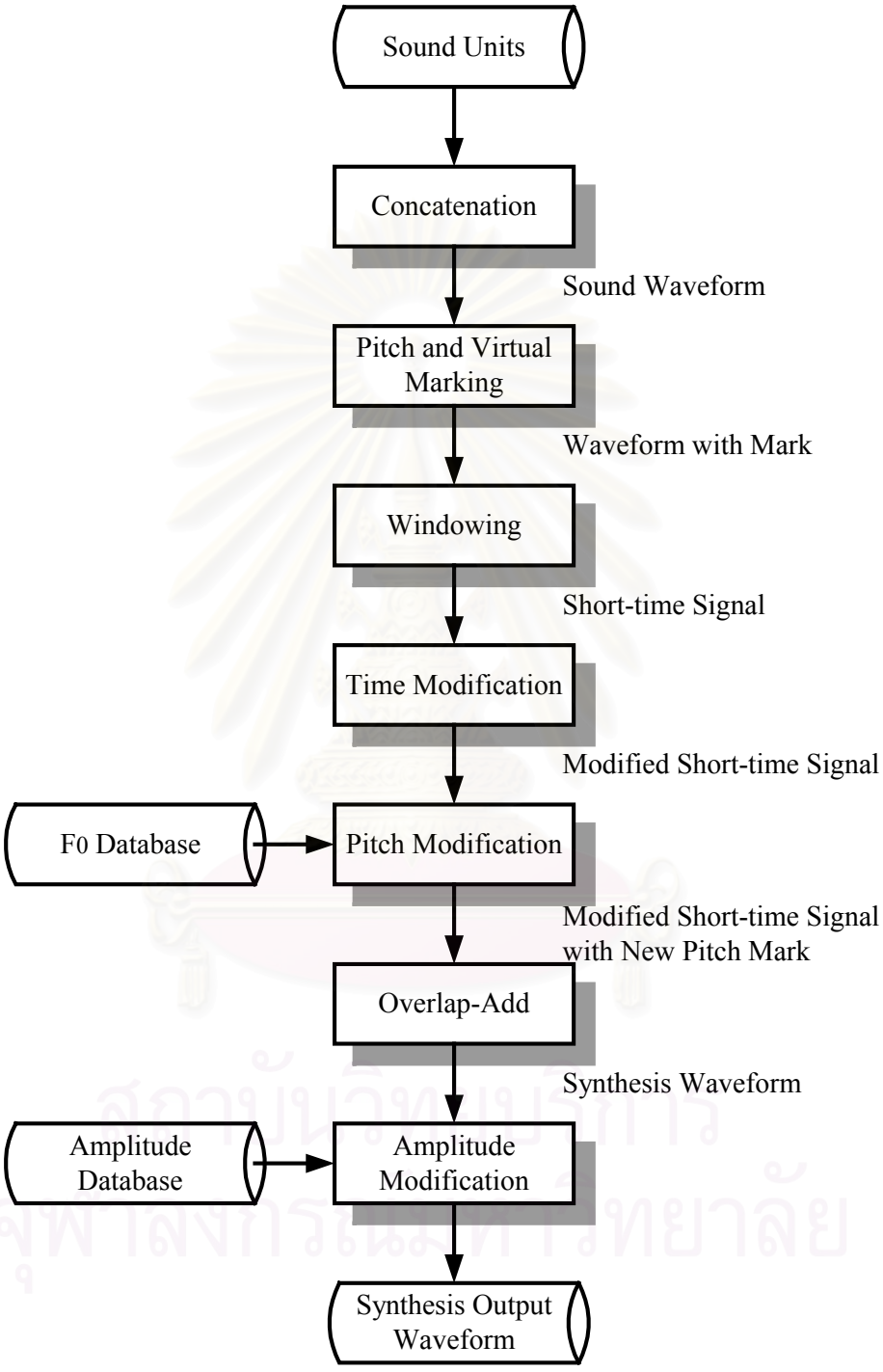
ในการเปลี่ยนช่วงเวลาและความถี่พิทช์นั้น ทำได้โดยใช้วิธี Time-Domain Pitch-Synchronous Overlap-Add (TD-PSOLA, ทีดี-พีเอสโอแอลเอ) (Charpentier and Moulines, 1989) วิธีนี้จะต้องระบุยอดพิทช์ (Pitch Marking) และวิเคราะห์ด้วยฟังก์ชันหน้าต่าง (Windowing) โดยที่กึ่งกลางของหน้าต่างอยู่ที่ตำแหน่งยอดพิทช์ โดยจะวิเคราะห์สัญญาณเสียงออกเป็นสัญญาณช่วงสั้น การเพิ่มหรือลดจำนวนสัญญาณช่วงสั้นจะเปลี่ยนแปลงช่วงเวลาสัญญาณเสียง และการเพิ่มหรือลดระยะห่างระหว่างพิทช์จะเปลี่ยนแปลงความถี่มูลฐาน จากนั้นจึงรวมสัญญาณช่วงสั้นที่ถูกปรับปรุงทั้งหมดกลับมาใหม่โดยการบวกซ้อน (Overlap-Add)

ด้วยรายละเอียดที่กล่าวมาข้างต้น บทนี้จะได้นำเสนอรายละเอียด ขั้นตอนและวิธีการสำหรับการสังเคราะห์พยางค์หนักเบา

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

## 4.1 ระบบการสังเคราะห์เสียงหนักเบาโดยรวม

ระบบการสังเคราะห์เสียงหนักเบาโดยรวม มีวิธีการและขั้นตอนดังรูปที่ 4.1




รูปที่ 4.1 ระบบสังเคราะห์พยางค์เสียงหนักและเบาโดยรวม

ระบบสังเคราะห์เสียงจะเริ่มจากการเชื่อมต่อหน่วยเสียงเป็นสัญญาณเสียง เพื่อนำมาปรับลักษณะทางสัทศาสตร์ต่างๆ โดยก่อนการปรับลักษณะทางสัทศาสตร์จะต้องทำการระบุพิทช์และพิทช์เสมือนเพื่อวิเคราะห์ด้วยฟังก์ชันหน้าต่างออกเป็นสัญญาณช่วงสั้น

ในกรณีที่พยางค์ที่ต้องการสังเคราะห์เป็นพยางค์เสียงเบาสัญญาณช่วงสั้นจะถูกเพิ่มหรือลบออกเพื่อปรับช่วงเวลาของพยางค์ และปรับระยะระหว่างพิทช์เพื่อปรับความถี่มูลฐานตามฐานข้อมูลความถี่มูลฐานที่ได้สร้างไว้ จากนั้นจึงนำสัญญาณช่วงสั้นและพิทช์ที่ผ่านการปรับช่วงเวลาและการปรับความถี่มูลฐานมาบวกซ้อนกลับเป็นสัญญาณเสียง และนำสัญญาณเสียงที่ได้มาปรับแอมพลิจูดตามฐานข้อมูลอัตราส่วนแอมพลิจูดระหว่างพยางค์เสียงหนักและพยางค์เสียงเบาได้เป็นสัญญาณเสียงสังเคราะห์พยางค์เสียงเบา

ในกรณีที่ต้องการสังเคราะห์เป็นพยางค์เสียงหนักจะปรับลักษณะทางสัทศาสตร์เพียงการปรับช่วงเวลา และนำมาบวกซ้อนออกเป็นสัญญาณเสียง

## 4.2 **การเชื่อมต่อหน่วยเสียง** (Concatenation)

ระบบสังเคราะห์เสียงนี้ใช้หน่วยเสียงแบบอัฒพยางค์ที่ใช้การเชื่อมต่อหน่วยเสียงโดยเชื่อมต่อหน่วยเสียงเริ่มกับหน่วยเสียงท้ายเป็นสัญญาณเสียง 1 พยางค์

## 4.3 **การระบุพิทช์** (Pitch marking)

การระบุยอดพิทช์จะแบ่งออกเป็นสองส่วน ส่วนแรกเป็นการหายอดพิทช์แรก และส่วนที่สองกล่าวถึงการหายอดพิทช์ต่อๆไปโดยการใช้วิธีการหาค่าอัตสหสัมพันธ์ โดยมีรายละเอียดดังนี้

### 4.3.1 **การหายอดพิทช์แรก**

การหายอดพิทช์แรกจะทำเฉพาะส่วนที่เป็นเสียงก้องโดยการนำสัญญาณเสียงก้องมาผ่านการขลิบกลาง (Central clipping) โดยที่ Deller et al. (1993) กำหนดการขริบกลางไว้ที่ 30 เปอร์เซ็นต์ของค่าสูงสุดของค่าสัมบูรณ์ของสัญญาณเสียง แต่เมื่อนำมาใช้ในวิทยานิพนธ์นี้พบว่า การขริบกลางที่ 30 เปอร์เซ็นต์ของค่าสูงสุดของค่าสัมบูรณ์ของสัญญาณเสียง จะทำให้เสียงนาสิกได้แก่ เสียง /m/, เสียง /n/ และ เสียง /ng/ ถูกขริบหายไป ไม่สามารถระบุยอดพิทช์ในช่วงเสียงนาสิกได้ เสียงสังเคราะห์ที่ได้จะรับฟังเพี้ยนไป ดังนั้นในวิทยานิพนธ์นี้จึงได้เลือกขริบสัญญาณที่ 10 เปอร์เซ็นต์ของค่าสูงสุดของค่าสัมบูรณ์ของสัญญาณเสียง หลังจากขริบสัญญาณเสียงแล้วกำหนดตำแหน่งยอดพิทช์แรกจากยอดพิทช์ที่คงอยู่หลังจากการขริบสัญญาณ โดยมีรายละเอียดวิธีหายอดพิทช์แรกดังนี้

(1) หาจุดเริ่มต้นของเสียงก้องด้วยการกรองสัญญาณเสียงที่ต้องการระบุยอดพิทช์ด้วยตัวกรองสัญญาณที่มีความถี่ตัด (Cutoff frequency) เท่ากับ 900 เฮิรตซ์ เสียงไม่ก้องส่วนใหญ่จะถูกกรองออกไป

(2) แบ่งสัญญาณเสียงที่ผ่านการกรองเป็นกรอบเสียงพูด แต่ละกรอบกว้าง 480 ตัวอย่าง สำหรับสัญญาณเสียงที่มีอัตราสุ่ม 16 กิโลเฮิรตซ์ หรือ 30 มิลลิวินาที โดยแต่ละกรอบเสียงพูดซ้อนทับกัน $3/4$ ของความกว้างกรอบเสียงพูด หาค่าพลังงานกำลังสอง (Square Energy) ของแต่ละกรอบเสียงพูด ตามสมการที่ 4.1

$$E_n = \sum_{i=1}^{N} S_n^2(i) \qquad (4.1)$$

โดยที่ $E_n$ แทน พลังงานกำลังสองของสัญญาณเสียงในกรอบเสียงพูดที่ n

$S_n(i)$ แทน สัญญาณเสียงในกรอบเสียงพูดที่ n

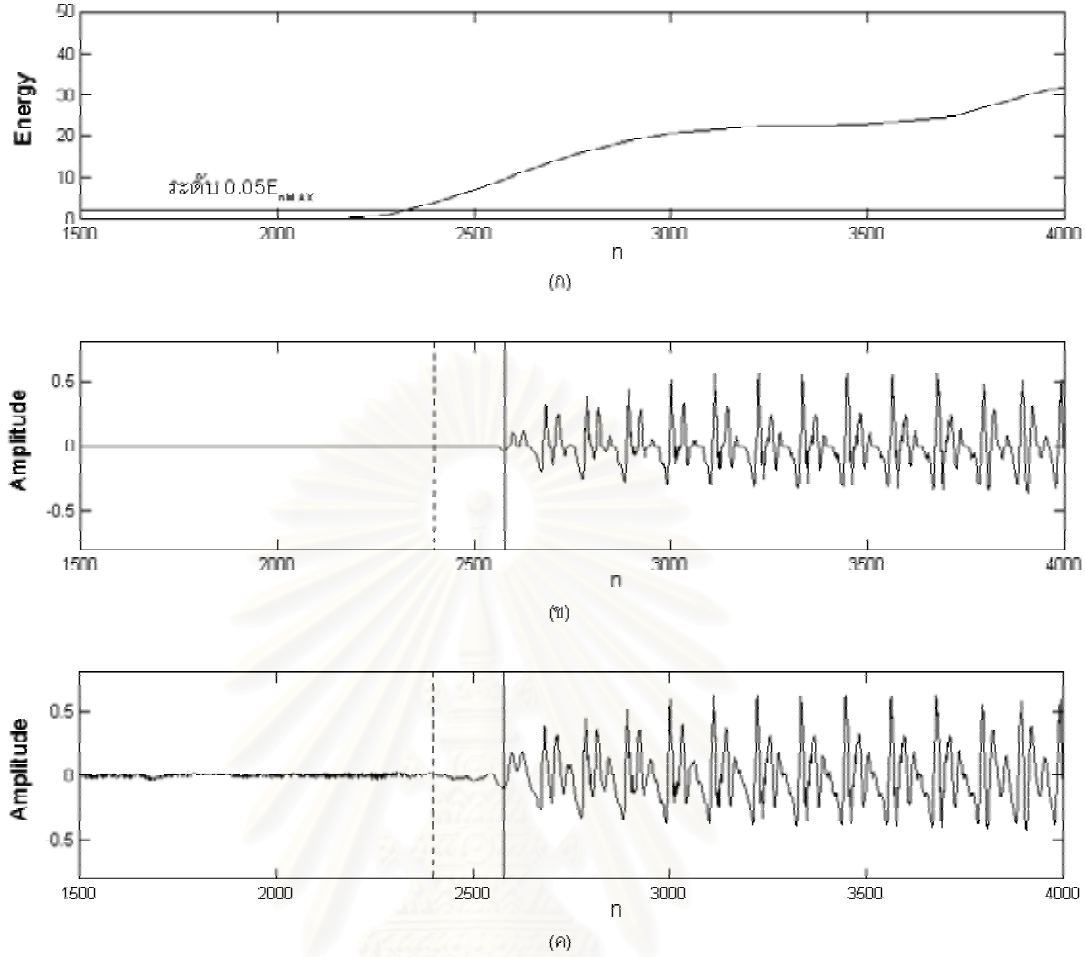
N แทน ความกว้างของกรอบเสียงพูด

(3) เลือกค่าพลังงานกำลังสองของกรอบที่มีพลังงานกำลังสองสูงที่สุด ($E_{nMAX}$) ใช้ค่า $0.05E_{nMAX}$ เป็นระดับกำหนดจุดเริ่มต้นเสียงก้อง โดยตำแหน่งกรอบสัญญาณแรกที่มีพลังงานกำลังสองมากกว่าค่า $0.05E_{nMAX}$ เป็นจุดเริ่มต้นเสียงก้อง

(4) ใช้การทำ Central clipping สัญญาณเสียงที่ 10 เปอร์เซ็นต์ของค่าสูงสุดสัมบูรณ์ ที่เลือกค่านี้เพราะเสียงตัวสะกดเสียงนาสิก ได้แก่ เสียง /m/, /n/ และ /ng/ มีขนาดสัญญาณเล็ก หากขลิบสัญญาณเสียงที่ 30 เปอร์เซ็นต์ ตัวสะกดเสียงนาสิกนี้จะถูกขลิบออกไป

(5) หายอดพิทช์ที่คงเหลืออยู่จากการทำ Central clipping ที่อยู่หลังจากจุดเริ่มต้นเสียงก้องเป็นยอดพิทช์ยอดแรกดังรูปที่ 4.2

จุฬาลงกรณ์มหาวิทยาลัย




รูปที่ 4.2 การเลือกยอดพิทช์แรก (ก) พลังงานกำลังสองของสัญญาณเสียงคำว่า "โซ่" ที่ผ่านการกรองความถี่ 900 เฮิรตซ์ และระดับ $0.05E_{nMAX}$ (ข) สัญญาณเสียงคำว่า "โซ่" หลังจากทำ Central clipping (ค) สัญญาณเสียงคำว่า "โซ่" เส้นประแนวดิ่งคือจุดแบ่งเสียงก้อง เส้นตรงแนวดิ่งคือตำแหน่งยอดพิทช์ยอดแรก

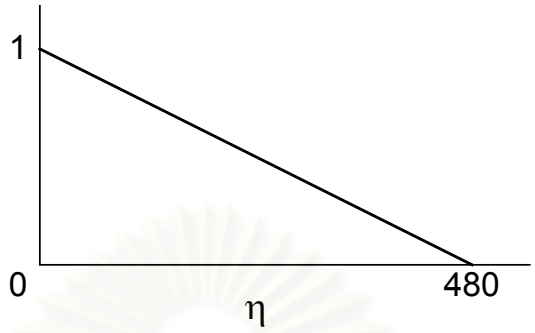
### 4.3.2 การหายอดพิทช์ต่อๆไป

การหายอดพิทช์ต่อๆไปจะใช้การทำนายด้วยการหาค่าอัตสหสัมพันธ์และการขลิบกลาง แล้วนำจุดยอดพิทช์ที่ทำนายได้มาหายอดพิทช์จริง โดยมีรายละเอียดดังนี้

(1) หาค่าอัตสหสัมพันธ์ของสัญญาณเสียงที่ผ่านการทำ Central clipping ที่ 10 เปอร์เซ็นต์ของค่าสูงสุดของค่าสัมบูรณ์ของสัญญาณเสียง โดยใช้ยอดพิทช์แรก ($M_1$) เป็นจุดอ้างอิง และเวลาหน่วง $\eta$ ตั้งแต่ 50 ถึง 320 ตัวอย่าง

(2) นำค่าอัตสหสัมพันธ์ที่ได้คูณกับฟังก์ชันน้ำหนัก (Dubnowski et al., 1976) ที่มีลักษณะเป็นเส้นตรงลาดเอียง เริ่มด้วยค่าเท่ากับ 1 ที่เวลาหน่วง $\eta = 0$ และ เท่ากับ 0 ที่เวลาหน่วง

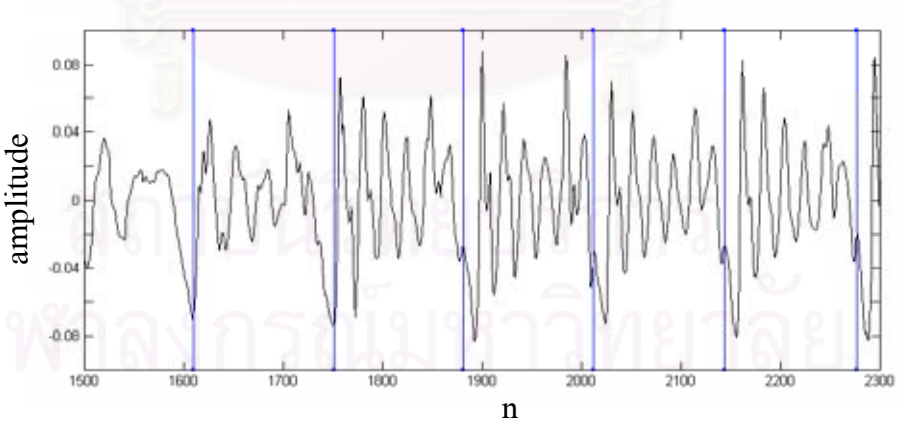
$\eta$ = 480 ตัวอย่าง ดังรูปที่ 4.3 ซึ่งเท่ากับ ระยะเวลา 30 มิลลิวินาที ที่อัตราสุ่ม 16 กิโลเฮิรตซ์




รูปที่ 4.3 ฟังก์ชันน้ำหนักของค่าอัตสหสัมพันธ์

(3) เลือกค่า $\eta$ ที่ให้ค่าอัตสหสัมพันธ์มากที่สุด ($\eta_{MAX}$) ตำแหน่งที่ห่างจากยอดพิทช์แรกเป็นระยะ $\eta_{MAX}$ จะเป็นตำแหน่งยอดพิทช์จากการทำนายซึ่งมีค่าเท่ากับ $M1+\eta_{MAX}$

(4) เนื่องจากตำแหน่งยอดพิทช์จากการทำนายอาจไม่ตรงกับตำแหน่งยอดพิทช์จริงหากใช้เป็นตำแหน่งยอดพิทช์ (pitch mark) และทำซ้ำวิธีดังกล่าวข้างต้น ยอดพิทช์ที่ได้จะผิดพลาดจากตำแหน่งยอดพิทช์จริงไปมากได้ การหาค่าอัตสหสัมพันธ์จึงไม่สามารถทำนายหาตำแหน่งยอดพิทช์ต่อไปได้ถูกต้องดังแสดงในรูปที่ 4.4




รูปที่ 4.4 ความผิดพลาดในการหายอดพิทช์โดยการหาค่าอัตสหสัมพันธ์ เมื่อทำซ้ำต่อไปหลายครั้ง (เส้นตรงในแนวดิ่งแทนตำแหน่งยอดพิทช์ที่ใช้วิธีการหาค่าอัตสหสัมพันธ์ทำนาย)

ฉะนั้นจะแก้ไขโดยพิจารณาจุดโดยรอบตำแหน่งยอดพิทช์จากการทำนาย $M1+\eta_{MAX}$ ในช่วง $M1+0.8(\eta_{MAX})$ ถึง $M1+1.2(\eta_{MAX})$ หาตำแหน่งสัญญาณเสียงที่มีแอมพลิจูดน้อยที่

สุดเป็นยอดพิทช์ (เนื่องจากในวิทยานิพนธ์นี้ใช้ยอดพิทช์ทางด้านลบจึงใช้ค่าน้อยที่สุด ถ้าต้องการใช้ยอดพิทช์ทางด้านบวกจะหาตำแหน่งสัญญาณเสียงที่มีค่ามากที่สุด)

(5) หายอดพิทช์ถัดไปโดยใช้ยอดพิทช์ล่าสุดเป็นจุดอ้างอิงในการทำนายด้วยวิธีข้อ (1) – (4)

### 4.4 **การระบุพิทช์เสมือน** (Virtual pitch marking)

การระบุยอดพิทช์เสมือนในเสียงไม่ก้อง ทำได้โดยแบ่งเสียงไม่ก้องเป็นส่วนที่มีระยะเท่ากัน ระยะระหว่างยอดพิทช์เสมือนจะใกล้เคียงกับค่าเฉลี่ยระยะห่างระหว่างยอดพิทช์ของเสียงก้อง นั่นคือ

$$D_m = \frac{M_N - M_1}{N-1} \tag{4.2}$$

โดยที่ $D_m$ แทน ค่าเฉลี่ยระยะห่างระหว่างยอดพิทช์ของเสียงก้อง

$M_N$ แทน ตำแหน่งของยอดพิทช์ที่ N

$M_1$ แทน ตำแหน่งของยอดพิทช์ที่ 1

N แทน จำนวนยอดพิทช์ทั้งหมด

เนื่องจากยอดพิทช์เสมือนจะต้องห่างเท่ากัน ค่าเฉลี่ยระยะห่างระหว่างยอดพิทช์ของเสียงก้อง ($D_m$) อาจแบ่งช่วงเวลาเสียงไม่ก้องเป็นช่วงเวลาเท่ากันไม่ได้ จึงใช้ระยะห่างระหว่างยอดพิทช์แบบเสมือน ดังสมการที่ 4.3

$$D_v = \frac{U}{round\left(\frac{U}{D_m}\right)} \tag{4.3}$$

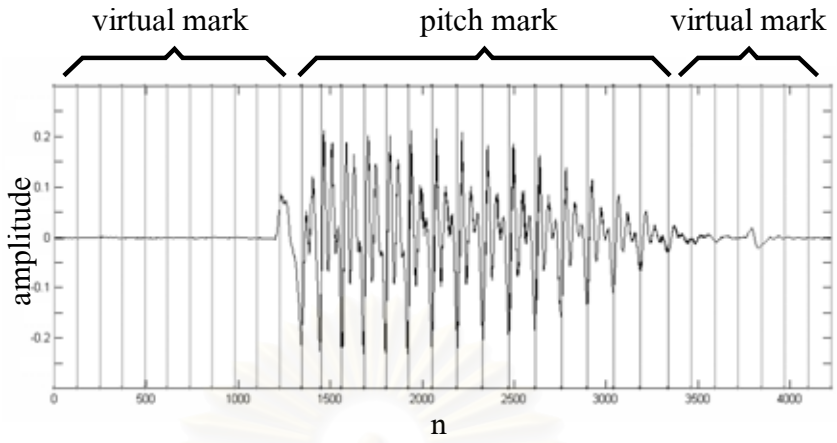
โดยที่ $D_v$ แทน ระยะห่างระหว่างยอดพิทช์เสมือน

U แทน ระยะของช่วงที่เป็นเสียงไม่ก้อง

Round () แทน ฟังก์ชันการประมาณค่า

เมื่อได้ค่าระยะระหว่างยอดพิทช์ในเสียงก้องและระยะระหว่างยอดพิทช์เสมือนจะสามารถระบุยอดพิทช์ทั้งหมดของสัญญาณเสียง 1 พยางค์ ได้ดังรูปที่ 4.5

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย




รูปที่ 4.5 ยอดพิทช์ล่างและยอดพิทช์เสมือนในสัญญาณเสียง

## 4.5 การวิเคราะห์ด้วยฟังก์ชันหน้าต่าง

ฟังก์ชันหน้าต่างใช้กำหนดกรอบสัญญาณเป็นช่วงสั้น เมื่อคูณสัญญาณเสียงด้วยฟังก์ชันหน้าต่างจะได้สัญญาณช่วงสั้นที่สามารถนำไปวิเคราะห์ต่อไปได้

