ระบบสอบถามรายนามผู้ใช้โทรศัพท์แบบอัตโนมัติโดยใช้ระบบรู้จำเสียงพูดและสังเคราะห์เสียงพูดภาษาไทย

นาย ทักษพล พึ่งยนต์

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2546
ISBN 974-17-4853-1
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

AUTOMATIC TELEPHONE DIRECTORY ASSISTANCE SYSTEM USING THAI SPEECH RECOGNITION AND SYNTHESIS

Mr.Taksapol Pungyont

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Computer Science
Department of Computer Engineering
Faculty of Engineering
Chulalongkorn University
Academic Year 2003

ISBN 974-17-4853-1

หัวข้อวิทยานิพนธ์	ระบบสอบถามรายนามผู้ใช้โทรศัพท์แบบอัตโนมัติโดยใช้ระบบรู้จำเสียงพูดและสังเคราะห์เสียงพูดภาษาไทย
โดย	นายทักษพล พึ่งยนต์
สาขาวิชา	วิทยาศาสตร์คอมพิวเตอร์
อาจารย์ที่ปรึกษา	ผู้ช่วยศาสตราจารย์ ดร. บุญเสริม กิจศิริกุล

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

.....คณบดีคณะวิศวกรรมศาสตร์

(ศาสตราจารย์ ดร.ดิเรก ลาวัณย์ศิริ)

คณะกรรมการสอบวิทยานิพนธ์

.....ประธานกรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.พรศิริ หมื่นไชยศรี)

..... อาจารย์ที่ปรึกษา

(ผู้ช่วยศาสตราจารย์ ดร.บุญเสริม กิจศิริกุล)

..... กรรมการ

(อาจารย์ ดร.ธรรถสิทธิ์ สุรฤกษ์)

..... กรรมการ

(อาจารย์ ดร.วิษณุ โคตรจรัส)

..... กรรมการ

(อาจารย์ ดร. เศรษฐา ปานงาม)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ทักษพล พึ่งยนต์ : ระบบสอบถามรายนามผู้ใช้โทรศัพท์แบบอัตโนมัติโดยใช้ระบบรู้จำเสียงพูดและสังเคราะห์เสียงพูดภาษาไทย. (AUTOMATIC TELEPHONE DIRECTORY ASSISTANCE SYSTEM USING THAI SPEECH RECOGNITION AND SYNTHESIS) อาจารย์ที่ปรึกษา : ผศ.ดร.บุญเสริม กิจศิริกุล, 89หน้า. ISBN 974-17-4853-1

วิทยานิพนธ์ฉบับนี้นำเสนอระบบสอบถามรายนามผู้ใช้โทรศัพท์แบบอัตโนมัติ โดยใช้เทคนิคของการรู้จำเสียงพูดและสังเคราะห์เสียงพูดภาษาไทยมาประยุกต์ใช้ในการพัฒนาระบบเพื่อให้ผู้ใช้งานได้เกิดความพึงพอใจและความสะดวกรวดเร็วในการให้บริการเพิ่มมากขึ้น ส่วนประกอบสำคัญของระบบประกอบด้วย ส่วนการรู้จำเสียงพูดตัวเลขต่อเนื่องภาษาไทยระดับพยางค์ และส่วนของการสังเคราะห์เสียงพูดชื่อและนามสกุลของผู้ใช้โทรศัพท์ภาษาไทย

ในขั้นตอนการตัดแบ่งพยางค์อัตโนมัติของเสียงพูดต่อเนื่อง ได้นำหลักเกณฑ์ของค่าพลังงานของเสียงและค่าอัตราการตัดผ่านระดับกำหนดมาใช้เป็นเกณฑ์ในการแบ่งพยางค์ ส่วนลักษณะสำคัญทางสวนศาสตร์ที่นำมาใช้ได้แก่ ราสต้า-พีแอลพี และอนุพันธ์อันดับที่หนึ่ง และเทคนิคการเรียนรู้ที่ใช้ในการรู้จำคือ ข่ายงานระบบประสาทเทียม ซึ่งใช้การฝึกแบบแพร่กระจายความผิดพลาดย้อนกลับ ส่วนของการสังเคราะห์เสียงพูดใช้วิธีการนำหน่วยเสียงย่อยที่ทำการเก็บไว้ในพจนานุกรมหน่วยเสียง แล้วนำมาต่อกันเป็นเสียงพูดชื่อและนามสกุล และได้นำเอาทฤษฎีเซตวิภัชนัยมาช่วยคำนวณหาค่าความใกล้เคียงของคำพ้องเสียง เพื่อลดจำนวนคำศัพท์ที่เพิ่มขึ้นในพจนานุกรมหน่วยเสียง ผลการทดลองของทั้งระบบปรากฏว่าให้ค่าความถูกต้องของระบบ 75 เปอร์เซ็นต์ และให้ค่าความถูกต้องระดับพยางค์ 94 เปอร์เซ็นต์

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาควิชา วิศวกรรมคอมพิวเตอร์ ลายมือชื่อนิสิต .....
สาขาวิชา วิทยาศาสตร์คอมพิวเตอร์ ลายมือชื่ออาจารย์ที่ปรึกษา.....
ปีการศึกษา 2546 ลายมือชื่ออาจารย์ที่ปรึกษาร่วม.....

4471416621 : MAJOR COMPUTER SCIENCE

KEY WORD : TELEPHONE DIRECTORY ASSISTANCE/.SPEECH RECOGNITION / NEURAL NETWORK / SPEECH SYNTHESIS / FUZZY SET

TAKSAPOL PUNGYONT : AUTOMATIC TELEPHONE DIRECTORY ASSISTANCE USING THAI SPEECH RECOGNITION AND SYNTHESIS. THESIS ADVISOR : ASST. PROF. BOONSERM KIJSIRIKUL, Ph.D., 89 pp. ISBN: 974-17-4853-1.

This thesis presents an Automatic Telephone Directory Assistance System which uses speech recognition and speech synthesis technique to increase satisfaction and convenience of the user. The system consists of speech recognition, based on connected word models, and speech synthesis of the Thai telephone directory.

The automatic continuous speech segmentation method uses Energy and Band Crossing Rate of sounds to segment syllables. The Rasta-PLP coefficients and its delta of the syllable are used as an acoustic feature vector. Artificial Neural Networks (ANNs) with the error-back propagation algorithm are employed for training the system. In the part of speech synthesis, Concatenation Synthesis method is employed. The Fuzzy Set is used for matching the approximated sounds. Experimental results show that the accuracies of 75% and 94% are obtained for the automatic telephone directory assistance system and for syllables , respectively.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Department Computer Engineering Student's signature……

Field of study Computer Science Advisor's signature……

Academic year 2003 Co-advisor's signature……..

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ ด้วยความช่วยเหลืออย่างดียิ่งของผู้ช่วยศาสตราจารย์ ดร.บุญเสริม กิจศิริกุล อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่ได้สละเวลาให้คำปรึกษา แนะนำแนวทางเกี่ยวกับงานวิจัยอย่างดีตลอดมาจนสำเร็จ และผู้วิจัยขอกราบขอบพระคุณ อาจารย์ ดร.ณัฐกร ทับทอง อาจารย์ประจำภาควิชาฟิสิกส์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ซึ่งเป็นผู้แนะนำแนวทางแก้ไขปัญหาที่เกิดขึ้นระหว่างการดำเนินงานวิจัย ผู้วิจัยขอกราบขอบพระคุณคณะกรรมการสอบวิทยานิพนธ์ทุกท่านที่ได้ให้คำแนะนำ ข้อคิดเห็นรวมถึงข้อเสนอแนะและแนวทางในการพัฒนางานวิจัยนี้

ขอขอบพระคุณ คุณอัจจิมา ตันสกุล สำหรับการบันทึกเสียงพูดของพจนานุกรมเพิ่มเติมและความเชื่อมั่นในความสำเร็จ ขอขอบคุณ คุณเพ็ญนที โรหิตเสถียร ในการสร้างคำอ่านชื่อและนามสกุลของรายนามผู้ใช้โทรศัพท์ และกำลังใจที่มีให้ผู้วิจัยมาโดยตลอด ขอขอบคุณเพื่อนและน้องในห้องปฏิบัติการอัจฉริยภาพเครื่องจักรและค้นพบความรู้(MIND LAB)ที่ได้ให้คำแนะนำในการแก้ไขปัญหาสำหรับงานวิจัยนี้ และขอขอบพระคุณผู้บอกภาษาและผู้ทดสอบระบบทุกท่านที่สละเวลาอันมีค่ามาช่วยในงานวิจัยนี้

และท้ายสุดนี้ ผู้วิจัยใคร่ขอกราบขอบพระคุณบิดา มารดาที่ให้การส่งเสริมในด้านการศึกษาอย่างต่อเนื่อง และให้กำลังใจแก่ผู้วิจัยตลอดมา

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญ

บทคัดย่อภาษาไทย ..... ง
บทคัดย่อภาษาอังกฤษ ..... จ
กิตติกรรมประกาศ ..... ฉ
สารบัญ ..... ช
สารบัญตาราง ..... ญ
สารบัญภาพ ..... ฎ
สารบัญภาพ ..... ฎ
บทที่
1 บทนำ
1.1 ความเป็นมาและความสำคัญของปัญหา 1
1.2 วัตถุประสงค์ของวิทยานิพนธ์ 3
1.3 ขอบเขตของวิทยานิพนธ์ 3
1.4 ขั้นตอนการวิจัย 3
1.5 ประโยชน์ที่คาดว่าจะได้รับจากงานวิจัย 4
1.6 โครงสร้างของวิทยานิพนธ์ 4
2 ทฤษฎีและการวิจัยที่เกี่ยวข้อง
2.1 ทฤษฎีทางภาษาศาสตร์ 5
2.1.1 อวัยวะที่ทำให้เกิดเสียง 5
2.1.2 กระบวนการทำให้เกิดเสียง 7
2.1.3 เสียงในภาษาไทย 8
2.2 การรู้จำเสียงพูด 12
2.2.1 การตัดแบ่งพยางค์และการหาจุดตัดหัวท้ายหน่วย 13
2.2.2 วิธีการปรับเรียบ 16
2.2.3 การทำนายเชิงเส้นแบบรับรู้ 16
2.2.4 การแปลงฟูเรียร์แบบเร็ว 18
2.2.5 การหาปริพันธ์ของแถบวิกฤตและการชักตัวอย่างใหม่ 19
2.2.6 โค้งความดังเทียบเท่า 20
2.2.7 กฎกำลังของการได้ยิน 21
2.2.8 การแปลงฟูเรียร์แบบไม่ต่อเนื่องผกผัน 21

2.2.9 การแก้ชุดสมการเชิงเส้น 21
2.2.10 การเวียนเกิดเซปสตรัม 22
2.2.11 ราสต้า-พีแอลพี 22
2.3 ข่ายงานระบบประสาทเทียม 25
2.3.1 เพอร์เซพตรอน 25
2.3.2 การฝึกข่ายงานระบบประสาทเทียม 27
2.4 การสังเคราะห์เสียงพูด 29
2.5 ทฤษฎีวิภัชนัย 30
2.6 เซตวิภัชนัย 30
2.6.1 ฟังก์ชันสมาชิกภาพแบบวิภัชนัย 32
2.7 งานวิจัยที่เกี่ยวข้อง 33
3 ขั้นตอนการดำเนินงานวิจัย
3.1 การตัดแบ่งพยางค์และการตัดหัวท้ายหน่วย 36
3.1.1 การตัดหัวท้ายหน่วย 36
3.1.2 การตัดแบ่งพยางค์ 37
3.2 การรู้จำเสียงพูด 41
3.2.1 ข้อมูลทางเสียง 42
3.2.2 การแบ่งชุดข้อมูล 43
3.2.3วิธีการทางสวนศาสตร์ 43
3.3 ข่ายงานระบบประสาทเทียม 44
3.4 การสืบค้นข้อมูลของรายนามผู้ใช้โทรศัพท์ 46
3.5 การสังเคราะห์เสียงพูดรายนามผู้ใช้โทรศัพท์ 46
3.5.1 การตัดคำภาษาไทย 49
3.5.2 หลักการออกเสียงภาษาไทย 49
3.5.3 พจนานุกรมหน่วยเสียง 49
3.6 การประมาณค่าความใกล้เคียงของคำพ้องเสียง 50
3.6.1 การกำหนดค่าตัววัดความใกล้เคียงของคำพ้องเสียง 51
3.6.2 ความสัมพันธ์แบบวิภัชนัยของการประมาณค่าความใกล้เคียงของคำพ้องเสียง . 51
3.7 การหาค่าอัตราความถูกต้องของระบบ 53
4 การทดลองและวิเคราะห์ผลการทดลอง
4.1 วิธีการทดลอง 54

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

4.2 ผลการทดลอง 56
4.3 วิเคราะห์ผลการทดลอง 56
5 สรุปผลการวิจัยและข้อเสนอแนะ
5.1 สรุปผลการวิจัย 58
5.2 ข้อเสนอแนะ 59
รายการอ้างอิง 61
ภาคผนวก
ภาคผนวก ก ลักษณะคลื่นเสียงพูดตัวเลขภาษาไทย 63
ภาคผนวก ข การแจกแจงผลการรู้จำของแต่ละชุดการทดลอง 65
ภาคผนวก ค พจนานุกรมหน่วยเสียง 71
ภาคผนวก ง ส่วนอธิบายโปรแกรม 85
ประวัติผู้เขียนวิทยานิพนธ์ 89

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญตาราง

2.1 การแบ่งแยกประเภทของการรู้จำเสียงพูด 12
3.1 การจัดกลุ่มเสียงพยัญชนะไทยเพื่อใช้ในการอ่านและพูด 50
3.2 การจัดกลุ่มของเสียงตัวสะกดภาษาไทยที่สอดคล้องกัน 8 กลุ่ม 51
4.1 ค่าพารามิเตอร์ที่ปรับเปลี่ยนของแต่ละชุดการทดลอง 55
4.2 อัตราความถูกต้องของแต่ละชุดการทดลอง 56
ข-1 ผลการทดลองชุดที่ 1 65
ข-2 ผลการทดลองชุดที่ 2 65
ข-3 ผลการทดลองชุดที่ 3 66
ข-4 ผลการทดลองชุดที่ 4 66
ข-5 ผลการทดลองชุดที่ 5 67
ข-6 ผลการทดลองชุดที่ 6 67
ข-7 ผลการทดลองชุดที่ 7 68
ข-8 ผลการทดลองชุดที่ 8 68
ข-9 ผลการทดลองชุดที่ 9 69

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญภาพ

1.1 การให้บริการสอบถามรายนามผู้ใช้โทรศัพท์ในปัจจุบัน....2

2.1 อวัยวะภายในของระบบการพูดของมนุษย์....5

2.2 การเกิดเรโซแนนท์ภานในแบบจำลองของช่องทางเดินเสียง....7

2.3 สเปกตัมของพลังงานเสียง....8

2.4 ขั้นตอนการคำนวณพีแอลพี....17

2.5 ผลกระทบจากแต่ละขั้นตอนของพีแอลพีที่มีต่อสเปกตัม....18

2.6 ตัวกรอบรูปสี่เหลี่ยมคางหมูของพีแอลพี....20

2.7 ขั้นตอนของราสต้า-พีแอลพี....24

2.8 ส่วนประกอบของเส้นใยประสาท....25

2.9 เพอร์เซพตรอนโดยฟังก์ชันซิกมอยด์....26

2.10 โครงสร้างของข่ายงานระบบประสาทเทียมที่มีเพอร์เซพตรอนแบบหลายชั้น....27

2.11 โครงสร้างขั้นตอนในการฝึกฝนแบบการแพร่กระจายย้อนกลับ....27

2.12 แผนภาพแสดงสมาชิก A,B และ $\bar{A}$, $\bar{B}$ 32

2.13 กราฟความสัมพันธ์ของค่าสมาชิกแบบวิภัชนัยกับค่าสมาชิกในเซต A' ซึ่งเป็นเซตของความสูงของคนที่จัดว่าสูงปานกลาง....32

3.1 การทำงานของระบบสอบถามรายนามผู้ใช้โทรศัพท์....35

3.2 สัญญาณเสียงพูดต่อเนื่องที่ทำการตัดหัวท้ายหน่วยแล้ว....37

3.3 ขั้นตอนการตัดแบ่งพยางค์....39

3.4 การใช้โปรแกรม SpeechViewตรวจสอบความถูกต้องในการแบ่งพยางค์....40

3.5 สัญญาณเสียงที่ได้จากการตัดแบ่งพยางค์....41

3.6 พลังงานของเสียงที่ได้จากการตัดแบ่งพยางค์....41

3.7 วิธีการกรองสัญญาณเสียงด้วยโปรแกรม Cool Edit.... 43

3.8 ขั้นตอนการรู้จำเสียงเมื่อใช้ข้อมูล 6 กรอบ....45

3.9 การใช้เทคนิคแบบแฮชชิ่งในการค้นหาข้อมูลรายนามผู้ใช้โทรศัพท์....46

3.10 ขั้นตอนของการสังเคราะห์เสียงพูด....48

3.11 ความสัมพันธ์แบบวิภัชนัยบน { M1, M2 }.... 52

ก-1 ลักษณะคลื่นเสียงพูดตัวเลขภาษาไทย ศูนย์ถึงสี่....63

ก-2 ลักษณะคลื่นเสียงพูดตัวเลขภาษาไทย ห้าถึงเก้า....64

สารบัญภาพ (ต่อ)

หน้า

ง-1 โปรแกรมระบบสอบถามรายนามผู้ใช้โทรศัพท์แบบอัตโนมัติ.....85
ง-2 โปรแกรมแสดงผลของการรู้จำหมายเลขโทรศัพท์.....86
ง-3 เมนูแสดงคำสั่งที่ใช้กับฐานข้อมูล.....87
ง-4 การเพิ่มหมายเลขโทรศัพท์ลงในฐานข้อมูล..... 87
ง-5 การลบหมายเลขโทรศัพท์ออกจากฐานข้อมูล.....87
ง-6 การตรวจสอบการอ่านออกเสียงชื่อและนามสกุล.....88

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 1

บทนำ

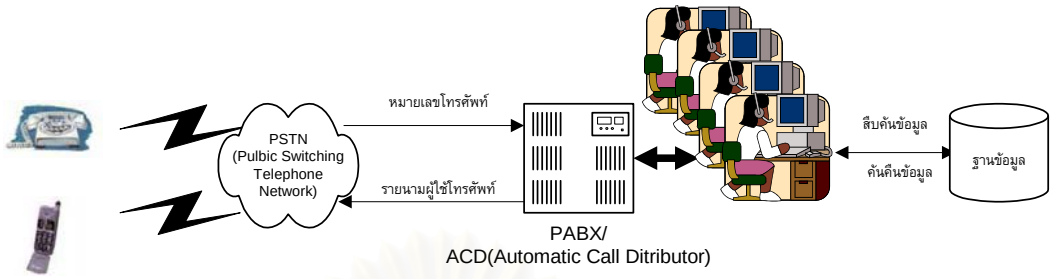
1.1 ความเป็นมาและความสำคัญของปัญหา

ในปัจจุบันการให้บริการทางด้านข้อมูลข่าวสารกับลูกค้าโดยผ่านทางระบบโทรศัพท์เริ่มเข้ามามีบทบาทสำคัญต่อการให้บริการกับลูกค้าเพื่อเพิ่มความพึงพอใจและความสะดวกรวดเร็วในการให้บริการ การให้บริการสอบถามรายนามและเลขหมายผู้ใช้โทรศัพท์ก็เป็นหนึ่งในบริการที่ผู้ให้บริการทางด้านโทรศัพท์ต้องให้บริการแก่ลูกค้า โดยข้อมูลของผู้ใช้โทรศัพท์ที่จะสอบถามได้จะต้องได้รับอนุญาตจากผู้ใช้โทรศัพท์ที่ต้องแจ้งให้ผู้ให้บริการทราบว่าสามารถเปิดเผยข้อมูลได้ การให้บริการสอบถามรายนามผู้ใช้โทรศัพท์ก็เป็นการให้บริการเมื่อผู้ใช้บริการต้องการทราบชื่อเจ้าของหมายเลขโทรศัพท์โดยแจ้งหมายเลขโทรศัพท์แก่พนักงาน พนักงานจะตอบกลับมาเป็นชื่อและนามสกุลของผู้ใช้โทรศัพท์ ทั้งนี้อาจเพื่อวัตถุประสงค์ต่างๆของผู้ใช้บริการ เช่น ใช้ในการตรวจสอบความถูกต้องของหมายเลขโทรศัพท์ว่าเจ้าของหมายเลขโทรศัพท์เป็นผู้ใด ส่วนการให้บริการสอบถามเลขหมายโทรศัพท์เป็นการให้บริการเมื่อผู้ใช้บริการต้องการทราบหมายเลขโทรศัพท์ของเจ้าของโทรศัพท์นั้น โดยแจ้งชื่อและนามสกุลแก่พนักงาน พนักงานจะตอบกลับมาเป็นเลขหมายโทรศัพท์

นอกจากนี้ยังมีบริการต่างๆ ที่คาดว่าทางผู้ให้บริการจะเปิดให้บริการผ่านทางระบบโทรศัพท์ต่อไปในอนาคต เช่น การให้บริการข่าวสารประจำวัน การให้บริการพยากรณ์อากาศ การให้บริการข่าวสารทางด้านตลาดหลักทรัพย์ เป็นต้น การบริการทั้งหมดที่กล่าวมานี้สามารถนำหลักการของระบบรู้จำเสียงพูดหรือระบบสังเคราะห์เสียงพูดมาประยุกต์ใช้งานได้เป็นอย่างดี เพื่อให้เกิดความรวดเร็วทันต่อเหตุการณ์และลดค่าใช้จ่ายในการเพิ่มประสิทธิภาพของระบบงาน ซึ่งระบบการรู้จำเสียงพูดแบบต่อเนื่องและการสังเคราะห์เสียงพูดภาษาไทยในปัจจุบันได้มีการพัฒนาไปได้ระดับหนึ่งแล้ว สามารถนำมาประยุกต์ ใช้งานกับชีวิตประจำวันของมนุษย์ได้จริงในปัจจุบัน