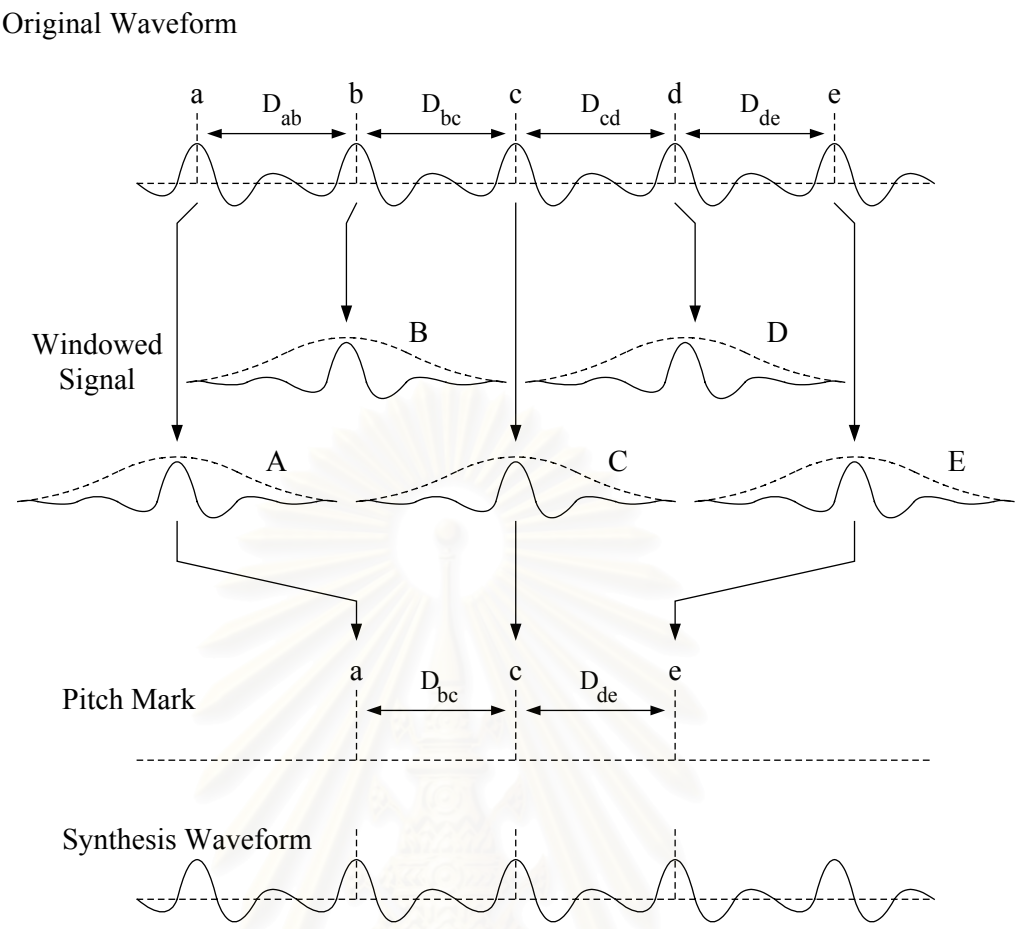
หน้าต่างที่นิยมใช้สำหรับการทำ Pitch-Synchronous Overlap-Add คือ หน้าต่างแฮนนิง (Charpentier, F. and Moulines, E., 1989) โดยขนาดของหน้าต่างในวิทยานิพนธ์นี้จะเลือกขนาดให้เท่ากับ 2 เท่าคาบเวลาของสัญญาณเสียงในแต่ละยอดพิทช์ เนื่องจากการใช้ขนาดของหน้าต่างที่ใหญ่เกินไปทำให้เสียงสังเคราะห์มีขนาดใหญ่ขึ้น และมีเสียงซ่าแทรกเข้ามาในสัญญาณเสียง

ฟังก์ชันหน้าต่างจะคูณกับสัญญาณโดยให้กึ่งกลางของฟังก์ชันหน้าต่างจะอยู่ที่ตำแหน่งยอดพิทช์หรือยอดพิทช์เสมือนทุกยอด สัญญาณเสียงจะถูกแบ่งออกเป็นสัญญาณช่วงสั้น พร้อมสำหรับการปรับช่วงเวลา การปรับความถี่มูลฐานและอื่นๆในขั้นตอนต่อไป

## 4.6 การปรับช่วงเวลา

### 4.6.1 การลดช่วงเวลา

การลดช่วงเวลาทำได้โดยตัดสัญญาณช่วงสั้นบางส่วนออก ดังแสดงในรูปที่ 4.6

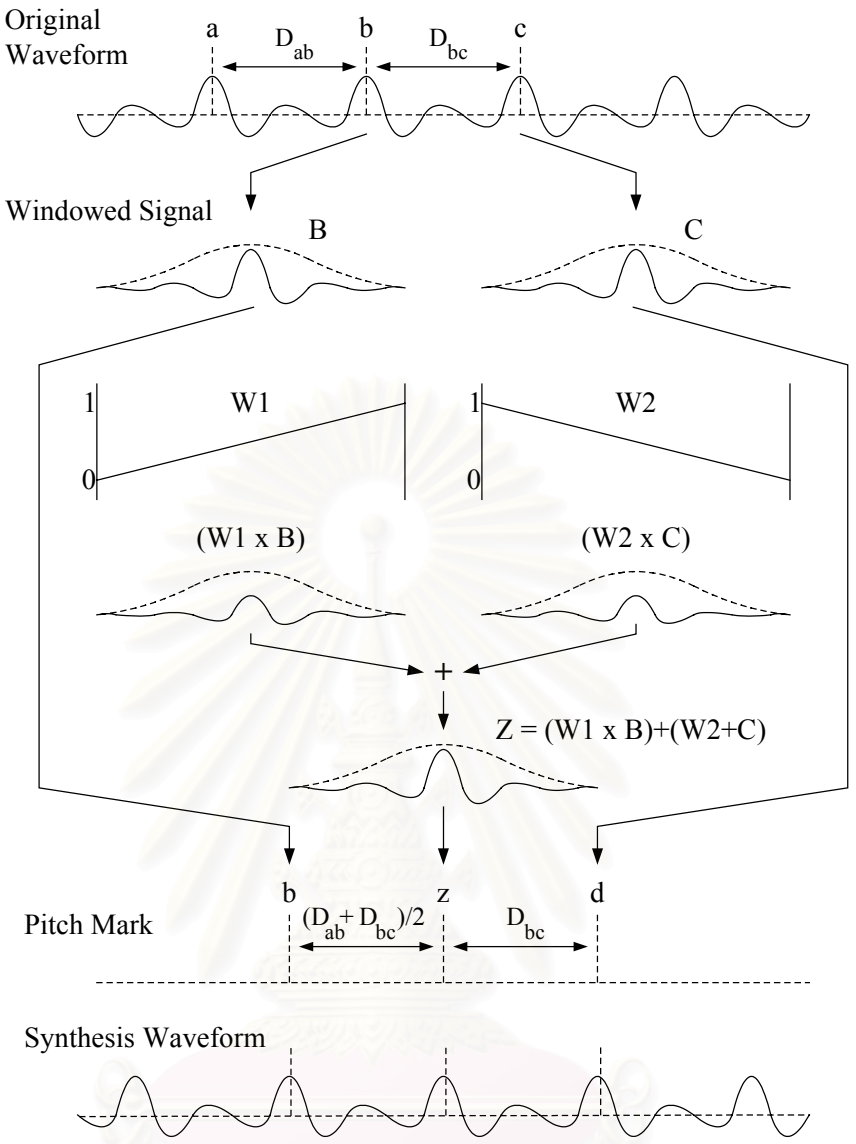



รูปที่ 4.6 การลดช่วงเวลา

จากรูปที่ 4.6 สัญญาณเสียงถูกแยกด้วยฟังก์ชันหน้าต่างที่ตำแหน่งยอดพิทช์ a, b, c, d และ e ได้สัญญาณช่วงสั้น A, B, C, D และ E ตามลำดับ โดยมีระยะห่างระหว่างยอดพิทช์ a กับ b, b กับ c, c กับ d และ d กับ e เป็น $D_{ab}$, $D_{bc}$, $D_{cd}$ และ $D_{de}$ ตามลำดับ จะลดช่วงเวลาโดยตัดสัญญาณช่วงสั้น B และ D ออก (โดยตัด $D_{ab}$ และ $D_{cd}$ ออก) ระยะห่างระหว่างยอดพิทช์ a และ c จะเท่ากับ $D_{bc}$ และระยะห่างระหว่างยอดพิทช์ c และ d จะเท่ากับ $D_{de}$ แล้วรวมสัญญาณช่วงสั้นกลับมาใหม่โดยการบวกซ้อน (Overlap-Add) ที่ตำแหน่งยอดพิทช์

### 4.6.2 การเพิ่มช่วงเวลา

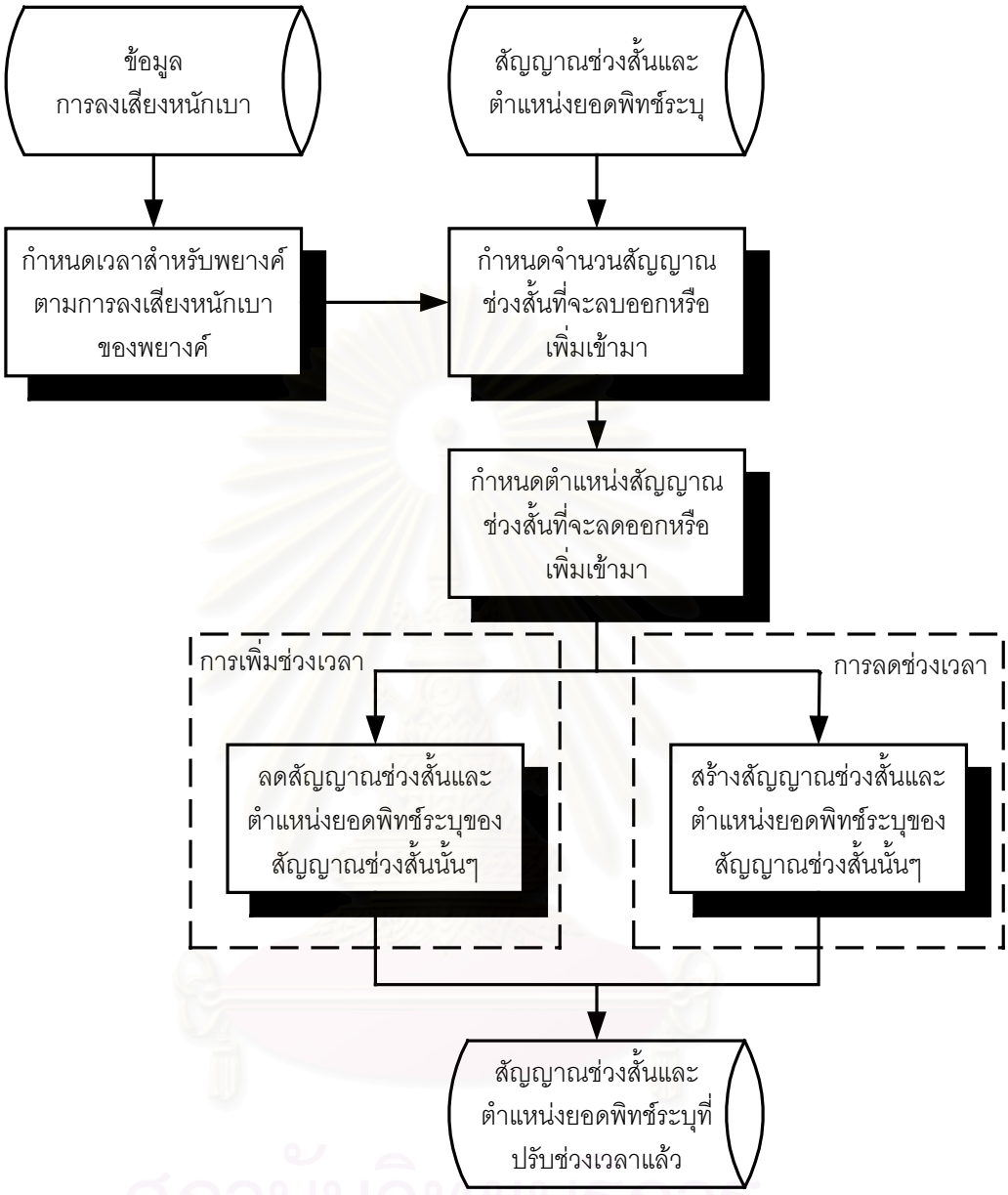
การเพิ่มช่วงเวลาทำได้โดยแทรกสัญญาณช่วงสั้นเพิ่มในสัญญาณเดิม โดยสัญญาณที่เพิ่มใช้วิธีการสร้างขึ้นโดยวิธีเปลี่ยนความเร็วเสียงพูด (อภิชาติ ตั้งทางธรรม, 2537) ซึ่งจะช่วยให้สัญญาณที่ได้มีความต่อเนื่องมากขึ้น นอกจากจะสร้างสัญญาณเพิ่มแล้วยังต้องสร้างยอดพิทช์ขึ้นใหม่ดังรูปที่ 4.7




รูปที่ 4.7 การเพิ่มช่วงเวลา

จากรูปที่ 4.7 ตำแหน่งยอดพิทช์ a, b และ c มีระยะห่างระหว่างยอดพิทช์ a กับ b และ b กับ c เท่ากับ $D_{ab}$ และ $D_{bc}$ ตามลำดับ สำหรับรูปที่ 4.7 จะแสดงการเพิ่มสัญญาณช่วงสั้น Z ระหว่างสัญญาณช่วงสั้น B และ C จะเพิ่มเวลาโดยแทรกสัญญาณช่วงสั้น Z ซึ่งสร้างจากสัญญาณช่วงสั้น B และ C โดยคูณสัญญาณช่วงสั้น B และ C ด้วยตัวน้ำหนักดังรูปที่ 4.7 ได้เป็น (W1xB) และ (W2xC) นำสัญญาณทั้งสองมาบวกกัน จะได้สัญญาณช่วงสั้น Z และสร้างตำแหน่งยอดพิทช์ z โดยมีระยะห่างระหว่างยอดพิทช์ก่อนหน้า (ยอดพิทช์ b) กับยอดพิทช์ z เท่ากับค่าเฉลี่ยของ $D_{ab}$ และ $D_{bc}$ การเพิ่มสัญญาณช่วงสั้นดังกล่าวนี้จะเพิ่มช่วงเวลาของสัญญาณเสียงได้

### 4.6.3 ขั้นตอนการเพิ่มและลบสัญญาณตัวอย่างสำหรับการเปลี่ยนช่วงเวลา




รูปที่ 4.8 แผนผังของการปรับช่วงเวลา

ขั้นตอนในการปรับช่วงเวลามีดังนี้

(1) กำหนดเวลาสำหรับพยางค์ตามการลงเสียงหนักเบาของพยางค์

ข้อมูลการลงเสียงหนักเบาของแต่ละพยางค์ใน 1 ประโยค จะถูกนำมาวิเคราะห์เป็นหน่วยจังหวะ ตามตัวอย่างที่แสดงในภาคผนวก ก. นำหน่วยจังหวะที่แบ่งได้มาเทียบหาค่าเวลาสำหรับแต่ละพยางค์ในหน่วยจังหวะ (Luangthongkum, 1977) ดังนี้

หน่วยจังหวะ 1 พยางค์หนัก

$|S|$ = 0.30 วินาที

หน่วยจังหวะ 1 พยางค์หนัก 1 พยางค์เบา

$|S\ U|$ = 0.40 วินาที หรือ $|0.25 : 0.15|$ วินาที

หน่วยจังหวะ 1 พยางค์หนัก 2 พยางค์เบา

$|S\ U\ U|$ = 0.49 วินาที หรือ $|0.23 : 0.13 : 0.13|$ วินาที

หน่วยจังหวะ 1 พยางค์หนัก 3 พยางค์เบา

$|S\ U\ U\ U|$ = 0.59 วินาที หรือ $|0.21 : 0.12 : 0.13 : 0.13|$ วินาที

(2) กำหนดจำนวนสัญญาณช่วงสั้นที่จะลดออกหรือเพิ่มเข้ามา

นำยอดพิทช์ (Pitch mark) และ ยอดพิทช์เสมือน (Virtual mark) ซึ่งมีจำนวนเท่ากับจำนวนสัญญาณช่วงสั้นทั้งหมด ($N_m$) มาหาค่าเฉลี่ยระยะห่างระหว่างยอดพิทช์ ($D_m$) ตามสมการ 4.2

นำค่าเฉลี่ยที่ได้มาคำนวณจำนวนสัญญาณช่วงสั้นที่ต้องการ ($N_r$) ตามค่าระยะเวลาของสัญญาณเสียงที่ต้องการ ($T_r$) ตามสมการที่ 4.4 และจำนวนสัญญาณช่วงสั้นที่จะลดหรือเพิ่ม ($N_t$) ตามสมการที่ 4.5

$$N_r = \frac{T_r}{D_m} \tag{4.4}$$

$$N_t = N_m - N_r \tag{4.5}$$

ถ้า $N_t$ เป็นลบ แสดงว่า จำนวนสัญญาณช่วงสั้นที่ต้องการมีค่ามากกว่าจำนวนสัญญาณช่วงสั้นทั้งหมดที่มีอยู่เดิม การปรับช่วงเวลาจะเป็นการเพิ่มสัญญาณช่วงสั้นจำนวนเท่ากับค่าสัมบูรณ์ของ $N_t$ ในทางกลับกันถ้าค่า $N_t$ เป็นจำนวนบวก การปรับช่วงเวลาจะเป็นการลดสัญญาณช่วงสั้นเท่ากับ $N_t$

(3) กำหนดตำแหน่งสัญญาณช่วงสั้นที่จะลดออกหรือเพิ่มเข้ามา

เนื่องจากสัญญาณช่วงสั้นมีทั้งส่วนที่เป็นสัญญาณช่วงสั้นของเสียงก้องและเสียงไม่ก้อง การเพิ่มหรือลดสัญญาณช่วงสั้นจะต้องแบ่งจำนวนสัญญาณช่วงสั้นที่จะเพิ่มหรือลบ

ตามอัตราส่วนของจำนวนสัญญาณช่วงสั้นที่เป็นเสียงก้องและสัญญาณช่วงสั้นที่เป็นเสียงไม่ก้อง

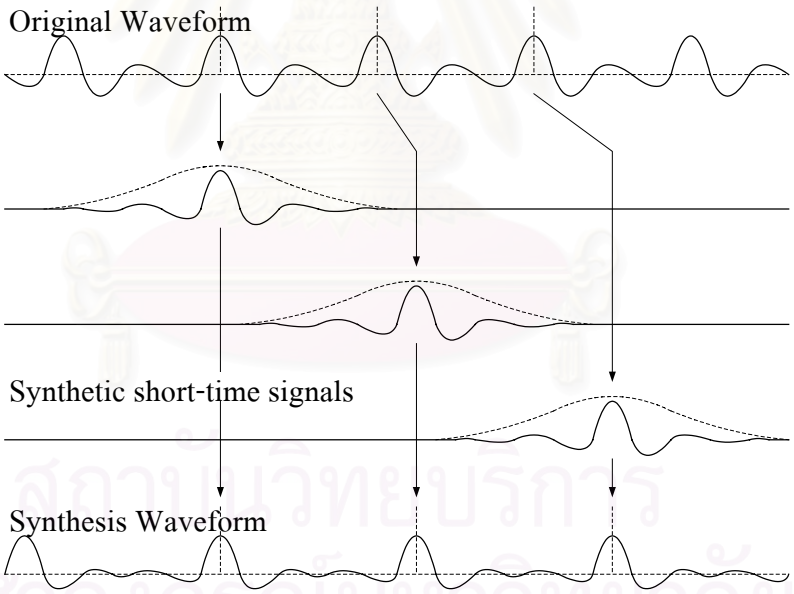
ตำแหน่งของสัญญาณช่วงสั้นสำหรับเพิ่มหรือลดสัญญาณออกจะเลือกแต่ละตำแหน่งให้กระจายออกมีระยะห่างเท่าๆกัน

(4) ลบหรือเพิ่มสัญญาณช่วงสั้นและตำแหน่งยอดพิทช์ของสัญญาณช่วงสั้นนั้น

การลบสัญญาณช่วงสั้นและตำแหน่งยอดพิทช์ทำได้โดยวิธีการที่ได้กล่าวไว้ในหัวข้อ 4.6.1

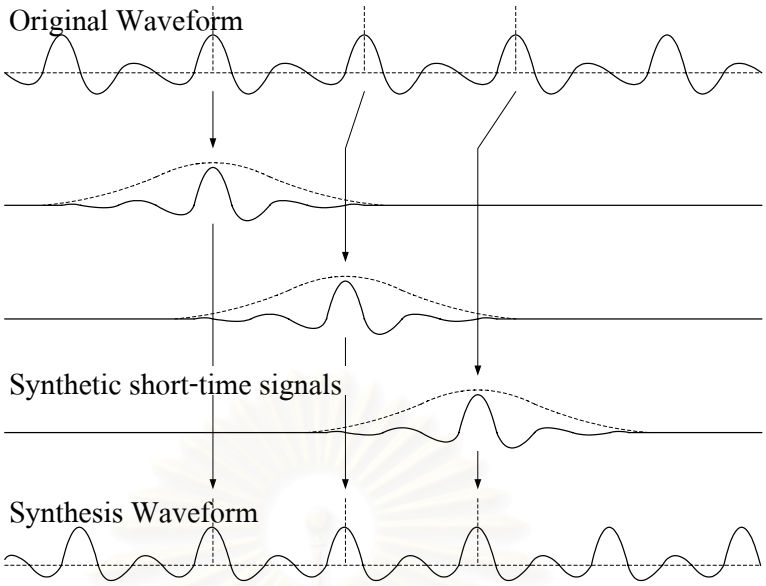
## 4.7 การปรับความถี่มูลฐาน

การปรับความถี่มูลฐานของสัญญาณเสียงเป็นการเปลี่ยนระดับเสียงสูงต่ำที่ทำได้โดยการเปลี่ยนระยะเวลาระหว่างยอดพิทช์ เมื่อระยะเวลานี้เพิ่มขึ้นความถี่จะลดลง ทำให้ระดับเสียงต่ำลง แต่เมื่อระยะเวลานี้น้อยลงความถี่จะเพิ่มขึ้น ทำให้ระดับเสียงสูงขึ้น ดังแสดงในรูปที่ 4.9




(ก)

รูปที่ 4.9 การเปลี่ยนระดับเสียงสูงต่ำของสัญญาณเสียงโดยใช้วิธี TD-PSOLA
(ก) การลดความถี่มูลฐาน (ข) การเพิ่มความถี่มูลฐาน



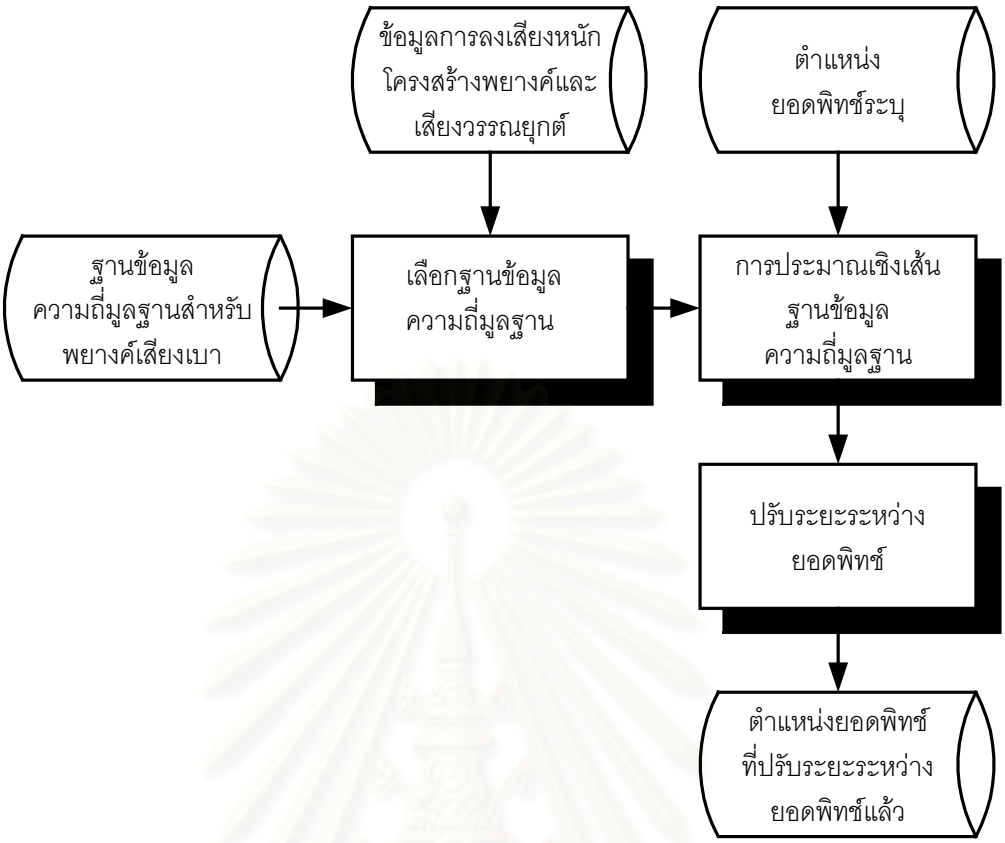

(ข)

รูปที่ 4.9 การเปลี่ยนระดับเสียงสูงต่ำของสัญญาณเสียงโดยใช้วิธี TD-PSOLA (ต่อ)

**ขั้นตอนการปรับความถี่มูลฐาน**

การปรับความถี่มูลฐานของสัญญาณเสียงจะปรับเฉพาะการสังเคราะห์พยางค์เสียงเบา ส่วนการสังเคราะห์พยางค์เสียงหนักไม่ต้องปรับเพราะหน่วยเสียงสำหรับการสังเคราะห์เสียงสร้างจากพยางค์เสียงหนัก การปรับความถี่มูลฐานเป็นการเพิ่มหรือลดระยะระหว่างยอดพิทช์ให้ได้รูปแบบความถี่มูลฐานตามต้องการ ซึ่งมีขั้นตอนดังรูปที่ 4.10

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย




รูปที่ 4.10 แผนผังการปรับความถี่มูลฐาน

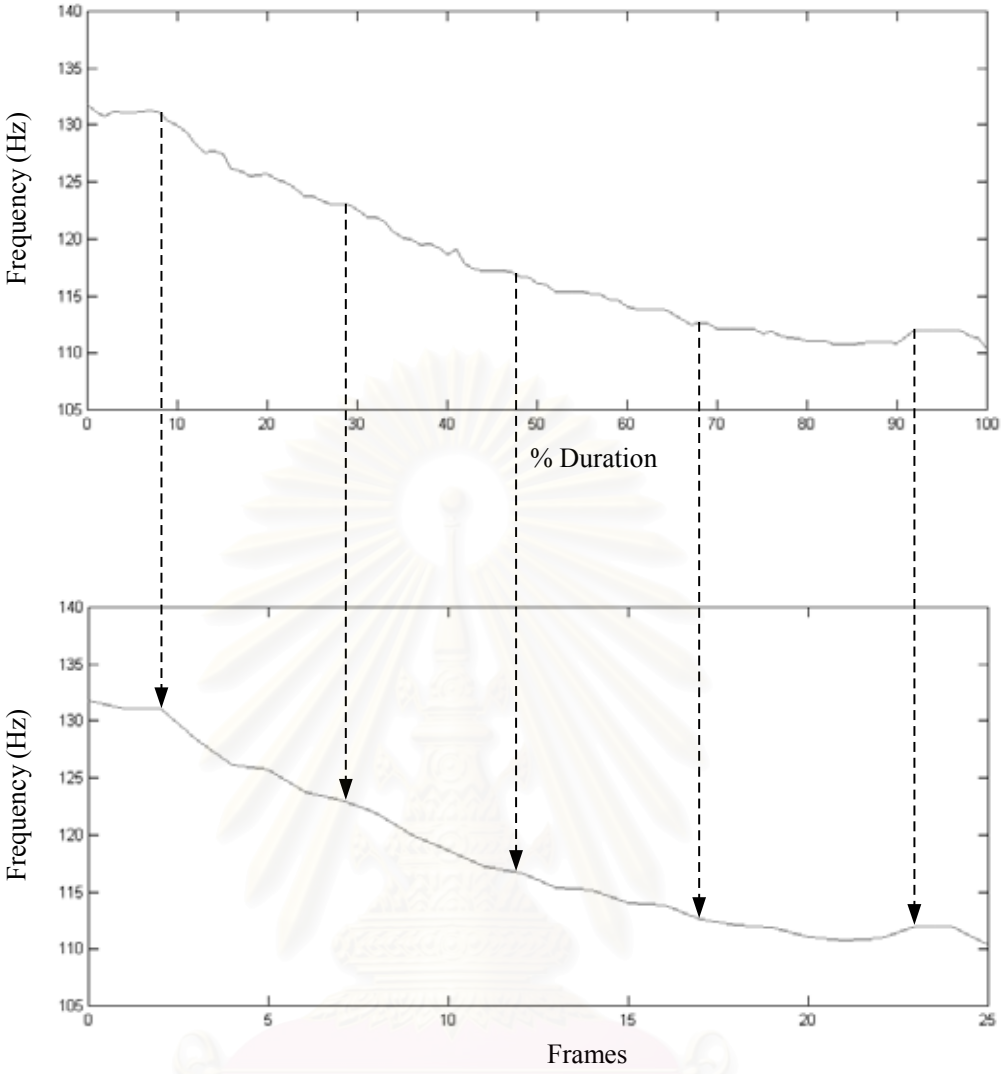
ขั้นตอนการปรับความถี่มูลฐาน มีดังนี้

(1) เลือกฐานข้อมูลความถี่มูลฐาน

เลือกฐานข้อมูลความถี่มูลฐานสำหรับพยางค์เสียงเบาตามโครงสร้างพยางค์และเสียงวรรณยุกต์ของพยางค์นั้น จากฐานข้อมูลความถี่มูลฐานสำหรับพยางค์เสียงเบา 14 แบบ แบ่งตามโครงสร้างพยางค์และเสียงวรรณยุกต์ที่มีรูปแบบการเก็บข้อมูลไว้จำนวน 100 ค่าตามร้อยละของช่วงเวลาเสียงก้องของพยางค์

(2) การประมาณเชิงเส้นฐานข้อมูลความถี่มูลฐาน

รูปแบบความถี่มูลฐานของเสียงวรรณยุกต์ที่เป็นเสียงเบาจะถูกนำมาประมาณค่าเชิงเส้นตามจำนวนยอดพิทช์เพื่อหาค่าพิทช์สำหรับปรับระยะระหว่างพิทช์ของยอดพิทช์แต่ละยอดพิทช์ดังรูปที่ 4.11




รูปที่ 4.11 การประมาณค่ารูปแบบความถี่มูลฐาน

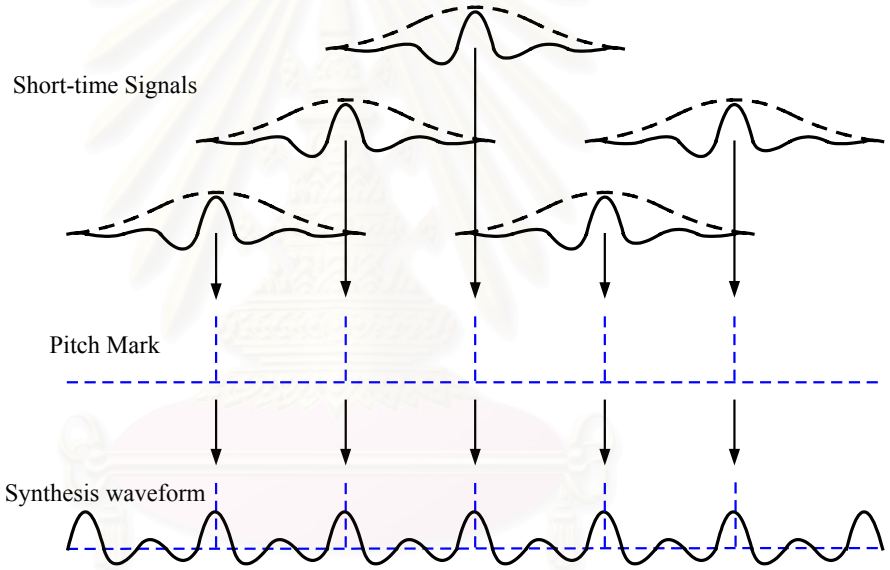
จากรูปที่ 4.11 ยอดพิทช์สัญญาณช่วงสั้นจำนวน 25 สัญญาณช่วงสั้นจะถูกปรับระยะห่างระหว่างยอดพิทช์เพื่อปรับความถี่มูลฐาน สัญญาณช่วงสั้นจะถูกนำมาหาความถี่มูลฐานสำหรับสัญญาณช่วงสั้นนั้นโดยพิจารณาว่าสัญญาณช่วงสั้นที่ 25 เทียบเป็น 100 % ระยะเวลา เมื่อดูค่าความถี่มูลฐานที่ฐานข้อมูลความถี่มูลฐานที่ 100 %ระยะเวลา จะได้ความถี่มูลฐานสำหรับยอดพิทช์ของสัญญาณช่วงสั้นที่ 25 ในทำนองเดียวกันเมื่อต้องการทราบความถี่มูลฐานสำหรับยอดพิทช์ของสัญญาณช่วงสั้นที่ 10 ซึ่งเทียบเป็น 40 %ระยะเวลา (จากการคิดสัญญาณช่วงสั้นที่ 25 เป็น 100 %ระยะเวลา) จะได้ความถี่มูลฐานของสัญญาณช่วงสั้นที่ 10 เท่ากับค่าความถี่มูลฐานที่ 40 %ระยะเวลาจากฐานข้อมูลความถี่มูลฐาน

(3) ปรับระยะห่างระหว่างยอดพิทช์

ค่าความถี่มูลฐานที่ประมาณเชิงเส้นได้จะถูกนำมาแปลงเป็นระยะห่างระหว่างยอดพิทช์ โดยระยะห่างระหว่างยอดพิทช์เท่ากับส่วนกลับของความถี่มูลฐานที่ประมาณได้ จากนั้นจึงปรับระยะห่างระหว่างยอดพิทช์

## 4.8 การบวกซ้อน (Overlap-Add)

หลังจากเพิ่มหรือลบจำนวนสัญญาณช่วงสั้นเพื่อปรับช่วงเวลาและระยะห่างระหว่างยอดพิทช์ขั้นตอนต่อไปจะเป็นการสังเคราะห์สัญญาณช่วงสั้นกลับเป็นสัญญาณเสียงด้วยวิธีการบวกซ้อน โดยบวกสัญญาณช่วงสั้นซ้อนกัน โดยที่กึ่งกลางของสัญญาณช่วงสั้นอยู่ที่ตำแหน่งยอดพิทช์ของสัญญาณช่วงสั้นนั้นๆ ดังรูปที่ 4.12

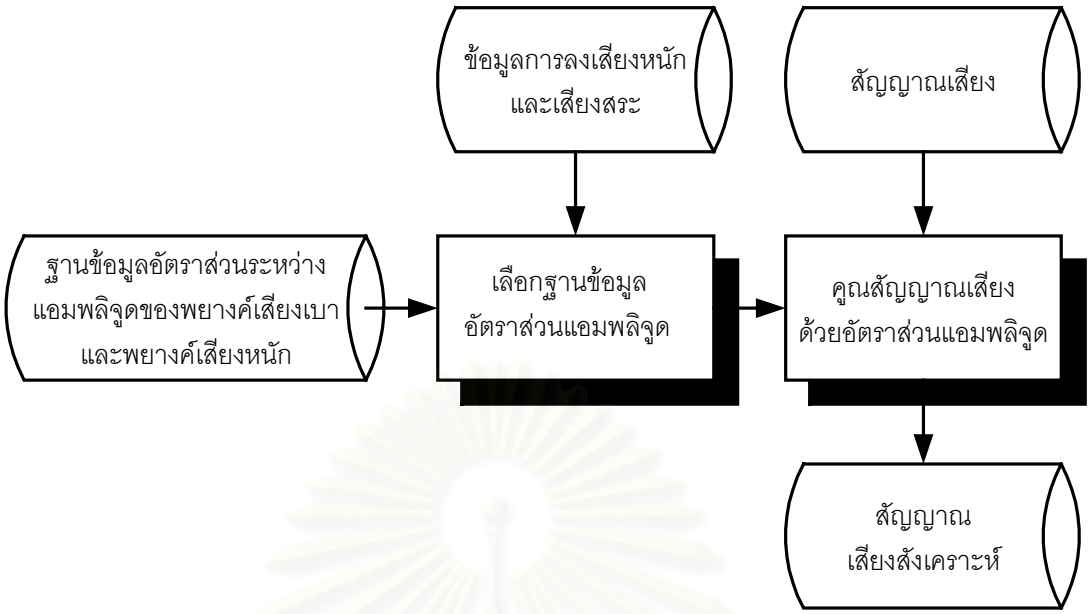



รูปที่ 4.12 การบวกซ้อนสัญญาณเสียง

## 4.9 การปรับแอมพลิจูดของสัญญาณเสียง

การปรับขนาดแอมพลิจูดของสัญญาณเสียงจะทำหลังจากการบวกซ้อนสัญญาณได้เป็นสัญญาณเสียงแล้ว

การปรับแอมพลิจูดของสัญญาณเสียงจะปรับเฉพาะการสังเคราะห์พยางค์เสียงเบา ส่วนการสังเคราะห์พยางค์เสียงหนักไม่ต้องปรับเพราะหน่วยเสียงสร้างจากพยางค์เสียงหนักอยู่แล้ว โดยมีขั้นตอนแสดงในรูปที่ 4.13 ดังนี้



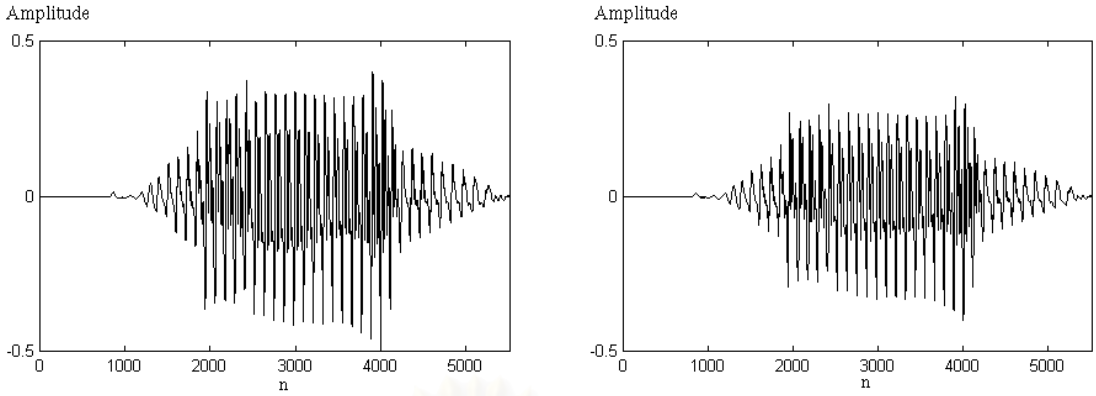

รูปที่ 4.13 แผนผังการปรับแอมพลิจูดของสัญญาณเสียง

(1) เลือกฐานข้อมูลอัตราส่วนแอมพลิจูดระหว่างพยางค์เสียงหนักและพยางค์เสียงเบา

ข้อมูลเสียงสระของพยางค์ที่ต้องการปรับแอมพลิจูดจะนำมาใช้เลือกฐานข้อมูลอัตราส่วนแอมพลิจูดระหว่างพยางค์เสียงหนักและพยางค์เสียงเบาซึ่งได้แบ่งตามเสียงสระไว้ทั้ง 24 ตัว ดังแสดงในตารางที่ 3.2

(2) คูณสัญญาณเสียงด้วยอัตราส่วนแอมพลิจูดที่ได้เลือกไว้ ได้สัญญาณเสียงดังรูปที่ 4.14

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย




(ก) (ข)

รูปที่ 4.14 การปรับแอมพลิจูดของสัญญาณเสียงสังเคราะห์คำว่า “นูน” (/nuun0/)
(ก) สัญญาณเสียงก่อนปรับแอมพลิจูด (ข) สัญญาณเสียงหลังปรับแอมพลิจูด

จากรูปที่ 4.14 แสดงตัวอย่างการปรับสัญญาณเสียงสังเคราะห์คำว่า ”นูน” (/nuun0/) ดังรูปที่ 4.14 ก ซึ่งมีเสียงสระ /uu/ ซึ่งอัตราส่วนแอมพลิจูดระหว่างพยางค์เสียงหนักและพยางค์เสียงเบาจากตารางที่ 3.2 มีค่าเท่ากับ 0.8 นำค่านี้คูณกับสัญญาณเสียงสังเคราะห์จะได้สัญญาณเสียงที่ผ่านการปรับแอมพลิจูดดังรูปที่ 4.14 ข

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

# บทที่ 5

# ผลการสังเคราะห์เสียง

## 5.1 รายละเอียดของเสียงสังเคราะห์สำหรับการทดสอบ

เสียงสังเคราะห์สำหรับการทดสอบในบทนี้ มีรายละเอียดดังนี้

1) หน่วยเสียงที่ใช้ในการสังเคราะห์เป็นหน่วยเสียงแบบอัฒพยางค์
2) สังเคราะห์เสียงด้วยวิธีการเชื่อมต่อหน่วยเสียง
3) ความละเอียดของสัญญาณเสียง 16 บิต
4) ความถี่สุ่ม 16 กิโลเฮิรตซ์

## 5.2 การทดสอบการปรับช่วงเวลา

### 5.2.1 ขั้นตอนทดสอบ

1) เลือกประโยคตัวอย่างสำหรับการสังเคราะห์เสียง
2) สังเคราะห์เสียงตามประโยคตัวอย่างที่ได้เลือกไว้ และทดลองวัดระยะเวลาของแต่ละพยางค์
3) วิเคราะห์การลงเสียงหนักเบาและชนิดของหน่วยจังหวะ (Luangthongkum, 1977) เพื่อหาค่าระยะเวลาพยางค์ของแต่ละพยางค์ในเสียงสังเคราะห์ ดังแสดงในหัวข้อ 2.6.1

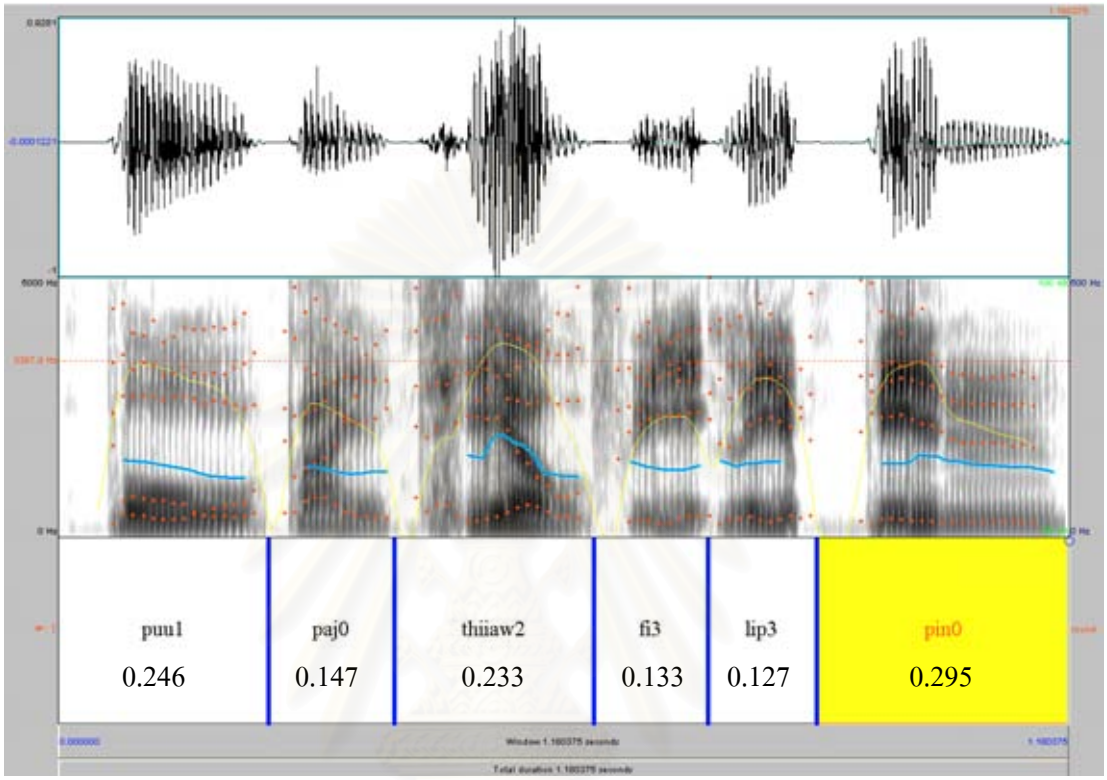
ประโยคทดสอบ : ปู่ไปเที่ยวฟิลิปปินส์

| | ปู่ | ไป | เที่ยว | ฟิ | ลิป | ปินส์ |
|---|---|---|---|---|---|---|
| การลงเสียงหนักเบา: | ‘puu1 | paj0 | ‘thiiaw2 | fi3 | lip3 | ‘pin0 |
| วิเคราะห์หน่วยจังหวะ: | ‘puu1 paj0 | | ‘thiiaw2 | fi3 | lip3 | ‘pin0 |
| | S | U | S | U | U | S |
| เวลา: | 0.25 | 0.15 | 0.23 | 0.13 | 0.13 | 0.30 วินาที |

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

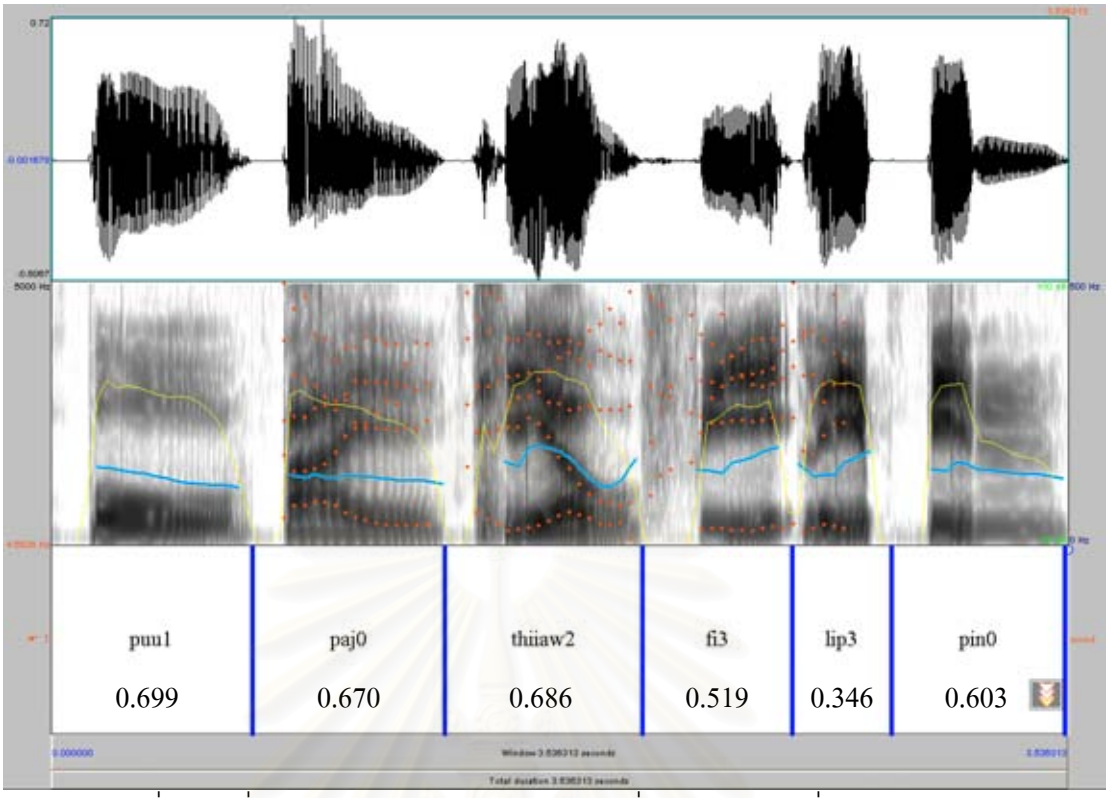
### 5.2.2 ผลการทดสอบ

หลังจากสังเคราะห์เสียงประโยคตัวอย่างซึ่งได้ปรับช่วงเวลาได้ผลดังรูปที่ 5.1 เปรียบเทียบกับเสียงสังเคราะห์ที่ไม่ได้ปรับช่วงเวลาดังรูปที่ 5.2




รูปที่ 5.1 คลื่นเสียงของเสียงสังเคราะห์ “ปู่ไปเที่ยวฟิลิปปินส์” ที่ปรับช่วงเวลา
(ตัวเลขในช่วงแบ่งพยางค์คือค่าระยะเวลา หน่วยเป็นวินาที)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย




รูปที่ 5.2 คลื่นเสียงของเสียงสังเคราะห์ “ปู่ไปเที่ยวฟิลิปปินส์” ที่ไม่ปรับช่วงเวลา
(ตัวเลขในช่วงแบ่งพยางค์คือค่าระยะเวลา หน่วยเป็นวินาที)

### 5.2.3 วิเคราะห์ผลการทดสอบปรับช่วงเวลา

ค่าเวลาจากการปรับช่วงเวลาใกล้เคียงกับค่าเวลาที่ได้กำหนดไว้สำหรับพยางค์สังเคราะห์ สาเหตุที่ระยะเวลาที่ได้ไม่เท่ากับค่าเวลาที่กำหนดไว้ เนื่องจากวิธีการที่ใช้ปรับช่วงเวลาเพิ่มหรือลดสัญญาณเป็นจำนวนเท่าของสัญญาณช่วงสั้น ระยะเวลาที่ได้จึงมีความคลาดเคลื่อนมากที่สุดได้เท่ากับระยะเวลาของสัญญาณช่วงสั้น 1 สัญญาณช่วงสั้นหรือเท่ากับระยะระหว่างยอดพิทช์ และการปรับความถี่มูลฐานซึ่งใช้วิธีการปรับระยะห่างระหว่างยอดพิทช์ทำให้ช่วงเวลาสั้นลงหรือยาวขึ้นได้อีกประการหนึ่งด้วย

## 5.3 การทดสอบการสังเคราะห์เสียงวรรณยุกต์สำหรับพยางค์เสียงเบา

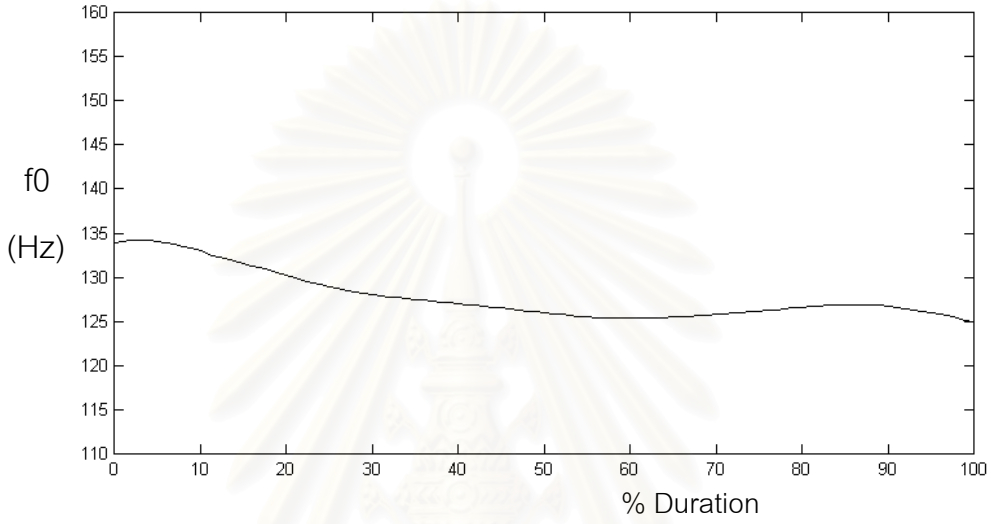
### 5.3.1 ขั้นตอนในการทดสอบ

1) เลือกคำทดสอบ 1 พยางค์ โดยเลือกคำให้สอดคล้องกับแต่ละโครงสร้างพยางค์และเสียงวรรณยุกต์นำคำทดสอบมาปรับความถี่มูลฐานตามฐานข้อมูลความถี่มูลฐานของแต่ละโครงสร้างพยางค์และเสียงวรรณยุกต์
2) นำเสียงสังเคราะห์ไปวิเคราะห์ความถี่มูลฐานโดยซอฟต์แวร์ Praat เพื่อหาความถี่มูลฐานสำหรับเปรียบเทียบกับฐานข้อมูลความถี่มูลฐานของแต่ละโครงสร้างพยางค์