งานวิจัยนี้ต้องการที่จะพัฒนาและเพิ่มประสิทธิภาพของระบบสอบถามรายนามผู้ใช้โทรศัพท์ผ่านทางระบบโทรศัพท์ให้สูงยิ่งขึ้น ซึ่งในปัจจุบันระบบสอบถามรายนามผู้ใช้โทรศัพท์ยังจำเป็นต้องมีพนักงานตอบรับโทรศัพท์คอยประจำอยู่ที่ศูนย์บริการ เพื่อทำหน้าที่รับโทรศัพท์จากผู้ใช้บริการที่ต้องการทราบรายนามผู้ใช้โทรศัพท์ตลอดทั้ง 24 ชั่วโมง ดังนั้นเมื่อผู้ใช้บริการโทรศัพท์ไปยังศูนย์บริการแล้วแจ้งหมายเลขโทรศัพท์ที่ต้องการทราบกับพนักงานรับโทรศัพท์ พนักงานรับโทรศัพท์จะค้นหารายนามผู้ใช้โทรศัพท์โดยพิมพ์ข้อมูลลงไปในระบบเพื่อจะนำข้อมูลนั้นไปทำการ

ค้นคืนในฐานข้อมูล เมื่อพบข้อมูลที่ต้องการแล้วจึงแจ้งให้ผู้ใช้บริการทราบดังแสดงขั้นตอนในรูปที่ 1 .1



รูปที่ 1.1 การให้บริการสอบถามรายนามผู้ใช้โทรศัพท์ในปัจจุบัน

ซึ่งขั้นตอนดังกล่าวทำให้เกิดความล่าช้าของเวลาที่ใช้ไปต่อหนึ่งการเรียกใช้บริการ เมื่อมีผู้ใช้บริการเข้ามาเรียกใช้งานในเวลาพร้อมกันทำให้เกิดแถวรอคอยเพื่อรับบริการเป็นจำนวนมาก เป็นเหตุให้ผู้ใช้บริการต้องรอคอยและอาจสร้างความไม่พึงพอใจขึ้นได้ ซึ่งระบบสอบถามรายนามผู้ใช้โทรศัพท์โดยใช้การรู้จำเสียงพูดและการสังเคราะห์เสียงพูดภาษาไทยแบบอัตโนมัตินี้จะสามารถลดระยะเวลาในการใช้บริการและลดค่าใช้จ่ายในการเพิ่มประสิทธิภาพของระบบงาน เนื่องจากไม่จำเป็นต้องเพิ่มจำนวนพนักงานตอบรับโทรศัพท์เพื่อที่จะรองรับให้เท่ากับปริมาณการเรียกใช้บริการที่ผู้ใช้เรียกเข้ามาสู่ระบบสูงสุด

การพัฒนาโปรแกรมต้นแบบของระบบสอบถามรายนามผู้ใช้โทรศัพท์แบบอัตโนมัติบนเครื่องไมโครคอมพิวเตอร์นั้นต้องประกอบด้วยส่วนประกอบ 2 ส่วนใหญ่ คือ

ก. การรู้จำเสียงพูดตัวเลขภาษาไทยแบบต่อเนื่องและไม่ขึ้นกับผู้พูด ซึ่งใช้ชุดหมายเลขโทรศัพท์ศูนย์ถึงเก้า

ข. สังเคราะห์เสียงพูดสำหรับชื่อเฉพาะภาษาไทย โดยพูดชื่อและนามสกุลจากรายนามผู้ใช้โทรศัพท์

สำหรับงานวิจัยนี้มุ่งเน้นในการพัฒนาเพื่อสร้างโปรแกรมต้นแบบของระบบสอบถามรายนามผู้ใช้โทรศัพท์แบบอัตโนมัติที่สามารถรู้จำเสียงพูดตัวเลขภาษาไทยศูนย์ถึงเก้าแบบต่อเนื่องที่เป็นหมายเลขโทรศัพท์ด้วยการรู้จำเสียงพูดต่อเนื่องภาษาไทยระดับพยางค์ โดยลักษณะสำคัญทางสวนศาสตร์ที่นำมาใช้ได้แก่ ราสต้า-พีแอลพี ส่วนเทคนิคของการเรียนรู้ที่นำมาใช้คือ ข่ายงานระบบประสาทเทียม แล้วนำเลขหมายที่ได้ไปทำการค้นคืนรายนามผู้ใช้โทรศัพท์ในฐานข้อมูลได้เป็นชื่อและนามสกุล แล้วทำการสังเคราะห์เสียงพูดออกมาเป็นภาษาไทยโดยใช้โมดูลการต่อหน่วยเสียงย่อย(Concatenation) โดยจำลองระบบบนเครื่องไมโครคอมพิวเตอร์ ใช้ไมโครโฟนเป็นอุปกรณ์รับข้อมูลขาเข้าและลำโพงเป็นอุปกรณ์ส่งข้อมูลขาออก ทำการบันทึกเสียงเก็บเป็นแฟ้มข้อ

มูลเสียง (WAV file) แบบโมโน 16 บิต ที่อัตราการชักข้อมูล (sampling rate) 11,025 เฮิรตซ์

1.2 วัตถุประสงค์ของวิทยานิพนธ์

1. เพื่อพัฒนาระบบสอบถามรายนามผู้ใช้โทรศัพท์โดยผ่านระบบโทรศัพท์
2. เพื่อศึกษาและกำหนดหาสถาปัตยกรรมโครงสร้างของระบบรู้จำเสียงพูดภาษาไทยและโครงสร้างของระบบการสังเคราะห์เสียงพูดภาษาไทยที่เหมาะสม

1.3 ขอบเขตของวิทยานิพนธ์

1. ใช้เครื่องไมโครคอมพิวเตอร์จำลองระบบโดยรับเสียงพูดที่เป็นภาษาไทยผ่านทางไมโครโฟน และอ่านชื่อและนามสกุลที่เป็นภาษาไทยออกทางลำโพงเท่านั้น
2. ทำการรู้จำและสังเคราะห์เสียงพูดที่เป็นภาษาไทยเท่านั้น
3. จำนวนข้อมูลผู้ใช้โทรศัพท์เป็นข้อมูลจริงที่ได้จากการสุ่มตัวอย่างมาจากเลขหมายของระบบโทรศัพท์พื้นฐานในเขตนครหลวงและปริมณฑล
4. สามารถทำการรู้จำเสียงพูดเป็นตัวเลขภาษาไทยแบบต่อเนื่อง
5. สามารถทำการรับเสียงพูดที่เป็นตัวเลข และทำการค้นคืนข้อมูลและออกเสียงเป็นชื่อและนามสกุลได้ในเวลาจริง

1.4 ขั้นตอนการวิจัย

1. ศึกษาสถาปัตยกรรมโครงสร้างของระบบการรู้จำเสียงพูดต่อเนื่องภาษาไทย
2. ศึกษาสถาปัตยกรรมโครงสร้างของระบบการสังเคราะห์เสียงพูดภาษาไทย
3. ศึกษาและพัฒนาฐานข้อมูลให้มีส่วนของคำอ่านภาษาไทยของชื่อและนามสกุล
4. สร้างโปรแกรมต้นแบบของระบบการสอบถามรายนามผู้ใช้โทรศัพท์ผ่านทางระบบโทรศัพท์
5. ทดสอบและประเมินประสิทธิภาพของระบบที่พัฒนาขึ้นกับข้อมูลจริงจำนวนหนึ่ง
6. สรุปผลการวิจัย และจัดทำรายงานวิทยานิพนธ์

1.5 ประโยชน์ที่คาดว่าจะได้รับจากงานวิจัย

1. สามารถนำไปประยุกต์ใช้กับระบบงานสอบถามรายนามผู้ใช้โทรศัพท์จริงได้
2. ช่วยให้ผู้ใช้โทรศัพท์มีความสะดวกในการใช้โทรศัพท์โดยไม่ต้องกดแป้นพิมพ์หมายเลขโทรศัพท์ และเพิ่มความพึงพอใจให้กับผู้ใช้บริการสอบถามรายนามผู้ใช้โทรศัพท์มากขึ้น
3. เป็นแนวทางในการพัฒนาในระบบงานสอบถามเลขหมายผู้ใช้โทรศัพท์ หรือระบบอื่นๆ ที่ผ่านทางระบบโทรศัพท์ เช่น ระบบตรวจรางวัลฉลากกินแบ่งรัฐบาล ระบบการจองโรงแรมแบบอัตโนมัติ การให้บริการพยากรณ์อากาศผ่านทางโทรศัพท์ เป็นต้น

1.6 โครงสร้างของวิทยานิพนธ์

เนื้อหาของวิทยานิพนธ์ฉบับนี้ถูกแบ่งออกเป็น 5 บท ดังนี้คือ บทที่ 1 เป็นบทนำ บทที่ 2 จะกล่าวถึงทฤษฎีและงานวิจัยที่เกี่ยวข้อง เช่น ลักษณะทางสวนศาสตร์ของเสียง การรู้จำเสียงพูด ข่ายงานระบบประสาทเทียม บทที่ 3 จะกล่าวถึงวิธีการดำเนินงานวิจัย โดยอธิบายเป็นขั้นตอนต่างๆทั้งทางส่วนการรู้จำเสียงและส่วนของการสังเคราะห์เสียง ส่วนในบทที่ 4 เป็นการทดลองและผลที่ได้จากการทดลองตามชุดการทดลองต่างๆ และท้ายสุดคือบทที่ 5 เป็นการสรุปผลการทดลองและข้อเสนอแนะของงานวิจัย ซึ่งอาจจะเป็นประโยชน์ต่องานวิจัยอื่นๆ ต่อไปในอนาคต

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีทางภาษาศาสตร์

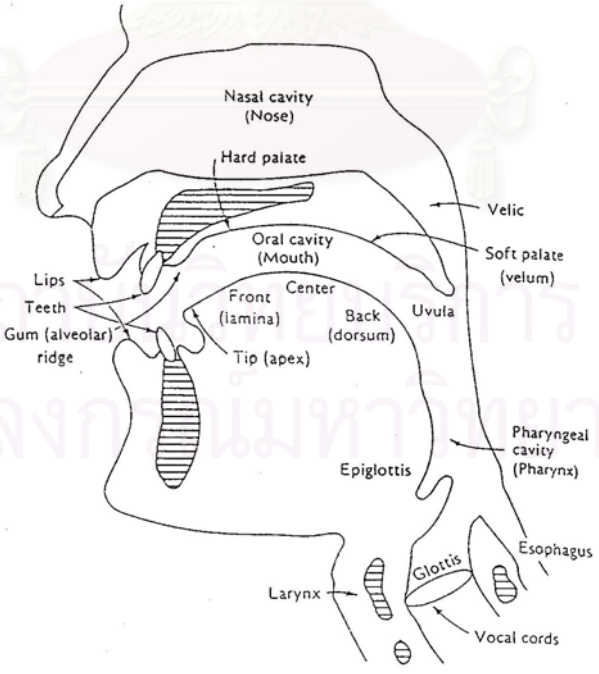
ในการศึกษาเกี่ยวกับเรื่องเสียงพูด แบ่งได้เป็น 2 ลักษณะ คือ

ก. สรีรสัทศาสตร์ (Articulatory Phonetics) เป็นการศึกษาเสียงพูดจากอวัยวะ และการเคลื่อนไหวของอวัยวะที่ทำให้เกิดเสียงพูด การอธิบายโดยอาศัยลักษณะและอาการเคลื่อนไหวของอวัยวะที่เกี่ยวข้องในการเปล่งเสียงพูดนั้น

ข. สวนสัทศาสตร์ (Acoustic Phonetics) เป็นการศึกษาเสียงพูดจากลักษณะของคลื่นเสียงที่ผู้พูดเปล่งออกมา และผู้ฟังได้ยินว่ามีลักษณะทางสวนศาสตร์เป็นอย่างใด การศึกษาตามแนวทางนี้ต้องอาศัยความรู้ทางฟิสิกส์ และคณิตศาสตร์ เข้าช่วยอธิบายลักษณะของคลื่นเสียง

2.1.1 อวัยวะที่ทำให้เกิดเสียง (The Organs of Speech)

ตำแหน่งของอวัยวะที่ทำให้เกิดเสียงของมนุษย์ได้แก่ ริมฝีปาก ฟัน เพดาน และลิ้น ซึ่งเรียงกันเป็นลำดับตลอดช่องกล้ามเนื้อช่องคอ ดังรูป 2.1 คำพูดที่เราเปล่งออกมาเป็นลำดับขั้นของความถี่คลื่นเสียงที่ได้คัดเลือกไว้ นั่นคือตำแหน่งของช่องทางเดินของเสียงมนุษย์จะมีการเปลี่ยนแปลงเป็นลำดับขั้นอย่างต่อเนื่อง



รูปที่ 2.1 อวัยวะภายในของระบบการพูดของมนุษย์ (LADEFOGED, 1975)

อวัยวะที่ใช้ในการออกเสียงมีดังต่อไปนี้

โพรงจมูก (Nasal cavity) ลักษณะของเสียงพูดที่เกิดขึ้นจะแปรผันไปตามการ ปิด-เปิดของช่องทางออกของอากาศที่จะออกสู่โพรงจมูก ซึ่งเป็นผลจากการยกขึ้นหรือเลื่อนลงของเพดานอ่อน

เพดานแข็ง (Hard palate) หมายถึง เพดานที่โค้งเป็นกระดูกแข็งเท่านั้น

เพดานอ่อน (Soft palate) คือ ส่วนของเพดานแข็งไปข้างใน มีลักษณะเป็นกระดูกอ่อนที่ขยับขึ้นลงได้ เวลาพูดส่วนใหญ่ปลายเพดานอ่านและลิ้นไก่ (Uvula) จะถูกยกขึ้นไปจรดกับหลังคอ

ลิ้น (Tongue) เป็นกล้ามเนื้อที่อ่อนตัวและเคลื่อนไหวมากที่สุดในการออกเสียงต่างๆ จึงแบ่งหน้าที่ของส่วนต่างๆของลิ้นได้ 3 ส่วนคือ 1) ส่วนปลายลิ้น (tip of the the tongue) สามารถยกขึ้นไปแตะอวัยวะส่วนต่างๆในปากตอนบนได้ง่าย 2) ส่วนหน้าลิ้น (front of the tongue) ที่อยู่ตรงข้ามกับเพดานแข็ง ขณะที่ว่างลิ้นราบกับปากตอนที่ไม่ได้พูด 3) ส่วนหลังลิ้น (back of the tongue) ที่อยู่ตรงข้ามกับเพดานอ่อนขณะที่วางลิ้นราบกับปาก

ช่องคอหอย (Pharynx) เป็นช่องระหว่างโคนลิ้นกับผนังหลังคอหอย อยู่ระหว่างกระบอกเสียงและปาก

ลิ้นปิดหลอดลม (Epiglottis) เป็นก้อนเนื้อเล็กๆ คล้ายลิ้นไก่ อยู่ต่อจากโคนลิ้นลงไปในลำคอมีหน้าที่ปิดช่องลมขณะพูดและรับประทานอาหาร

เส้นเสียง (Vocal Cords) เป็นอวัยวะสำคัญที่กำเนิดเสียง มีลักษณะเป็นกล้ามเนื้อเอ็นสองแผ่นอยู่คู่กันและปิดขวางจากปากช่องหลอดลมจากด้านหลังมาด้านหน้า อยู่ในอวัยวะลูกกระเดือก(Adam's apple)

ช่องว่างระหว่างเส้นเสียง (Glottis) ระหว่างเส้นเสียงจะมีช่องว่างซึ่งเป็นทางให้ลมผ่านเข้าไปถึงปอดและออกจากปอดมาได้ และช่องว่างระหว่างเส้นเสียงทั้งสองสามารถดึงให้ออกจากกันหรือดึงเข้ามาประชิดกันได้

กระบอกเสียง (Larynx) เป็นอวัยวะส่วนเหนือของหลอดลมตอนที่ขยายกว้างและยื่นออกมาตรงคอหอยมีผนังเป็นกระดูกอ่อนบุด้วยกล้ามเนื้อและพังผืดซึ่งก็คือลูกกระเดือก ที่สามารถเลื่อนขึ้นลงได้ประมาณครึ่งนิ้ว โดยอาจใช้มือคลำขณะออกเสียงสระต่างๆ

กระบังลม (Diaphragm) และกล้ามเนื้อหน้าอก (Chest muscles) ซึ่งควมคุมลมหายใจเข้าออกปอด

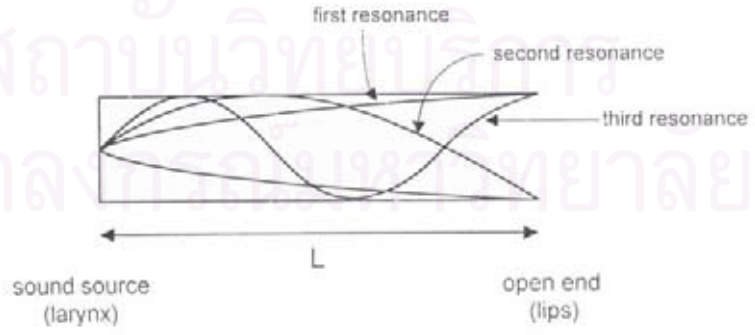
ฟัน (Teeth) เป็นอวัยวะที่เกิดของเสียงหลายชนิด เช่น เมื่อฟันบนกดฟันล่าง ลมที่ผ่านออกมาจะทำให้เกิดเสียงที่เรียกว่า เสียงเสียดแทรกที่เกิดที่ฟัน เป็นต้น

ริมฝีปาก (Lips) เป็นอวัยวะที่สามารถเคลื่อนไหวได้มาก และทำให้เสียงแตกต่างกันได้มาก โดยจะทำหน้าที่ในการทำเสียงขัดสี (Frictional sounds) บางเสียงโดยบางทีใช้กับฟันบนหรือริมฝีปากด้วยกัน

2.1.2 กระบวนการทำให้เกิดเสียง (Speech Production)

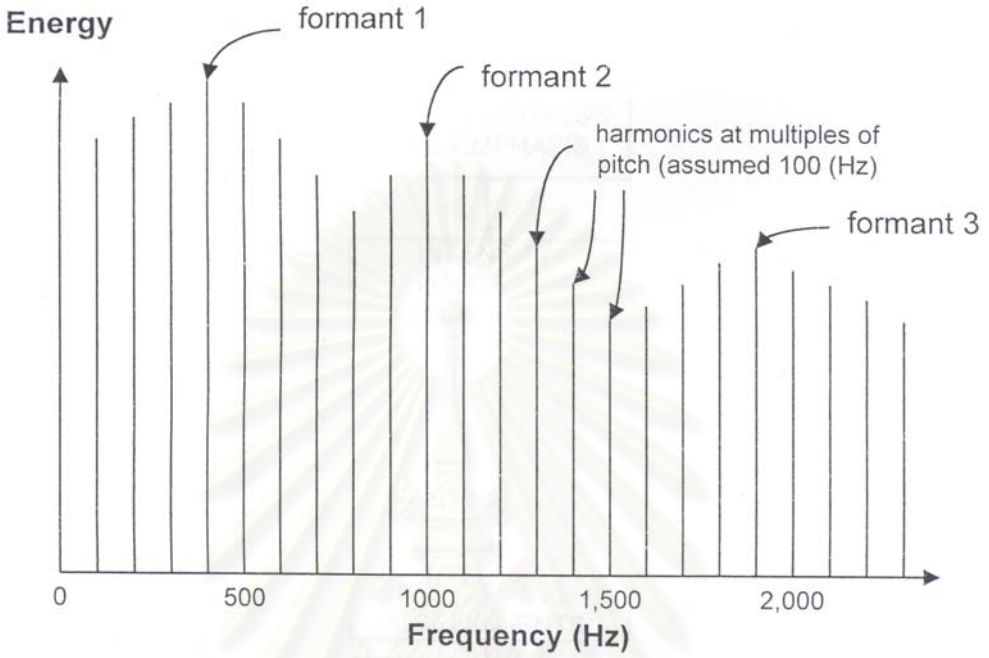
ระบบเสียงพูดสามารถพิจารณาได้ว่า ประกอบด้วยลำดับของท่อ และช่องที่ต่อออกมาจากปอดไปยังปากและจมูก ท่อและช่องนี้จะมีความยาวโดยประมาณ 7 นิ้ว เส้นเสียงจะอยู่ในตำแหน่งปลายที่ตรงข้ามกับขั้วปอด ทำหน้าที่ควบคุมการไหลของลมให้ผ่านปอดเข้าสู่ช่องทางเดินเสียง ภายใต้การควบคุมของกล้ามเนื้อส่วนประกอบของช่องทางเดินเสียงที่มีลักษณะเป็นท่อจะสามารถเปลี่ยนแปลงรูปร่างได้ในอัตราถึง 10 ครั้งต่อวินาที ส่วนเส้นเสียงนั้นจะสามารถเปิด-ปิดด้วยอัตราเร็วประมาณ 100-300 ครั้งต่อวินาที การเปลี่ยนแปลงรูปร่างของช่องทางเดินเสียงและรูปร่างและตำแหน่งของสื่อกลางที่ทำให้เกิดเสียงดังกล่าวนี้รวมเรียกว่า กระบวนการทำให้เกิดเสียง