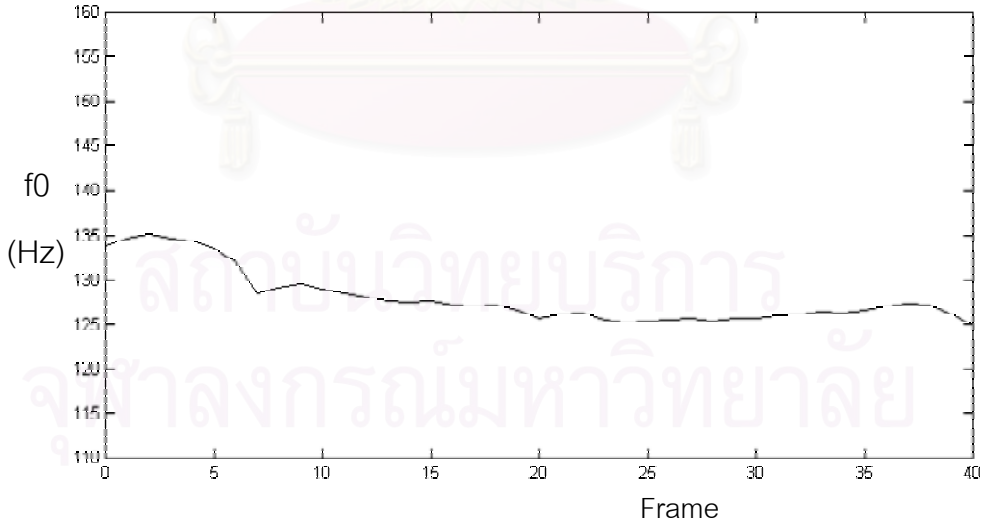
### 5.3.2 ผลการทดสอบ

1. พยางค์เป็นเสียงสามัญ

เลือกคำทดสอบว่า “โอ” (/?oo0/) ทดลองปรับเป็นพยางค์เสียงเบาตามฐานข้อมูลความถี่มูลฐานดังแสดงในรูปที่ 5.3 ก และนำเสียงสังเคราะห์ไปวิเคราะห์ความถี่มูลฐานได้ความถี่มูลฐานดังรูปที่ 5.3 ข




(ก)



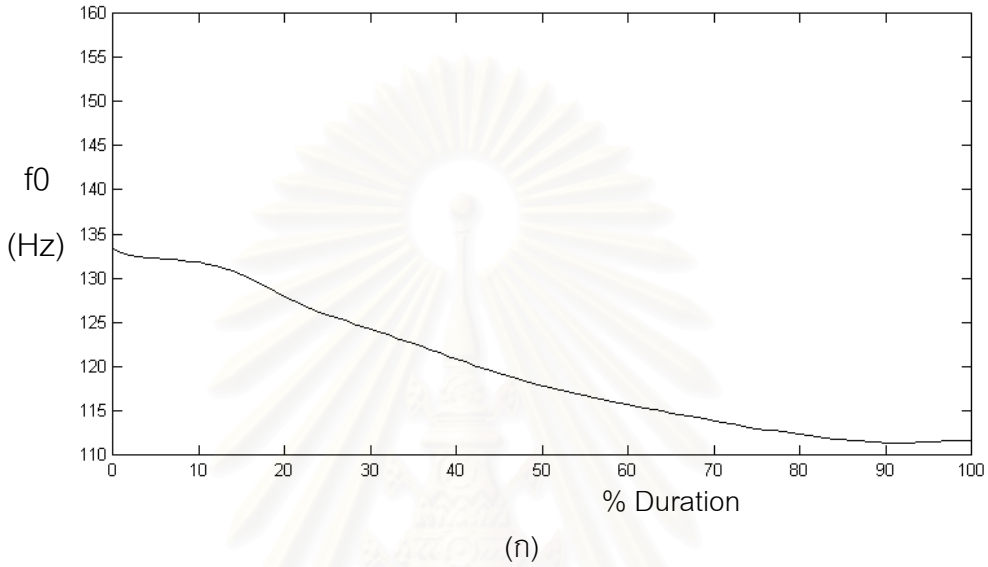

(ข)

รูปที่ 5.3 ผลการสังเคราะห์เสียงวรรณยุกต์สามัญสำหรับพยางค์เสียงเบาในพยางค์เป็น (ก) รูปแบบความถี่มูลฐานจากฐานข้อมูล (ข) ความถี่มูลฐานของคำสังเคราะห์

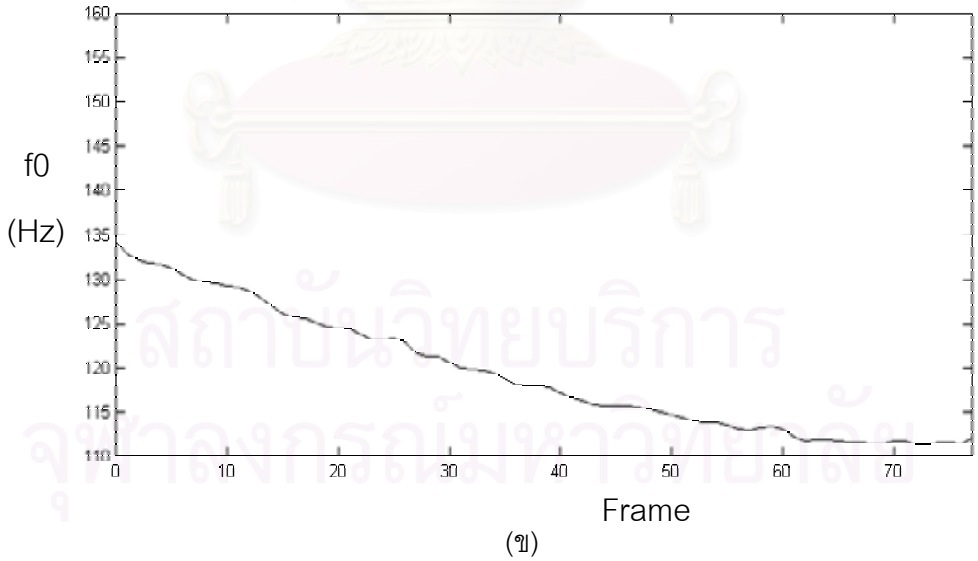
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

2. พยางค์เป็นเสียงเอก

เลือกคำทดสอบว่า “ต่อ” (/t@@1/) ทดลองปรับเป็นพยางค์เสียงเบาตามฐานข้อมูลความถี่มูลฐานดังแสดงในรูปที่ 5.4 ก และนำเสียงสังเคราะห์ไปวิเคราะห์ความถี่มูลฐานได้ความถี่มูลฐานดังรูปที่ 5.4 ข




(ก)

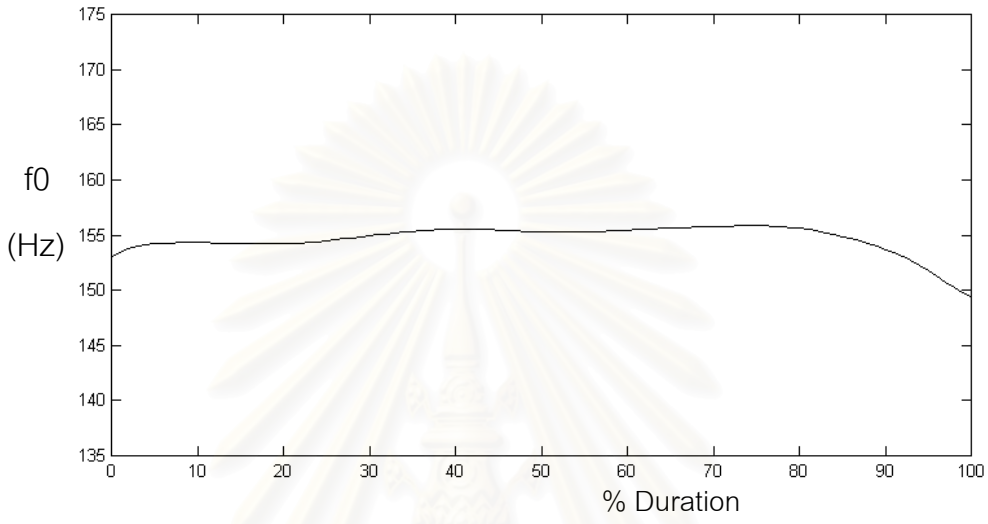



(ข)

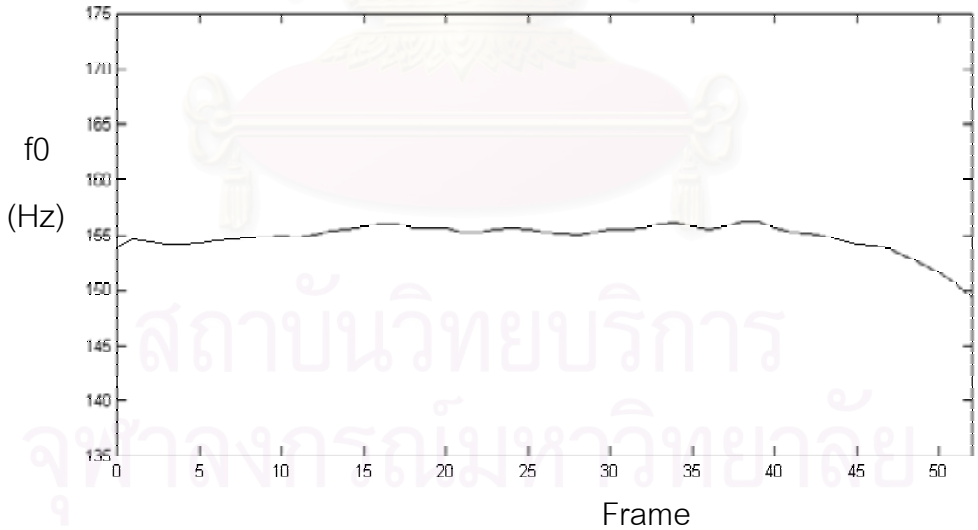
รูปที่ 5.4 ผลการสังเคราะห์เสียงวรรณยุกต์เอกสำหรับพยางค์เสียงเบาในพยางค์เป็น (ก) รูปแบบความถี่มูลฐานจากฐานข้อมูล (ข) ความถี่มูลฐานของคำสังเคราะห์

3. พยางค์เป็นเสียงโท

เลือกคำทดสอบว่า “โก้” (/koo2/) ทดลองปรับเป็นพยางค์เสียงเบาตามฐานข้อมูลความถี่มูลฐานดังแสดงในรูปที่ 5.5 ก และนำเสียงสังเคราะห์ไปวิเคราะห์ความถี่มูลฐานได้ความถี่มูลฐานดังรูปที่ 5.5 ข




(ก)



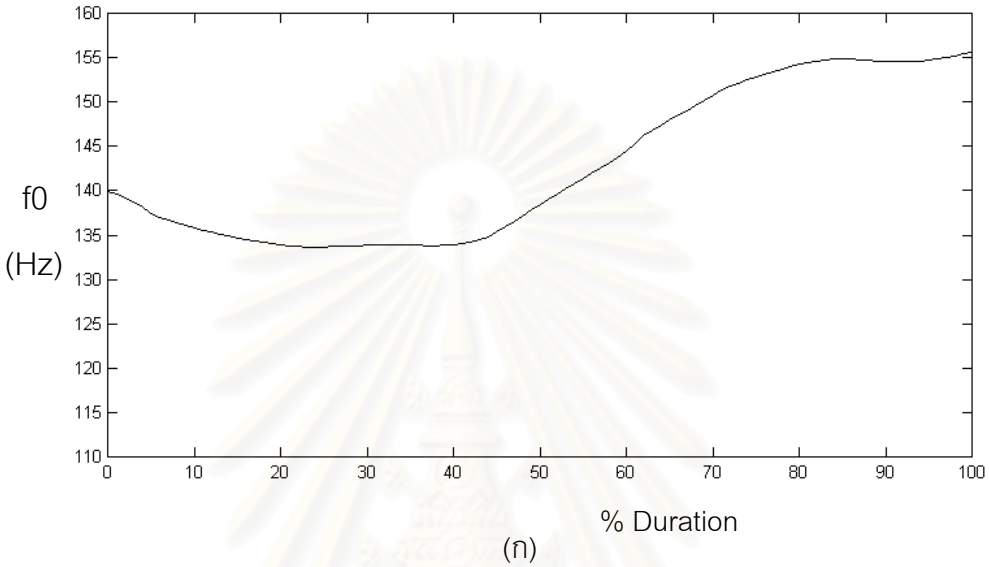

(ข)

รูปที่ 5.5 ผลการสังเคราะห์เสียงวรรณยุกต์โทสำหรับพยางค์เสียงเบาในพยางค์เป็น (ก) รูปแบบความถี่มูลฐานจากฐานข้อมูล (ข) ความถี่มูลฐานของคำสังเคราะห์

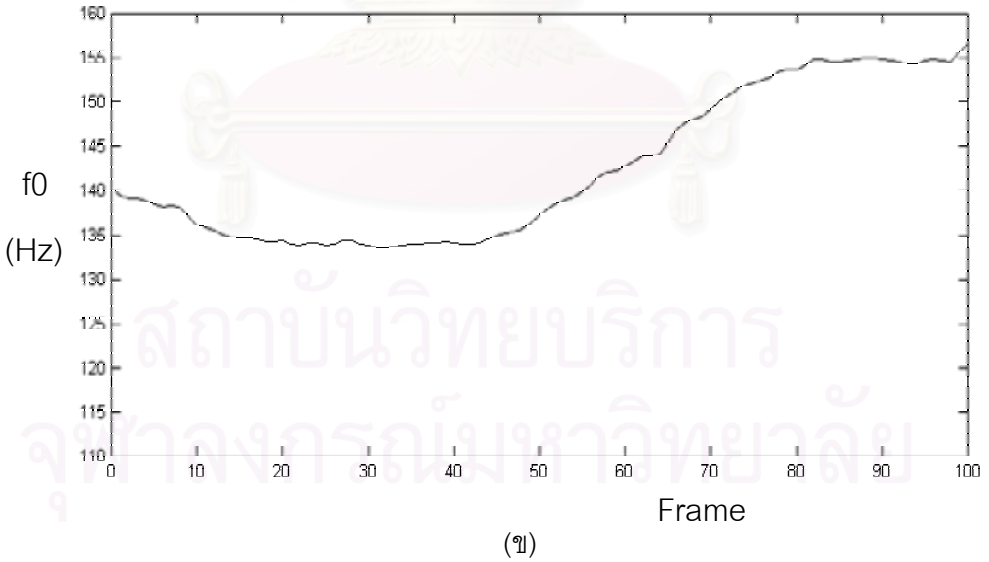
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

4. พยางค์เป็นเสียงตรี

เลือกคำทดสอบว่า “โม้” (/moo3/) ทดลองปรับเป็นพยางค์เสียงเบาตามฐานข้อมูลความถี่มูลฐานดังแสดงในรูปที่ 5.6 ก และนำเสียงสังเคราะห์ไปวิเคราะห์ความถี่มูลฐานได้ความถี่มูลฐานดังรูปที่ 5.6 ข




(ก)



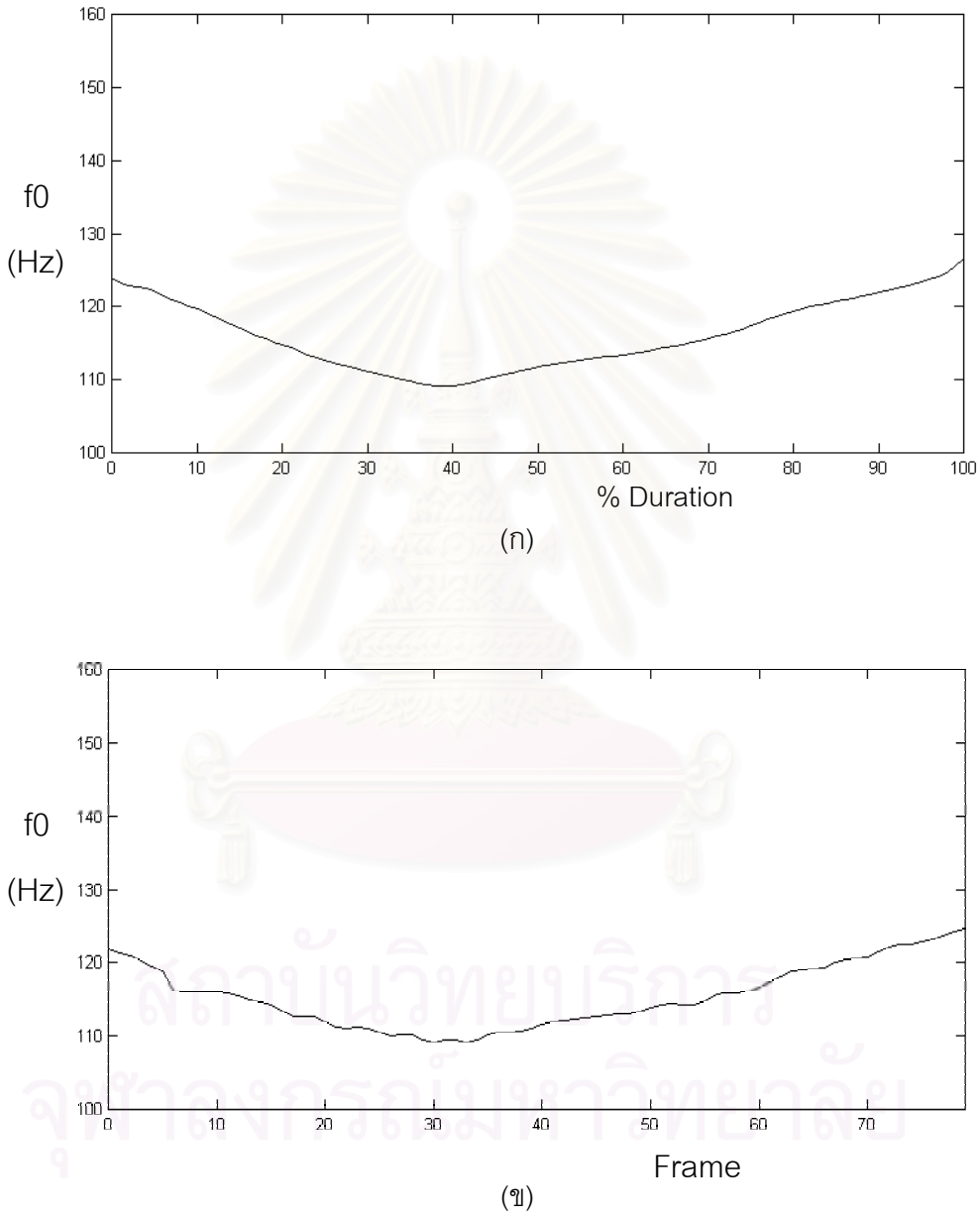

(ข)

รูปที่ 5.6 ผลการสังเคราะห์เสียงวรรณยุกต์ตรีสำหรับพยางค์เสียงเบาในพยางค์เป็น (ก) รูปแบบความถี่มูลฐานจากฐานข้อมูล (ข) ความถี่มูลฐานของคำสังเคราะห์

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

5. พยางค์เป็นเสียงจัตวา

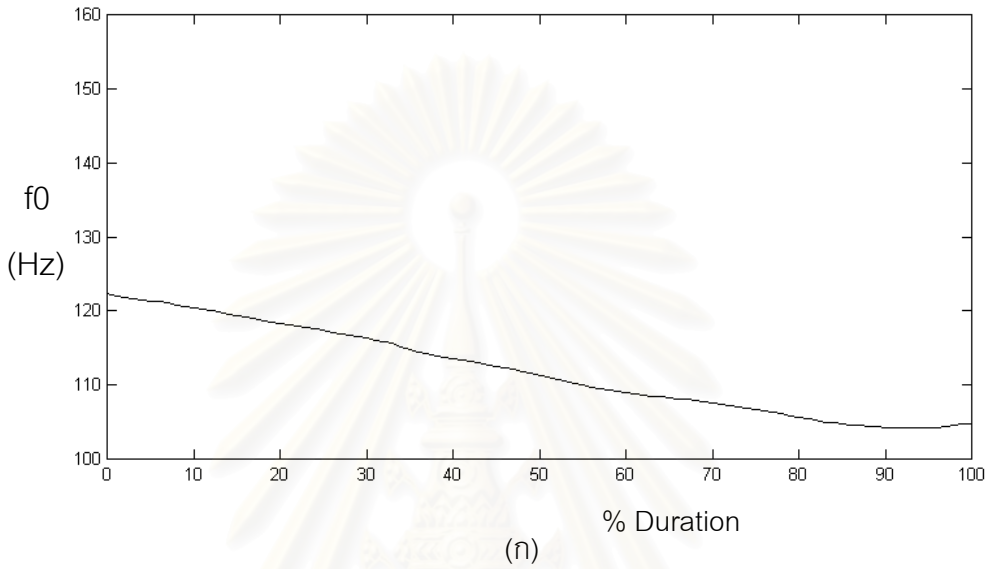
เลือกคำทดสอบว่า “โถ” (/thoo4/) ทดลองปรับเป็นพยางค์เสียงเบาตามฐานข้อมูลความถี่มูลฐานดังแสดงในรูปที่ 5.7 ก และนำเสียงสังเคราะห์ไปวิเคราะห์ความถี่มูลฐานได้ความถี่มูลฐานดังรูปที่ 5.7 ข



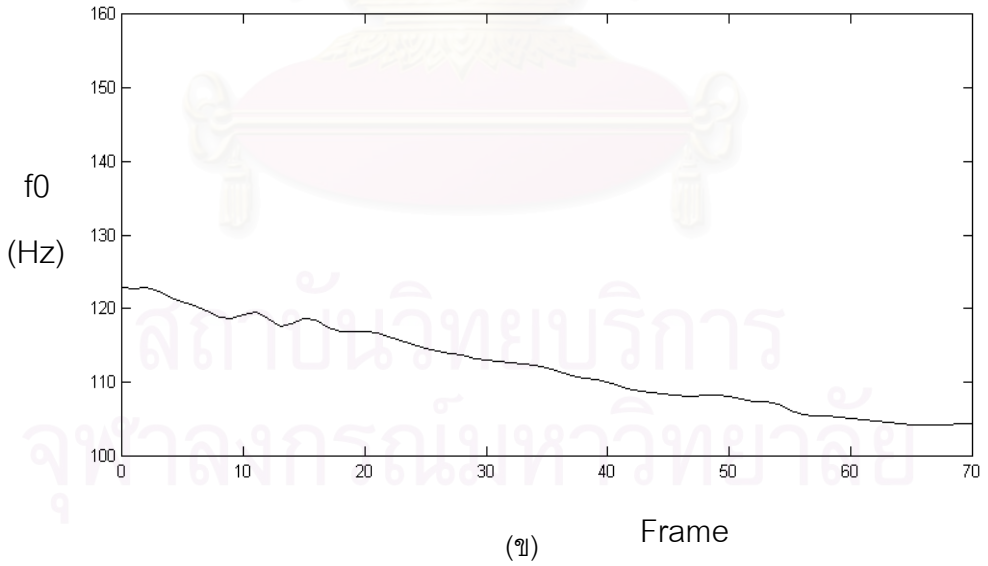

รูปที่ 5.7 ผลการสังเคราะห์เสียงวรรณยุกต์จัตวาสำหรับพยางค์เสียงเบาในพยางค์เป็น (ก) รูปแบบความถี่มูลฐานจากฐานข้อมูล (ข) ความถี่มูลฐานของคำสังเคราะห์

6. พยางค์ตายสระเสียงยาวเสียงเอก

เลือกคำทดสอบว่า “โสด” (/soot1/) ทดลองปรับเป็นพยางค์เสียงเบาตามฐานข้อมูลความถี่มูลฐานดังแสดงในรูปที่ 5.8 ก และนำเสียงสังเคราะห์ไปวิเคราะห์ความถี่มูลฐานได้ความถี่มูลฐานดังรูปที่ 5.8 ข




(ก)

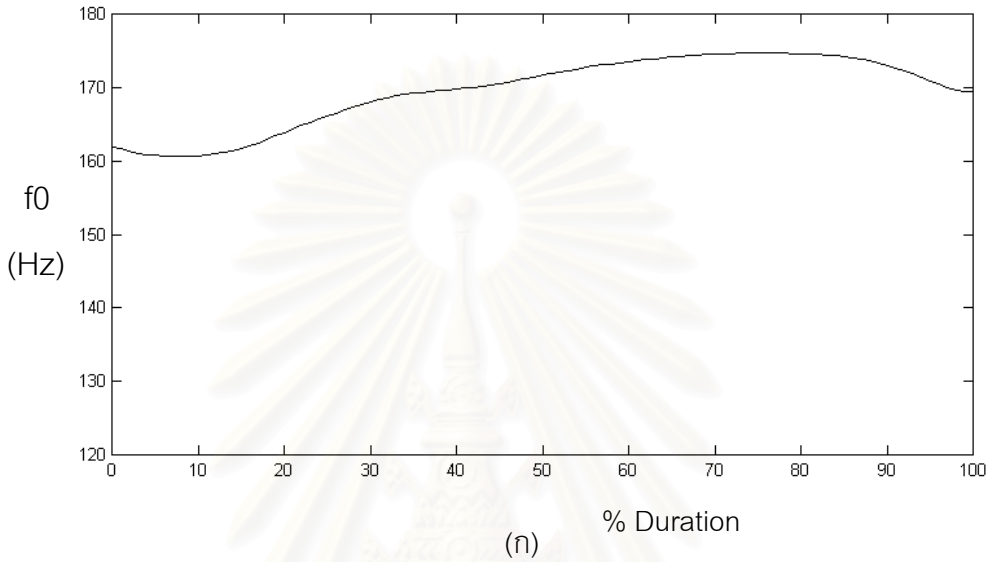



(ข)

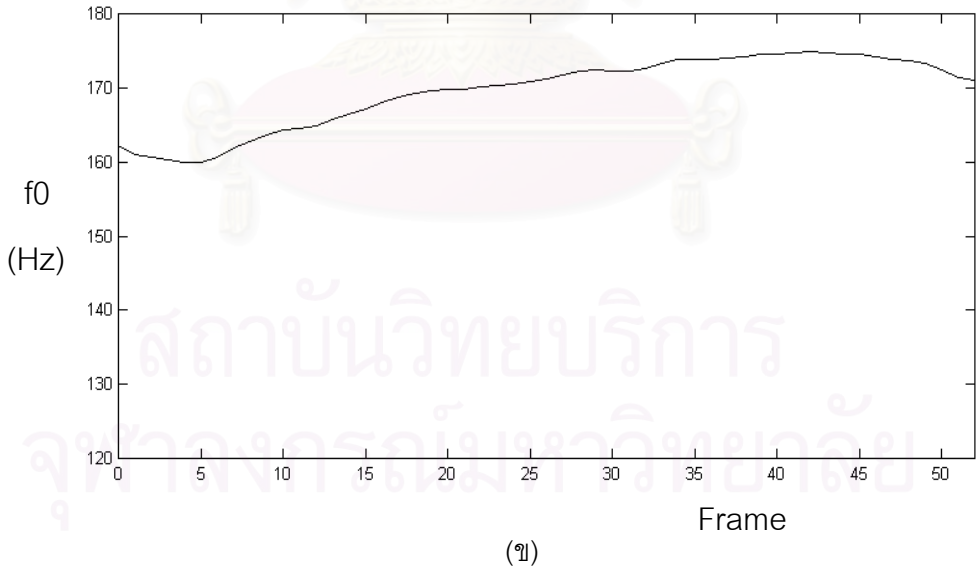
รูปที่ 5.8 ผลการสังเคราะห์เสียงวรรณยุกต์เอกสำหรับพยางค์เสียงเบาในพยางค์ตายสระเสียงยาว (ก) รูปแบบความถี่มูลฐานจากฐานข้อมูล (ข) ความถี่มูลฐานของคำสังเคราะห์

7. พยางค์ตายสระเสียงยาวเสียงโท

เลือกคำทดสอบว่า “โคก” (/khook2/) ทดลองปรับเป็นพยางค์เสียงเบาตามฐานข้อมูลความถี่มูลฐานดังแสดงในรูปที่ 5.9 ก และนำเสียงสังเคราะห์ไปวิเคราะห์ความถี่มูลฐานได้ความถี่มูลฐานดังรูปที่ 5.9 ข




(ก)

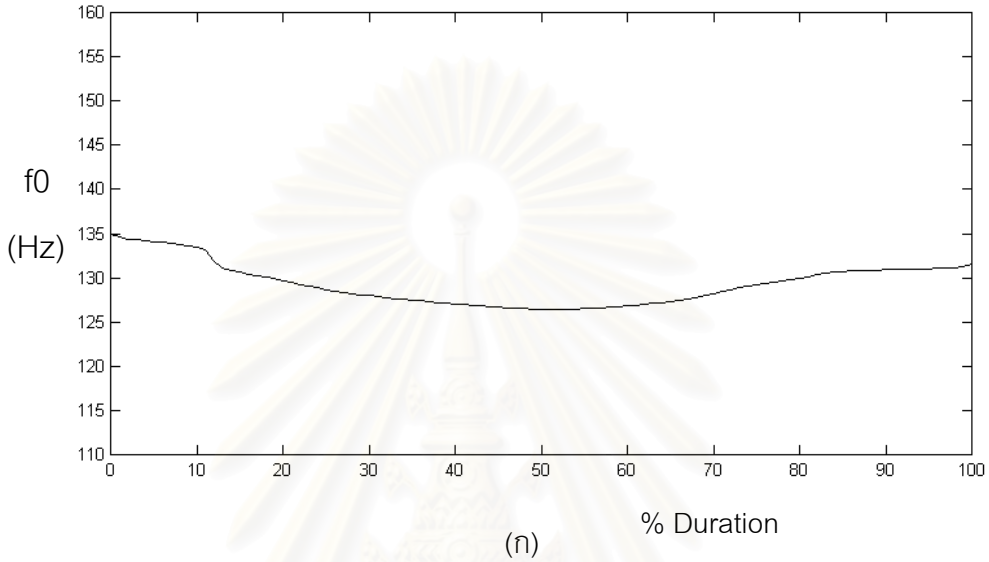



(ข)

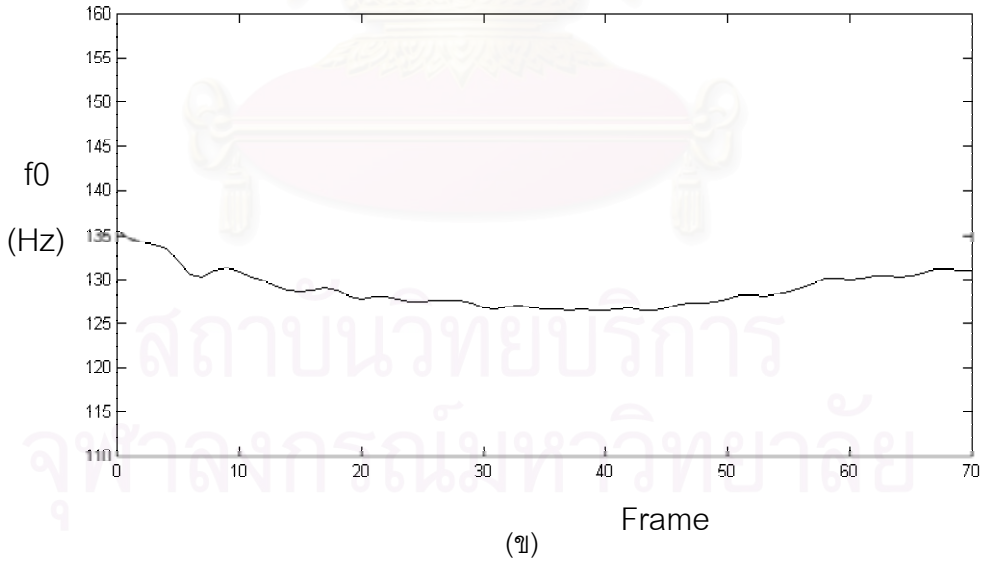
รูปที่ 5.9 ผลการสังเคราะห์เสียงวรรณยุกต์โทสำหรับพยางค์เสียงเบาในพยางค์ตายสระเสียงยาว (ก) รูปแบบความถี่มูลฐานจากฐานข้อมูล (ข) ความถี่มูลฐานของคำสังเคราะห์

8. พยางค์ตายสระเสียงยาวเสียงตรี

เลือกคำทดสอบว่า “โจ๊ก” (/cook3/) ทดลองปรับเป็นพยางค์เสียงเบาตามฐานข้อมูลความถี่มูลฐานดังแสดงในรูปที่ 5.10 ก และนำเสียงสังเคราะห์ไปวิเคราะห์ความถี่มูลฐานได้ความถี่มูลฐานดังรูปที่ 5.10 ข




(ก)



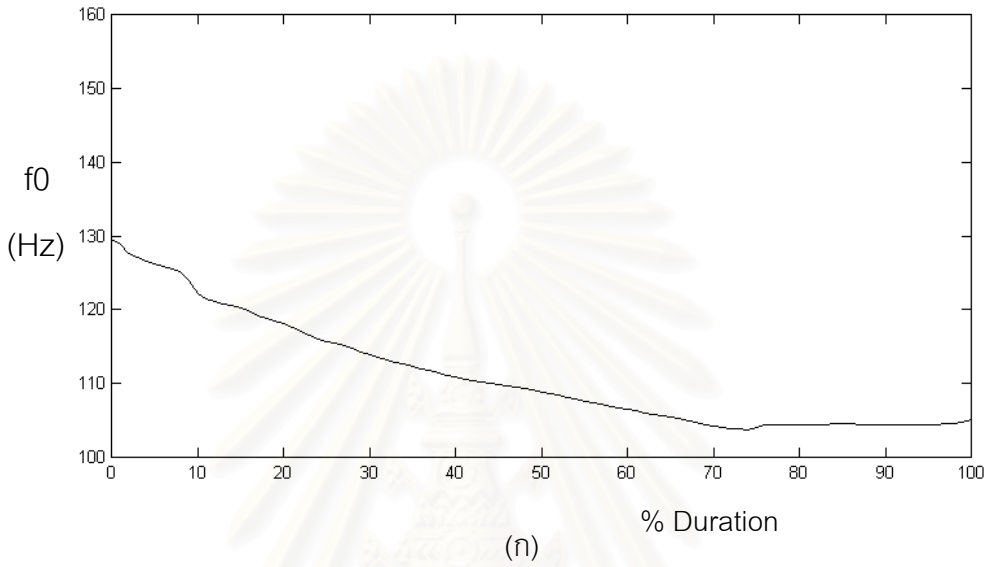

(ข)

รูปที่ 5.10 ผลการสังเคราะห์เสียงวรรณยุกต์ตรีสำหรับพยางค์เสียงเบาในพยางค์ตายสระเสียงยาว (ก) รูปแบบความถี่มูลฐานจากฐานข้อมูล (ข) ความถี่มูลฐานของคำสังเคราะห์

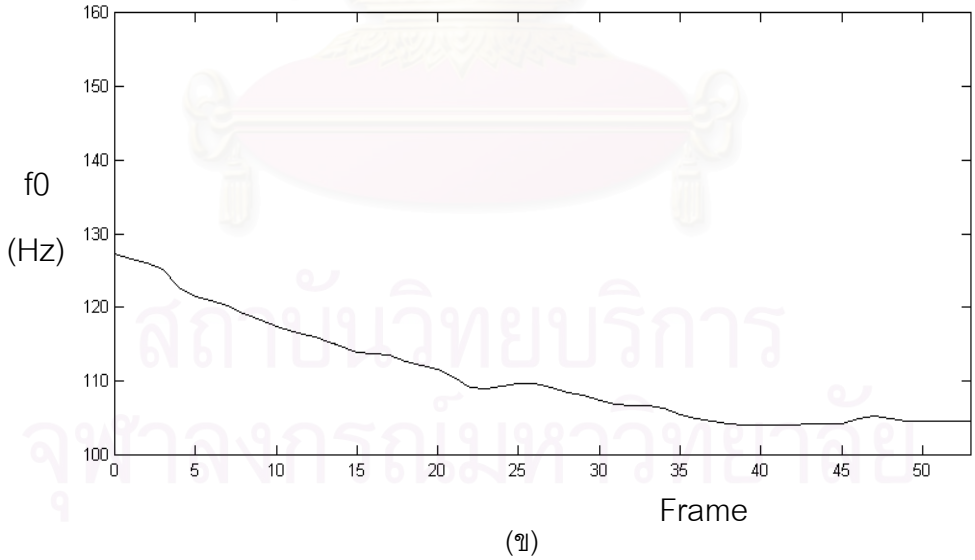
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

9. พยางค์ตายสระเสียงสั้นลงท้ายด้วยเสียงกักที่เส้นเสียงเสียงเอก

เลือกคำทดสอบว่า “โปะ” (/po1/) ทดลองปรับเป็นพยางค์เสียงเบาตามฐานข้อมูลความถี่มูลฐานดังแสดงในรูปที่ 5.11 ก และนำเสียงสังเคราะห์ไปวิเคราะห์ความถี่มูลฐานได้ความถี่มูลฐานดังรูปที่ 5.11 ข




(ก)

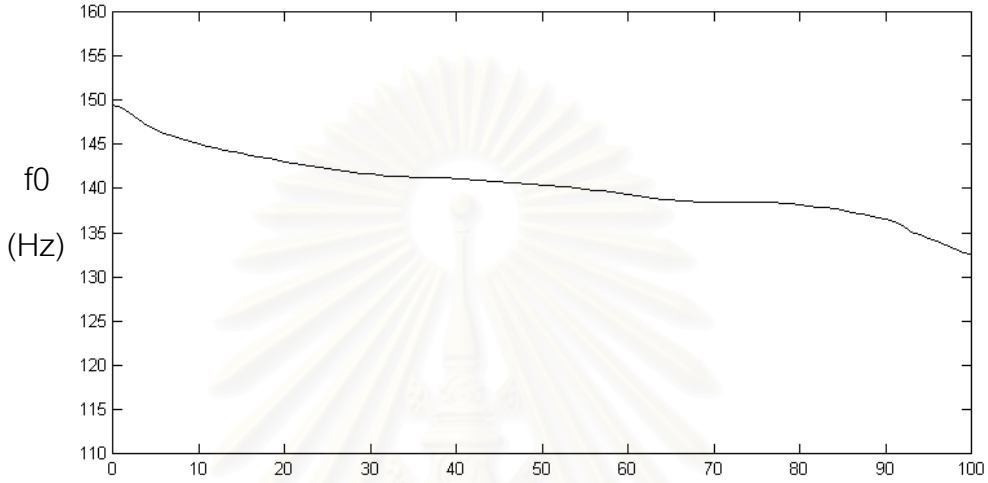



(ข)

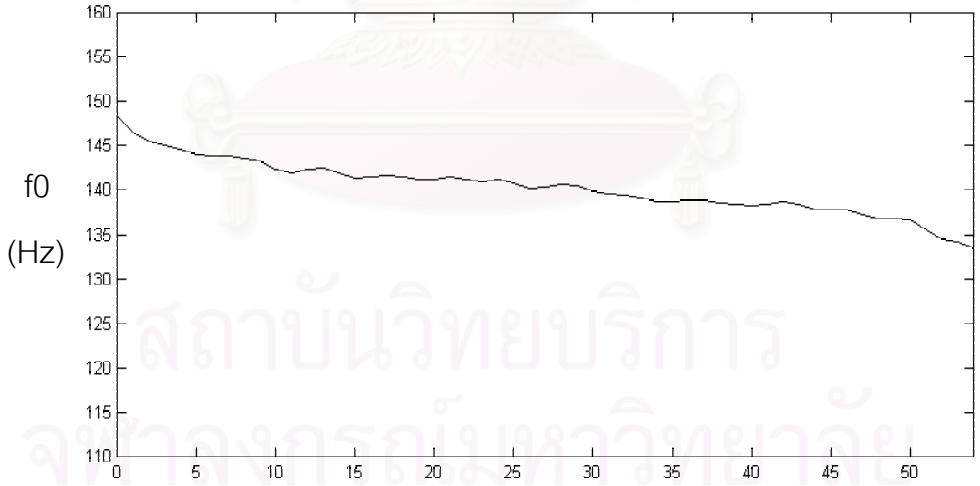
รูปที่ 5.11 ผลการสังเคราะห์เสียงวรรณยุกต์เอกสำหรับพยางค์เสียงเบา
ในพยางค์ตายสระเสียงสั้นลงท้ายด้วยเสียงกักที่เส้นเสียง
(ก) รูปแบบความถี่มูลฐานจากฐานข้อมูล (ข) ความถี่มูลฐานของคำสังเคราะห์

10. พยางค์ตายสระเสียงสั้นลงท้ายด้วยเสียงกักที่เส้นเสียงเสียงโท

เลือกคำทดสอบว่า “น่ะ” (/na2/) ทดลองปรับเป็นพยางค์เสียงเบาตามฐานข้อมูล ความถี่มูลฐานดังแสดงในรูปที่ 5.12 ก และนำเสียงสังเคราะห์ไปวิเคราะห์ความถี่มูลฐานได้ความถี่มูลฐานดังรูปที่ 5.12 ข




(ก)




(ข)

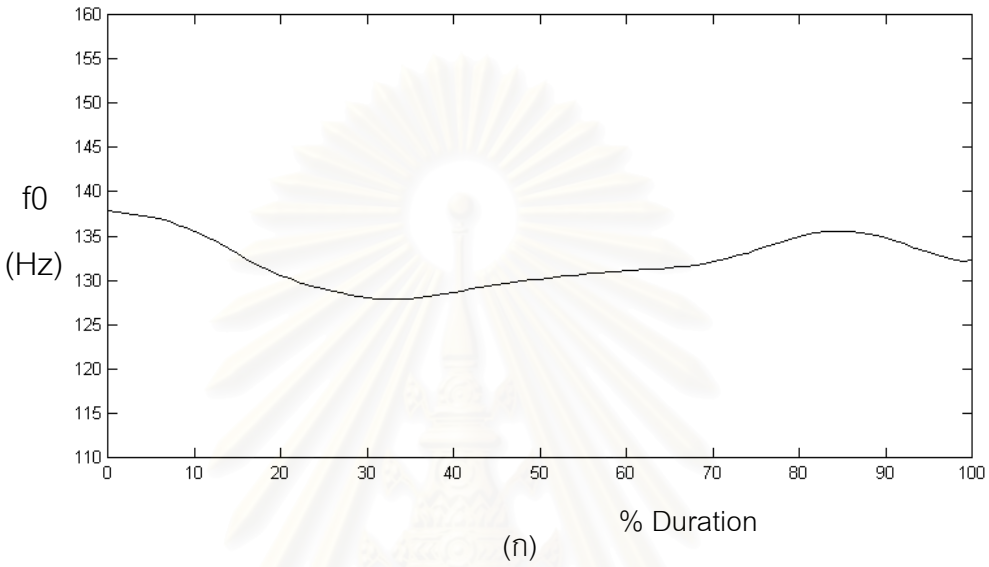
รูปที่ 5.12 ผลการสังเคราะห์เสียงวรรณยุกต์โทสำหรับพยางค์เสียงเบา ในพยางค์ตายสระเสียงสั้นลงท้ายด้วยเสียงกักที่เส้นเสียง

(ก) รูปแบบความถี่มูลฐานจากฐานข้อมูล (ข) ความถี่มูลฐานของคำสังเคราะห์

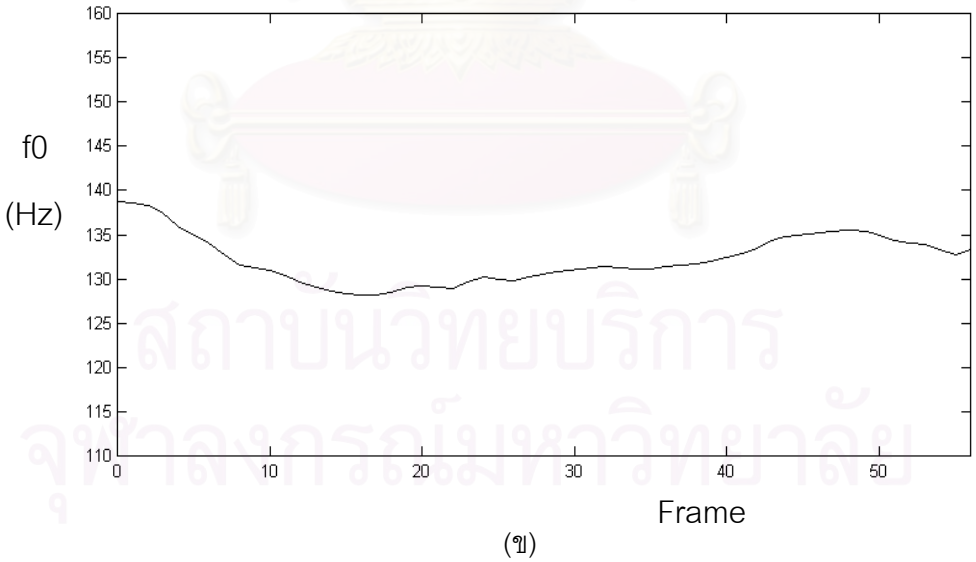
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

11. พยางค์ตายสระเสียงสั้นลงท้ายด้วยเสียงกักที่เส้นเสียงเสียงตรี

เลือกคำทดสอบว่า “โต๊ะ” (/to3/) ทดลองปรับเป็นพยางค์เสียงเบาตามฐานข้อมูล ความถี่มูลฐานดังแสดงในรูปที่ 5.13 ก และนำเสียงสังเคราะห์ไปวิเคราะห์ความถี่มูลฐานได้ความถี่มูลฐานดังรูปที่ 5.13 ข




(ก)




(ข)

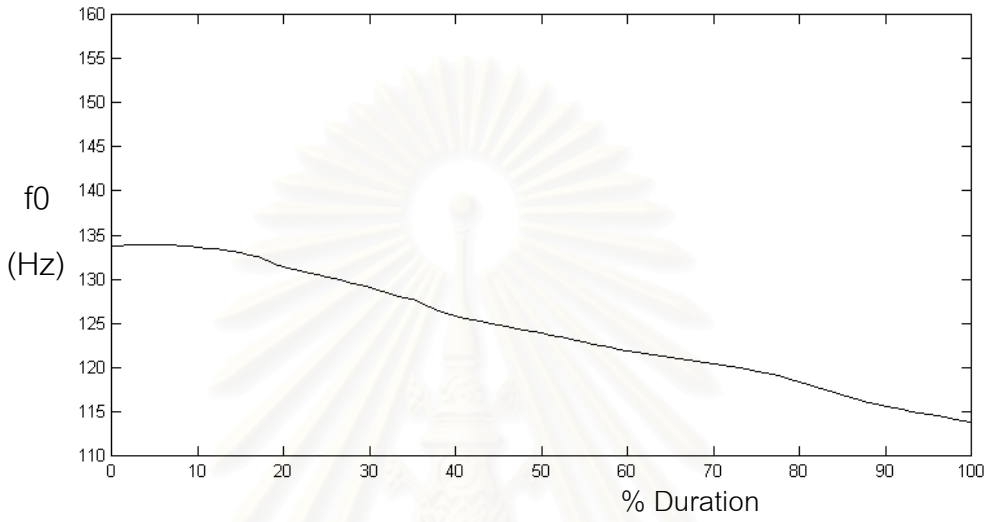
รูปที่ 5.13 ผลการสังเคราะห์เสียงวรรณยุกต์ตรีสำหรับพยางค์เสียงเบา ในพยางค์ตายสระเสียงสั้นลงท้ายด้วยเสียงกักที่เส้นเสียง

(ก) รูปแบบความถี่มูลฐานจากฐานข้อมูล (ข) ความถี่มูลฐานของคำสังเคราะห์

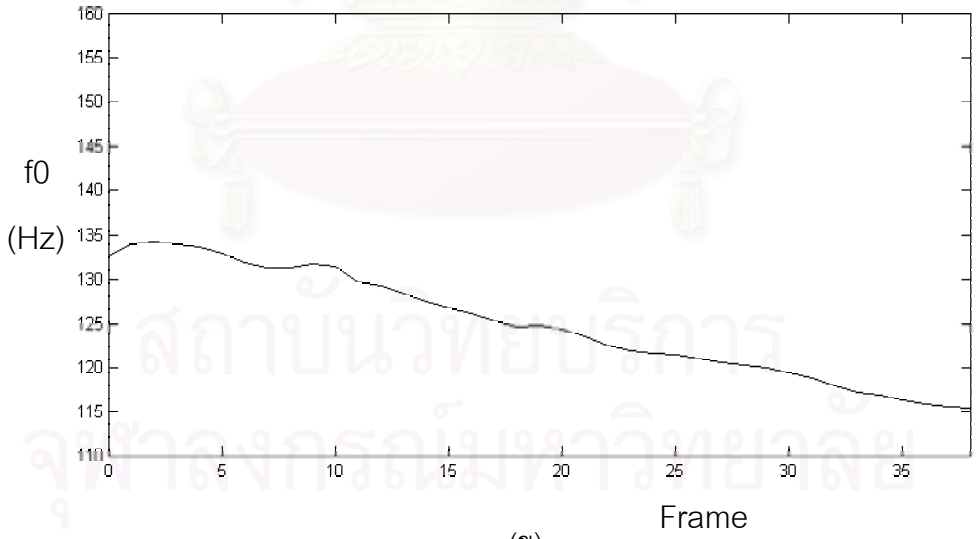
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

12. พยางค์ตายสระเสียงสั้นลงท้ายด้วยเสียงกักอื่นๆเสียงเอก

เลือกคำทดสอบว่า “ปก” (/pok1/) ทดลองปรับเป็นพยางค์เสียงเบาตามฐานข้อมูลความถี่มูลฐานดังแสดงในรูปที่ 5.14 ก และนำเสียงสังเคราะห์ไปวิเคราะห์ความถี่มูลฐานได้ความถี่มูลฐานดังรูปที่ 5.14 ข




(ก)



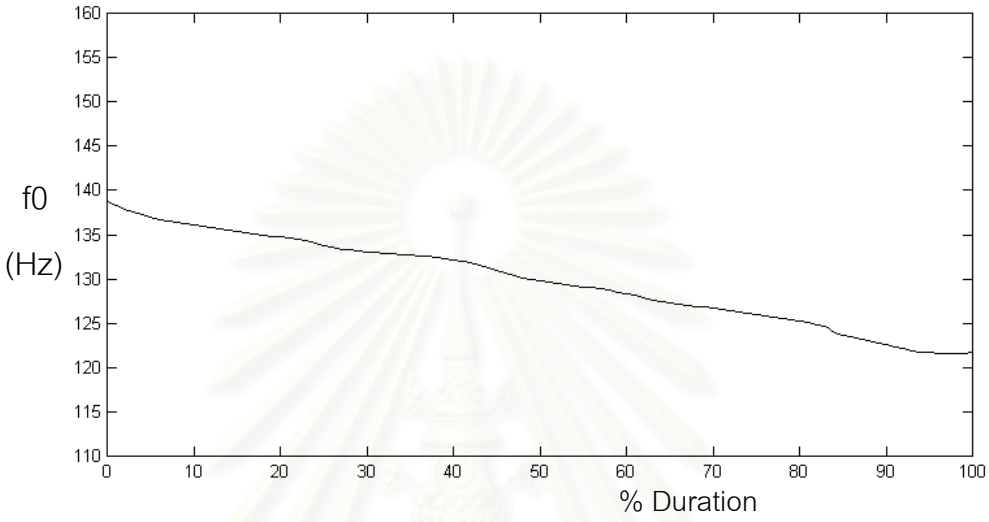

(ข)

รูปที่ 5.14 ผลการสังเคราะห์เสียงวรรณยุกต์เอกสำหรับพยางค์เสียงเบา
ในพยางค์ตายสระเสียงสั้นลงท้ายด้วยเสียงกักที่อื่นๆ
(ก) รูปแบบความถี่มูลฐานจากฐานข้อมูล (ข) ความถี่มูลฐานของคำสังเคราะห์