รูปแบบจำลองอย่างง่ายของช่องทางเดินเสียง ก็อาจมองได้เป็นลักษณะของท่อทรงกระบอกที่มีต้นกำเนิดเสียงอยู่ที่ปลายข้างหนึ่ง (ส่วนของกล่องเสียง) และปลายอีกข้างหนึ่งจะเปิด (ส่วนของปาก) ดังรูปที่ 2.2 ดังนั้นมันจะเกิดเรโซแนนท์ภายในท่อได้ที่ความยาวเท่ากับ 4L, 4L/3, 4L/5 เฮิรตซ์ โดยที่ L คือความยาวท่อ ถ้าคิดเป็นความถี่ที่เกิดเรโซแนนท์จะได้ความถี่ที่ c/4L, 3c/4L, 5c/4Lเฮิรตซ์ โดยที่ c คือ ค่าความเร็วของเสียงในอากาศ และถ้าจะคำนวณหาความถี่ในการเรโซแนนท์ของช่องทางเดินเสียงของคน ซึ่งปกติช่องทางเดินเสียงของคนเราจะมีความยาวประมาณ 7 นิ้ว หรือ 17 เซนติเมตร และ c มีค่าเท่ากับ 340 เมตรต่อวินาที ดังนั้นจึงมีเร-โซแนนท์ที่ความถี่ประมาณ 500 เฮิรตซ์, 1500 เฮิรตซ์, 2500 เฮิรตซ์ เป็นต้น



รูปที่ 2.2 การเกิดเรโซแนนท์ภายในแบบจำลองของช่องทางเดินเสียง (WITTEN, 1982)

เมื่อกล่องเสียงกระตุ้นให้เกิดเคลื่อนที่ประกอบไปด้วยฮาร์มอนิก (harmonic) ต่าง ๆ มากมาย เวโซแนนท์ของช่องทางเดินเสียงนี้จะสร้างรูปคลื่นที่มียอดสูงเด่นขึ้นมา เมื่อดูจากสเปกตรัมพลังงานของรูปคลื่น ซึ่งเรียกว่า ฟอร์แมนท์ (formants) ของเสียง ดังรูปที่ 2.3



รูปที่ 2.3 สเปกตัมของพลังงานเสียง (WITTEN, 1982)

ฟอร์แมนท์ที่มีความถี่ต่ำที่สุดจะเรียกว่า ฟอร์แมนท์ที่หนึ่ง ซึ่งจะมีค่าประมาณ 200-1000 เฮิรตซ์ ทั้งนี้ขึ้นอยู่กับขนาดของช่องทางเดินเสียงด้วย ส่วนฟอร์แมนท์ที่สองที่อยู่ถัดไปก็จะมีค่าประมาณ 500-2500 เฮิรตซ์ และฟอร์แมนท์ที่สามมีค่าประมาณ 1500-3500 เฮิรตซ์ เป็นต้น โดยที่ฟอร์แมนท์ที่หนึ่งและที่สองนี้ จะเป็นตัวที่แสดงคุณลักษณะที่สำคัญมากอันหนึ่งในการพิจารณาการประยุกต์ใช้ในด้านการวิเคราะห์สัญญาณเสียงทางด้านสเปกตรัม

2.1.3 **เสียงในภาษาไทย**

เสียงในภาษาไทย หมายถึง เสียงที่เปล่งออกมาเพื่อใช้ในการสื่อสาร และสื่อความหมายระหว่างมนุษย์ด้วยกัน ซึ่งอย่างน้อยจะประกอบด้วยหน่วยพยางค์เพื่อนำมาประสมให้กิดเป็นคำพูด ลักษณะพื้นฐานของเสียงในภาษาไทยและลักษณะสำคัญที่เกี่ยวข้อง เสียงที่เปล่งออกมาเป็นคำหนึ่งในภาษาไทย อาจประกอบไปด้วย

2.1.3.1 **เสียงพยัญชนะ**

เสียงพยัญชนะในภาษาไทย เป็นเสียงที่เกิดจากลมที่ผ่านเส้นเสียงแล้วถูกดัดแปลงด้วยอวัยวะต่างๆภายในปากจนทำให้เกิดเสียงขึ้น มีคุณสมบัติที่สำคัญ คือ ความก้องของเสียง ซึ่ง

คุณสมบัติดังกล่าวเป็นคุณสมบัติสำคัญที่ใช้ในการแบ่งแยกเสียงพยัญชนะโฆษะ (Voiced, เสียงก้อง) ออกจากพยัญชนะอโฆษะ (Voiceless,เสียงไม่ก้อง) โดยในการออกเสียงพยัญชนะโฆษะนั้นเส้นเสียงจะสั่นสะเทือนในขณะที่เราเปล่งเสียงออกมา แต่ในกรณีของเสียงอโฆษะนั้นจะไม่มีการสั่นสะเทือนของเส้นเสียง เสียงพยัญชนะในภาษาไทยประกอบด้วย

ก.) เสียงพยัญชนะหยุด

เสียงพยัญชนะหยุดแบ่งออกเป็น 2 ลักษณะย่อย คือ เสียงพยัญชนะระเบิด (Plosives) และเสียงพยัญชนะกัก (Stop) โดยเสียงพยัญชนะระเบิด คือ เสียงพยัญชนะที่เกิดจากลมที่เปล่งออกมาถูกกัก ณ ที่ใดที่หนึ่งในปาก แล้วเปิดช่องที่กักลมนั้นให้ลมพุ่งออกมาโดยแรง เสียงระเบิดนี้ถ้ามีลมออกมาด้วยจะเรียกว่า เสียงระเบิดที่มีลม (Aspirated Plosives) และเสียงระเบิดที่ไม่มีลม (Unaspirated Plosives) เพิ่มมาอีก สำหรับเสียงพยัญชนะอีกพวกหนึ่ง จะมีลักษณะการเกิดในระยะต้นเช่นเดียวกับพยัญชนะระเบิดแต่แทนที่ลมนั้นจะถูกกักไว้จะระเบิดออกมา กลับถูกกลืนกลับลงไปใหม่ หรืออาจเป็นลมธรรมดาที่ออกไปทางจมูกซึ่งลักษณะนี้เรียกว่า เสียงกัก

ข.) เสียงพยัญชนะเสียดแทรก (Fricatives) คือ เสียงพยัญชนะที่เกิดขึ้นเมื่อมีลมที่ทำให้เกิดเสียงซ่าขึ้น

ค.) เสียงพยัญชนะนาสิก (Nasals) ในการเปล่งเสียงนาสิก ลมจากปอดจะผ่านขึ้นไปยังจมูก แต่จะไม่สามารถผ่านออกทางช่องปากได้เพราะทางเดินของลมในช่องปากถูกปิดกั้นสนิท

ง.) เสียงพยัญชนะข้างลิ้น (Lateral) เป็นเสียงพยัญชนะก้อง ซึ่งเปล่งออกมาโดยใช้ลิ้นปิดบริเวณปุ่มเหงือก และตรงส่วนกลางเพดานไว้ ปล่อยให้ลมออกข้างลิ้น

จ.) เสียงพยัญชนะรัว (Rolled) เป็นเสียงพยัญชนะก้อง ซึ่งในขณะที่มีการเปล่งเสียงต้องมีการกระดกลิ้นร่วมด้วย

ฉ.) เสียงพยัญชนะกึ่งสระหรืออรรธสระ (Semi-Vowel) เป็นเสียงเลื่อนที่เกิดขึ้นระหว่างเสียงสระ 2 เสียง โดยในการเปล่งเสียงอรรธสระนั้นอวัยวะที่ใช้ในการออกเสียงจะอยู่ในตำแหน่งของสระใดสระหนึ่งก่อนแล้วจึงเปล่งเสียงออกมา

หน่วยเสียงพยัญชนะต้น หมายถึง เสียงที่เปล่งออกก่อนเสียงอื่นในพยางค์ อาจเป็นเสียงพยัญชนะเดี่ยว เช่น การ (ก เป็นหน่วยเสียงพยัญชนะต้น) หรืออาจเป็นเสียงพยัญชนะควบกล้ำ เช่น กราบ (กร เป็นหน่วยเสียงพยัญชนะต้น ซึ่งเป็นเสียงควบกล้ำ)

หน่วยเสียงพยัญชนะท้ายพยางค์ หมายถึง เสียงพยัญชนะที่ปิดท้ายพยางค์ตามเสียงสระหรือ เรียกว่า มาตราตัวสะกด มีทั้งหมด 8 เสียง 8 มาตรา ดังนี้คือ

1) มาตราตัวสะกด แม่กน มี น เป็นตัวสะกด และมีพยัญชนะอื่นทำหน้าที่เป็นตัวสะกดแทนได้ ได้แก่ ญ ณ ร ล ฬ
2) มาตราตัวสะกด แม่กง มี ง เป็นตัวสะกด
3) มาตราตัวสะกด แม่กก มี ก เป็นตัวสะกด และมีพยัญชนะอื่นทำหน้าที่เป็นตัวสะกดแทนได้ ได้แก่ ข ค ฆ
4) มาตราตัวสะกด แม่กด มี ด เป็นตัวสะกด และมีพยัญชนะอื่นทำหน้าที่เป็นตัวสะกดแทนได้ ได้แก่ จ ฉ ช ซ ฌ ฎ ฏ ฐ ฑ ฒ ต ถ ท ธ ศ ษ ส
5) มาตราตัวสะกด แม่กบ มี บ เป็นตัวสะกด และมีพยัญชนะอื่นทำหน้าที่เป็นตัวสะกดแทนได้ ได้แก่ ป พ ฟ ภ
6) มาตราตัวสะกด แม่กม มี ม เป็นตัวสะกด
7) มาตราตัวสะกด แม่เกย มี ย เป็นตัวสะกด
8) มาตราตัวสะกด แม่เกอว มี ว เป็นตัวสะกด

2.1.3.2 **เสียงสระ**

เสียงสระในภาษาไทย เป็นส่วนสำคัญอย่างยิ่งในการรู้จำเสียงพูดต่อเนื่องภาษาไทย ทั้งนี้เพราะพยางค์หนึ่งๆของคำพูดต่อเนื่องภาษาไทยจะมีช่วงระยะเวลาของเสียงสระยาวกว่าส่วนของเสียงพยัญชนะ นอกจากนี้เสียงสระยังมีความถี่ต่างๆแฝงอยู่ ซึ่งความถี่ต่างๆนี้เองเป็นสิ่งที่มีความจำเป็นอย่างมากในการแยกแยะเสียงของคำ พยางค์และหน่วยเสียงต่างๆออกจากกัน ลักษณะสำคัญของหน่วยเสียงสระสามารถสรุปได้เป็นดังนี้

- เป็นเสียงที่ผ่านออกจากช่องปากได้อย่างสะดวก ไม่ถูกอวัยวะในช่องปากปิดกั้น
- เป็นเสียงก้องทุกเสียง
- สามารถออกเสียงได้ยาวนาน และชัดเจน
- มีทั้งเสียงสั้นและเสียงยาว
- ลักษณะของเสียงขึ้นอยู่กับลิ้น ริมฝีปาก ปริมาณของเสียงที่พุ่งออกมาและระยะเวลาที่ใช้ออกเสียง

เสียงสระในภาษาไทยจำแนกได้เป็น

- สระเสียงสั้น ได้แก่ สระ อะ อิ อึ อุ เอะ แอะ โอะ เอาะ เอียะ เอือะ อัวะ ฤ ฦ อำ ใอ ไอ เอา
- สระเสียงยาว ได้ แก่ สระ อา อี อือ อู เอ แอ โอ ออ เออ เอีย เอือ อัว ฤๅ ฦๅ

หรืออาจสามารถแบ่งตามเสียงสระหนักเบาได้เป็น

- สระเสียงเบา ได้แก่ สระ อะ อิ อึ อุ เอะ แอะ โอะ เอาะ เออะ เอียะ เอือะ อัวะ ฤ ฦ

- สระเสียงหนัก ได้แก่ สระ อา อี อือ อู เอ แอ โอ ออ เออ เอีย เอือ อัว ฤๅ ฦๅ อำ ใอ ไอ เอา

2.1.3.3 **เสียงวรรณยุกต์**

เสียงวรรณยุกต์ในภาษาไทย คือ เสียงสูงต่ำในภาษาที่เกิดจากการสั่นสะเทือนของเส้นเสียงในอัตราความถี่ที่ต่างกันไป ดังนั้นเสียงวรรณยุกต์จะปรากฏในส่วนของเสียงสระเป็นส่วนใหญ่ เพราะเสียงสระเป็นเสียงที่เกิดจากการสั่นสะเทือนของเส้นเสียง นอกจากนี้อาจปรากฏอยู่บ้างในบางส่วนของเสียงพยัญชนะ แต่เพียงเฉพาะพยัญชนะที่เป็นเสียงก้องหรือเสียงพยัญชนะนาสิกเท่านั้น เสียงวรรณยุกต์ในภาษาไทยประกอบด้วย

ก.) เสียงวรรณยุกต์ระดับ (Level Tone) ประกอบด้วย เสียงวรรณยุกต์สามัญ เสียงวรรณยุกต์เอก และเสียงวรรณยุกต์ตรี

ข.) เสียงวรรณยุกต์เปลี่ยนระดับ (Contour Tone) ประกอบด้วย เสียงวรรณยุกต์โท และเสียงวรรณยุกต์จัตวา

2.1.3.4 **พยางค์** (Syllable) เป็นหน่วยโครงสร้างทางเสียงที่เปล่งออกมาครั้งหนึ่ง มีเสียงดังเด่น 1 เสียง และเสียงที่อยู่ข้างเคียงอย่างน้อย 2 เสียง พยางค์อาจจะเป็นคำได้ในกรณีที่พยางค์นั้นมีความหมาย หน่วยเสียงที่ดังเด่นกว่าหน่วยเสียงอื่น ซึ่งปรากฏในเสียงพูดเป็นแกนกลางพยางค์ ส่วนหน่วยเสียงอื่นๆ ที่ไม่ดังเด่นก็ทำหน้าที่เป็นส่วนประกอบหรือเสริมพยางค์นั้น ความดังเด่นของหน่วยเสียงขึ้นอยู่กับลักษณะเฉพาะตัวของหน่วยเสียง เช่น หน่วยเสียงสระมีลักษณะประจำตัวเป็นเสียงก้อง จึงดังเด่นกว่าหน่วยเสียงอื่นของการออกเสียงพูดในแต่ละครั้ง โดยปกติหน่วยเสียงสระจะเป็นส่วนสำคัญที่ทำให้เกิดพยางค์

พยางค์ในภาษาไทยสามารถเขียนในรูปของโครงสร้างได้ดังนี้

$$S = C\ (C)\ V\ (V)\ (C)^{T}$$

โดยที่

S = พยางค์

C = พยัญชนะ

V = สระ

T = วรรณยุกต์ ซึ่งมี 5 เสียง ได้แก่

0 = เสียงสามัญ

1 = เสียงเอก

2 = เสียงโท

3 = เสียงตรี

4 = เสียงจัตวา

(V) = สระเสียงยาว

(C) = พยัญชนะควบกล้ำ

2.2 การรู้จำเสียงพูด (Speech Recognition)

ในปัจจุบันการรู้จำเสียงพูดแบบอัตโนมัติมีการพัฒนามาอย่างต่อเนื่อง มีการแบ่งแยกประเภทโดยขึ้นอยู่กับประเด็นที่ใช้ในการแยก [2] ดังตารางที่ 2.1

ตารางที่ 2.1 การแบ่งแยกประเภทของการรู้จำเสียงพูด

ประเด็นที่ใช้ในการแบ่งแยก	ชนิดของการรู้จำ
การขึ้นกับผู้พูด	1. ขึ้นกับผู้พูด (Speaker Dependent) 2. ไม่ขึ้นกับผู้พูด(Speaker Independent)
ระดับของเสียงที่รู้จำ	1. การรู้จำคำพูดเดี่ยว (Isolated Word) 2. การรู้จำคำพูดแบบต่อเนื่อง (Continuous Speech) - คำต่อเนื่อง (Connected Word) - พูดเป็นประโยค (Conversational Speech)
หน่วยของเสียงที่ใช้เป็นตัวทาบเทียบ (Reference Template)	1. ใช้คำเป็นตัวทาบเทียบ (Word Based) 2. ใช้หน่วยเสียงเป็นตัวทาบเทียบ (Phoneme Based) 3. ใช้หน่วยที่อยู่ระหว่างคำกับหน่วยเสียง เช่น พยางค์ (Syllable Based) หน่วยเสียงคู่ หรือเสียงสามหน่วยเป็นตัวทาบเทียบ

การรู้จำเสียงพูดแบบขึ้นกับผู้พูดหมายถึงผู้พูดหรือผู้ที่ทดสอบนั้นจะต้องเป็นบุคคลเดียวกับที่ฝึกสอนให้รู้จำเสียงนั้นเท่านั้น ส่วนการรู้จำเสียงพูดแบบไม่ขึ้นกับผู้พูดนั้นก็สามารถให้บุคคลใดก็ได้เป็นผู้พูดหรือผู้ทดสอบ การรู้จำโดยใช้คำเป็นตัวทาบเทียบ (Word Based Recognition) นั้นมีข้อดีคือ สามารถหลีกเลี่ยงปัญหาความต่อเนื่องระหว่างหน่วยเสียงได้ เหมาะสำหรับการรู้จำคำศัพท์จำนวนไม่มาก (น้อยกว่า 1,000 คำศัพท์) แต่ก็ไม่สามารถหลีกเลี่ยงปัญหาความต่อเนื่อง

คำ ในการรู้จำเสียงพูดแบบต่อเนื่อง (Continuous Speech) ได้ การรู้จำโดยใช้หน่วยเสียงเป็นตัวทาบเทียบ สามารถเพิ่มจำนวนคำศัพท์ที่จะจดจำได้มากขึ้น ปัญหาใหญ่ในการใช้หน่วยเสียงเป็นตัวทาบเทียบคือ ปัญหาความต่อเนื่องระหว่างหน่วยเสียง ทำให้ยากต่อการตัดแบ่งและมีผลทำให้ยากต่อการรู้จำ และอีกวิธีหนึ่งก็คือการใช้หน่วยที่ใหญ่กว่าหน่วยเสียงขึ้นมา แต่อาจจะเล็กกว่าคำเป็นตัวทาบเทียบ เช่น หน่วยเสียงคู่ (Diphone) เสียงสามหน่วย (Triphone) หรือพยางค์ (Syllable) ซึ่งใช้ในงานวิจัยนี้ด้วย เป็นต้น

2.2.1 การตัดแบ่งพยางค์และการหาจุดตัดหัวท้ายหน่วย(Syllable Segmentation and Endpoint Detection)

การตัดแบ่งพยางค์จะหาจุดแบ่งระหว่างพยางค์ในคำ มีประโยชน์อย่างมากในการตรวจสอบจำนวนพยางค์ และการรู้จำเสียงแบบต่อเนื่องระดับพยางค์ ส่วนการหาจุดตัวหัวท้ายหน่วยที่ใช้ในการรู้จำเป็นกระบวนการที่จะแยกส่วนที่เป็นคำพูด ออกจากส่วนที่ไม่ใช่คำพูดหรือส่วนที่เป็นเสียงพื้นหลัง (Background Sound) การหาจุดตัดหัวท้ายหน่วยถือว่าเป็นกระบวนการที่สำคัญ มีผลต่ออัตราการรู้จำค่อนข้างมากเช่นเดียวกัน

วิธีในการหาจุดตัดหัวท้ายหน่วยที่ใช้ในการรู้จำจะใช้หลักการเดียวกับการตัดแบ่งพยางค์ ซึ่งในปัจจุบันมีวิธีที่นำมาช่วยในการปรับแต่งหลายวิธี ดังนี้ [9]

2.2.1.1 ใช้ค่าแอมปลิจูด (Amplitude)

โดยเริ่มจากการหาค่าสัมบูรณ์ของแอมปลิจูดของสัญญาณเสียงแต่ละจุดที่เวลา n ใดๆ ดังสมการที่ 2.1

$$A_{(j)} = |s(n)| \tag{2.1}$$

หลักการหาจุดตัดหัวท้ายหน่วยที่ใช้ในการรู้จำคือ เมื่อสัญญาณมีค่า *A(j)* มากกว่าค่าที่กำหนด (Threshold) ติดกันมากกว่าจำนวนที่กำหนดไว้ (Duration Threshold) ถือได้ว่าเป็นช่วงเริ่มต้นของคำ และจะทำเช่นเดียวกันในส่วนท้ายของสัญญาณเสียง เพื่อหาจุดสิ้นสุดคำ วิธีการนี้อาจมีข้อผิดพลาดได้ ถ้ามีสัญญาณรบกวน (Noise) มากพอจนทำให้เห็นเป็นคำได้ วิธีแก้ไขวิธีหนึ่งคือแทนที่จะพิจารณาค่าแอมปลิจูดทีละจุด ก็เปลี่ยนเป็นแบ่งสัญญาณเสียงเป็นเฟรม แล้วหาผลรวมของค่าสัมบูรณ์ของแอมปลิจูดของสัญญาณในส่วนย่อยที่ j ใดๆ ได้ดังสมการที่ 2.2

$$A_{(j)} = \sum_{n \in frame_j} |s(n)| \tag{2.2}$$

โดยมีหลักการในการหาจุดตัดหัวท้ายหน่วยที่ใช้ในการรู้จำคล้ายกับวิธีแรกคือ เมื่อสัญญาณมีค่า *A(j)* มากกว่าค่าที่กำหนดติดกันมากกว่าจำนวนส่วนย่อยที่กำหนด (Frame Duration Threshold)