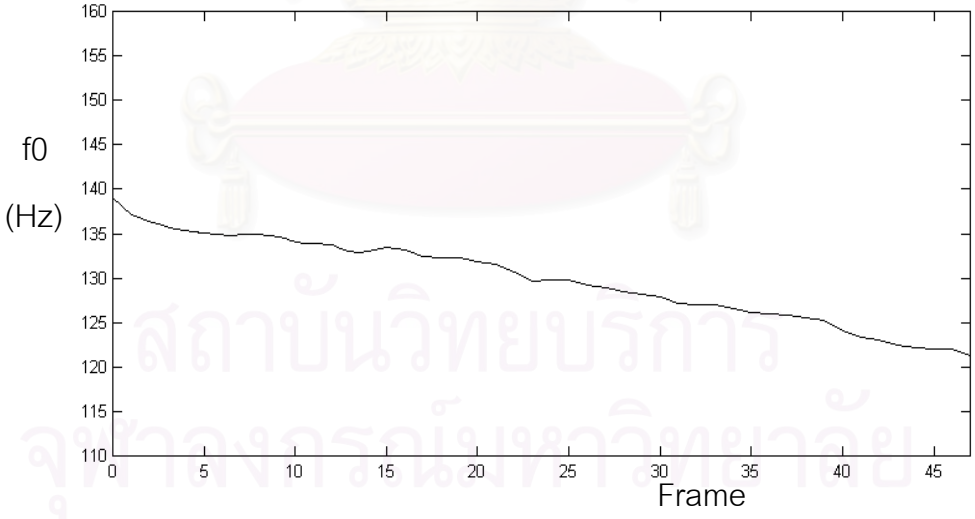
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

13. พยางค์ตายสระเสียงสั้นลงท้ายด้วยเสียงกักอื่นๆเสียงโท

เลือกคำทดสอบว่า “ว่อก” (/w@k2/) ทดลองปรับเป็นพยางค์เสียงเบาตามฐานข้อมูลความถี่มูลฐานดังแสดงในรูปที่ 5.15 ก และนำเสียงสังเคราะห์ไปวิเคราะห์ความถี่มูลฐานได้ความถี่มูลฐานดังรูปที่ 5.15 ข




(ก)



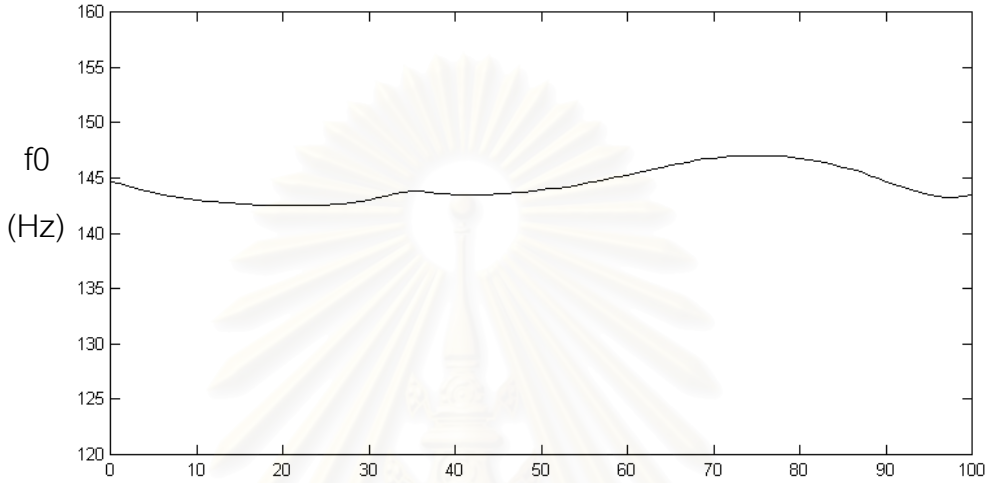

(ข)

รูปที่ 5.15 ผลการสังเคราะห์เสียงวรรณยุกต์โทสำหรับพยางค์เสียงเบา
ในพยางค์ตายสระเสียงสั้นลงท้ายด้วยเสียงกักที่อื่นๆ
(ก) รูปแบบความถี่มูลฐานจากฐานข้อมูล (ข) ความถี่มูลฐานของคำสังเคราะห์

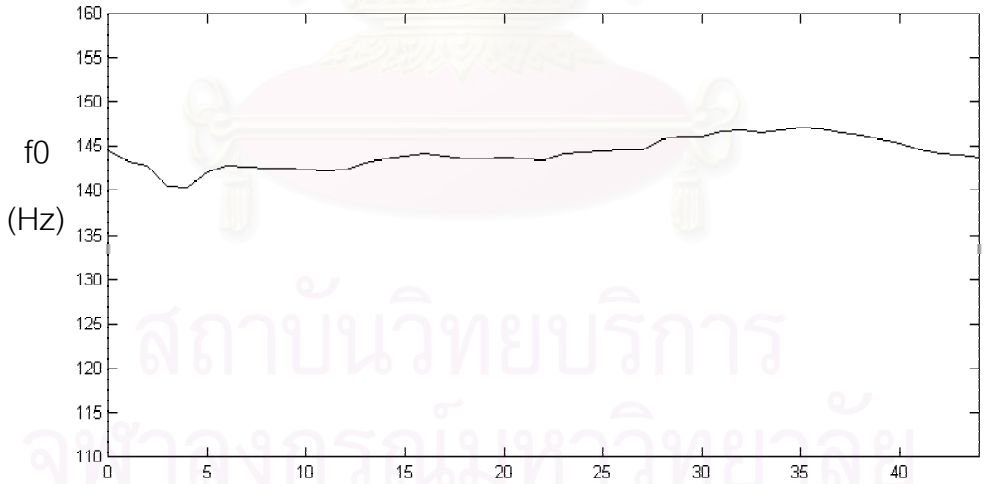
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

14. พยางค์ตายสระเสียงสั้นลงท้ายด้วยเสียงกักอื่นๆเสียงตรี

เลือกคำทดสอบว่า “รก” (/rok3/) ทดลองปรับเป็นพยางค์เสียงเบาตามฐานข้อมูล ความถี่มูลฐานดังแสดงในรูปที่ 5.16 ก และนำเสียงสังเคราะห์ไปวิเคราะห์ความถี่มูลฐานได้ความถี่มูลฐานดังรูปที่ 5.16 ข




(ก)




(ข)

รูปที่ 5.16 ผลการสังเคราะห์เสียงวรรณยุกต์ตรีสำหรับพยางค์เสียงเบา ในพยางค์ตายสระเสียงสั้นลงท้ายด้วยเสียงกักที่อื่นๆ

(ก) รูปแบบความถี่มูลฐานจากฐานข้อมูล (ข) ความถี่มูลฐานของคำสังเคราะห์

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

### 5.3.3 วิเคราะห์ผลการทดสอบการสังเคราะห์เสียงวรรณยุกต์สำหรับพยางค์เสียงเบา

ความถี่มูลฐานของพยางค์เสียงเบาของคำสังเคราะห์ให้ผลใกล้เคียงกับฐานข้อมูลความถี่มูลฐานสำหรับพยางค์เสียงเบาที่ได้สร้างไว้ ความคลาดเคลื่อนเกิดจากปัจจัยต่างๆดังนี้

- การประมาณเชิงเส้นสำหรับประมาณระยะห่างระหว่างยอดพิทช์ในการปรับความถี่มูลฐาน จะทำให้ความถี่มูลฐานคลาดเคลื่อนไปเล็กน้อย
- การต่อคลื่นเสียง เนื่องจากวิธีการสังเคราะห์เสียงด้วยวิธีการต่อคลื่นเสียงสัญญาณเสียงบริเวณรอยต่อเกิดความไม่ต่อเนื่องของสัญญาณ ทำให้ความถี่มูลฐานหลังจากการปรับความถี่มูลฐานมีลักษณะไม่ต่อเนื่อง ดังจะเห็นได้ในความถี่มูลฐานของคำว่า “อา“ (รูปที่ 5.3 ข), “โถ“ (รูปที่ 5.7 ข), “โสด“ (รูปที่ 5.9 ข), “โคก” (รูปที่ 5.10 ข), “รก“ (รูปที่ 5.16 ข)
- การระบุยอดพิทช์ การระบุยอดพิทช์อาจเกิดการผิดพลาดในกรณีที่สัญญาณเสียงมีการแกว่งของยอดคลื่นมาก จะทำให้การระบุยอดพิทช์มีความคลาดเคลื่อนได้

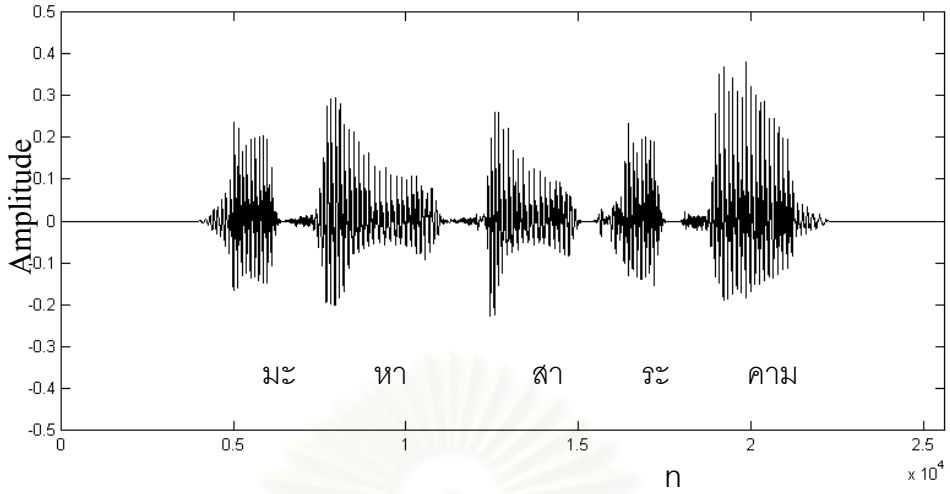
## 5.4 การทดสอบการสังเคราะห์เสียงเมื่อปรับช่วงเวลาความถี่และความดัง

### 5.4.1 วิธีการทดสอบ

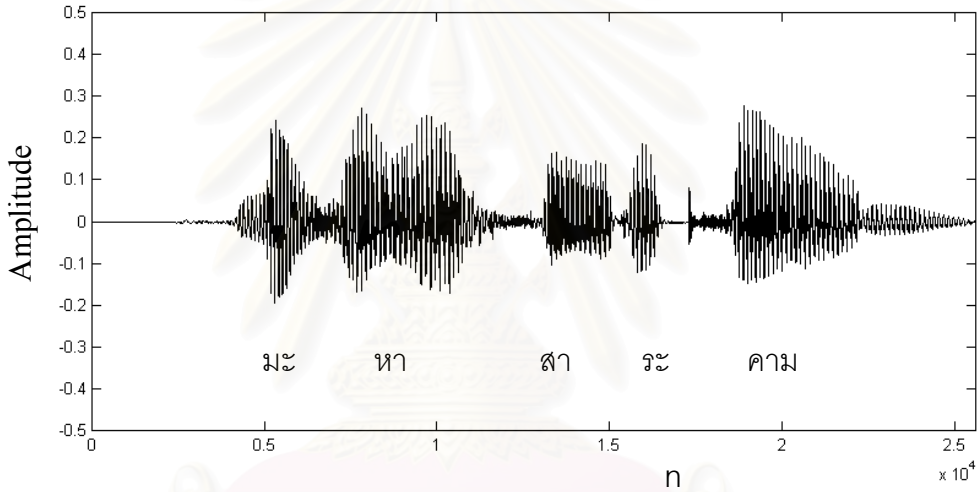
ทดสอบโดยสังเคราะห์เสียงคำว่า “มหาสารคาม” (/ma3 ‘haa4 ‘saa4 ra3 ‘khaam0/) ได้ดังรูป 5.17 ก เปรียบเทียบกับคำเดียวกันที่เป็นเสียงพูดจริงดังรูปที่ 5.17 ข

### 5.4.2 วิเคราะห์ผลทดสอบการสังเคราะห์เสียงเมื่อปรับช่วงเวลาความถี่และความดัง

จากการทดสอบการสังเคราะห์เสียงเมื่อปรับช่วงเวลาความถี่และความดัง สัญญาณเสียงจากการสังเคราะห์มีลักษณะใกล้เคียงกับเสียงพูดจริง โดยเฉพาะค่าระยะเวลาของพยางค์ยกเว้นในส่วนของพยางค์สุดท้ายซึ่งในการพูดปกติจะลงเสียงหนักเป็นพิเศษ ซึ่งจะทำให้ค่าระยะเวลาของพยางค์ยาวกว่าพยางค์ที่ลงเสียงหนักทั่วไป และพิจารณาเสียงพูดจริงและเสียงสังเคราะห์จะเห็นว่าอัตราส่วนระหว่างส่วนที่เป็นเสียงก้องและเสียงไม่ก้องจะไม่เท่ากัน เนื่องจากการปรับช่วงเวลาสำหรับวิทยานิพนธ์นี้จะลดหรือเพิ่มช่วงเวลาสำหรับในที่ส่วนที่เป็นเสียงก้องและเสียงไม่ก้องด้วยอัตราส่วนที่เท่ากัน จึงควรมีการวิจัยหาอัตราส่วนในการลดหรือเพิ่มช่วงเวลาของเสียงก้องและเสียงไม่ก้องต่อไป




(ก)




(ข)

รูปที่ 5.17 การเปรียบเทียบเสียงสังเคราะห์กับเสียงพูดจริงจากคำว่า “มหาสารคาม”
(ก) เสียงสังเคราะห์ (ข) เสียงพูดจริง

## 5.5 **เอ็มโอเอส** (Mean Opinion Score: MOS)

วิธีนี้ใช้วัดคุณภาพเสียงโดยรวมเพื่อวัดความเป็นธรรมชาติของเสียง การทดสอบโดยให้อาสาสมัครฟังเสียงเป็นคำและประโยค โดยที่จะประเมิณผลการสังเคราะห์พยางค์หนักเบาในระดับคำและในระดับประโยค

### 5.5.1 **การประเมินผล**

กลุ่มตัวอย่างประเมินความเป็นธรรมชาติของเสียงสังเคราะห์ โดยแบ่งการประเมินเป็นระดับคำและระดับประโยค เพื่อดูผลการประเมินคำสังเคราะห์ในระดับคำและระดับ

ประโยคให้ผลแตกต่างกันอย่างไร ในการประเมินกลุ่มตัวอย่างให้คะแนนแบ่งระดับความประทับใจเป็น 5 ระดับคือ ดีมาก (Excellent) ดี (Good) ปานกลาง (Fair) แย่ (Poor) และแย่มาก (Bad) โดยแต่ละระดับมีน้ำหนักคะแนนเป็น 5 4 3 2 และ 1 ตามลำดับ นำคะแนนที่ได้มาเฉลี่ยซึ่งคะแนนดังกล่าวเรียกว่า MOS

### 5.5.2 **การสร้างแบบทดสอบ**

1) เลือกคำทดสอบจำนวน 10 คำ
2) สร้างประโยคทดสอบโดยนำคำทดสอบแต่ละคำมาแต่งเป็นประโยค
3) สังเคราะห์เสียงตามคำทดสอบ โดยแบ่งเป็น 2 ชุด
   - ชุดที่ 1 สังเคราะห์เสียงด้วยวิธีเชื่อมต่อหน่วยเสียงโดยไม่ปรับลักษณะทางสัทศาสตร์
   - ชุดที่ 2 สังเคราะห์เสียงด้วยวิธีเชื่อมต่อหน่วยเสียงโดยปรับลักษณะทางสัทศาสตร์
4) สังเคราะห์เสียงตามประโยคทดสอบ โดยแบ่งเป็น 2 ชุด
   - ชุดที่ 1 สังเคราะห์เสียงด้วยวิธีเชื่อมต่อหน่วยเสียงโดยไม่ปรับลักษณะทางสัทศาสตร์
   - ชุดที่ 2 สังเคราะห์เสียงด้วยวิธีเชื่อมต่อหน่วยเสียงโดยปรับลักษณะทางสัทศาสตร์
5) สร้างแบบทดสอบจากคำสังเคราะห์โดยแบ่งเป็น 2 ส่วน ดังภาคผนวก ข. ส่วนแรกจะเป็นการประเมินผลฟังคำสังเคราะห์ในระดับคำ และส่วนที่สองจะเป็นการประเมินผลในระดับประโยค เนื่องจากวิธีการในการสร้างแบบทดสอบของทั้งสองส่วนจะเหมือนกัน ดังนั้นจึงขอกล่าวเฉพาะการสร้างแบบทดสอบของส่วนแรก
   - กลุ่มตัวอย่างจะทราบคำสังเคราะห์ที่ได้รับฟังล่วงหน้า
   - มีจำนวน 30 ข้อ จากคำทดสอบ 10 คำ ซึ่งจะเปิดซ้ำ 3 ครั้ง
   - ในหนึ่งข้อ จะเปิดให้ฟังเสียงสังเคราะห์ที่ไม่ปรับลักษณะทางสัทศาสตร์หนึ่งครั้ง และเสียงสังเคราะห์ที่ปรับลักษณะทางสัทศาสตร์อีกหนึ่งครั้ง

### 5.5.3 **ผลการประเมิน**

การประเมินได้ผลดังตารางที่ 5.1

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

**ตารางที่** 5.1 **ผลการประเมินด้วยวิธี** MOS

| | ระดับคำ | | ระดับประโยค | |
|---|---|---|---|---|
| | ปรับลักษณะทางสัทศาสตร์ | ไม่ปรับลักษณะทางสัทศาสตร์ | ปรับลักษณะทางสัทศาสตร์ | ไม่ปรับลักษณะทางสัทศาสตร์ |
| คะแนนเฉลี่ย (ค่าเอ็มโอเอส) | 3.66 | 3.23 | 3.92 | 3.07 |
| ส่วนเบี่ยงเบนมาตรฐานของคะแนนเฉลี่ย | 0.39 | 0.71 | 0.21 | 0.88 |
| คะแนนเฉลี่ยต่ำที่สุด | 2.93 | 2.20 | 3.50 | 2.00 |
| คะแนนเฉลี่ยสูงที่สุด | 4.13 | 4.27 | 4.17 | 4.43 |
| พิสัย | 1.2 | 2.07 | 0.67 | 2.43 |

**การวิเคราะห์ผลเอ็มโอเอส**

1) คะแนนเฉลี่ย (ค่าเอ็มโอเอส) ของคำสังเคราะห์ที่ปรับลักษณะทางสัทศาสตร์มีค่าสูงกว่าคำสังเคราะห์ที่ไม่ปรับลักษณะทางสัทศาสตร์ ทั้งในระดับคำและในระดับประโยค เพราะเสียงสังเคราะห์ที่ปรับลักษณะทางสัทศาสตร์มีการลงเสียงหนักเบาเช่นเดียวกับเสียงพูดจริง ซึ่งแตกต่างจากเสียงสังเคราะห์ที่ไม่ปรับลักษณะทางสัทศาสตร์ คำสังเคราะห์จะมีลักษณะพูดเป็นคำๆ และค่อนข้างช้าและน่าเบื่อ
2) คะแนนเฉลี่ยของคำสังเคราะห์ที่ปรับลักษณะทางสัทศาสตร์ในระดับประโยคจะมีค่าสูงกว่าในระดับคำ เนื่องจากการพูดในระดับคำหลายๆครั้งที่ผู้พูดจะพูดเป็นคำๆเพื่อให้ผู้ฟังเข้าใจได้ง่ายและชัดเจน แต่ในระดับประโยคเสียงพูดตามธรรมชาติจะไม่มีลักษณะการพูดเป็นคำๆจะพูดโดยมีการลงเสียงหนักเบา
3) คะแนนเฉลี่ยต่ำที่สุดให้ผลที่สอดคล้องกับคะแนนเฉลี่ย (ค่าเอ็มโอเอส) แต่คะแนนเฉลี่ยสูงที่สุดให้ผลไม่สอดคล้อง เนื่องจากการสังเคราะห์พยางค์หนักเบานี้จะทำให้คุณภาพเสียงลดลงประกอบกับคุณภาพของหน่วยเสียงบางหน่วยเสียงไม่ดี และเสียงสังเคราะห์ที่ไม่ปรับลักษณะทางสัทศาสตร์มีลักษณะช้ารับฟังได้ชัดเจน กลุ่มตัวอย่างบางส่วนจึงพึงพอใจกับเสียงสังเคราะห์ที่ไม่ปรับลักษณะทางสัทศาสตร์มากกว่า
4) ค่าพิสัยและส่วนเบี่ยงเบนมาตรฐานให้ผลที่สอดคล้องกัน โดยทั้งระดับคำหรือระดับประโยค คำสังเคราะห์ที่ปรับลักษณะทางสัทศาสตร์มีค่าพิสัยและส่วนเบี่ยงเบนมาตรฐานต่ำกว่า ซึ่งแสดงถึงความเห็นพ้องต้องกันต่อความพึงพอใจในเสียงสังเคราะห์ของกลุ่มตัวอย่าง คะแนน

เฉลี่ยของกลุ่มตัวอย่างใกล้เคียงกัน แต่คำสังเคราะห์ที่ไม่ปรับลักษณะทางสัทศาสตร์ให้ค่าพิสัยและส่วนเบี่ยงเบนมาตรฐานที่สูง แสดงถึงกลุ่มตัวอย่างมีความเห็นแตกต่างกัน บางส่วนพึงพอใจและบางส่วนไม่พอใจ



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

# บทที่ 6

# ข้อสรุปและข้อเสนอแนะ

## 6.1 ข้อสรุป

งานวิจัยทางด้านการสังเคราะห์เสียงพูดของมนุษย์ที่นำเสนอในวิทยานิพนธ์ฉบับนี้เป็นวิธีการสังเคราะห์เสียงหนักเบาให้กับพยางค์สำหรับเสียงสังเคราะห์ภาษาไทยโดยการปรับลักษณะทางสัทศาสตร์ของสัญญาณเสียงที่เกี่ยวข้อง ได้แก่ ความถี่มูลฐาน ช่วงเวลา และแอมพลิจูดของสัญญาณ โดยการนำวิธี Time-Domain Pitch-Synchronous Overlap-Add (TD-PSOLA) รวมกับวิธีการอื่นๆเพื่อประยุกต์ใช้ในการปรับความถี่มูลฐานและการปรับช่วงเวลา และในส่วนของแอมพลิจูดสามารถปรับโดยการคูณสัญญาณเสียงทั้งหมดด้วยตัวคูณที่ต้องการ

ระบบสังเคราะห์เสียงที่ใช้สำหรับการสังเคราะห์เสียงในวิทยานิพนธ์นี้ใช้การต่อคลื่นเสียง (Concatenation) โดยเลือกใช้หน่วยเสียงอัฒพยางค์เป็นหน่วยเสียงสำหรับการต่อคลื่นเสียง ซึ่งมีจำนวน 1,178 หน่วยเสียง แบ่งเป็นหน่วยเสียงเริ่ม 396 หน่วยเสียง หน่วยเสียงท้าย 782 หน่วยเสียง โดยหน่วยเสียงบันทึกจากผู้บอกภาษาเพศชายที่มีคุณภาพเสียงที่อัตราสุ่ม 16,000 เฮิรตซ์ และมีความละเอียดสัญญาณ 16 บิตต่อหนึ่งตัวอย่าง

การวิเคราะห์เป็นพื้นฐานที่สำคัญการสังเคราะห์พยางค์หนักเบา โดยต้องวิเคราะห์ลักษณะทางสัทศาสตร์ของพยางค์เสียงหนักและพยางค์เสียงเบาเพื่อนำมาสร้างเป็นฐานข้อมูลสำหรับการสังเคราะห์พยางค์เสียงหนักและพยางค์เสียงเบาอันได้แก่ ฐานข้อมูลความถี่มูลฐานสำหรับเสียงวรรณยุกต์ในพยางค์เสียงเบา และฐานข้อมูลอัตราส่วนแอมพลิจูดระหว่างพยางค์เสียงเบาและพยางค์เสียงหนัก ในส่วนของค่าระยะเวลาสำหรับแต่ละพยางค์ได้ใช้ค่าที่ได้จากงานวิจัยของ ธีรพันธ์ เหลืองทองคำ (Luangthongkum, 1977)

ฐานข้อมูลความถี่มูลฐานสำหรับเสียงวรรณยุกต์ในพยางค์เสียงเบามีจำนวน 14 รูปแบบความถี่มูลฐานแบ่งตามโครงสร้างพยางค์และเสียงวรรณยุกต์ ฐานข้อมูลอัตราส่วนแอมพลิจูดระหว่างพยางค์เสียงเบาและพยางค์เสียงหนักมีจำนวน 24 ค่าแบ่งตามเสียงสระของพยางค์

การทดสอบความพึงพอใจในการรับฟังจากกลุ่มตัวอย่างจำนวน 10 คน ในระดับคำ เมื่อปรับลักษณะทางสัทศาสตร์ของเสียงสังเคราะห์ได้ค่าเอ็มโอเอสเท่ากับ 3.66 เมื่อไม่ปรับลักษณะทางสัทศาสตร์ของเสียงสังเคราะห์ได้ค่าเอ็มโอเอสเท่ากับ 3.23 ในระดับประโยค เมื่อปรับลักษณะทางสัทศาสตร์ของเสียงสังเคราะห์ได้ค่าเอ็มโอเอสเท่ากับ 3.92 เมื่อไม่ปรับลักษณะทางสัทศาสตร์ของเสียงสังเคราะห์ได้ค่าเอ็มโอเอสเท่ากับ 3.07 การปรับลักษณะทางสัทศาสตร์จึงช่วยให้เสียงสังเคราะห์รับฟังเป็นธรรมชาติมากขึ้น

วิธีการสังเคราะห์พยางค์หนักเบาในวิทยานิพนธ์นี้สามารถใช้ปรับลักษณะทางสัทศาสตร์ของพยางค์ได้เป็นอย่างดี เสียงสังเคราะห์จะรับฟังได้เป็นธรรมชาติมากขึ้น แต่เสียงที่รับฟังก็อาจผิดเพี้ยนบ้าง ต้องนำความรู้ทางภาษาศาสตร์มาช่วยในการปรับปรุงให้การสังเคราะห์พยางค์หนักเบามีคุณภาพดีขึ้น

## 6.2 ข้อเสนอแนะ

1. ในการปรับลักษณะทางสัทศาสตร์ด้วยการใช้วิธี Time-Domain Pitch-Synchronous Overlap-Add (TD-PSOLA) สำหรับส่วนพยัญชนะต้นอาจทำให้เสียงที่รับฟังได้ผิดพลาดได้ เนื่องจากเสียงพยัญชนะต้นมีส่วนสำคัญในการรับรู้คำ จึงควรละเว้นไม่ปรับลักษณะทางสัทศาสตร์ของเสียงพยัญชนะต้น
2. วิเคราะห์อัตราส่วนช่วงเวลาของส่วนที่เป็นเสียงก้องและเสียงไม่ก้องในพยางค์เสียงหนักและพยางค์เสียงเบา เพื่อนำมาหาสัดส่วนในการปรับระยะเวลาให้ได้ลักษณะของเสียงสังเคราะห์ใกล้เคียงกับลักษณะของเสียงตามธรรมชาติมากยิ่งขึ้น
3. สร้างหน่วยเสียงโดยที่เก็บหน่วยเสียงในลักษณะเป็นสัญญาณช่วงสั้นซึ่งเกิดจากการวิเคราะห์สัญญาณเสียงด้วยฟังก์ชันหน้าต่าง เพราะจะช่วยให้ประหยัดเวลาในการระบุยอดพิทช์และการวิเคราะห์ด้วยฟังก์ชันหน้าต่าง สามารถดูยอดพิทช์และปรับตำแหน่งยอดพิทช์หากยอดพิทช์ระบุคลาดเคลื่อนได้ แต่จะทำให้ขนาดของหน่วยเสียงใหญ่ขึ้น
4. การทำให้ราบเรียบ (Smoothing) บริเวณรอยต่อระหว่างพยางค์ในกรณีที่เสียงตัวสะกดของพยางค์หน้าและเสียงพยัญชนะต้นของพยางค์หลังที่เป็นเสียงก้อง จะช่วยให้เสียงสังเคราะห์รับฟังเป็นธรรมชาติมากขึ้น
5. ในการปรับความถี่มูลฐาน ในขั้นตอนการเปลี่ยนจากค่าความถี่มูลฐานเป็นระยะระหว่างยอดพิทช์ ค่าระยะระหว่างยอดพิทช์ที่ได้จะต้องปัดเศษเป็นเลขจำนวนเต็ม ในวิทยานิพนธ์นี้ใช้การประมาณค่าใกล้เคียง ซึ่งในบางกรณีอาจทำให้ความถี่มูลฐานของพยางค์ที่ปรับความถี่มูลฐานมีลักษณะแกว่งเป็นลูกคลื่น ดังรูป 5.12 ข. การปัดเศษขึ้นหรือลงจะควรดูแนวโน้มของความถี่มูลฐานจุดที่ใกล้เคียงด้วย

# รายการอ้างอิง

**ภาษาไทย**


กฤษดา เรเยส. 2530. ต้นแบบเครื่องสังเคราะห์เสียงพูดด้วยวิธีเข้ารหัสแบบลิเนียร์พรีดิกทีฟ. วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

ชัชวาล หาญสกุลบันเทิง. 2542. การสังเคราะห์พยางค์ภาษาไทยด้วยวิธีการสังเคราะห์แบบวิเคราะห์โดยใช้คู่เส้นสเปกตรัม. วิทยานิพนธ์วิศวกรรมศาสตร์มหาบัณฑิต ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

ปิยฉัตร ปานโรจน์, 2534. ลักษณะเชิงกลสัทศาสตร์ของวรรณยุกต์ในภาษาไทยกรุงเทพฯ : การแปรตามกลุ่มอายุ. วิทยานิพนธ์อักษรศาสตร์มหาบัณฑิต ภาควิชาภาษาศาสตร์ คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

ผณินทรา ธีรานนท์. 2543. หน่วยจังหวะกับการแปรของวรรณยุกต์ในคำพูดต่อเนื่องภาษาไทย. วิทยานิพนธ์อักษรศาสตร์มหาบัณฑิต ภาควิชาภาษาศาสตร์ คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

วิเชียร แซ่โล้ว. 2539. อ่านคำไทยพยางค์เดี่ยวแบบปรับความเร็วได้. วิทยานิพนธ์วิศวกรรมศาสตร์มหาบัณฑิต ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

รุจนา พินิจารมณ์. 2533. ลักษณะเชิงกลสัทศาตร์ของวรรณยุกต์ในภาษาไทยกรุงเทพฯ. วิทยานิพนธ์อักษรศาสตร์มหาบัณฑิต ภาควิชาภาษาศาสตร์ คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

สุกัลยา สุรินทร์ไพบูลย์. 2528. ระบบพยางค์หนักเบาของคำหลายพยางค์ในภาษาไทย. วิทยานิพนธ์อักษรศาสตร์มหาบัณฑิต ภาควิชาภาษาศาสตร์ คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

อภิชาติ ตั้งทางธรรม. 2537. การเปลี่ยนความเร็วของเสียงพูด. การประชุมวิชาการทางวิศวกรรมไฟฟ้า ครั้งที่17.

เอกพล อนุสุเรนทร์. 2541. การสังเคราะห์เสียงวรรณยุกต์และสระในพยางค์เปิดภาษาไทย โดยใช้หน่วยเสียงอนุภาค. วิทยานิพนธ์วิศวกรรมศาสตร์มหาบัณฑิต ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

**ภาษาอังกฤษ**

Brosnahan, L. F. and B. Malmberg. 1970. Introduction to Phonetics. Cambridge: W. Heffer & Sons Ltd.

Boersma, P. Available from: http://www.fon.hum.uva.nl/praat/ [2000,October 27]

Charpentier, F. and Moulines, E. 1989. Pitch-Synchronous Waveform Processing Techniques for Text-to-Speech Synthesis using Diphones. European Conference on Speech Communication and Technology. pp. 013-019.

Charpentier, F. and Stella, M.G. 1986. Diphone Synthesis Using an Overlap-Add Technique for Speech Waveforms Concatenation. Proceedings of ICASSP'86 International Conference on Acoustic, Speech and Signal Processing. pp. 2015-2018.

Craig C. Available from: http://www.goldwave.com.

Deller Jr., J., Proakis, J. and Hansen, J. 1993. Discrete-Time Processing of Speech Signals. Prentice- Hall Inc., New Jersey.

Dubnowski J.J., Schafer R.W., and Rabiner L.R. 1976. Real-time Digital Hardware Pitch Detector. IEEE Transactions on Acoustic, Speech, and Signal Processing, Vol. ASSP-24, No.1.

Dutoit, T., 1997. An Introduction to Text-to-Speech Synthesis. Kluwar academic publishers., Dordrecht.

Fry, D. B. 1958. Experiments in the Perception of Stress, in Language and Speech 1.

Hiranburana, S. 1971. The Role of Accent in Thai Grammar. Doctor of Philosophy Thesis, University of London.

Hiranburana, S. 1972. Change in the Pitch Contours of Unaccented Syllables in Spoken Thai, In J. G. Harris and R. B. Noss (eds.). Tai Phonetics and Phonology. CIEL, Bangkok. pp. 23-27.

Luangthongkum, T. 1977. Rhythm in Standard Thai. Doctor of Philosophy Thesis, University of Edinburgh.

Luksaneeyanawin, S. 1983. Intonation in Thai. Doctor of Philosophy Thesis, University of Edinburgh.

Luksaneeyanawin, S. 1989. A Thai text to speech system. Proceedings of the Regional Workshop on Computer Processing of Asian Languages (CPAL). Asian Institute of Technology. 305-15.

Luksaneeyanawin, S. 1990. Syllable and demisyllable base speech synthesis in Thai Proceeding of the Conference on Electronics and Computer Research and Development. NECTEC, Ministry of Science and Technology Volume I. 23 pages (In Thai).

Luksaneeyanawin, S. 1991 a. Speech synthesis Technology. Paper presented at the Conference on Audio Technology. Faculty of science. Chulalongkorn University. (In Thai).

Luksaneeyanawin, S. 1991 b. Problem in Thai text to speech system. Proceeding of the Conference on Electronics and Computer Research and Development. NECTEC, Ministry of Science and Technology Vol. I. 41-62. (In Thai).

Luksaneeyanawin, S. 1993. Speech Computing and Speech Technology in Thailand. Proceedings of the Symposium on Natural Language Processing in Thailand. pp. 276-321.

Luksaneeyanawin, S. 1998. Intonation system in Thai. In D. Hirst and A. Di Cristo(eds.). Intonation Systems: A Survey of Twenty Languages. Cambridge University Press. pp. 376-394.

O'Connor , J. D. 1973. Phonetics. Hazell & Viney Ltd., Great Britain.

Rabiner, L.R. 1977. On the Use of Autocorrelation Analysis for Pitch Detection. IEEE Transactions on Acoustic, Speech, and Signal Processing, Vol. ASSP-25, No.1

Schoknecht C. 2000. Appendix: Three system of Thai Transcription. In D. Burnham, S. Luksaneeyanawin, C. Davis and M. Lafourcade(eds.). Interdisciplinary Approaches to Language Processing. NECTEC, Bangkok. pp. 329-336.

Sondhi M.M. 1968. New Methods of Pitch Extraction. IEEE Transactions on Audio and Electroacoustics, Vol. AU-16, No.2.

# ภาคผนวก

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

# ภาคผนวก ก.

## การแบ่งหน่วยจังหวะ

หน่วยจังหวะจะเริ่มต้นจากพยางค์ที่ลงเสียงหนักซึ่งอาจเป็นพยางค์ปกติที่สามารถได้ยินหรือพยางค์เงียบไปจนถึงพยางค์สุดท้ายก่อนพยางค์ที่ได้รับการลงเสียงหนักถัดไป ดังตัวอย่าง

ประโยคตัวอย่าง : ปู่ไปเที่ยวฟิลิปปินส์

| | น้อง | ไป | เที่ยว | แค | นา | ดา |
|---|---|---|---|---|---|---|
| การลงเสียงหนักเบา | ‘n@@ng3 | paj0 | ‘thiiaw2 | khxx0 | naa0 | ‘daa0 |
| | S | U | S | U | U | S |

(สัทอักษรนี้ใช้ตาม ระบบการเขียนแบบ LRU ซึ่งกล่าวไว้ใน Schoknecht, 2000))

จากตัวอย่างนี้จะทราบว่าพยางค์ที่ลงเสียงหนักได้แก่ “น้อง” (/n@@ng3/), “เที่ยว” (/thiiaw2/), “ดา” (/daa0/)

ดังนั้นหน่วยจังหวะแรกเริ่มตั้งแต่ “น้อง” (/n@@ng3/) ไปจนถึง “ไป” (/paj0/) มีสองพยางค์ หน่วยจังหวะแรกจึงเป็นหน่วยจังหวะ 1 พยางค์หนัก 1 พยางค์เบา

หน่วยจังหวะที่สองเริ่มจาก “เที่ยว” (/thiiaw2/), “แค” (/khxx0/) เละ “นา” (/naa0/) รวม 3 พยางค์ หน่วยจังหวะที่สองเป็นหน่วยจังหวะ 1 พยางค์หนัก 2 พยางค์เบา

หน่วยจังหวะที่สามเริ่มจาก “ดา” (daa0) แลไม่มีคำใดต่อ จึงเป็นหน่วยจังหวะ 1 พยางค์หนัก

ฉะนั้นเราจึงแบ่งหน่วยจังหวะได้ดังนี้

| ‘n@@ng3 paj0 | ‘thiiaw2 khxx0 naa0 | ‘daa0 |

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

# ภาคผนวก ข.

## แบบทดสอบสำหรับการประเมินผลด้วยวิธีเอ็มโอเอส



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ชื่อ วันที่

**ส่วนที่ 1** จงประเมินคำสังเคราะห์ตามระดับความพอใจของท่าน(5=ดีมาก, 4=ดี, 3=ปานกลาง, 2=แย่, 1=แย่มาก )

| ข้อที่ | คำ | คำชุดที่ 1 | | | | | คำชุดที่ 2 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 5 | 4 | 3 | 2 | 1 | 5 | 4 | 3 | 2 | 1 |
| ตัวอย่าง | คำสังเคราะห์ | | | | | | | | | | |
| 1 | คาทอลิก | | | | | | | | | | |
| 2 | ดาวกระจาย | | | | | | | | | | |
| 3 | โซปราโน | | | | | | | | | | |
| 4 | ซามูไร | | | | | | | | | | |
| 5 | เอราวัณ | | | | | | | | | | |
| 6 | จุฬาลงกรณ์ | | | | | | | | | | |
| 7 | สุริโยไท | | | | | | | | | | |
| 8 | ฟิลิปปินส์ | | | | | | | | | | |
| 9 | เจนีวา | | | | | | | | | | |
| 10 | คอมมานโด | | | | | | | | | | |
| 11 | ดาวกระจาย | | | | | | | | | | |
| 12 | เอราวัณ | | | | | | | | | | |
| 13 | สุริโยไท | | | | | | | | | | |
| 14 | คอมมานโด | | | | | | | | | | |
| 15 | โซปราโน | | | | | | | | | | |
| 16 | คาทอลิก | | | | | | | | | | |
| 17 | เจนีวา | | | | | | | | | | |
| 18 | ซามูไร | | | | | | | | | | |
| 19 | ฟิลิปปินส์ | | | | | | | | | | |
| 20 | จุฬาลงกรณ์ | | | | | | | | | | |
| 21 | คอมมานโด | | | | | | | | | | |
| 22 | โซปราโน | | | | | | | | | | |
| 23 | เจนีวา | | | | | | | | | | |
| 24 | เอราวัณ | | | | | | | | | | |
| 25 | คาทอลิก | | | | | | | | | | |
| 26 | สุริโยไท | | | | | | | | | | |
| 27 | ฟิลิปปินส์ | | | | | | | | | | |
| 28 | ซามูไร | | | | | | | | | | |
| 29 | ดาวกระจาย | | | | | | | | | | |
| 30 | จุฬาลงกรณ์ | | | | | | | | | | |

ชื่อ วันที่

**<u>ส่วนที่ 2</u>** จงประเมินคำสังเคราะห์ตามระดับความพอใจของท่าน(5=ดีมาก, 4=ดี, 3=ปานกลาง, 2=แย่, 1=แย่มาก )

| ข้อที่ | คำ | คำชุดที่ 1 | | | | | คำชุดที่ 2 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 5 | 4 | 3 | 2 | 1 | 5 | 4 | 3 | 2 | 1 |
| 1 | เพื่อนของฉันเป็นคาทอลิก | | | | | | | | | | |
| 2 | นินจาปาดาวกระจาย | | | | | | | | | | |
| 3 | เธอร้องเสียงโซปราโน | | | | | | | | | | |
| 4 | ติดตามชมซามูไรพ่อลูกอ่อน | | | | | | | | | | |
| 5 | อุทยานแห่งชาติเอราวัณ | | | | | | | | | | |
| 6 | พระคุณของแหล่งเรียนมาจุฬาลงกรณ์ | | | | | | | | | | |
| 7 | วีรสตรีศรีสุริโยไท | | | | | | | | | | |
| 8 | ปู่ไปเที่ยวฟิลิปปินส์ | | | | | | | | | | |
| 9 | สนธิสัญญาเจนีวา | | | | | | | | | | |
| 10 | เขาสังกัดหน่วยคอมมานโด | | | | | | | | | | |
| 11 | นินจาปาดาวกระจาย | | | | | | | | | | |
| 12 | อุทยานแห่งชาติเอราวัณ | | | | | | | | | | |
| 13 | วีรสตรีศรีสุริโยไท | | | | | | | | | | |
| 14 | เขาสังกัดหน่วยคอมมานโด | | | | | | | | | | |
| 15 | เธอร้องเสียงโซปราโน | | | | | | | | | | |
| 16 | เพื่อนของฉันเป็นคาทอลิก | | | | | | | | | | |
| 17 | สนธิสัญญาเจนีวา | | | | | | | | | | |
| 18 | ติดตามชมซามูไรพ่อลูกอ่อน | | | | | | | | | | |
| 19 | ปู่ไปเที่ยวฟิลิปปินส์ | | | | | | | | | | |
| 20 | พระคุณของแหล่งเรียนมาจุฬาลงกรณ์ | | | | | | | | | | |
| 21 | เขาสังกัดหน่วยคอมมานโด | | | | | | | | | | |
| 22 | เธอร้องเสียงโซปราโน | | | | | | | | | | |
| 23 | สนธิสัญญาเจนีวา | | | | | | | | | | |
| 24 | อุทยานแห่งชาติเอราวัณ | | | | | | | | | | |
| 25 | เพื่อนของฉันเป็นคาทอลิก | | | | | | | | | | |
| 26 | วีรสตรีศรีสุริโยไท | | | | | | | | | | |
| 27 | ปู่ไปเที่ยวฟิลิปปินส์ | | | | | | | | | | |
| 28 | ติดตามชมซามูไรพ่อลูกอ่อน | | | | | | | | | | |
| 29 | นินจาปาดาวกระจาย | | | | | | | | | | |
| 30 | พระคุณของแหล่งเรียนมาจุฬาลงกรณ์ | | | | | | | | | | |

# ภาคผนวก ค.

**บทความที่ได้รับการตีพิมพ์ในการประชุมวิชาการทางวิศวกรรมไฟฟ้าครั้งที่ 25**

25th Electrical Engineering Conference (EECON-25)



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

# การสังเคราะห์พยางค์เสียงหนักและพยางค์เสียงเบาในภาษาไทย
# Synthesis of Stressed and Unstressed Syllables in Thai

นัฐพล พานสมบัติ และ เอกชัย ลีลารัศมี
ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย แขวงวังใหม่ เขตปทุมวัน กรุงเทพฯ 10330
โทร. 0-2218-6537 โทรสาร 0-2218-6488 E-mail : rhythm@digital.ee.eng.chula.ac.th
สุดาพร ลักษณียนาวิน
ภาควิชาภาษาศาสตร์ คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย แขวงวังใหม่ เขตปทุมวัน กรุงเทพฯ 10330

**บทคัดย่อ**

บทความนี้นำเสนอวิธีการเปลี่ยนสัทลักษณะของพยางค์เสียงหนักและพยางค์เสียงเบาในภาษาไทย โดยใช้วิธี Time Domain Pitch Synchronous Overlapped Add(TD-PSOLA) ในการเปลี่ยนช่วงเวลาและความถี่มูลฐานของสัญญาณเสียงตามสัทลักษณะของพยางค์เสียงหนักและพยางค์เสียงเบา จะช่วยให้คำพูดที่ได้จากการสังเคราะห์เสียงฟังเป็นธรรมชาติมากขึ้น สำหรับการปรับช่วงเวลาสามารถปรับได้ทั้งเพิ่มและลดช่วงเวลา และการปรับความถี่มูลฐานสามารถปรับได้ตามฐานข้อมูลความถี่มูลฐานของพยางค์เสียงเบาที่ได้ทำไว้

**Abstract**

This paper presents acoustic characteristic modification method of stressed and unstressed syllables by using Time Domain Pitch Synchronous Overlapped Add (TD-PSOLA). Modifying time and fundamental frequency of acoustic characteristics will help the synthesized speech sounds more naturally. Additionally, duration can be increased and lessen, as well as fundamental frequency can also be modified according to unstressed syllable database .

**Keywords :** Speech Synthesis, Stressed and Unstressed Syllable, Time-Scale Modification, Pitch-Scale Modification, Time Domain Pitch Synchronous Overlap-Add

## 1. บทนำ

ในภาษาไทยนั้น มีการลงเสียงหนักเบาในคำพูดที่พูดออกมา ดังจะเห็นได้จากงานวิจัยที่ศึกษาลักษณะร่วมทางสัทศาสตร์ของพยางค์เสียงหนักและพยางค์เสียงเบาออกมา จากงานวิจัยของ สุดาพร ลักษณียนาวิน[1],[2] สุกัลยา สุรินทร์ไพบูลย์[3] และ รุจนา พินิจารมย์[4] สามารถสรุปได้ว่า การลงเสียงหนักเกี่ยวพันกับลักษณะร่วมทางสัทศาสตร์ทั้ง 4 คือ ความดัง ความยาว ระดับเสียงสูงต่ำ คุณสมบัติของสระและวรรณยุกต์

การลงเสียงหนักเบานั้น นอกจากจะทำให้คำพูดที่ออกมาดูเป็นธรรมชาติแล้ว ยังเป็นส่วนสำคัญในการเข้าใจความหมายของคำพูดที่พูดออกมาอีกด้วย เช่น คำว่า “เสือหมอบ” ในประโยคว่า “เสือหมอบอยู่ใต้ต้นไม้” ความหมายแรกหมายถึง เสือ ทำกริยา หมอบ อยู่ใต้ต้นไม้ ในความหมายนี้จะลงเสียงหนักทั้งคำว่า ”เสือ” และคำว่า ”หมอบ” อีกความหมาย หมายถึง จักรยานเสือหมอบอยู่ที่ใต้ต้นไม้ ความหมายนี้จะลงเสียงหนักที่คำว่า ”หมอบ” ส่วนคำว่า “เสือ” จะลงเสียงเบา

ดังนั้นบทความนี้จึงนำเสนอการสังเคราะห์พยางค์เสียงหนักและพยางค์เสียงเบา สำหรับคำพยางค์เดี่ยวที่ได้จากการสังเคราะห์ เพื่อให้คำจากการสังเคราะห์รับฟังเป็นธรรมชาติมากขึ้นและเข้าใจไม่ผิดพลาด

## 2. ทฤษฎีทางภาษาศาสตร์

### 2.1 ระบบเสียงหนักเบาในภาษาไทย

ในการพูดแต่ละคำมีการลงเสียงหนักหรือลงเสียงเบา โดยมีเกณฑ์การกำหนดตำแหน่งของพยางค์เสียงหนักในคำพยางค์เดียว จะมีกฎการลงเสียงหนัก ดังนี้

คำหลัก(Content Word) ได้แก่ คำนาม คำกริยา คำวิเศษณ์ และคำคุณศัพท์ จะลงเสียงหนักในการพูดปกติ

คำไวยากรณ์(Grammatical Word) ได้แก่คำประเภทอื่นที่ไม่ใช่คำเนื้อหา จะลงเสียงเบาในการพูดปกติ ยกเว้น ในการพูดเพียงลำพังหรือมีการเน้น

ตัวอย่างการลงเสียงหนักของคำพยางค์เดียว ยกตัวอย่างคำว่า”กัน”

| | | | | |
|---|---|---|---|---|
| - | อย่า | กัน | ดี | กว่า |
| | ja:1 | kan0 | di:0 | khwa:1 |
| | (ปฏิเสธ) | (กริยา) | (วิเศษณ์) | |
| | U | S | U | S |
| - | หย่า | กัน | ดี | กว่า |
| | ja:1 | kan0 | di:0 | khwa:1 |
| | (กริยา) | (สรรพนามแสดงการปฏิภาค) | (วิเศษณ์) | |
| | S | U | U | S |

โดย S แทน การลงเสียงหนัก , U แทน การลงเสียงเบา

### 2.2 ลักษณะร่วมทางโสตสัทศาสตร์ของพยางค์เสียงหนักและพยางค์เสียงเบาในภาษาไทย

#### 2.2.1 ระดับเสียงสูงต่ำ(Pitch)

สุดาพร ลักษณียนาวิน[1],[2] สุกัลยา สุรินทร์ไพบูลย์[3] และ รุจนา พินิจารมย์[4] ได้ศึกษาการเปลี่ยนแปลงของระดับเสียงสูงต่ำในพยางค์เสียงเบาในพยางค์ 2 ประเภท สรุปได้ดังนี้

ประเภทพยางค์เชื่อม

เสียงสูง เมื่อไม่ลงเสียงหนักส่วนมากจะปรากฏเป็นเสียงกลางระดับ แต่บางส่วนปรากฏเป็นเสียงสูงระดับ

เสียงต่ำ เมื่อไม่ลงเสียงหนักส่วนมากจะปรากฏเป็นเสียงกลางระดับ แต่บางส่วนปรากฏเป็นเสียงต่ำระดับ

ประเภทพยางค์ไม่เชื่อม

เสียงตก(เสียงโท) เมื่อไม่ลงเสียงหนักจะปรากฏเป็นได้ทั้งเสียงสูงระดับและเสียงสูงตกที่มีช่วงการตกสั้นลง

เสียงขึ้น(เสียงจัตวา)เมื่อไม่ลงเสียงหนักจะปรากฏเป็นเสียงขึ้น

เสียงสูง(เสียงตรี) เมื่อไม่ลงเสียงหนักส่วนมากจะปรากฏเป็นเสียงสูงระดับ แต่บางส่วนปรากฏเป็นเสียงกลางระดับ

เสียงกลาง(เสียงสามัญ) เมื่อไม่ลงเสียงหนักจะปรากฏเป็นเสียงกลางระดับ

เสียงต่ำ(เสียงเอก) เมื่อไม่ลงเสียงหนักจะปรากฏเป็นเสียงต่ำระดับ แต่บางส่วนปรากฏเป็นเสียงกลางระดับ

### 2.2.2 ความยาว(Length)

ธีรพันธ์ เหลืองทองคำ[5] ได้ทำการศึกษาความยาวของช่วงเวลาโดยการวัดค่าระยะเวลา แบ่งหน่วยเวลาตามจำนวนพยางค์ในหน่วยจังหวะ ตั้งแต่ 1 พยางค์หนัก( | S | ) จนถึง 1 พยางค์หนัก 3 พยางค์เบา( | S U U U| ) ดังนี้

1. $|S| = 3$
2. $|S\ U| = 2:1$
3. $|S\ U\ U| = 1\frac{1}{2}:\frac{3}{4}:\frac{3}{4}$
4. $|S\ U\ U\ U| = 1:\frac{2}{3}:\frac{2}{3}:\frac{2}{3}$

เมื่อ | | แทน ขอบเขตของหน่วยจังหวะ S แทน พยางค์ที่ลงเสียงหนัก U แทน พยางค์ที่ลงเสียงเบา

ช่วงเวลาของหน่วยจังหวะที่จำนวนพยางค์ไม่เท่ากัน โดยมีค่าเฉลี่ยเวลาของโครงสร้างหน่วยจังหวะแต่ละแบบ และค่าเฉลี่ยช่วงเวลาของพยางค์ในโครงสร้างหน่วยจังหวะแต่ละแบบ แสดงดังนี้

$|S| = 0.30$ sec.