ถือว่าเป็นช่วงเริ่มต้นของคำ และใช้หลักการเดียวกันในการหาจุดสิ้นสุดของคำ ข้อดีของการใช้ค่าแอมปลิจูดในการหาจุดตัดหัวท้ายหน่วยที่ใช้ในการรู้จำคือ คำนวณง่าย ใช้เวลาน้อย แต่มีข้อเสียคือ มีโอกาสตัดผิดพลาดได้ ถ้ามีสัญญาณรบกวนที่มีค่าแอมปลิจูดสูง

2.2.1.2 **ใช้ค่าพลังงาน** (Energy)

พลังงานของสัญญาณเสียงพูดเป็นค่าคุณลักษณะที่ถูกใช้ในการวิเคราะห์เสียงพูดมานานและนิยมใช้กันอย่างแพร่หลาย เนื่องจากคำนวณค่าได้ง่ายและรวดเร็ว พลังงานของสัญญาณเสียงพูดเป็นคุณสมบัติที่แสดงให้เห็นว่ามีสัญญาณเสียง(รวมกับสัญญาณรบกวน) เกิดขึ้น ณ เวลานั้นหรือไม่ การคำนวณหาค่าพลังงานจะทำทีละกรอบเสียงพูดโดย *E(j)* คือค่าพลังของกรอบเสียงพูดที่จุด *j* และในแต่ละกรอบเสียงพูดจะมีตัวอย่างสัญญาณเสียงพูดจำนวน *n* ตัวอย่าง ซึ่งมีวิธีคำนวณหาค่าพลังงานของสัญญาณเสียงพูดหลายวิธีดังเช่นตัวอย่างต่อไปนี้ [9]

2.2.1.2.1 พลังงานแบบรากที่สองของกำลังสองเฉลี่ย(Root Mean Square Energy)

เป็นการหาค่าพลังงานของสัญญาณเสียงพูดจากรากที่สองของค่าเฉลี่ยของค่ากำลังสองของสัญญาณเสียงพูด ซึ่งส่งผลให้การตรวจสอบระดับของ *E(j)* เพื่อหาจุดตัดหัวท้ายหน่วยหรือระหว่างพยางค์มีความถูกต้องมากขึ้น ซึ่งในงานวิจัยนี้ได้ใช้การคำนวณค่าพลังงานเป็นส่วนช่วยในการตัดพยางค์และการหาจุดตัดหัวท้ายหน่วย โดยใช้สมการดังนี้

$$E_{(j)} = \sqrt{\frac{\sum_{i=j}^{j+n} s_i^2}{n}} \qquad (2.3)$$

โดย E_j คือ ค่าพลังงานจุดที่ j

s_i คือ ค่าแอมปลิจูด ของสัญญาณเสียงที่ตำแหน่ง i

n คือ จำนวนจุดที่นำมาคิดเป็นพลังงาน

2.2.1.2.2 พลังงานความถี่และเวลา (Frequency-Time Energy)

เป็นการนำพลังงานแถบความถี่ (Band Frequency Energy) มารวมเข้ากับพลังงานในเชิงเวลาพลังงานแถบความถี่เฉลี่ย ได้มาจากการแปลงสัญญาณเสียงพูดในเชิงเวลา (Time Domain) ให้อยู่ในเชิงความถี่ (Frequency Domain) โดยใช้การแปลงฟูเรียร์แบบเร็ว (Fast Fourier Transform) แล้วนำขนาดของสัญญาณในช่วงความถี่ 250-3000 เฮิรตซ์ มาหาค่าพลังงานในเชิงเวลาได้มาจากค่า log ของพลังงานเฉลี่ย

2.2.1.2.3 พลังงานกำลังสอง (Square Energy)

เป็นการวัดค่าพลังงานจากเสียงพูดยกกำลังสอง ทำให้ค่าพลังงานมีความไวต่อสัญญาณเสียงพูดที่มีขนาดใหญ่ ดังสมการที่ 2.4

$$E_n = \sum_{i=1}^{K} S_n^2[i] \tag{2.4}$$

2.2.1.3 การใช้ค่าอัตราการตัดผ่านแกนศูนย์ (Zero-crossing Rate)

อัตราการตัดผ่านแกนศูนย์ [9] เป็นการวัดจำนวนครั้งที่สัญญาณเสียงพูดตัดผ่านแกนเวลาที่ระดับศูนย์หรือจำนวนครั้งของการเปลี่ยนเครื่องหมายทางคณิตศาสตร์ของสัญญาณเสียงพูด อัตราการตัดผ่านแกนศูนย์นี้เป็นเครื่องมืออย่างง่าย ที่สามารถนำมาใช้อธิบายได้ว่าสัญญาณเสียงพูดนั้นเป็นเสียงก้อง (Voiced) หรือเสียงไม่ก้องเสียดแทรก (Unvoiced Fricative) สัญญาณเสียงพูดที่มีอัตราการตัดผ่านแกนศูนย์ต่ำจะเป็นเสียงก้อง ส่วนสัญญาณเสียงพูดที่มีอัตราการตัดผ่านแกนศูนย์สูงจะเป็นเสียงไม่ก้องเสียดแทรก ซึ่งนั้นหมายความว่าโดยปกติสัญญาณเสียงในช่วงที่เงียบจะมีค่าอัตราการตัดผ่านแกนศูนย์มากกว่าช่วงที่มีเสียงดัง อัตราการตัดผ่านแกนศูนย์ [9] สามารถหาได้จากสมการที่ 2.5

$$Z_n = \frac{1}{K}\sum_{i=1}^{k} \frac{\left|\operatorname{sgn}\{S_n[i]\} - \operatorname{sgn}\{S_n[i-1]\}\right|}{2} \tag{2.5}$$

เมื่อ

$$\operatorname{sgn}\{S[i]\} = \begin{cases} +1 & ,S[i] \geq 0 \\ -1 & ,S[i] < 0 \end{cases}$$

โดยที่ Z_n คือ อัตราการตัดผ่านแกนศูนย์ที่กรอบเสียงพูด n

$S_n[i]$ คือ สัญญาณเสียงพูดที่ i ในกรอบเสียงพูด n

K คือ ความกว้างของกรอบสียงพูด(จำนวนตัวอย่างเสียงพูดในแต่ละกรอบเสียงพูด)

ต่อมาได้มีการปรับปรุงอัตราการตัดผ่านแกนศูนย์เป็นอัตราการตัดผ่านระดับกำหนด (Band Crossing Rate) [9] ดังแสดงในสมการที่ 2.6 ซึ่งในงานวิจัยนี้ได้ใช้คุณลักษณะนี้ในการตัดแบ่งพยางค์เนื่องจากสามารถกำหนดความกว้างของระดับที่จะพิจารณาได้

$$B_n = \sum_{i=1}^{K} \left|\operatorname{sgn}\{S_n[i]\} - \operatorname{sgn}\{S_n[i-1]\}\right| \tag{2.6}$$

เมื่อ

$$\operatorname{sgn}\{S[i]\} = \begin{cases} +1 & ,S[i] \geq L \\ \operatorname{sgn}\{S[i-1]\} & ,-L \leq S[i] < L \\ -1 & ,S[i] < -L \end{cases}$$

ซึ่ง L คือค่าความกว้างของระดับที่พิจารณา

2.2.1.4 **ใช้ค่าความถี่มูลฐาน** (Fundamental Frequency)

เนื่องจากสัญญาณเสียงในช่วงที่เป็นเสียงสระ (Vowel) จะมีลักษณะเป็นรายคาบ (Periodic) ซึ่งจะมีค่าความถี่มูลฐาน และเสียงพยางค์ในภาษาไทยจะประกอบด้วยส่วนที่เป็นสระเสมอ ดังนั้นช่วงของความถี่มูลฐานช่วงหนึ่งจะตรงกับพยางค์หนึ่งพยางค์

2.2.1.5 **ระยะเวลาในการเปล่งเสียง** (Time Duration)

เราสามารถเก็บข้อมูลเพื่อตรวจสอบว่าพยางค์ในภาษาไทยควรมีระยะเวลาในการเปล่งเสียงนานเท่าเป็นส่วนช่วยเสริมความถูกต้องของการตัดแบ่งพยางค์

2.2.2 **วิธีการปรับเรียบ** (Smoothing)

เนื่องจากการตัดแบ่งพยางค์และการหาจุดตัดหัวท้ายหน่วยต้องอาศัยแผนภูมิเส้นระดับพลังงาน อัตราการตัดผ่านระดับที่กำหนด ความถี่มูลฐานของสัญญาณเสียงพูด จึงจำเป็นต้องปรับให้แผนภูมิเส้นคุณลักษณะเหล่านี้ของสัญญาณเสียงเรียบขึ้น เพื่อปรับปรุงให้เหมาะสมก่อนนำไปวิเคราะห์หาจุดแบ่งพยางค์หรือจุดตัดหัวท้ายหน่วยต่อไปซึ่งกรรมวิธีปรับเรียบที่น่าสนใจมี 2 วิธี คือ

ก. วิธีการปรับเรียบโดยค่ากลาง (Median Smoothing)

ทำได้โดยการกำหนดจุดข้อมูลที่เราสนใจ ทำการกำหนดหน้าต่างขึ้น โดยมีจุดข้อมูลที่เราสนใจเป็นจุดศูนย์กลาง ทำการหาค่ากลาง (Median: Middle Value) ของข้อมูลภายในกรอบหน้าต่างที่กำหนดขึ้น จากนั้นก็แทนค่ากลางที่ได้ลงที่จุดกึ่งกลางของหน้าต่าง กรรมวิธีนี้ทำให้ข้อมูลที่ผ่านการปรับเรียบแล้วไม่เกิดการกระโดด (Excite Value)

ข. วิธีการปรับเรียบโดยค่าเฉลี่ยเคลื่อนไหว (Moving Average Smoothing)

ทำได้โดยการกำหนดจุดข้อมูลที่เราสนใจ ทำการกำหนดหน้าต่างขึ้นเหมือนกับวิธีการปรับเรียบโดยค่ากลาง แต่แทนที่จะหาค่ากลาง เราจะหาค่าเฉลี่ยของข้อมูลภายในกรอบหน้าต่างแทนแล้วแทนค่าเฉลี่ยที่ได้ลงที่จุดกึ่งกลางของหน้าต่าง ดังแสดงในสมการที่ 2.7

$$E_n = \frac{1}{m1+m2+1}\sum_{i=n-m1}^{n+m2} E_i \qquad (2.7)$$

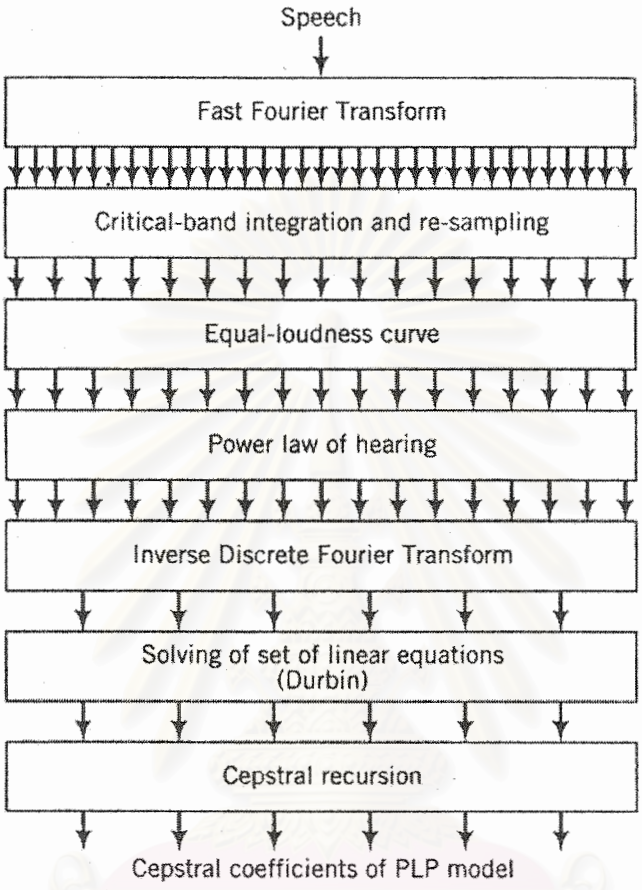
โดยที่ $m1$ คือ ความกว้างของกรอบหน้าต่างครึ่งซ้าย

$m2$ คือ ความกว้างของกรอบหน้าต่างครึ่งขวา

2.2.3 **การทำนายเชิงเส้นแบบรับรู้** (Perceptual Linear Prediction: PLP)

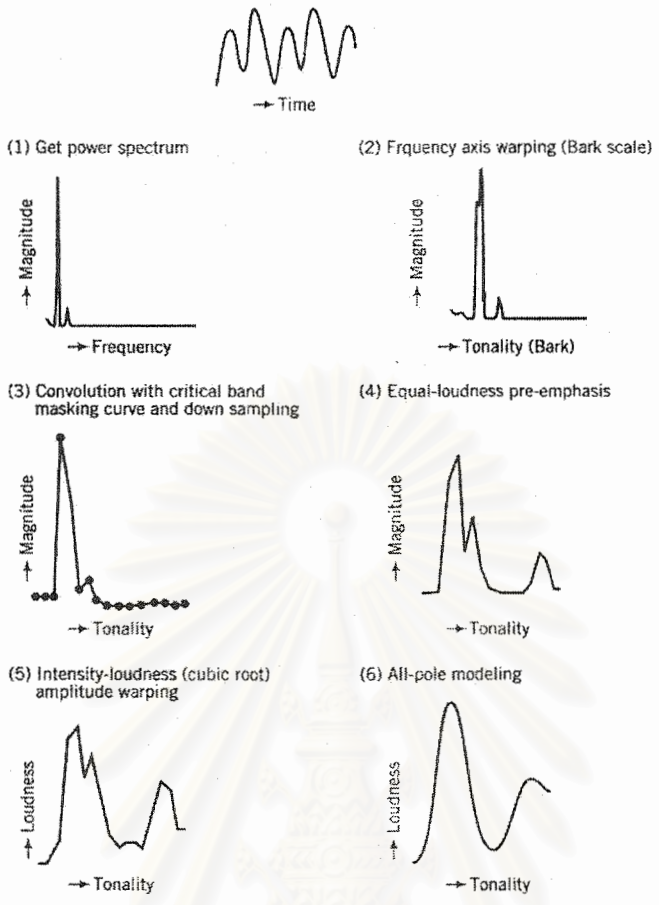
จากระบบการได้ยินของมนุษย์ที่ว่า การรับรู้เสียงในความถี่ต่างๆของหูมนุษย์นั้นไม่ได้มีสเกลเป็นเชิงเส้นตามหน่วยเฮิรตซ์ แต่แท้จริงแล้วเป็นสเกลลอการิทึม (Logarithmic scale) ซึ่งความรู้นี้สามารถที่จะนำไปประยุกต์ใช้ในการรู้จำเสียงพูดได้เป็นอย่างดี ดังนั้น Hermansky (1990) จึงได้เสนอลักษณะสำคัญ คือ การทำนายเชิงเส้นแบบรับรู้ : พีแอลพี (Perceptual Linear

Prediction: PLP) ซึ่งจะใช้เป็นตัวแทนสัญญาณเสียง เพื่อประยุกต์ใช้ในคำพูดและจะนำมาประยุกต์ใช้กับราสต้า ในการสร้างระบบการรู้จำคำพูดตัวเลขต่อเนื่องของผู้วิจัย ดังนี้



รูปที่ 2.4 ขั้นตอนการคำนวณพีแอลพี (Gold และ Morgan, 2000)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 2.5 ผลกระทบจากแต่ละขั้นตอนของพีแอลพี ที่มีต่อสเปกตรัม (GOLD และ MORGAN, 2000)

ในที่นี้เราจะกล่าวถึงขั้นตอนพื้นฐานอย่างสังเขปของการวิเคราะห์พีแอลพี ในแต่ละขั้นตอน และเราจะใช้รูปที่ 2.4 ซึ่งแสดงถึงกระบวนการของพีแอลพี ในการอธิบายและ รูปที่ 2.5 ในการอธิบายถึงตัวอย่างของผลกระทบที่ได้จากการดำเนินการในแต่ละขั้นตอน ซึ่งมีรายละเอียดดังต่อไปนี้

2.2.4 การแปลงฟูเรียร์แบบเร็ว (Fast Fourier Transform)

เริ่มต้นคำนวณค่าประมาณของสเปกตรัมกำลัง (power spectral estimate) สำหรับหน้าต่างวิเคราะห์ (analysis window) ซึ่งทำโดยการเอาหน้าต่างไปใส่ในขอบเขตที่จะวิเคราะห์ โดยการคูณแต่ละค่าของสัญญาณในกรอบข้อมูลเสียงด้วยค่าฟังก์ชันกรอบ เช่น หน้าต่าง Hamming ดังสมการ

$$W(n) = 0.54 + 0.46\cos\left(\frac{2\pi n}{N-1}\right) \qquad (2.8)$$

โดยที่ N คือ จำนวนข้อมูลต่อ 1 กรอบ

นำสัญญาณที่ได้ผ่านกระบวนการแปลงฟูเรียร์แบบเร็ว (Fast Fourier Transform: FFT) และคำนวณขนาดกำลังสองของมันพร้อมทั้งสเปกตรัมกำลัง (power spectrum) ดังสมการที่ 2.9

$$P(\omega) = \mathrm{Re}[S(\omega)]^2 + \mathrm{Im}[S(\omega)]^2 \tag{2.9}$$

โดยที่ $S(\omega)$ คือ สัญญาณในโดเมนความถี่ที่ผ่านการแปลงฟูเรียร์

ขั้นตอนนี้จะเทียบเท่ากับขั้นตอนแรกในรูปที่ 2.4 และผลที่ได้ก็คือสเปกตรัมกำลังซึ่งแสดงดังรูปย่อยที่ (1) ของรูปที่ 2.5

2.2.5 การหาปริพันธ์ของแถบวิกฤตและการชักตัวอย่างใหม่ (Critical-Band Integration and Re-sampling)

ตัวกรองซึ่งถูกทำให้เป็นรูปสี่เหลี่ยมคางหมู (ดังรูปที่ 2.6) จะถูกนำมาใช้ที่ช่วงห่างประมาณ 1 บาร์ก (Bark) ซึ่งแกนของบาร์กนั้นจะได้มาจากแกนความถี่ โดยใช้ฟังก์ชันวอร์ป (warping function) ของ Schroeder และ $P(\omega)$ ในแกนความถี่เฮริตซ์ (Hertz) จะถูกแปลงในอยู่ในแกนความถี่บาร์ก (Bark) โดยใช้สมการ ดังนี้

$$\Omega(\omega) = 6\ln\left\{\frac{\omega}{1200\pi} + \left[\left(\frac{\omega}{1200\pi}\right)^2 + 1\right]^{0.5}\right\} \tag{2.10}$$

โดยที่ ω คือ ความถี่เชิงมุม ในหน่วยของเรเดียนต่อวินาที (rad/s)

สำหรับหน้าต่างรูปสี่เหลี่ยมคางหมูนั้นก็คือการประมาณสเปกตรัมกำลังของเส้นโค้งแถบวิกฤต (critical band curve) ซึ่งจะเป็นดังสมการ 2.11

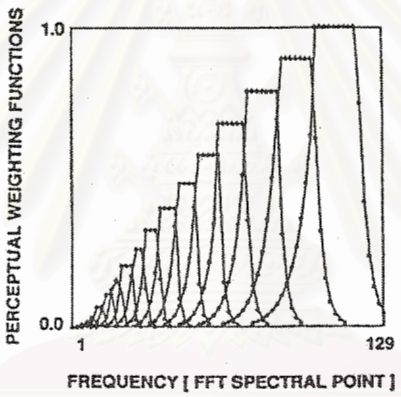
$$\Psi(\Omega) = \begin{cases} 0 & for \quad \Omega < -1.3 \\ 10^{2.5(\Omega+0.5)} & for \quad -1.3 \le \Omega \le -0.5 \\ 1 & for \quad -0.5 < \Omega < 0.5 \\ 10^{-1.0(\Omega-0.5)} & for \quad 0.5 \le \Omega \le 2.5 \\ 0 & for \quad \Omega > 2.5 \end{cases} \tag{2.11}$$

$P(\omega)$ จะถูกกระทำด้วยเส้นโค้งแถบวิกฤตโดยใช้สมการ 2.12 การกระทำนี้เพื่อลดความละเอียดของ $\Theta(\Omega)$ เมื่อเทียบกับ $P(\omega)$

$$\Theta(\Omega_i) = \sum_{\Omega=-1.3}^{2.5} P(\Omega - \Omega_i)\Psi(\Omega) \qquad (2.12)$$

$\Theta(\Omega)$ ที่ได้เรียกว่าสเปกตรัมกำลังแถบวิกฤต (critical-band power spectrum) ซึ่งจะมีทั้งหมด 18 ค่า ครอบคลุม 0-16.9 บาร์ก (0-5 กิโลเฮิรตซ์) และแต่ละค่าจะปรากฏที่ตำแหน่งต่างกัน 0.994 บาร์ก

สำหรับผลกระทบสุทธินั้นจะใช้เพื่อลดความไวทางความถี่ของการประมาณค่าสเปกตรัมดั้งเดิม โดยเฉพาะอย่างยิ่งที่ความถี่สูง ซึ่งความถี่สูงๆนั้นจะถูกเน้นย้ำโดยใช้ความกว้างแถบกรอง (filter bandwidth) ที่กว้างขึ้น ซึ่งผลที่ได้ดังแสดงในรูปที่2.6



รูปที่ 2.6 ตัวกรองรูปสี่เหลี่ยมคางหมูของพีแอลพี

2.2.6 โค้งความดังเทียบเท่า (Equal-Loudness Curve)

ทำการเน้นสเปกตรัมอีกครั้งหนึ่งเพื่อประมาณค่าความไวที่ไม่สมดุลของการได้ยินของมนุษย์ ณ ความถี่ต่างๆกัน ขั้นตอนนี้จะถูกทำให้เกิดผลด้วยการถ่วงน้ำหนักในของส่วนของสเปกตรัมแถบวิกฤต ซึ่งขั้นตอนนี้ก็เทียบได้กับขั้นตอนของโค้งความดังเทียบเท่า (equal-loudness curve) ในรูปที่ 2.4

$\Theta[\Omega(\omega)]$ จะถูกเน้นสัญญาณโดยเส้นโค้งความดังเท่าจำลอง (simulated equal-loudness curve) โดยใช้สมการ 2.13

$$\Xi[\Omega(\omega)] = E(\omega)\Theta[\Omega(\omega)] \qquad (2.13)$$

โดยที่ $E(\omega)$ คือ ค่าประมาณของความไวในการรับเสียงของหูมนุษย์ที่ความถี่ต่างๆ ซึ่งถูกจำลองที่ความดัง –40 เดซิเบล และมีรูปแบบดังสมการ 2.14

$$E(\omega) = \frac{[(\omega^2 + 56.8\times10^6)\omega^4]}{[\omega^2 + 6.3\times10^6)^2 \times (\omega^2 + 0.38\times10^9)(\omega^6 + 9.58\times10^{26})]} \qquad (2.14)$$

2.2.7 กฎกำลังของการได้ยิน (Power Law of Hearing)

เป็นการบีบแอมปลิจูดรากที่สาม (cubic-root amplitude compression) ทำการบีบอัดขนาดของสเปกตรัม ซึ่งโดยทั่วไปแล้วจะมีการหาค่าลอการิทึมหลังจากการหาปริพันธ์ มักจะใช้การหารากที่สามแทนการหาค่าลอการิทึม แสดงได้ดังสมการที่ 2.15

$$\Phi(\Omega) = \Xi(\Omega)^{0.33} \qquad (2.15)$$

การคำนวณนี้เป็นการประมาณด้วยกฎกำลัง (power law) เพื่อที่จะจำลองความสัมพันธ์แบบไม่เชิงเส้นระหว่างความเข้มของเสียง (intensity) และความรู้สึกถึงความดังของเสียง (loudness) ซึ่งกระบวนการนี้จะช่วยลดความแปรปรวนของขนาด (amplitude) ของสเปกตรัมแถบวิกฤต (critical-band spectrum) หรือการสั่นพ้องของสเปกตรัม (spectral resonances)

2.2.8 การแปลงฟูเรียร์แบบไม่ต่อเนื่องผกผัน (Inverse Discrete Fourier Transform)

ทำการแปลงฟูเรียร์แบบไม่ต่อเนื่องผกผัน (Discrete Fourier Transform: DFT) เนื่องจากค่าลอการิทึมไม่ได้ถูกคำนวณ ดังนั้นผลที่ได้จึงมักจะคล้ายกับค่าสัมประสิทธิ์อัตสหสัมพันธ์ (autocorrelation coefficients) มากกว่าถึงแม้ว่ามันจะมาจากสเปกตรัมที่ถูกบีบอัดก็ตาม และเนื่องจากค่าสเปกตรัมกำลังนั้นเป็นจำนวนจริงและเป็นเลขคู่ ดังนั้นจึงไม่จำเป็นที่จะต้องทำการคำนวณส่วนประกอบโคไซน์ของการผกผันผลของการแปลงฟูเรียร์ก็ได้

$\Phi(\Omega)$ จะถูกประมาณโดยสเปกตรัมของแบบจำลองทุกขั้ว (all-pole) ด้วยวิธีอัตสหสัมพันธ์ (autocorrelation)

2.2.9 การแก้ชุดสมการเชิงเส้น (Solving of Set of Linear Equations)

การหาปริพันธ์นั้นจะมีประโยชน์ต่อการลดผลกระทบของต้นกำเนิดซึ่งไม่เกี่ยวข้องกับทางภาษาศาสตร์ของความแปรปรวนในสัญญาณเสียงพูดได้แบบจำลองอัตราถดถอย

(autoregressive model)จะได้มาจากผลเฉลยของสมการเชิงเส้น ซึ่งถูกสร้างขึ้นจากอัตสหสัมพันธ์ของขั้นตอนก่อนหน้านี้นั่นเอง โดยแบบจำลองอัตตาถดถอยนี้จะถูกใช้ในการขัดเกลาสเปกตรัมแถบวิกฤตซึ่งถูกบีบอัดแล้ว เหมือนกับในการทำ LPC แบบดั้งเดิม โดยสเปกตรัมผลลัพธ์ซึ่งถูกขัดเกลาแล้วจะเหมาะสมกับยอดสเปกตรัม (spectral peak) มากกว่าเหวสเปกตรัม (spectral valley) นักวิจัยหลายท่านได้พบแล้วว่าวิธีนี้จะนำไปสู่ความทนทานต่อเสียงรบกวน และการไม่ขึ้นกับผู้พูดได้ดีกว่าการตัดปลายเซปสตรัม สำหรับรูปย่อยที่ (6) ในรูปที่ 2.5 นั้นจะแสดงถึงตัวแทนซึ่งถูกขัดเกลาแล้วของยอดซึ่งเป็นชนิดหนึ่งของผลที่ได้มาจากกระบวนการพีแอลพี นั่นเอง

2.2.10 การเวียนเกิดเซปสตรัล (Cepstral Recursion) [16]

ทำการใช้ตัวแทนเชิงตั้งฉาก (Orthogonal representation) สัมประสิทธิ์อัตตาถดถอย (autoregressive coefficients) จะถูกแปลงให้เป็นตัวแปรเซปสตรัมแทน

2.2.11 ราสต้า-พีแอลพี (RASTA Speech Analysis)

เนื่องจากเทคนิคของการวิเคราะห์เสียงพูดแบบพีแอลพีนั้นตั้งอยู่บนพื้นฐานของสเปกตรัมแบบช่วงสั้นของเสียงพูด ทำให้มีจุดอ่อนในด้านความคนทนต่อสภาวะแวดล้อม โดยเฉพาะอย่างยิ่งเมื่อค่าของสเปกตรัมถูกรบกวนโดยการตอบสนองความถี่ของช่องสัญญาณคมนาคม หรือพูดอย่างง่ายๆก็คือพีแอลพีนั้นจะได้ผลไม่ดีถ้าเกิดมีเสียงรบกวนในระหว่างการบันทึกเสียง เช่น บันทึกเสียงจากการพูดทางโทรศัพท์ หรือบันทึกเสียงโดยส่งสัญญาณเสียงผ่านสายไฟยาวๆหรือสายไฟที่มีราคาถูก เป็นต้น ดังนั้นจึงได้มีการคิดค้นเทคนิคของสเปกตรัมสัมพันธ์ : ราสต้า (RelAtive SpecTrAl: RASTA) ขึ้นมา ซึ่งเทคนิคนี้สามารถที่จะช่วยให้พีแอลพี นั้นมีความคงทนต่อการบิดเบี้ยวของสเปกตรัมเชิงเส้นได้ดีขึ้น ดังนั้นเราจึงเรียกวิธีการของพีแอลพี ซึ่งนำเทคนิคของราสต้าเข้ามาช่วยนี้ว่าราสต้า-พีแอลพี [16]

ขั้นตอนในการทำราสต้า-พีแอลพี สำหรับแต่ละกรอบที่จะวิเคราะห์มีดังต่อไปนี้

1. คำนวณสเปกตรัม แถบวิกฤตเหมือนในพีแอลพี โดยใช้ขั้นตอนวิธีเดียวกัน
2. ประมาณค่าอนุพันธ์เวลาของสเปกตรัมแถบวิกฤตเชิงล็อก ด้วยการใช้เส้นของการถดถอย (regression line) โดยให้ผ่านค่าสเปกตรัมที่ต่อเนื่องกัน 5 ค่า การประมวลผลแบบไม่เชิงเส้น เช่น การกรองขีดจำกัดหรือมัธยฐาน สามารถที่จะทำได้ในโดเมนนี้ แต่อาจจะทำหรือไม่ทำก็ได้ ขึ้นอยู่กับความเหมาะสม
3. ทำการหาปริพันธ์ (integrate) ของค่าอนุพันธ์เวลาแถบวิกฤตเชิงล็อกอีกครั้งหนึ่งโดยใช้ระบบ IIR อันดับที่หนึ่ง โดยตำแหน่งขั้วของระบบนี้สามารถที่จะจัดให้อยู่ในขนาด

หน้าต่างที่เหมาะสมได้ โดยทั่วไปมักจะตั้งค่านี้เป็น 0.98 และทำให้ได้หน้าต่างปริพันธ์ของการยกกำลังมีจุดที่ 3 เดซิเบลหลังจาก 34 กรอบผ่านไป

4. ในทำนองเดียวกับพีแอลพี ทั่วๆไป เราเพิ่มโค้งความดังเทียบเท่าแล้วคูณด้วย 0.33 เพื่อกระตุ้นพลังแห่งการได้ยิน
5. หาค่าล็อกผกผัน (ซึ่งก็คือฟังก์ชันยกกำลัง) ของสเปกตรัมล็อกที่เกี่ยวพัน เพื่อให้ได้สเปกตรัมการได้ยินที่เกี่ยวพัน
6. คำนวณแบบทุกขั้วของสเปกตรัมนี้ ตามเทคนิคพีแอลพี ทั่วๆไป

จะเห็นได้ว่าถ้าอนุพันธ์ในขั้นตอนที่ 2 ถูกประมาณด้วยผลต่างที่หนึ่งอย่างง่าย หรือถ้าการหาปริพันธ์แบบเต็มในขั้นตอนที่ 4 ทำเสร็จเรียบร้อย (ขั้วที่ $z = 1.0$) แล้ว เทอมทั้งหมดในระหว่างนี้ก็จะถูกยกเลิก และเทคนิคนี้ก็จะเหมือนกับการลบสเปกตรัมล็อกของการวิเคราะห์กรอบแรกออกจากกรอบใหม่แต่ละกรอบ ซึ่งในกรณีพิเศษนี้ เทคนิคราสต้า ก็จะเหมือนกับการลบสเปกตรัมหรือการใช้เทคนิคสังวัตนาการบอด (blind convolution)

อย่างไรก็ตาม ในกรณีทั่วๆไปซึ่งแสดงในที่นี้ กระบวนการหาอนุพันธ์และปริพันธ์ทั้งหมดก็เทียบเท่ากับการกรองผ่านแถบของช่องความถี่แต่ละช่องด้วยตัวกรอง IIR โดยใช้ฟังก์ชันโอนย้ายดังสมการที่ 2.16

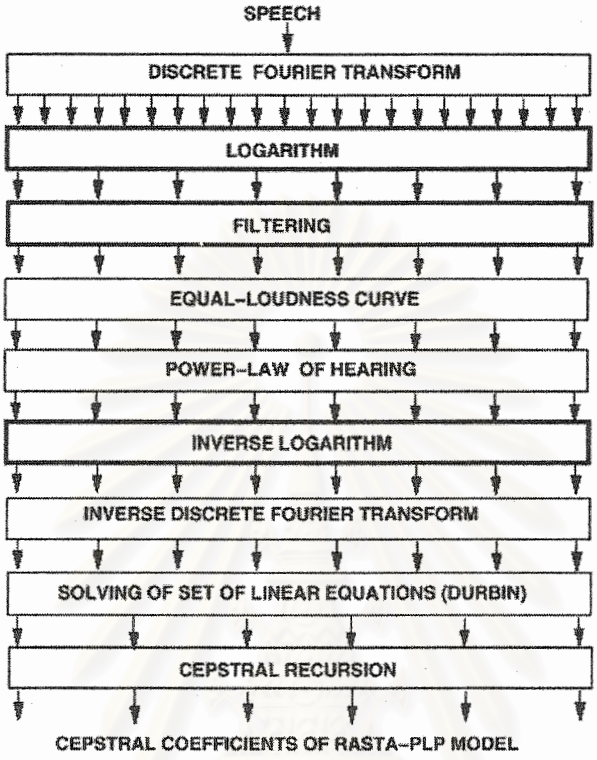
$$H(z) = 0.1 \times \frac{2 + z^{-1} - z^{-3} - 2z^{-4}}{z^{-4} \times (1 - 0.98z^{-1})} \qquad (2.16)$$

ความถี่ตัดออกด้านต่ำของตัวกรองจะเป็นตัวตัดสินการเปลี่ยนแปลงสเปกตรัมที่เร็วที่สุดของสเปกตรัมล็อกซึ่งถูกทิ้งไปในข้อมูลออก ในขณะที่ความถี่ตัดออกด้านสูงจะเป็นตัวตัดสินการเปลี่ยนแปลงสเปกตรัมที่เร็วที่สุดซึ่งถูกสงวนเอาไว้

การบิดเบี้ยวเชิงเส้นซึ่งมีสาเหตุมาจากช่องสัญญาณในการโทรคมนาคม หรือจากการใช้ไมโครโฟนที่แตกต่างกันไปนั้น จะปรากฏเป็นค่าบวกในสเปกตรัมล็อก และส่วนผ่านสูงของตัวกรองผ่านแถบเทียบเท่าสามารถที่จะลดผลกระทบของเสียงรบกวนสังวัตนาการ (convolutional noise) ซึ่งก่อกำเนิดขึ้นในช่องสัญญาณได้ นอกจากนั้นการกรองผ่านต่ำจะช่วยในการปรับการเปลี่ยนแปลงสเปกตรัมแบบกรอบต่อกรอบอย่างรวดเร็วบางส่วนซึ่งเกิดขึ้นในการประมาณสเปกตรัมแบบช่วงสั้น เนื่องจากเครื่องมือที่ใช้การวิเคราะห์ให้เรียบยิ่งขึ้น สำหรับในสมการที่ 2.16 นั้น ความถี่ตัดออกด้านต่ำจะเป็น 0.26 เฮิรตซ์ และความชันในการกรองจะเอียงลาดลงด้วยอัตรา 6 เดซิเบลต่อ oct จาก 12.8 เฮิรตซ์ด้วยศูนย์ฉับพลันที่ 28.9 เฮิรตซ์และที่ประมาณ 50 เฮิรตซ์ จะไม่มีเหตุผลพิเศษ (ยกเว้นในทางประวัติศาสตร์) สำหรับการใช้ตัวกรองเฉพาะของสมการ 2.16 และตัวกรองเดียวกันนี้เอง จะไม่เป็นที่ต้องการสำหรับช่องสัญญาณความถี่ทั้งหมด ยิ่งไปกว่านั้น การกรองยังไม่จำเป็นที่จะต้องเป็นแบบผ่านแถบ หรือแม้แต่แบบเชิงเส้นก็ได้

ผลที่ได้โดยทั่วไปนั้นจะขึ้นกับจุดเริ่มต้นของการวิเคราะห์ ซึ่งในการประยุกต์ใช้ส่วนใหญ่ที่นิยมกันนั้นก็มักจะเริ่มการวิเคราะห์ที่ส่วนเงียบ ซึ่งจะอยู่ก่อนหน้าเสียงพูด

ขั้นตอนของราสต้า-พีแอลพี ทั้งหมดสามารถที่จะแสดงได้ดังรูปที่ 2.7



รูปที่ 2.7 ขั้นตอนของราสต้า-พีแอลพี (GOLD และ MORGAN, 2000)

อนุพันธ์อันดับที่หนึ่ง (Delta)

สำหรับอนุพันธ์อันดับที่หนึ่ง พีแอลพี และ ราสต้า นั้นจะมีรายละเอียดของสมการดังนี้

อนุพันธ์อันดับที่หนึ่งของ พีแอลพี และ ราสต้า (Delta PLP and Delta RASTA)

สำหรับอนุพันธ์อันดับที่หนึ่งของ พีแอลพี และอนุพันธ์อันดับที่หนึ่งของ ราสต้า นั้นจะใช้สมการเดียวกัน ซึ่งสามารถที่จะแสดงได้ดังสมการที่ 2.15

$$\Delta\hat{c}_{\ell}(m)=\left[\sum_{k=-K}^{K} k\hat{c}_{\ell-k}(m)\right];1\le m\le Q \qquad (2.17)$$

โดยที่ Q = จำนวนอันดับ (order) ของ พีแอลพี หรือ ราสต้า

$2K+1$ = จำนวนกรอบที่ใช้ในการคำนวณอนุพันธ์

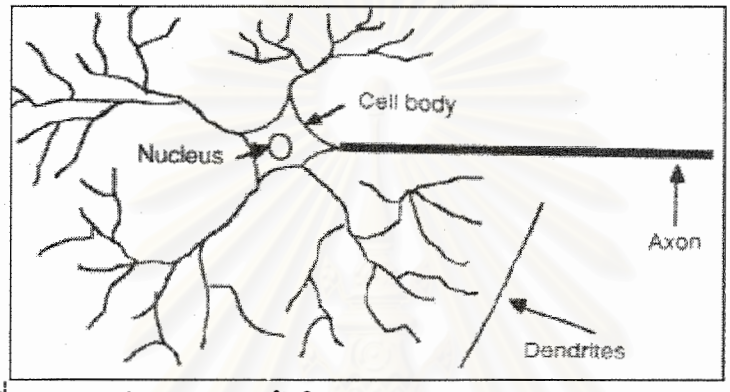
ℓ = กรอบปัจจุบัน (current frame) ที่นำมาคำนวณ

$\hat{c}$ = สัมประสิทธิ์ พีแอลพี หรือ ราสต้า

$\Delta\hat{c}$ = อนุพันธ์อันดับที่หนึ่งของสัมประสิทธิ์ พีแอลพี หรือ ราสต้า

2.3 ข่ายงานระบบประสาทเทียม (Artificial Neural Network: ANN)

ข่ายงานระบบประสาทเทียม เป็นชื่อของแขนงหนึ่งในสาขาวิทยาการคอมพิวเตอร์ ในการพยายามที่จะเลียนแบบการทำงานของสมองมนุษย์ เฉพาะงานที่คอมพิวเตอร์ทำไม่ได้ หรือทำได้ไม่ดี เช่น งานด้านการรู้จำภาพและรู้จำคำพูด เป็นต้น

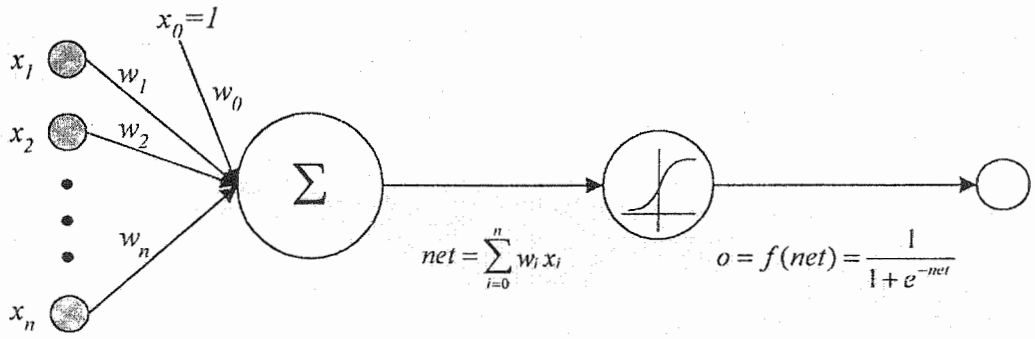


รูปที่ 2.8 ส่วนประกอบของเส้นใยประสาท(Gold และ Morgan,2000)

ข่ายงานระบบประสาทเทียมมีอยู่ด้วยหลายประเภท สำหรับข่ายงานระบบประสาทเทียมที่ใช้ในการทดลองเป็นเพอร์เซพตรอนแบบหลายชั้น (Multi-Layer Perceptron, MLP) โดยใช้ระเบียบวิธีการเรียนรู้แบบแพร่กระจายย้อนกลับ (Back-propagation Algorithm) ซึ่งมีข้อดี คือ การคำนวณง่ายไม่ซับซ้อน และให้ผลรู้จำที่ดี

2.3.1 เพอร์เซพตรอน (Perceptron)

เพอร์เซพตรอนถูกออกแบบมาเพื่อจำลองการทำงานของเซลล์ประสาท แสดงดังรูป 2.9 ข้อมูลเข้า (input) จะมาจากข้อมูลออก (output) ของเซลล์ประสาทตัวอื่น หรือจากสัญญาณภายนอก โดยขั้นแรกนั้น สัญญาณข้อมูลเข้าจะถูกคูณเข้ากับน้ำหนักของการเชื่อมต่อ (connection weight) จนครบทุกคู่ แล้วนำมารวมกันเพื่อบวกกับน้ำหนักอคติ (bias weight) น้ำหนักอคติจะเหมือนกับน้ำหนักของการเชื่อมต่อเพียงแต่มีข้อมูลเท่ากับ 1 ตลอดเวลา เมื่อผลรวมทั้งหมดแล้วจะนำมาผ่านฟังก์ชันแปลง (transform function) ซึ่งเป็นฟังก์ชันไม่เชิงเส้น (nonlinear function) ที่สามารถหาค่าอนุพันธ์ได้ สิ่งที่ได้คือสัญญาณข้อมูลออก (output) เพื่อส่งต่อไปยังเซลล์ประสาทตัวอื่นต่อไป



รูปที่ 2.9 เพอร์เซพตรอนโดยใช้ฟังก์ชันซิกมอยด์

จากรูป 2.9 เป็นการหาค่าข้อมูลออก (output; o) โดยมีข้อมูลเข้า (input, x) จำนวน n ตัว และใช้ฟังก์ชัน sigmoid เป็นฟังก์ชันการแปลง แสดงได้ดังสมการ 2.18

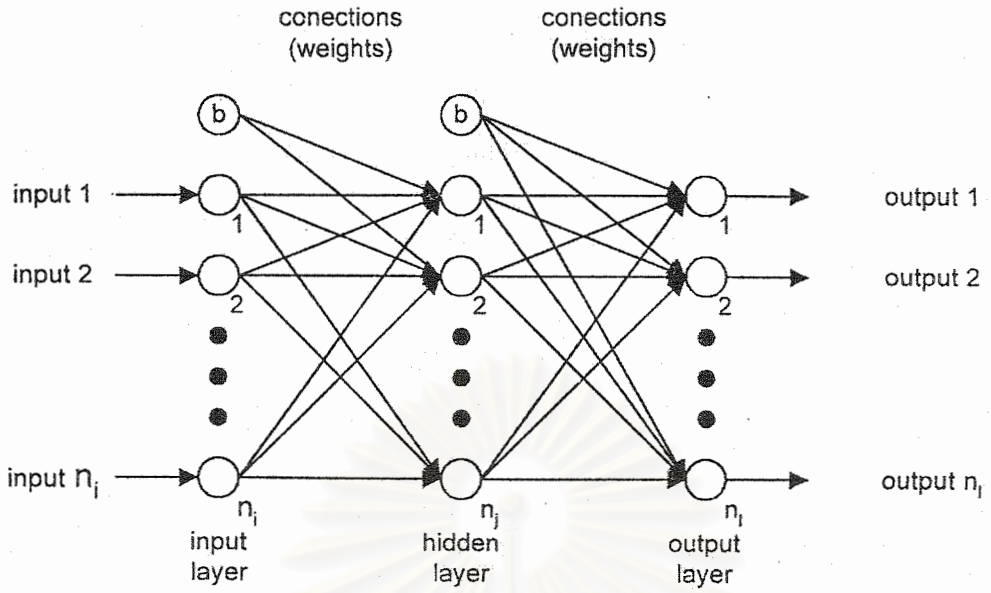
$$o = f(net) = \frac{1}{1+e^{-net}} \qquad (2.18)$$

เมื่อ $net = \sum_{i=0}^{n} w_i x_i$

เพอร์เซพตรอนแบบหลายชั้น (Multi-Layer Perceptron, MLP)

ข่ายงานระบบประสาทเทียมที่ใช้ในงานวิจัยนี้เป็นเพอร์เซพตรอนแบบ 3 ชั้น ประกอบด้วยชั้นข้อมูลเข้า (input layer) ชั้นข้อมูลซ่อนตัว (hidden layer) และชั้นข้อมูลออก (output layer) แต่ละชั้นจะประกอบด้วยเพอร์เซพตรอนหลายตัวเชื่อมต่อกันข่ายงานระบบประสาทเทียมที่ใช้ในการทดลองนี้เป็นเพอร์เซพตรอนแบบ 3 ชั้น ประกอบด้วยชั้นข้อมูลเข้า (input layer) ชั้นข้อมูลซ่อนตัว (hidden layer) และชั้นข้อมูลออก (output layer) เป็นโครงข่ายดังรูปที่ 2.10

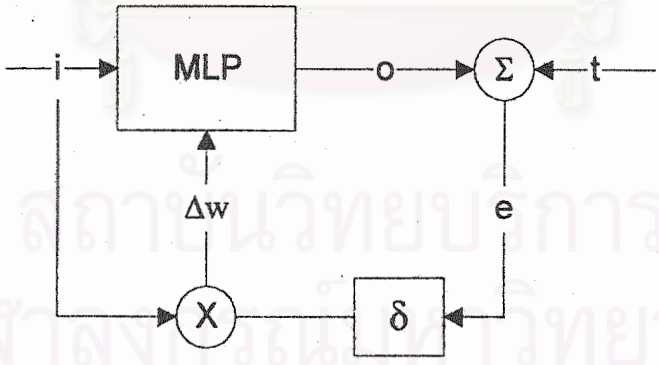
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 2.10 โครงสร้างของข่ายงานระบบประสาทเทียมที่ประกอบด้วยเพอร์เซพตรอนแบบหลายชั้น (GOLD และ MORGAN, 2000)

2.3.2 การฝึกข่ายงานระบบประสาทเทียม (Artificial Neural Network Training)

การฝึกข่ายงานระบบประสาทเทียม คือ การปรับน้ำหนักของการเชื่อมต่อเพื่อให้ข้อมูลออกมีค่าใกล้เคียงหรือตรงกับค่าเป้าหมาย (target) สรุป คือ ฝึกเพื่อหาค่าน้ำหนักของการเชื่อมต่อที่ทำให้ความผิดพลาดระหว่างค่าเป้าหมายและค่าข้อมูลออกกำลังสองน้อยที่สุด แสดงดังรูปที่ 2.11



รูปที่ 2.11 โครงสร้างขั้นตอนในการฝึกฝนแบบการแพร่กระจายย้อนกลับ

นิยามค่าความผิดพลาดดังกล่าวได้ดังสมการ

$$E[\vec{w}] = \frac{1}{2}\sum_{d\in D}(t_d - o_d)^2 \qquad (2.19)$$

โดยที่ D คือ เซตของตัวอย่างที่ต้องการฝึก

t_d คือ ค่าเป้าหมายตัวที่ d

o_d คือ ค่าข้อมูลออกตัวที่ d

การคำนวณการปรับน้ำหนักสำหรับแต่ละค่าน้ำหนักทำได้ดังนี้

$$\Delta w_{i,j}(n) = \eta \delta_j x_{i,j} + \alpha \Delta w_{i,j}(n-1) \quad (2.20)$$

โดยที่ η คือ อัตราการเรียนรู้ (learning rate)

α คือ โมเมนตัม (momentum) มีหน้าที่รักษาน้ำหนักให้ไปในทิศทางเดิม ซึ่งทำให้ลู่เข้าเร็วขึ้น

สำหรับเส้นเชื่อมต่อระหว่างชั้นข้อมูลซ่อนตัวกับชั้นข้อมูลออก

$$\delta_k = o_k(1-o_k)(t_k - o_k) \quad (2.21)$$

สำหรับเส้นเชื่อมต่อระหว่างชั้นข้อมูลเข้ากับชั้นข้อมูลซ่อนตัว

$$\delta_k = o_h(1-o_h)\sum_{k \in outputs} w_{h,k}\delta_k \quad (2.22)$$

การปรับค่าน้ำหนักของเส้นเชื่อมจากบัพ i ไปยังบัพ j ใช้สมการ 2.23

$$w_{i,j}(n+1) = w_{i,j}(n) + \Delta w_{i,j}(n) \quad (2.23)$$

เมื่อปรับค่าน้ำหนักการเชื่อมต่อหมดทุกค่าแล้ว จะทำการแพร่กระจายค่าข้อมูลขาเข้าอีกครั้ง โดยใช้ค่าน้ำหนักการเชื่อมต่อค่าใหม่ที่ได้จากการปรับ กระทำเช่นนี้ต่อไปเรื่อย ๆ จนกระทั่งค่าความผิดพลาด (E) น้อยกว่าค่าความผิดพลาดที่ยอมรับได้ (Error Threshold) หรือกระทำจนครบจำนวนรอบที่ต้องการ เป็นอันว่าสิ้นสุดการฝึกฝนข่ายงานระบบประสาทเทียม และจะเก็บค่าน้ำหนักการเชื่อมต่อสุดท้ายไว้สำหรับการทดสอบการรู้จำ

ขั้นตอนการรู้จำ

1. ป้อนค่าข้อมูลขาเข้าของชุดทดสอบเข้าไปยังระดับชั้นข้อมูลเข้า
2. แพร่กระจายข้อมูลขาเข้า เข้าไปในข่ายงานระบบประสาทเทียมทำนองเดียวกับการฝึก โดยอาศัยค่าน้ำหนักการเชื่อมต่อที่ได้จากการฝึกฝน จนกระทั่งได้ค่าข้อมูลขาออกที่ระดับชั้นข้อมูลออก
3. ทำการตัดสินใจโดยใช้กฎเกณฑ์การตัดสินใจ (Decision rule) โดยการดูจากค่าข้อมูลขาออกที่มากที่สุดในกลุ่ม ดังสมการ 2.24

$$Unknown\ pattern = pattern_i, \quad o_i = \max(o_1, o_2, \ldots, o_c) \quad (2.24)$$

โดยที่ c คือจำนวนบัพที่ระดับชั้นข้อมูลขาออก

2.4 การสังเคราะห์เสียงพูด (Speech Synthesis)

โครงสร้างของระบบสังเคราะห์เสียงโดยทั่วไป สามารถแบ่งการทำงานได้เป็น 3 ส่วนได้แก่

2.4.1 ส่วนการวิเคราะห์ข้อความ (Text Analysis)

ส่วนนี้จะมีหน้าที่วิเคราะห์ข้อความอินพุทเพื่อแปลงข้อมูล เสียงอ่าน (phoneme) ของคำนั้น และส่งต่อให้ส่วนของการสังเคราะห์เสียง (Speech Synthesis) ต่อไป นอกจากนี้ส่วนนี้ยังทำหน้าที่อย่างอื่น เช่น การแบ่งประโยคจากข้อความที่ยาว (Sentence breaking) การหาขอบเขตของวลีของการอ่านในประโยค

2.4.2 ส่วนการวิเคราะห์สัทสัมพันธ์ (Prosody Analysis)

ส่วนนี้ทำหน้าที่ในการวิเคราะห์และสังเคราะห์ข้อมูล สัทสัมพันธ์ (prosody) ของประโยคใดๆจากข้อมูลเสียงอ่าน และข้อความ ข้อมูลสัทสัมพันธ์ที่วิเคราะห์ออกมาได้ในระบบทั่วไป เช่น

Segment Duration หมายถึง ความยาวของเสียงย่อยที่ต้องการสังเคราะห์ ค่านี้จะมีผลต่อจังหวะของเสียงที่ทำการสังเคราะห์ เช่น ถ้ากำหนดให้ค่าความยาวของเสียงย่อยที่ต้องการสังเคราะห์มีขนาดสั้น เสียงที่ทำการสังเคราะห์ก็จะเหมือนกับการพูดเร็ว

Pitch Contour หมายถึง ค่าความสัมพันธ์ของความถี่มูลฐานกับเวลา ค่านี้จะมีผลต่อเสียงสูงต่ำ

2.4.3 ส่วนการสังเคราะห์เสียง (Speech Synthesis)

ส่วนนี้ทำหน้าที่ในการสร้างสัญญาณคลื่นเสียง จากข้อมูลเสียงอ่าน (phonetic transcription) และข้อมูลสัทสัมพันธ์(Prosody Transcription) จากส่วนของการวิเคราะห์ข้อความและส่วนของการวิเคราะห์สัทสัมพันธ์ และส่งออกสู่ลำโพง เพื่อให้เราได้ยินเสียงพูดประโยคนั้นๆ โดยทั่วไป ส่วนนี้สามารถแบ่งตามเทคนิควิธีการสังเคราะห์เสียง ได้ 3 ประเภท คือ

2.4.3.1 Formant Synthesis

เทคนิคการสังเคราะห์วิธีแบบนี้ ข้อมูลเสียงอ่านใดๆจะถูกกำหนด ไว้อยู่ในรูปของความถี่ฟอร์แมนท์ต่างๆ (F1,F2,F3) ของเสียงนั้นๆ เมื่อต้องการสังเคราะห์เสียงใดๆก็นำข้อมูลเหล่านี้มาทำการสังเคราะห์ให้เป็นสัญญาณเสียง ซึ่งวิธีการนี้จะมีข้อดีที่สามารถควบคุมค่าความเปลี่ยนแปลงของความถี่ฟอร์แมนท์ (Formant transition) ที่บริเวณรอยต่อระหว่างเสียงได้ง่าย แต่มีข้อเสียคือ การจะแทนเสียงใดๆด้วยค่าฟอร์แมนท์ทำได้ยากจะต้องมีกฎในการสังเคราะห์เสียงจำนวนมาก และเสียงที่สังเคราะห์ออกมาได้จะไม่ค่อยเป็นธรรมชาติ

2.4.3.2 Articulation Synthesis

สำหรับวิธีนี้ข้อมูลเสียงที่ต้องการสังเคราะห์ จะอยู่ในรูปของค่าพารามิเตอร์ของโครงสร้างทางกายภาพของการเคลื่อนไหวของอวัยวะในช่องปากที่ทำให้เกิดเสียงต่างๆ วิธีนี้ค่อนข้างยากในแง่การโมเดลเสียงต่างๆ ซึ่งจะต้องศึกษาจากอวัยวะในการออกเสียงจริงๆ

2.4.3.3 Concatenation Synthesis

เสียงที่ทำการสังเคราะห์ขึ้นเกิดจากการนำหน่วยเสียงย่อย ที่ทำการเก็บไว้ก่อนแล้วมาต่อกันเป็นเสียงพูดที่ต้องการ โดยทั่วไปหน่วยเสียงย่อยที่ทำเก็บไว้จะอยู่ระดับต่ำกว่าคำ เช่น หน่วยเสียงพยางค์ หน่วยเสียงครึ่งพยางค์ (demesyllable) หน่วยเสียงของเสียงคู่เสียง (diphone) เป็นต้น

2.5 ทฤษฎีวิภัชนัย (Fuzzy Theory)

ในงานทั่วๆไปที่ต้องมีการตัดสินใจ โดยที่ปัจจัยที่นำมาพิจารณาในการตัดสินใจมีความซับซ้อน กำกวม ยากต่อการตัดสินใจ จะพบว่ามนุษย์สามารถทำการตัดสินใจได้ดีกว่าเครื่องจักรกล ตัวอย่างเช่น ต้องการจะคัดส้มเพื่อแบ่งแยกราคาขาย ปัจจัยที่ใช้ในการพิจารณาไม่ใช่เพียงขนาดของส้มเท่านั้น แต่ยังต้องพิจารณาสีผิวภายนอก น้ำหนักหรือความแน่นของเนื้อส้มและน้ำส้มในผลส้ม นอกจากนี้ยังอาจจะมีปัจจัยนอกเหนือจากที่คาดคิดออกไปได้อีก เช่น สภาวะเศรษฐกิจ สถานะทางบ้าน เป็นต้น ในกรณีดังกล่าวนี้เราไม่สามารถสร้างสมการทางคณิตศาสตร์ที่ครอบคลุมปัจจัยเหล่านี้เพื่อใช้ในการตัดสินใจได้ง่าย ทฤษฎีทางด้านวิภัชนัยถูกสร้างขึ้นมาเพื่อกรณีเหล่านี้

ข้อมูลที่พบอยู่ในโลกปัจจุบันสามารถแบ่งออกเป็นข้อมูลที่แน่นอน (Certain Information) และข้อมูลที่มีความไม่แน่นอน (Uncertain Information) ซึ่งทฤษฎีเซตวิภัชนัยเป็นทฤษฎีทางคณิตศาสตร์ที่ใช้กับข้อมูลที่มีความไม่แน่นอน เช่นเดียวกับทฤษฎีทางสถิติ (Statistic Theory)

2.6 เซตวิภัชนัย (Fuzzy Set)

เซตวิภัชนัย คือ เซตที่มีการสนใจระดับมากน้อยของการเป็นสมาชิกแต่ละตัวต่างกับเซตคริสป์ (Crisp Set) ที่สนใจเพียงว่าสมาชิกที่สนใจอยู่ในเซตหรือไม่ โดยจะอธิบายดังต่อไปนี้

เซตแบบแน่นอน (Crisp Set)

กำหนดให้ A และ B เป็นเซตที่ประกอบด้วยสมาชิก (Member) ดังนี้

$$A = \{a, b, c\} \tag{2.25}$$

$$B = \{b, c, d\}$$

กำหนดให้ $\mu_x(y) \in \{0,1\}$ เป็นค่าสมาชิกภาพ (Membership Value) ของ y ในเซต X ซึ่งมีความหมายคือ $\mu_x(y)$ เป็นตัวเลขที่แสดงความเป็นสมาชิกของ y ในเซต X ถ้า $\mu_x(y)$ มีค่ามาก

แสดงว่า *y* มีความเป็นสมาชิกในเซต *X* มาก เมื่อพิจารณาเซตแบบแน่นอน *A* และ *B* จะพบว่า *a*, *b* และ *c* เป็นสมาชิกของเซต *A* แน่นอน และ *b*,*c* และ *d* เป็นสมาชิกของเซต *B* แน่นอน ดังนั้น ค่าสมาชิกภาพจะเป็นดังนี้

$$\begin{aligned} \mu_A(a)&=1 \qquad & \mu_B(a)&=0 \\ \mu_A(b)&=1 \qquad & \mu_B(b)&=1 \\ \mu_A(c)&=1 \qquad & \mu_B(c)&=1 \\ \mu_A(d)&=0 \qquad & \mu_B(d)&=1 \end{aligned} \tag{2.26}$$

หรือสามารถเขียนรวมอยู่ในเซต ได้ดังนี้

$$\begin{aligned} A&=\left\{\frac{a}{1},\frac{b}{1},\frac{c}{1},\frac{d}{0}\right\} \\ B&=\left\{\frac{a}{0},\frac{b}{1},\frac{c}{1},\frac{d}{1}\right\} \end{aligned} \tag{2.27}$$

ดังนั้นสำหรับเซตแบบแน่นอน ค่าสมาชิกภาพ $\mu_x(y)$ ของสมาชิกในเซตจะมีเพียง 1 หรือ 0 เท่านั้น สามารถเขียนสมการแสดงค่าสมาชิกอยู่ในรูปทั่วไปดังนี้

$$\mu_x(y)=\begin{cases}1, & y\in X \\ 0, & y\notin X\end{cases} \tag{2.28}$$

เซตวิภัชนัย (Fuzzy Set) ประกอบด้วยสมาชิกที่มีค่าสมาชิกภาพ $\mu_x(y)\in\{0,1\}$ ยกตัวอย่างเช่น

$$\begin{aligned} \tilde{A}&=\left\{\frac{a}{1},\frac{b}{1},\frac{c}{0.5},\frac{d}{0}\right\} \\ \tilde{B}&=\left\{\frac{a}{0},\frac{b}{0.4},\frac{c}{0.9},\frac{d}{1}\right\} \end{aligned} \tag{2.29}$$

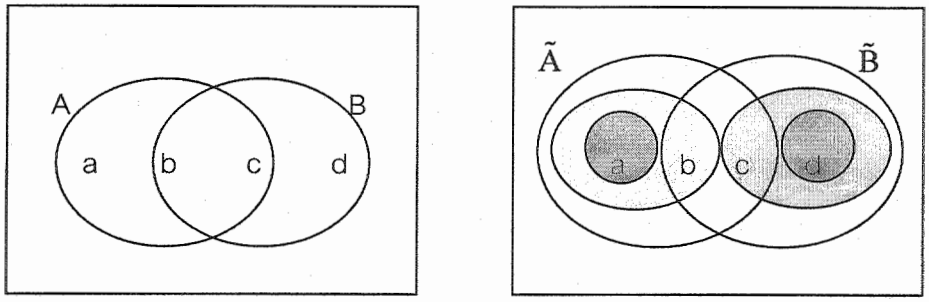
ซึ่งสามารถเขียนเซตวิภัชนัยอยู่ในรูปทั่วไปได้ดังนี้

$$\tilde{X}=\left\{\frac{y}{\mu_{\tilde{x}}(y)}\right\}, \qquad \mu_{\tilde{x}}(y)\in\{0,1\}$$

(2.30)

และสามารถเขียนเป็นแผนภาพของเซต *A*, *B* ซึ่งเป็นเซตแบบแน่นอน และเซต $\tilde{A},\tilde{B}$ ซึ่งเป็นเซตวิภัชนัยได้ดังรูปที่ 2.12

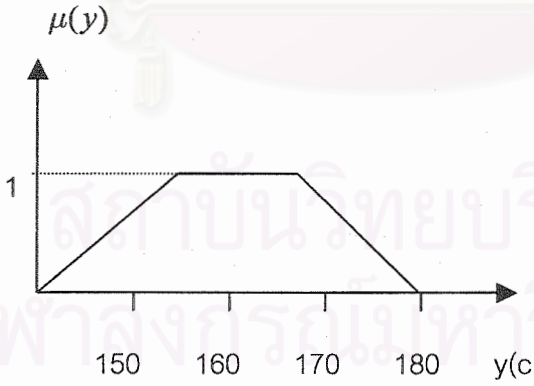
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 2.12 แผนภาพแสดงสมาชิกในเซต A,B และ $\tilde{A}$, $\tilde{B}$

2.6.1 **ฟังก์ชันสมาชิกภาพแบบวิภัชนัย** (Fuzzy Membership Function)

ฟังก์ชันสมาชิกภาพแบบวิภัชนัย คือ ฟังก์ชันที่ใช้หาค่าสมาชิกภาพ $\mu_{\tilde{x}}(y)$ ของสมาชิก y ในเซตแบบวิภัชนัย $\tilde{X}$ ในกรณีที่สมาชิกในเซตมีจำนวนจำกัด และไม่มากเกินไปสามารถเขียนแบบแจกแจงพร้อมทั้งแสดงค่าสมาชิกภาพได้ ดังเช่นในสมการที่ 2.28 แต่ถ้าสมาชิกในเซตเป็นค่าต่อเนื่อง เช่น เซต $\tilde{A}$ เป็นเซตของความสูงเป็นเซนติเมตรของคนที่จัดว่าเป็นคนสูงปานกลาง สมาชิกของเซตนี้อาจจะมีค่าสมาชิกภาพเป็นศูนย์ในช่วงความสูงน้อยกว่า 150 เซนติเมตร และความสูงมากกว่า 180 เซนติเมตร ส่วนในช่วงความสูงระหว่าง 150 ถึง 180 เซนติเมตร จัดอยู่ในเซตนี้ด้วยค่าสมาชิกภาพต่างๆกัน ซึ่งนิยมแสดงฟังก์ชันสมาชิกภาพแบบวิภัชนัยเป็นรูปกราฟความสัมพันธ์ของค่าสมาชิกแบบวิภัชนัยเทียบกับค่าของสมาชิกทั้งหมดในเซต ดังแสดงกราฟในรูปที่ 2.13



รูปที่ 2.13 กราฟความสัมพันธ์ของค่าสมาชิกแบบวิภัชนัยกับค่าของสมาชิกในเซต $\tilde{A}$ ซึ่งเป็นเซตของความสูงของคนที่จัดว่าสูงปานกลาง

2.7 งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องกับระบบสอบถามรายนามผู้ใช้โทรศัพท์แบบอัตโนมัติที่นำมาพัฒนาและประยุกต์ใช้นั้นมีดังต่อไปนี้

สิริขวัญ เลิศวิไล และสิริวงศ์ วงศ์ทองเสริม [7] ได้ศึกษาและทำเรื่องการรู้จำคำพูดตัวเลขเดี่ยวภาษาไทยแบบทันกาล โดยใช้เทคนิคข่ายงานระบบประสาทเทียม ซึ่งใช้การฝึกแบบแพร่กระจายความผิดพลาดย้อนกลับ และใช้ลักษณะสำคัญทางสวนศาสตร์ได้แก่ ราสต้า-พีแอลพีอันดับ 12 และอนุพันธ์อันดับหนึ่ง โดยทำการคำนวณทั้งหมดจำนวน 6 เฟรม เฟรมละ 25 มิลลิวินาที ต่อหนึ่งสัญญาณเสียง ผลการทดลองปรากฏว่าสามารถรู้จำเสียงโดยมีผลความถูกต้อง 95 เปอร์เซ็นต์ ซึ่งในงานวิจัยดังกล่าวนี้เป็นการรู้จำคำพูดตัวเลขเดี่ยวภาษาไทยที่ทำการตัดหัวและท้ายของหน่วยพยางค์เท่านั้น ในการออกเสียงพูดเพื่อทดสอบจะต้องออกเสียงทีละพยางค์และเว้นช่วงความเงียบไว้เพื่อให้สามารถตัดหัวท้ายพยางค์ได้ ดังนั้นจึงนำเอางานวิจัยนี้มาพัฒนาต่อเป็นส่วนของการรู้จำหมายเลขโทรศัพท์โดยเพิ่มส่วนของการตัดแบ่งพยางค์เพื่อให้สามารถทำการรู้จำเสียงพูดต่อเนื่องภาษาไทยได้

อัจจิมา ตันสกุล [11] ได้ศึกษาและทำวิจัยเรื่องการสังเคราะห์ข้อความเสียงพูดภาษาไทยสำหรับคำทับศัพท์ภาษาอังกฤษและคำนามเฉพาะ โดยใช้หลักเกณฑ์การทับศัพท์ภาษาอังกฤษในการตัดแบ่งพยางค์ภาษาไทย และทำการค้นหาเสียงที่มีความใกล้เคียงทางเสียงมากที่สุดกับเสียงพยางค์ที่ตัดไว้ โดยใช้การประมาณค่าความใกล้เคียงกับเสียงที่จะการค้นหา จากพจนานุกรมหน่วยเสียง ซึ่งรูปแบบของเสียงที่ใช้จะเป็นการตัดหน่วยเสียง และนำหน่วยเสียงมาต่อรวมกันด้วยเทคนิคการสังเคราะห์เสียงโดยการต่อหน่วยเสียง และผลการทดลองพบว่าค่าเฉลี่ยของค่าความถูกต้องของการค้นหาคำ และค่าความถูกต้องของการสังเคราะห์เสียงเป็นค่าที่สูงเกิน 90 เปอร์เซ็นต์ ซึ่งได้นำงานวิจัยนี้มาประยุกต์ใช้กับการสังเคราะห์เสียงพูดชื่อและนามสกุลของระบบสอบถามรายนามผู้ใช้โทรศัพท์นี้ และได้ปรับปรุงวิธีการประมาณค่าความใกล้เคียงกับเสียงที่จะค้นหา ไม่ได้นำเอาค่าของสระและวรรณยุกต์เข้ามาประกอบการพิจารณาด้วย นั่นคือใช้การประมาณค่าความใกล้เคียงของพยัญชนะต้นและตัวสะกดเท่านั้น ซึ่งทำให้ความใกล้เคียงของคำพ้องเสียงที่ได้ลดลงไป

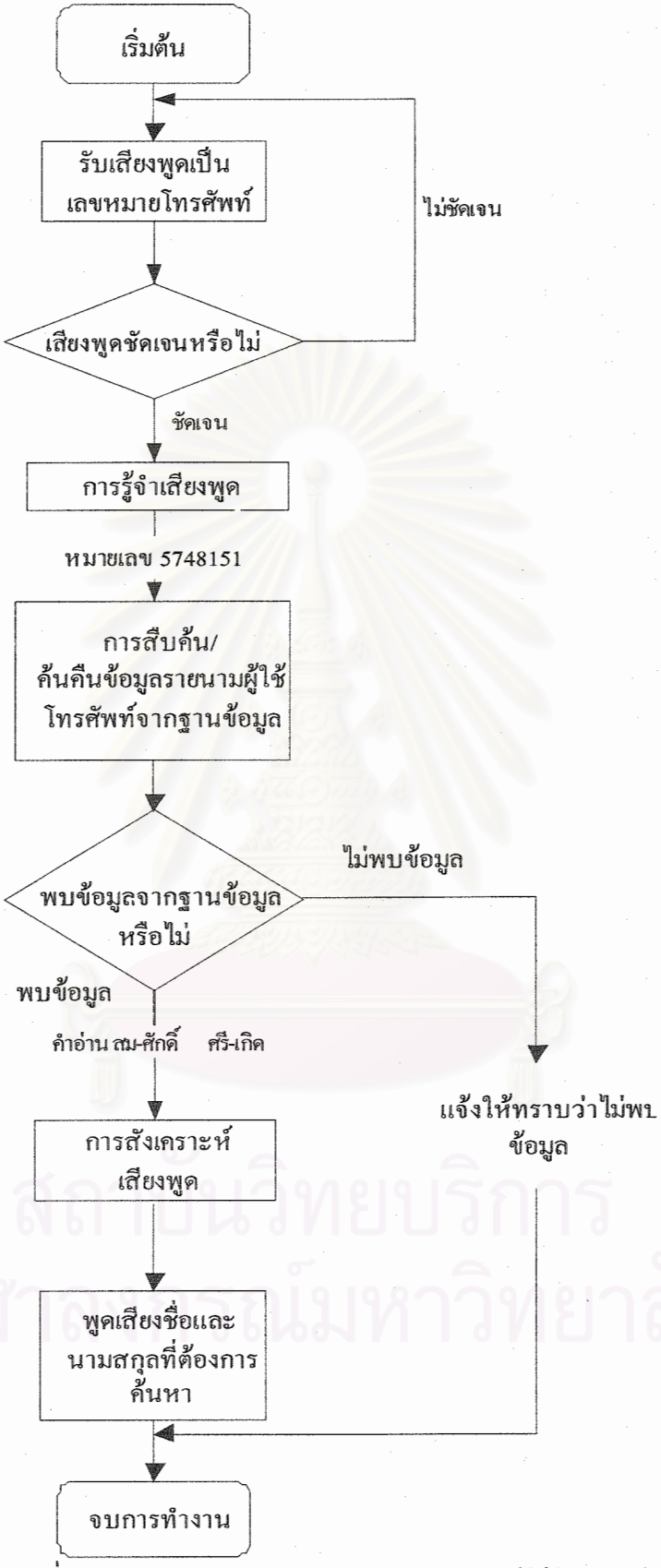
บทที่ 3

ขั้นตอนการดำเนินงานวิจัย

การเพิ่มความสามารถและประสิทธิภาพให้กับระบบสอบถามรายนามผู้ใช้โทรศัพท์โดยให้สามารถทำงานแบบอัตโนมัติได้นั้นต้องอาศัยเทคโนโลยีหลักสองอย่างเข้ามาใช้ร่วมกัน กล่าวคือ การรู้จำเสียงพูด (Speech Recognition) และ การสังเคราะห์เสียงพูด (Speech Synthesis) ซึ่งในงานวิจัยนี้ได้จำลองระบบบนเครื่องไมโครคอมพิวเตอร์ โดยมีข้อมูลขาเข้าเป็นเสียงพูดตัวเลขผ่านทางไมโครโฟนและข้อมูลขาออกเป็นเสียงพูดชื่อและนามสกุลผ่านทางลำโพง เพื่อใช้เป็นต้นแบบ (Prototype) ซึ่งมีรายละเอียดการใช้งานต่างๆดังแสดงในภาคผนวก ง ระบบดังกล่าวมีส่วนประกอบหลัก ดังต่อไปนี้

1. เริ่มต้นเมื่อรับเสียงพูดตัวเลขต่อเนื่องภาษาไทยที่ต้องการทราบถึงรายนามผู้ใช้โทรศัพท์ผ่านทางไมโครโฟนเมื่อผู้พูดพูดจบประโยค เสียงที่ได้ทั้งหมดจะผ่านส่วนที่ทำการตัดหัวท้ายหน่วยหลังจากนั้นก็นำเสียงที่ได้มาผ่านส่วนการตัดแบ่งพยางค์ออกมาเป็นตัวเลขเดี่ยว ในกรณีที่เสียงพูดนั้นไม่ชัดเจนอาจจะเกิดจากระบบไม่สามารถตัดแบ่งพยางค์ได้ตามจำนวนหลักของตัวเลขที่ระบบกำหนด ระบบจะให้ผู้ใช้พูดเสียงตัวเลขซ้ำใหม่อีกครั้งหนึ่ง
2. เมื่อรับเสียงพูดที่ชัดเจน(มีจำนวนพยางค์ครบตามจำนวนหลักของตัวเลขที่ระบบกำหนด)มาแล้ว จากนั้นก็นำเสียงพูดที่ได้นั้นมาผ่านส่วนของการรู้จำเสียงพูดภาษาไทยซึ่งจะให้ผลคือได้หมายเลขโทรศัพท์ออกมา และแสดงออกมาทางจอภาพเพื่อช่วยให้ผู้ใช้สามารถตรวจสอบความถูกต้องของหมายเลขโทรศัพท์ที่พูดไปได้อีกทางหนึ่ง
3. ส่วนของการค้นคืนรายนามผู้ใช้โทรศัพท์จากฐานข้อมูล ในกรณีที่พบรายนามผู้ใช้โทรศัพท์ก็จะพูดชื่อและนามสกุลที่เป็นคำอ่านของหมายเลขโทรศัพท์นั้นออกมาจากฐานข้อมูล ในกรณีที่ไม่พบรายนามผู้ใช้โทรศัพท์ระบบก็จะแจ้งให้ทราบว่าไม่พบหมายเลขโทรศัพท์นั้นในฐานข้อมูลที่มีอยู่
4. ส่วนการสังเคราะห์เสียงพูดภาษาไทย ซึ่งจะนำเอาคำอ่านจากข้อ 3 มาแล้วทำการค้นคืนในพจนานุกรมเสียงที่มีอยู่แล้วออกเสียงทางลำโพง

การทำงานของระบบสอบถามรายนามผู้ใช้โทรศัพท์แบบอัตโนมัติสามารถแสดงดังรูปที่ 3.1 ต่อไปนี้



รูปที่ 3.1 การทำงานของระบบสอบถามรายนามผู้ใช้โทรศัพท์

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ดังนั้นขั้นตอนในการพัฒนาระบบสอบถามรายนามผู้ใช้โทรศัพท์อัตโนมัติดังกล่าว จึงขอแบ่งออกเป็น 4 ขั้นตอนหลักๆคือ 1) ทำการตัดแบ่งพยางค์และตัดหัวท้ายของหน่วยซึ่งขั้นตอนการทำการตัดหัวท้ายของหน่วยเสียงได้นำวิธีการของ สิริขวัญและสิริวงศ์ [7] มาใช้และปรับปรุงให้สามารถรับและเก็บเสียงพูดได้จนกว่าผู้พูดจะหยุดพูดนานเกินกว่าเวลาที่กำหนด ซึ่งแส่ดงว่าผู้พูดพูดจบประโยคแล้ว เพื่อให้สามารถรับเสียงครบตามจำนวนของชุดหมายเลข แล้วทำการพัฒนาการแบ่งพยางค์โดยพิจารณาจากค่าพลังงานและค่าอัตราตัดผ่านระดับกำหนด เพื่อให้ได้พยางค์ของตัวเลขเดี่ยวออกมาเพื่อจะนำไปทำการรู้จำในระดับพยางค์ 2) จากนั้นนำเอาเสียงที่ได้เป็นพยางค์ของตัวเลขเดี่ยวมาผ่านขั้นตอนการรู้จำเสียงพูด โดยผ่านวิธีการทางสวนศาสตร์คือ ราสต้า-พีแอลพี [17] เพื่อหาคุณลักษณะสำคัญของเสียง จากนั้นนำมาผ่านขั้นตอนข่ายงานระบบประสาทเทียม [21] เพื่อทำการเรียนรู้และฝึกฝนพยางค์ของเสียงนั้นๆ 3) พัฒนาการสืบค้นรายนามผู้ใช้โทรศัพท์ โดยออกแบบโครงสร้างข้อมูลของฐานข้อมูลเป็นแบบแฮชชิ่งเนื่องจากต้องการจำกัดเวลาในการค้นหาข้อมูล เพื่อไม่ให้เป็นตัวแปรที่สำคัญในเรื่องเวลาของระบบทั้งหมดที่ใช้คือตั้งแต่พูดหมายเลขโทรศัพท์จนถึงพูดชื่อและนามสกุลของผู้ใช้ออกมา และ 4) การสังเคราะห์เสียงพูด คือจะนำชื่อที่ได้ผ่านกระบวนการสังเคราะห์เสียงโดยใช้วิธีการตัดคำและแบ่งพยางค์ รวมถึงฐานข้อมูลของพจนานุกรมเสียงของ อัจจิมา [11] มาพัฒนาโดยทำการเพิ่มเติมคำศัพท์และพจนานุกรมหน่วยเสียง และได้ปรับปรุงวิธีการประมาณค่าความใกล้เคียงของเสียงซึ่งทำให้สามารถค้นหาคำพ้องเสียงในพจนานุกรมได้แม้คำนั้นจะไม่ได้บรรจุไว้ในพจนานุกรม ซึ่งจะกล่าวขั้นตอนโดยละเอียดดังต่อไปนี้

3.1 **การตัดหัวท้ายหน่วยและการตัดแบ่งพยางค์**(Endpoint Detection and Syllable Segmentation)

ส่วนของการตัดหัวท้ายหน่วยและการตัดแบ่งพยางค์เป็นส่วนที่สำคัญและจำเป็นในการรู้จำเสียงพูดต่อเนื่องภาษาไทย เนื่องจากในงานวิจัยนี้ได้ใช้หน่วยเสียงพยางค์เป็นตัวทาบเทียบ (Syllable Based) โดยทั่วไปเสียงพูดต่อเนื่องจะประกอบด้วยพยางค์หลายๆพยางค์ติดกันทำให้เกิดความหลากหลายของเสียงพูดและทำให้ยากต่อการฝึกฝนและการรู้จำของคอมพิวเตอร์ ดังนั้นจึงจำเป็นต้องมีการตัดแบ่งพยางค์เพื่อให้สามารถแบ่งคำพูดที่ต่อเนื่องออกเป็นพยางค์เดี่ยวแล้วก็นำเอาพยางค์นั้นเข้าไปฝึกและทำการรู้จำได้ง่ายขึ้น

3.1.1 **การตัดหัวท้ายหน่วย** (Endpoint Detection)

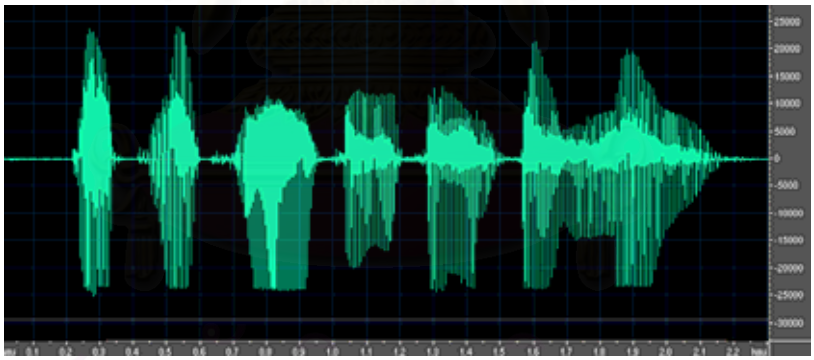
ขั้นตอนในการตัดหัวท้ายหน่วยสามารถแบ่งเป็นขั้นตอนได้ดังนี้

1.) รับสัญญาณเสียงพูดที่เป็นตัวเลขต่อเนื่องเข้ามาโดยจะรับเสียงเข้ามาจากไมโครโฟน เพื่อประมวลผล โดยพิจารณาพลังงานเสียงที่เข้ามาในแต่ละ 200 มิลลิวินาที ว่ามีสัญญาณเสียงหรือไม่ ถ้าไม่ใช่จะคำนวณไว้เป็นค่าสภาวะแวดล้อม

2.) ถ้าเป็นสัญญาณเสียงเข้ามาซึ่งก็จะพิจารณาได้จากค่าพลังงานของเสียงที่คำนวณได้จากสมการที่ 2.3 นั้นมีค่าเกินกว่าค่าที่กำหนด ก็จะทำการรวมสัญญาณเสียงที่เข้ามาใหม่นี้กับรอบก่อนหน้านี้เพื่อรอสัญญาณเสียงต่อไป

3.) ตรวจสอบว่าเสียงนั้นจบประโยคแล้วหรือไม่ โดยจะพิจารณาจากค่าพลังงานของเสียงที่เข้ามาว่าน้อยกว่าค่าที่กำหนดและนานเกินกว่าระยะเวลาที่กำหนด(ในระบบกำหนดระยะเวลาไว้เท่ากับ 600 มิลลิวินาที)

4.) ทำวนรอบซ้ำไปจนกระทั่งพบว่าไม่มีสัญญาณเสียงพูด แล้วเก็บเสียงพูดที่ได้ลงในหน่วยความจำสำรองและหน่วยความจำหลักเพื่อที่จะไปทำการตัดแบ่งพยางค์ต่อไป ซึ่งแฟ้มข้อมูลที่ได้จากการเก็บลงหน่วยความจำหลักเมื่อนำมาเปิดดูเพื่อตรวจสอบความถูกต้องของการตัดสามารถแสดงดังรูปที่ 3.2



รูปที่ 3.2 สัญญาณเสียงพูดต่อเนื่องที่ทำการตัดหัวท้ายหน่วยแล้ว

3.1.2 **การตัดแบ่งพยางค์** (Syllable Segmentation)

ขั้นตอนการตัดแบ่งพยางค์มีขั้นตอนดังต่อไปนี้

1.) รับสัญญาณเสียงพูดที่เป็นประโยคคำพูดต่อเนื่องเข้ามาแล้ววิเคราะห์หาค่าคุณลักษณะ ของสัญญาณเสียงพูด โดยกำหนดคุณลักษณะที่ใช้ในระบบนี้คือ พลังงานเฉลี่ย ตามสมการที่ 2.3 และอัตราการตัดผ่านระดับกำหนด ตามสมการที่ 2.6 ซึ่งกำหนดค่าระดับกำหนด (L) ไว้มีค่าเท่ากับ 0

2.) ทำการปรับเรียบคุณลักษณะของสัญญาณเสียงพูด โดยใช้วิธีปรับเรียบด้วยค่าเฉลี่ยเคลื่อนไหวตามสมการที่ 2.7

3.) พิจารณาหาจุดแบ่งพยางค์ โดยใช้ค่าพลังงานเฉลี่ยและอัตราการตัดผ่านระดับกำหนดเป็นตัวพิจารณา

การหาจุดแบ่งพยางค์สามารถแบ่งเป็นขั้นตอนย่อยๆ ได้ดังนี้

(3.1) ทำการหาจุดต้น (Beginning Point) และจุดปลาย(Ending Point) ของประโยค โดยพิจารณาจากค่าพลังงานของเสียงทั้งหมด ค่าของพลังงานเสียงที่มากกว่าค่าที่กำหนดระดับให้เป็นจุดต้นของประโยคจะเป็นจุดเริ่มต้นของประโยคนั้น ส่วนค่าพลังงานเสียงที่น้อยกว่าค่าที่กำหนดระดับให้เป็นจุดปลายของประโยคก็จะเป็นจุดปลายของประโยคนั้น โดยการกำหนดค่าระดับของจุดต้นกับจุดปลายประโยคนั้นจะต้องทำการหาและปรับค่าบรรทัดฐานจากระบบจริง เนื่องจากค่าที่ได้จะไม่เท่ากันถ้าสิ่งแวดล้อมของระบบเปลี่ยนไป

(3.2) หาจุดพลังงานสูงสุดภายในระยะเวลาพยางค์ คือ การหาค่าพลังงานที่มากที่สุดในระยะเวลาพยางค์ โดยจุดที่ทำการพิจารณาจะเป็นจุดสูงสุดในระยะเวลาพยางค์ ถ้าจุดนั้นมีค่าพลังงานมากที่สุดในกรอบหน้าต่างเท่ากับระยะเวลาพยางค์ โดยมีจุดดังกล่าวเป็นจุดกึ่งกลางกรอบหน้าต่าง

(3.3) หาจุดพลังงานต่ำสุดภายในระยะเวลาพยางค์ คือ การหาจุดที่มีค่าพลังงานต่ำสุดระหว่างจุดสูงสุดภายในระยะเวลาพยางค์ 2 จุดที่อยู่ต่อเนื่องกัน โดยเมื่อทำการหาจุดต่ำสุดภายในระยะเวลาพยางค์ได้แล้ว ให้เลื่อนจุดสูงสุดภายในระยะเวลาพยางค์ไปทางขวาทีละจุด แล้วพิจารณาหาจุดต่ำสุดภายในระยะเวลาพยางค์ตัวต่อไป

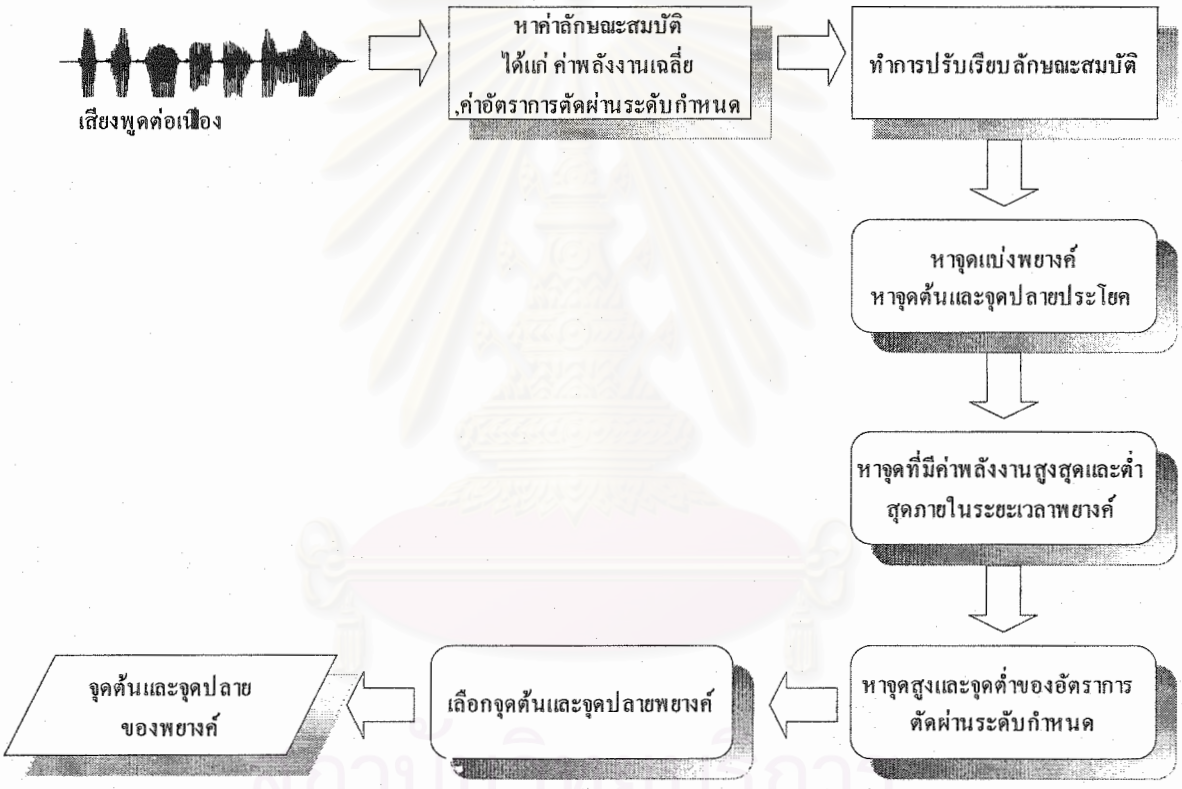
(3.4) หาจุดสูงและจุดต่ำของอัตราการตัดผ่านระดับกำหนด โดยจุดต่ำของอัตราการตัดผ่านระดับกำหนดคือ จุดที่มีอัตราการตัดผ่านระดับกำหนดน้อยที่สุดที่อยู่ระหว่างหาจุดพลังงานสูงสุดภายในระยะเวลาพยางค์ 2 จุดที่อยู่ต่อเนื่องกัน ส่วนจุดสูงของอัตราการตัดผ่านระดับกำหนด คือ จุดที่มีค่าอัตราการตัดผ่านระดับกำหนดมากที่สุดที่อยู่ระหว่างจุดพลังงานสูงสุดภายในระยะเวลาพยางค์กับจุดต่ำของอัตราการตัดผ่านระดับกำหนด

(3.5) พิจารณาเลือกจุดต้นและจุดปลายของพยางค์ โดยจะพิจารณาจากอัตราการตัดผ่านระดับกำหนดของจุดที่พิจารณาเป็นอันดับแรก ถ้ามีค่าน้อยกว่าระดับที่กำหนดไว้ให้ไปพิจารณาค่าความต่างของอัตราการตัดผ่านระดับกำหนด ถ้ามากเกินกว่าระดับที่กำหนดไว้ แสดงว่าจุดนั้นคือจุดปลายพยางค์ ถ้าไม่มากเท่า

ระดับที่กำหนดไว้ให้ไปพิจารณาค่าความต่างของพลังงานระหว่างจุดพลังงานสูงสุดและจุดพลังงานต่ำสุดที่จุดที่พิจารณาอยู่ระหว่างจุดทั้งสอง หากมีค่ามากกว่าระดับที่กำหนดไว้ จุดนั้นก็เป็นจุดปลายพยางค์ ซึ่งเมื่อหาจุดปลายพยางค์ได้แล้วจุดนั้นก็จะเป็นจุดต้นของพยางค์ต่อไป

4.) ทำการปรับปรุงขอบเขตของพยางค์ โดยการตรวจสอบว่ามีพยางค์ไหนบ้างที่ไม่สมบูรณ์ เช่น มีระยะเวลาพยางค์สั้นเกินกว่าค่าที่กำหนด ก็ให้ตัดพยางค์นั้นออกไป แล้วนำไปรวมกับพยางค์ก่อนหน้านี้หนึ่งพยางค์

ขั้นตอนและวิธีในการตัดแบ่งพยางค์แสดงได้ดังรูปที่ 3.3 ต่อไปนี้



รูปที่ 3.3 ขั้นตอนการตัดแบ่งพยางค์

ในการตัดแบ่งพยางค์แต่ละพยางค์จะได้ค่าจุดต้นและจุดปลายของพยางค์ซึ่งมีหน่วยเป็นเวลา เพื่อให้เป็นการง่ายและสะดวกในการตรวจสอบความถูกต้องในการตัดแบ่งพยางค์ จึงได้ใช้โปรแกรม SpeechView ซึ่งอยู่ในชุดโปรแกรม CSLU toolkit รุ่น 2.0.0 [13] พัฒนาโดย Center for Spoken Language Understanding, Oregon Graduate Institute of Science and Technology เป็นเครื่องมือที่ช่วยในการตรวจสอบความถูกต้องในการตัดแบ่งพยางค์ โดย

โปรแกรมดังกล่าวจะเก็บค่าจุดต้นและจุดปลายของพยางค์ไว้ในแฟ้มที่มีนามสกุล .phn และรูปแบบภายในแฟ้มมีลักษณะดังนี้

MillisecondsPerFrame: 1.0

END OF HEADER

100 420 A1

420 740 A2

740 1100 A3

1100 1390 A4

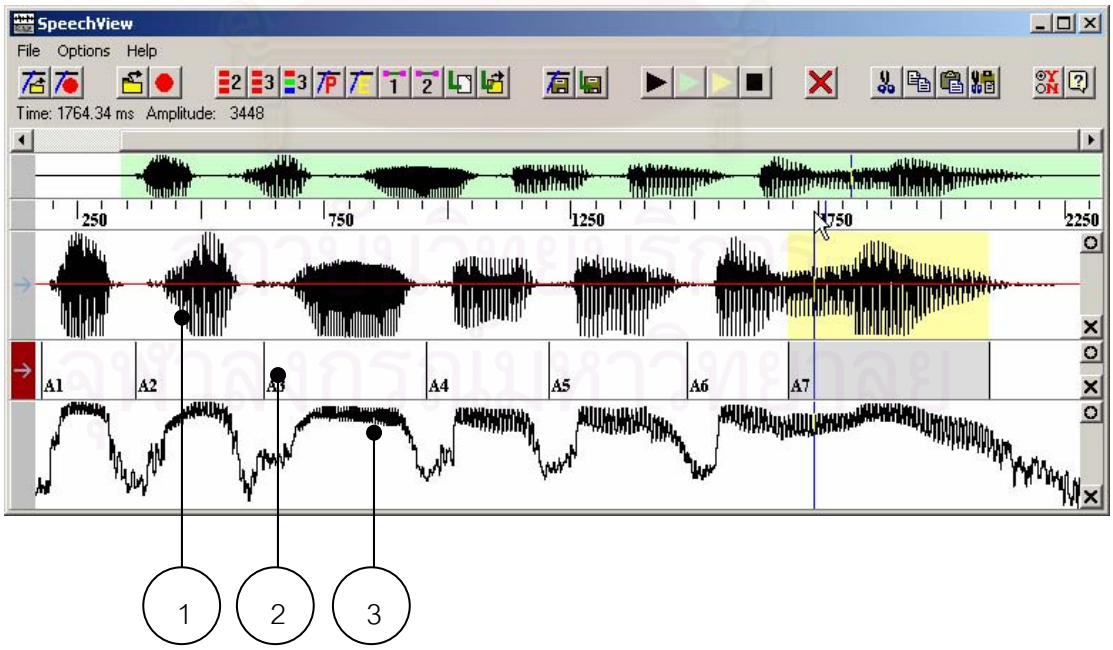
1390 1620 A5

1620 1980 A6

1980 2520 A7

ภายในแฟ้มประกอบด้วยส่วนของหัวของแฟ้ม(header) ซึ่งแสดงถึงหน่วยเวลาที่ใช้ต่อหนึ่งเฟรมมีหน่วยเป็นมิลลิวินาที และส่วนที่สองแสดงถึงเวลาของจุดต้นและจุดปลายของพยางค์ในแต่ละพยางค์นั้นๆ ซึ่งสดมภ์แรกเป็นเวลาของจุดต้นของพยางค์แต่ละพยางค์ สดมภ์ที่สองเป็นเวลาของจุดปลายของพยางค์นั้นๆ และสดมภ์ที่สามเป็นชื่อของพยางค์แต่ละพยางค์

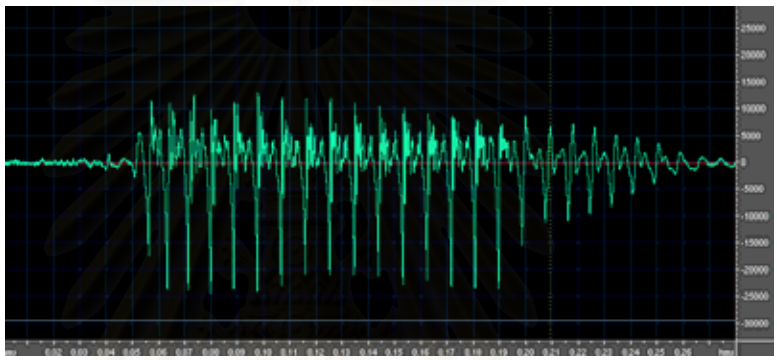
การใช้งานโปรแกรม SpeechView [14] เพื่อเป็นเครื่องมือที่ช่วยในการตรวจสอบความถูกต้องในการตัดแบ่งพยางค์ แสดงดังรูปที่ 3.4



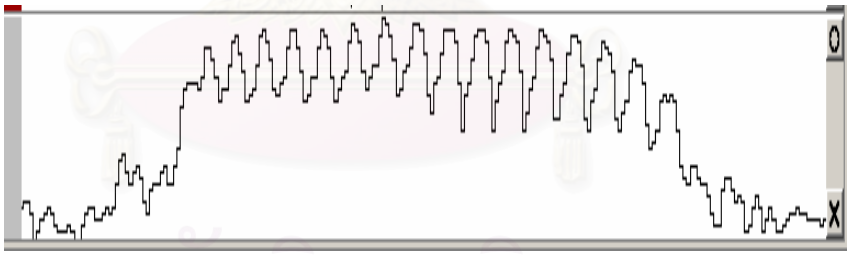
รูปที่ 3.4 การใช้โปรแกรม SpeechViewเป็นเครื่องมือตรวจสอบความถูกต้องในการแบ่งพยางค์

จากรูปที่ 3.4 ในส่วนที่ 1 แสดงค่าของแอมปลิจูดของเสียงที่ทำการตัดหัวท้ายหน่วยแล้ว ส่วนที่ 2 แสดงชื่อของพยางค์และขอบเขตของพยางค์แต่ละพยางค์ ซึ่งนำค่าของจุดต้นและจุดปลายที่อ่านค่าจากแฟ้มนามสกุล .phn มาใช้งาน และส่วนที่ 3 แสดงค่าพลังงานของเสียงทั้งหมด

แฟ้มข้อมูลของเสียงที่ได้จากการตัดแบ่งพยางค์จะถูกเก็บไว้ในทั้งหน่วยความจำสำรองและหน่วยความจำหลักเพื่อการตรวจสอบความถูกต้องจะทำได้ง่ายขึ้น ดังนั้นเราจึงสามารถนำเอาแฟ้มข้อมูลของเสียงที่ตัดแบ่งพยางค์แล้วมาเปิดดูเพื่อตรวจสอบความถูกต้องว่าตัดแบ่งพยางค์ได้ถูกต้องมากเพียงใดหรืออาจตรวจสอบจากค่าพลังงานของเสียงที่ตัดแบ่งพยางค์แล้วนั้นๆ ดังรูปที่ 3.5 และ รูปที่ 3.6



รูปที่ 3.5 สัญญาณเสียงที่ได้จากการตัดแบ่งพยางค์



รูปที่ 3.6 พลังงานของเสียงที่ได้จากการตัดพยางค์

3.2 **การรู้จำเสียงพูด** (Speech Recognition)

การรู้จำเสียงพูดนั้นจัดเป็นการเปรียบเทียบรูปแบบ (Pattern matching) ชนิดหนึ่งแต่ค่อนข้างจะมีความซับซ้อนมาก และต้องอาศัยการเปรียบเทียบหลายขั้นตอน โดยจะนำสัญญาณทางสวนศาสตร์ (acoustic signal) มาตรวจสอบและจัดโครงสร้างให้เป็นลำดับขั้น (hierarchy) ของหน่วยเสียง คำ พยางค์ โดยในแต่ละขั้นก็จะค่อยๆเพิ่มข้อจำกัด (constraint) เข้าไปเรื่อยๆ เช่น การออกเสียงของคำที่รู้จัก หรือการเรียงคำที่ถูกหลักไวยากรณ์ ซึ่งจะช่วยให้สามารถทดแทนกับความผิดพลาดหรือความไม่แน่นอนในระดับที่ต่ำกว่าได้ โดยลำดับขั้นของข้อจำกัดนี้จะเป็น

ประโยชน์ต่อการตัดสินใจโดยอาศัยหลักของความน่าจะเป็น ณ ระดับที่ต่ำกว่าทั้งหมด ทำให้ได้ผลลัพธ์ที่ถูกต้องที่สุดในระดับที่สูงที่สุด

ในงานวิจัยนี้ได้นำเสนอการรู้จำเสียงพูดภาษาไทยโดยใช้ข่ายงานระบบประสาทเทียมเป็นเทคนิคของการเรียนรู้ที่นำมาใช้

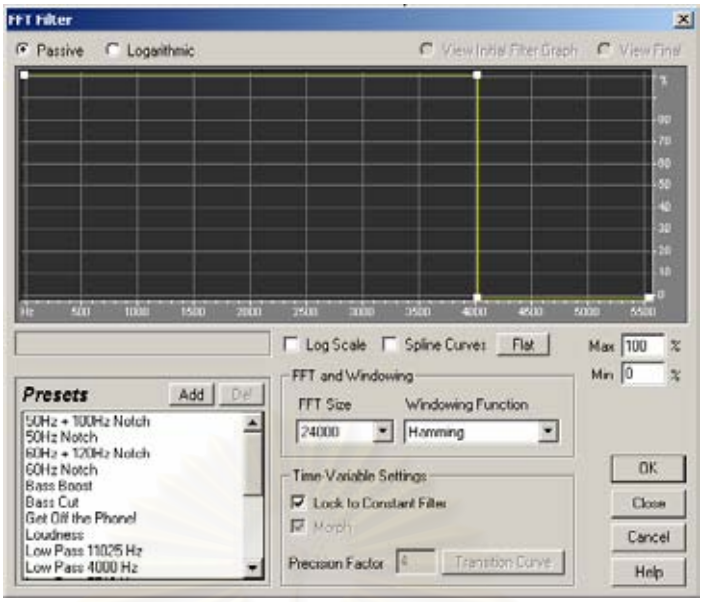
3.2.1 ข้อมูลทางเสียง

ข้อมูลเสียงที่ใช้ในการฝึกข่ายงานระบบประสาทเทียมนำมาจากฐานข้อมูลของ Thubthong และ Kijsirikul(1999) ซึ่งมีขั้นตอนในการเก็บข้อมูลดังนี้

1.) เก็บข้อมูลเสียงพูดจากผู้พูดที่ใช้ภาษาไทย ไม่จำกัดสำเนียง ทั้งชายและหญิงอายุระหว่าง 18 –45 ปี จำนวน 50 คน เพศละ 25 คน แต่ละคนให้อ่านหมายเลขจากชุดข้อมูลตัวเลข 7หลัก ที่สร้างเองชุดละ 5 หมายเลข โดยใช้รูปแบบการพูดปกติ 2 รอบ ดังนั้นข้อมูลจะประกอบด้วยหมายเลขโทรศัพท์ทั้งสิ้น 500 ชุดหมายเลข (3,500 ตัวเลข)

2.) บันทึกเสียงลงเครื่องคอมพิวเตอร์ Packard Bell รุ่น Legend 1008T AP โดยใช้โปรแกรม Cool Edit Pro รุ่น 2.1 [13] (ปัจจุบันเปลี่ยนชื่อเป็นโปรแกรม Adobe Audition) ของบริษัท Adobe Systems Incorporated จัดเก็บเป็นแฟ้มข้อมูลเสียง (Sound File) ในรูปแบบของ Wav 1 แฟ้ม แบบโมโน 16 บิต ด้วยอัตราการชักข้อมูล (Sampling Rate) 11,025 เฮิรตซ์ โดยแต่ละแฟ้มจะประกอบด้วยเสียงตัวเลขต่อเนื่อง 1 ชุด ถ้ามีประโยคใดที่ผู้พูดออกเสียงไม่ดีหรือมีความผิดพลาดก็จะอัดทับเฉพาะคำนั้นๆใหม่

3.) หลังจากทำการเก็บบันทึกเสียงแล้ว จะทำการกรองสัญญาณที่ไม่ได้ต้องการทิ้ง ด้วย FFT Filter โดยกำหนดค่าพารามิเตอร์ต่างๆ คือ Hamming Window ขนาด 24000 และช่วงความถี่ที่ผ่านได้ (Band Pass) คือ 60-4000 เฮิรตซ์ ซึ่งจะช่วยทำให้การนำไปเรียนรู้นั้นได้ผลดียิ่งขึ้น การทำ FFT Filter นั้นจะใช้โปรแกรม Cool Edit Pro เช่นเดียวกับการบันทึกเสียง



รูปที่ 3.7 วิธีการกรองสัญญาณเสียงด้วยโปรแกรม Cool Edit

3.2.2 การแบ่งชุดข้อมูล

ข้อมูลเสียงพูดที่เก็บไว้ประกอบด้วยข้อมูลจากผู้พูดทั้งหมด 50 คน เป็นชาย 25 คนและหญิง 25 คน จากนั้นทำการแบ่งข้อมูลเป็น 2 ชุด คือ ชุดฝึก และชุดทดสอบ

ชุดฝึก (Training set) ประกอบด้วยข้อมูลจากผู้พูดทั้งหมด 40 คน แบ่งเป็นชาย 20 คน หญิง 20 คน รวมทั้งสิ้น 2,800 ตัวเลข

ชุดทดสอบ (Test set) ประกอบด้วยข้อมูลจากผู้พูดทั้งหมด 10 คน แบ่งเป็นชาย 5 คน หญิง 5 คน รวมทั้งสิ้น 700 ตัวเลข

3.2.3 วิธีการทางสวนศาสตร์

การหาคุณลักษณะที่สำคัญของสัญญาณเสียงที่ใช้ในงานวิจัยนี้คือ ราสต้า-พีแอลพี และอนุพันธ์อันดับหนึ่ง ซึ่งมีค่าพารามิเตอร์ดังนี้

1. อันดับของราสต้า-พีแอลพี ซึ่งใช้ราสต้า-พีแอลพีอันดับ 12
2. จำนวนกรอบ เพื่อให้ทราบว่าจำนวนกรอบเท่าไรจึงให้ค่าความถูกต้องได้ดีที่สุด จำนวนกรอบที่เลือกใช้ได้แก่ 6, 9, 12, 15 กรอบ โดยแต่ละกรอบจะมีความยาว 25 มิลลิวินาที ตำแหน่งในการเลือกกรอบมีดังนี้

   จำนวนกรอบ 6 กรอบ เลือกกรอบจากตำแหน่งที่ 5, 23, 41, 59, 77, 95 เปอร์เซ็นต์ตามแกนเวลาของข้อมูลเสียง

   จำนวนกรอบ 9 กรอบ เลือกกรอบจากตำแหน่งที่ 5, 16, 28, 39, 50, 61, 73, 84, 95 เปอร์เซ็นต์ ตามแกนเวลาของข้อมูลเสียง

จำนวนกรอบ 12 กรอบ เลือกกรอบจากตำแหน่งที่ 5, 13, 21, 30, 38, 46, 54, 62, 70, 79, 87, 95 เปอร์เซ็นต์ ตามแกนเวลาของข้อมูลเสียง