$|S\ U| = 0.40$ sec. หรือ $|0.25 : 0.15|$

$|S\ U\ U| = 0.49$ sec. หรือ $|0.23 : 0.13 : 0.13|$

$|S\ U\ U\ U| = 0.59$ sec. หรือ $|0.21 : 0.12 : 0.13 : 0.13|$

## 3. ฐานข้อมูลความถี่ของพยางค์เสียงเบา

ในการเปลี่ยนความถี่มูลฐานในพยางค์เสียงเบาจะเปลี่ยนโดยใช้ฐานข้อมูลที่ได้สร้างไว้ โดยเก็บข้อมูลความถี่มูลฐานตามโครงสร้างพยางค์สำหรับวรรณยุกต์ต่างๆดังนี้

3.1 วรรณยุกต์สามัญ ใน พยางค์เป็น

3.1 วรรณยุกต์เอก ใน พยางค์เป็น พยางค์ตายสระเสียงยาว พยางค์ตายสระเสียงสั้นที่ลงท้ายด้วยเสียงกักที่เส้นเสียงและเสียงกักอื่นๆ

3.2 วรรณยุกต์โท ใน พยางค์เป็นและพยางค์ตายสระเสียงยาว

3.4 วรรณยุกต์ตรี ใน พยางค์เป็น พยางค์ตายสระเสียงสั้นที่ลงท้ายด้วยเสียงกักที่เส้นเสียง และเสียงกักอื่นๆ

3.5 วรรณยุกต์จัตวา ใน พยางค์เป็น

คำตัวอย่างมีโครงสร้างพยางค์ต่างๆกัน โดยมีตัวอย่างโครงสร้างละ 3 คำ เลือกประโยคที่ชุดคำนี้ลงเสียงเบาไปบันทึกเสียง และวิเคราะห์หาความถี่มูลฐาน จากนั้นนำความถี่มูลฐานที่ได้มาแบ่งเป็น 100% ตามช่วงเวลา จากนั้นนำความถี่มูลฐานของแต่ละโครงสร้างพยางค์มาหาค่าเฉลี่ยเพื่อเก็บเป็นฐานข้อมูลของพยางค์เสียงเบาสำหรับโครงสร้างพยางค์นั้นๆต่อไป

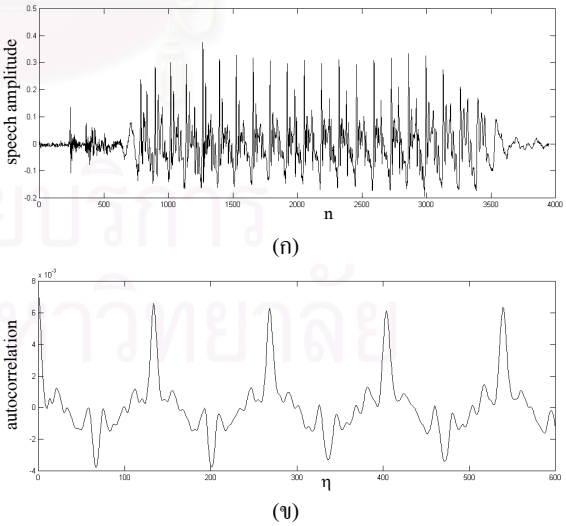
## 4. วิธีการในการสังเคราะห์เสียงหนักเบา

### 4.1 การระบุพิทช์(Pitch marking)

การระบุยอดพิทช์(pitch marking)ใช้สำหรับการวิเคราะห์แบบหน้าต่างให้เป็นสัญญาณช่วงสั้น เพื่อที่จะปรับช่วงเวลาและระดับเสียงสูงต่ำของสัญญาณเสียง โดยใช้การหาค่าอัตสหสัมพันธ์(Autocorrelation) หาตำแหน่งยอดพิทช์ โดยค่าอัตสหสัมพันธ์สามารถหาได้ดังสมการ

$$r_x(\eta;m) = \frac{1}{N}\sum_{n=-\infty}^{\infty} s(n)w(m-n)s(n-|\eta|)w(m-n+|\eta|) \qquad (1)$$

ในการหาค่าอัตสหสัมพันธ์ของสัญญาณเสียง จะได้ค่าอัตสหสัมพันธ์สูงสุดที่ค่า η เป็นจำนวนเท่าของคาบ[6] ดังรูปที่ 1



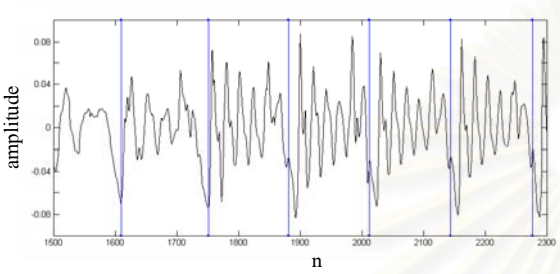

รูปที่ 1 ค่าอัตสหสัมพันธ์ ของสัญญาณเสียงคำว่า “กา” (kaa0) ณ จุดตัวอย่างที่ 2055(m=2055) (ก)สัญญาณเสียง (ข)ค่าอัตสหสัมพันธ์

จากรูปที่1 รูปที่1(ก)แสดงสัญญาณเสียงคำว่า “กา”(kaa0) รูปที่1(ข) ได้แสดงค่าอัตสหสัมพันธ์ของสัญญาณเสียง ณ จุดตัวอย่างที่ 2055

ค่าอัตสหสัมพันธ์จะมีลักษณะซ้ำเป็นคาบ และระยะห่างระหว่างค่าสูงสุดจะเท่ากัน ซึ่งเท่ากับคาบของสัญญาณเสียง

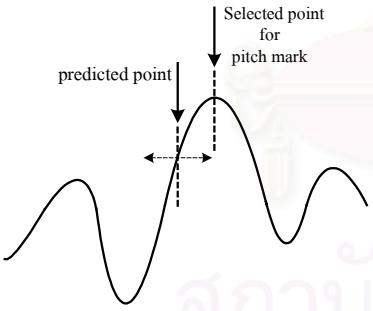
ด้วยคุณสมบัตินี้จึงใช้การหาค่าอัตสหสัมพันธ์ ช่วยในการระบุพิทช์ เมื่อได้ตำแหน่งยอดพิทช์ถัดไป จะใช้ยอดพิทช์นั้นเป็นจุดอ้างอิงสำหรับหาค่าอัตสหสัมพันธ์ของยอดพิทช์ถัดไป

แต่เมื่อใช้การหาค่าอัตสหสัมพันธ์หายอดพิทช์ต่อไปหลายครั้งยอดพิทช์ที่ได้อาจเกิดการผิดพลาดจากตำแหน่งยอดพิทช์จริง การหาค่าอัตสหสัมพันธ์ไม่สามารถหายอดพิทช์ต่อไปได้พอดีดังรูปที่ 2




รูปที่ 2 ความผิดพลาดในการหายอดพิทช์โดยการหาค่าอัตสหสัมพันธ์เมื่อทำซ้ำต่อไปหลายครั้ง

ในการหาค่าอัตสหสัมพันธ์ประมาณตำแหน่งยอดพิทช์ถัดไปจะใช้ยอดพิทช์ที่ได้เป็นจุดหลักโดยพิจารณาจุดโดยรอบจุดนั้นในช่วงเวลาที่กำหนดขึ้นแล้วหาจุดที่มีค่ามากที่สุด เพื่อใช้เป็นจุดอ้างอิงสำหรับหายอดพิทช์ต่อไป ดังรูปที่ 3

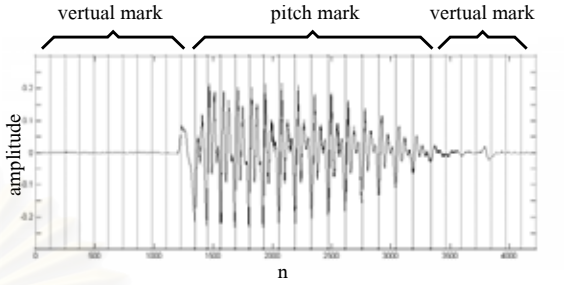



รูปที่ 3 การเลือกยอดพิทช์ จากยอดพิทช์ที่ประมาณได้จากการหาค่าอัตสหสัมพันธ์

### 4.2 การระบุพิทช์แบบเสมือน(Virtual pitch marking)

สัญญาณเสียงประกอบด้วย ส่วนที่เป็นเสียงก้อง และเสียงไม่ก้อง เสียงไม่ก้องนั้นจะอยู่ที่ต้นพยางค์และที่ท้ายพยางค์ การระบุพิทช์เป็นการหายอดพิทช์ในส่วนของพยางค์ที่เป็นเสียงก้องเท่านั้น ซึ่งเพียงพอในการเปลี่ยนระดับเสียงสูงต่ำ แต่สำหรับการเปลี่ยนช่วงเวลาของพยางค์นั้น จะต้องเพิ่มหรือลดสัญญาณทั้งในส่วนที่เป็นเสียงก้องและเสียงไม่ก้อง แต่ในส่วนของเสียงไม่ก้องไม่มีการระบุพิทช์ จึงต้องทำการระบุพิทช์แบบเสมือนสำหรับใช้ในการลดหรือเพิ่มเวลาของสัญญาณเสียง

สำหรับส่วนที่เป็นเสียงก้อง ระยะห่างระหว่างจุดระบุจะเท่ากับระยะระหว่างยอดพิทช์ ดังนั้นเราจะใช้ค่าเฉลี่ยระยะห่างระหว่างยอดพิทช์ในการระบุพิทช์แบบเสมือน เพื่อใช้ในการเพิ่มหรือลดค่าเวลาของสัญญาณ




รูปที่ 4 ยอดพิทช์และยอดพิทช์เสมือนในสัญญาณเสียง
(รูปนี้ ระบุยอดพิทช์ที่ยอดล่างของสัญญาณ)

### 4.3 การวิเคราะห์แบบหน้าต่าง(Windowing)

การวิเคราะห์แบบหน้าต่างเป็นการวิเคราะห์สัญญาณเสียงโดยกำหนดกรอบสัญญาณเป็นช่วงสั้นๆเท่าๆกัน โดยการคูณสัญญาณเสียงด้วยฟังก์ชันหน้าต่าง สัญญาณช่วงสั้นที่ได้จะสามารถนำไปวิเคราะห์ต่อไปได้

ฟังก์ชันหน้าต่างที่เลือกใช้คือ หน้าต่างแฮนนิ่ง (Hanning window) หน้าต่างแฮนนิ่ง w(n) มีค่าดังสมการที่ 2

$$w(n) = \begin{cases} 0.5\times\left[1-\cos\left(\frac{2\pi n}{N-1}\right)\right] & , \quad n = 0, 1, 2, \ldots, N-1 \\ 0 & , \quad n \quad \text{ค่าอื่นๆ} \end{cases} \qquad (2)$$

### 4.4 วิธีเปลี่ยนลักษณะทางกลสัทศาสตร์ของพยางค์

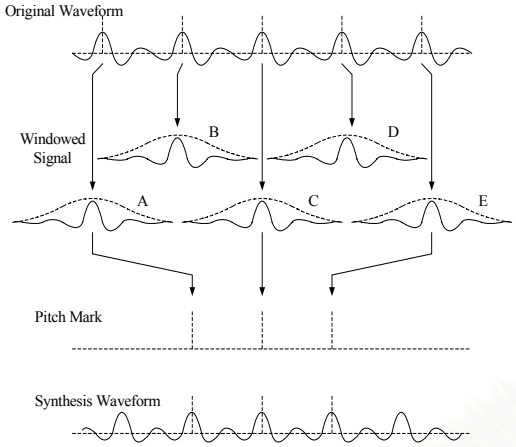
การเปลี่ยนลักษณะทางกลสัทศาสตร์ จะใช้วิธี Time Domain Pitch Synchronized Overlapped Add (TD-PSOLA)[7] ซึ่งเป็นวิธีที่ใช้ในการปรับทำนองเสียงโดยการปรับตำแหน่งพิทช์และสามารถปรับช่วงเวลาโดยการทำซ้ำหรือลบสัญญาณออก

#### 4.4.1 การเปลี่ยนช่วงเวลา

##### 4.4.1.1 การลดช่วงเวลา

การลดช่วงเวลาทำได้โดยตัดสัญญาณช่วงสั้นบางส่วนและยอดพิทช์ของสัญญาณนั้นออก ดังแสดงในรูปที่ 5

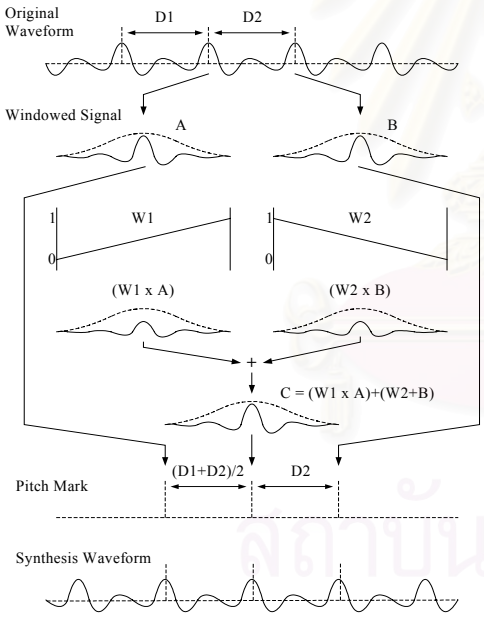
รูปนี้แสดงให้เห็นว่าสัญญาณเสียงที่วิเคราะห์แบบหน้าต่างจะเป็นสัญญาณช่วงสั้นเช่น สัญญาณช่วงสั้น A, B, C, D และ E เราสามารถลดช่วงเวลาโดยตัดสัญญาณช่วงสั้น B และ D ออก แล้วรวมสัญญาณช่วงสั้นกลับมาใหม่โดยการบวกซ้อน(Overlap-Add) ที่ตำแหน่งยอดพิทช์




รูปที่ 5 การลดช่วงเวลา

#### 4.4.1.2 การเพิ่มช่วงเวลา

การเพิ่มช่วงเวลาทำได้โดยแทรกสัญญาณช่วงสั้นในสัญญาณเดิม โดยสัญญาณที่แทรกจะสร้างขึ้นโดยวิธีการเปลี่ยนความเร็วเสียงพูด (อภิชาต[8]) และ สร้างยอดพิทช์ขึ้นใหม่

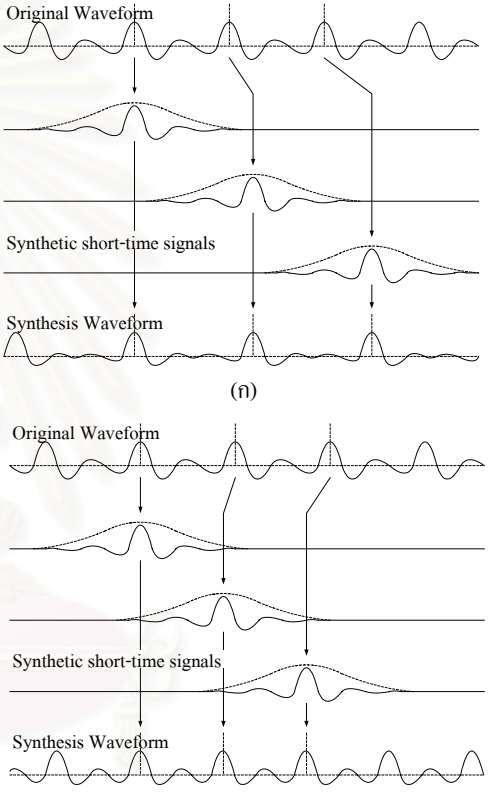



รูปที่ 6 การเพิ่มช่วงเวลา

รูปที่ 6 แสดงให้เห็นว่าสัญญาณเสียงถูกวิเคราะห์แบบหน้าต่างจะเป็นสัญญาณช่วงสั้นเช่น สัญญาณช่วงสั้น A, B เราสามารถเพิ่มช่วงเวลาโดยแทรกสัญญาณช่วงสั้น C ซึ่งสร้างจากสัญญาณช่วงสั้น A และB โดยคูณสัญญาณช่วงสั้น A และ B ด้วยตัวน้ำหนักดังรูปที่ 6 ได้เป็น (W1xA) และ (W2xB) นำสัญญาณที่ได้มาบวกกัน จะได้สัญญาณช่วงสั้น C และพิทช์สำหรับสัญญาณช่วงสั้น C หาได้โดยนำระยะระหว่างยอดพิทช์ของสัญญาณช่วงสั้น A กับสัญญาณช่วงสั้นก่อนหน้าสัญญาณช่วงสั้น A และระยะระหว่างยอดพิทช์ของสัญญาณช่วงสั้น B กับสัญญาณช่วงสั้น A มาหาค่าเฉลี่ยกัน และรวมสัญญาณช่วงสั้นกลับมาใหม่โดยการบวกซ้อน(Overlap-Add) ที่ตำแหน่งยอดพิทช์ของแต่ละสัญญาณช่วงสั้น

#### 4.4.2 การเปลี่ยนระดับเสียงสูงต่ำ

การเปลี่ยนระดับเสียงสูงต่ำเป็นการปรับความถี่มูลฐานของสัญญาณเสียงสามารถกระทำได้โดยการเปลี่ยนระยะเวลาระหว่างยอดพิทช์ เมื่อเพิ่มระยะเวลาระหว่างยอดพิทช์จะทำให้ความถี่ลดลง ทำให้เสียงต่ำลง แต่เมื่อลดระยะระหว่างยอดพิทช์จะทำให้ความถี่เพิ่มขึ้น ทำให้เสียงสูงขึ้น ดังแสดงในรูปที่ 7




รูปที่ 7 การเปลี่ยนแปลงระดับเสียงสูงต่ำของสัญญาณเสียงโดยใช้วิธี TD-PSOLA (ก) การลดความถี่มูลฐาน (ข) การเพิ่มความถี่มูลฐาน

## 5. การทดสอบการฟังเสียงสังเคราะห์

การทดสอบการฟังเสียงสังเคราะห์ใช้คำทดสอบจากการสังเคราะห์ด้วยหน่วยเสียงแบบอัตเพยางค์(Demisyllables) โดยจะเปรียบเทียบคำสังเคราะห์โดยแบ่งเป็นกรณี ดังนี้

5.1 คำสังเคราะห์ไม่ปรับลักษณะร่วมทางกลสัทศาสตร์(W0)

5.2 คำสังเคราะห์ที่ปรับลักษณะร่วมทางกลสัทศาสตร์(W1)

5.3 คำสังเคราะห์เมื่ออยู่ในประโยคสังเคราะห์ที่ไม่ปรับลักษณะร่วมทางกลสัทศาสตร์(S0)

5.4 คำสังเคราะห์เมื่ออยู่ในประโยคสังเคราะห์ที่ปรับลักษณะร่วมทางกลสัทศาสตร์(S1)

แต่ละกรณีใช้คำทดสอบ 10 คำ โดยแต่ละคำมีตัวอย่างคำทดสอบ 3 ชุด ชุดละ 30 คำ รวม 4 ชุด เป็นจำนวนตัวอย่างคำทดสอบทั้งหมด 120 คำ โดยให้กลุ่มตัวอย่างประเมินระดับความพอใจจากการรับฟังคำทดสอบ(5 = ดีมาก, 4 = ดี, 3 = ปานกลาง, 2 = แย่, 1 = แย่มาก)โดยใช้กลุ่มตัวอย่าง 9 คนได้ผลสรุปดังตารางที่ 1

ตารางที่ 1 ผลการทดสอบระดับความพอใจของผู้ฟัง

| กรณีของคำทดสอบ | W0 | W1 | S0 | S1 |
|---|---|---|---|---|
| คะแนนรวมเฉลี่ย | 3.29 | 3.00 | 3.37 | 3.54 |
| คะแนนเฉลี่ยของผู้เข้าร่วมทดสอบ สูงสุด-ต่ำสุด | 4.1 - 2.2 | 3.8 - 1.8 | 4.2 - 2.5 | 4.6 - 2.7 |

จากผลการทดลองจะพบว่าการปรับลักษณะร่วมทางกลสัทศาสตร์ในระดับคำ ไม่ได้ช่วยให้การรับฟังดีขึ้น แต่ในระดับประโยค การปรับลักษณะร่วมทางกลสัทศาสตร์ จะให้ผลที่ดีกว่าทั้งคะแนนรวมเฉลี่ย และค่าสูงสุด-ต่ำสุด ของคะแนนเฉลี่ยของผู้เข้าร่วมทดสอบ เนื่องจากในระดับประโยค การปรับลักษณะร่วมมีผลต่อท่วงทำนองเสียงของประโยค

## 6. สรุป

บทความนี้นำเสนอวิธีการสังเคราะห์พยางค์เสียงหนักและพยางค์เสียงเบาด้วยการปรับลักษณะทางกลศาสตร์อันได้แก่ ช่วงเวลาและระดับเสียงสูงต่ำ โดยใช้วิธี Time Domain Pitch Synchronized Overlapped Add(TD-PSOLA) เป็นวิธีการหลักในการปรับลักษณะร่วมทางกลสัทศาสตร์นี้ ซึ่งช่วยให้เสียงจากการสังเคราะห์เป็นประโยครับฟังได้เป็นธรรมชาติและผู้รับฟังพึงพอใจมากขึ้น

นอกจากช่วงเวลาและระดับเสียงสูงต่ำแล้ว ยังมีลักษณะร่วมทางกลสัทศาสตร์ที่เกี่ยวกับพยางค์เสียงหนักและพยางค์เสียงเบาอื่นๆอีก คือ ความดังและคุณสมบัติของสระและวรรณยุกต์ จึงควรมีการทำวิจัยเรื่องนี้ต่อไปเพื่อให้คุณภาพเสียงจากการสังเคราะห์ดียิ่งขึ้น

## 7. กิตติกรรมประกาศ

งานวิจัยนี้ได้รับทุนสนับสนุนจาก โครงการทุนบัณฑิตศึกษาภายในประเทศ สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ และกองทุนรัชดาภิเษกสมโภช จุฬาลงกรณ์มหาวิทยาลัย

## เอกสารอ้างอิง

[1] Luksaneeyanawin, S. “Intonation in Thai.” Doctor of Philosophy Thesis, University of Edinburgh, 1983.

[2] Luksaneeyanawin, S. “Intonation in Thai.” D.Hirst and A.Di Cristo(eds). Intonation Systems: A Survey of Twenty Languages. Cambridge University Press. pp 376-394.

[3] สุกัลยา สุรินทร์ไพบูลย์. ระบบพยางค์หนักเบาของคำหลายพยางค์ในภาษาไทย. วิทยานิพนธ์อักษรศาสตร์มหาบัณฑิต จุฬาลงกรณ์มหาวิทยาลัย,2528

[4] รุจนา พินิจารมณ์ ลักษณะเชิงกลสัทศาสตร์ของวรรณยุกต์ในภาษาไทยกรุงเทพฯ. วิทยานิพนธ์อักษรศาสตร์มหาบัณฑิต จุฬาลงกรณ์มหาวิทยาลัย,2533

[5] Luangthongkum, T. “Rhythm in Standard Thai.” Doctor of Philosophy Thesis, University of Edinburgh, 1977.

[6] Deller Jr.,J., Proakis,J. and Hansen,J. Discrete-Time Processing of Speech Signals. Prentice- Hall Inc., New Jersey, 1993

[7] Charpentier,F. and Moulines,E. “Pitch-Synchronous Waveform Processing Techniques for Text-to-Speech Synthesis Using Diphones.” Eurospeech’89, European Conference on Speech Communication and Technology, Paris,1989

[8] อภิชาติ ตั้งทางธรรม. การเปลี่ยนความเร็วของเสียงพูด. การประชุมวิชาการทางวิศวกรรมไฟฟ้า ครั้งที่17,2537

นัฐพล พานสมบัติ สำเร็จการศึกษาระดับปริญญาตรีสาขาวิศวกรรมไฟฟ้า จากจุฬาลงกรณ์มหาวิทยาลัย ในปี พ.ศ. 2542 ปัจจุบันกำลังศึกษาต่อในระดับปริญญาโท ณ สถาบันเดียวกัน

เอกชัย ลีลารัศมี สำเร็จการศึกษาระดับปริญญาตรีสาขาวิศวกรรมไฟฟ้า จากจุฬาลงกรณ์มหาวิทยาลัย ในปี พ.ศ. 2517 ในปี พ.ศ. 2519 ถึง พ.ศ. 2525 ได้รับทุนอานันทมหิดล เพื่อไปศึกษาต่อในระดับปริญญาโทและเอก ณ University of California at Berkeley ประเทศสหรัฐอเมริกา ปัจจุบันดำรงตำแหน่งรองศาสตราจารย์ ประจำภาควิชาวิศวกรรมไฟฟ้า จุฬาลงกรณ์มหาวิทยาลัย งานวิจัยที่อยู่ในความสนใจคือ โปรแกรมจำลองการทำงานของวงจรไฟฟ้า, การสังเคราะห์เสียงภาษาไทย, ระบบคำบรรยายภาพในเครื่องรับโทรทัศน์ และการออกแบบวงจรรวม

สุดาพร ลักษณียนาวิน สำเร็จการศึกษาระดับปริญญาตรีและโท จากคณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในปี พ.ศ. 2511 และ 2512 ได้รับทุน Ford Foundation ไปศึกษาที่ University of Leeds ในสาขาสัทศาสตร์ภาษาอังกฤษ ต่อมาสำเร็จการศึกษาระดับปริญญาเอกสาขาภาษาศาสตร์ จาก University of Edinburge เป็น Postdoctoral Fellow ทาง Speech Computing and Speech Technology ที่มหาวิทยาลัยเดียวกัน ปัจจุบันดำรงตำแหน่งผู้อำนวยการศูนย์วิจัยการประมวลผลภาษาและวัจนะ คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

# ประวัติผู้เขียนวิทยานิพนธ์

นายนัฐพล พานสมบัติ เกิดวันที่ 30 พฤศจิกายน พ.ศ. 2520 ที่อำเภอเมืองนนทบุรี จังหวัดนนทบุรี สำเร็จการศึกษาปริญญาตรีวิศวกรรมศาสตรบัณฑิต สาขาวิศวกรรมไฟฟ้า จากคณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2541 และเข้าศึกษาต่อในหลักสูตรวิศวกรรมศาสตรมหาบัณฑิต สาขาวิศวกรรมไฟฟ้า(แขนงวิชาระบบเชิงเลข) ที่คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย เมื่อ พ.ศ. 2542 ในระหว่างการศึกษาได้รับรางวัลรองชนะเลิศอันดับสองจากการประกวดแข่งขันออกแบบวงจรรวมแห่งประเทศไทยครั้งที่ 1 ประเภทดิจิทัล ซึ่งจัดขึ้นโดยศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ และเคยได้รับการตีพิมพ์บทความทางวิชาการในการประชุมวิชาการทางวิศวกรรมไฟฟ้า ครั้งที่ 25

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย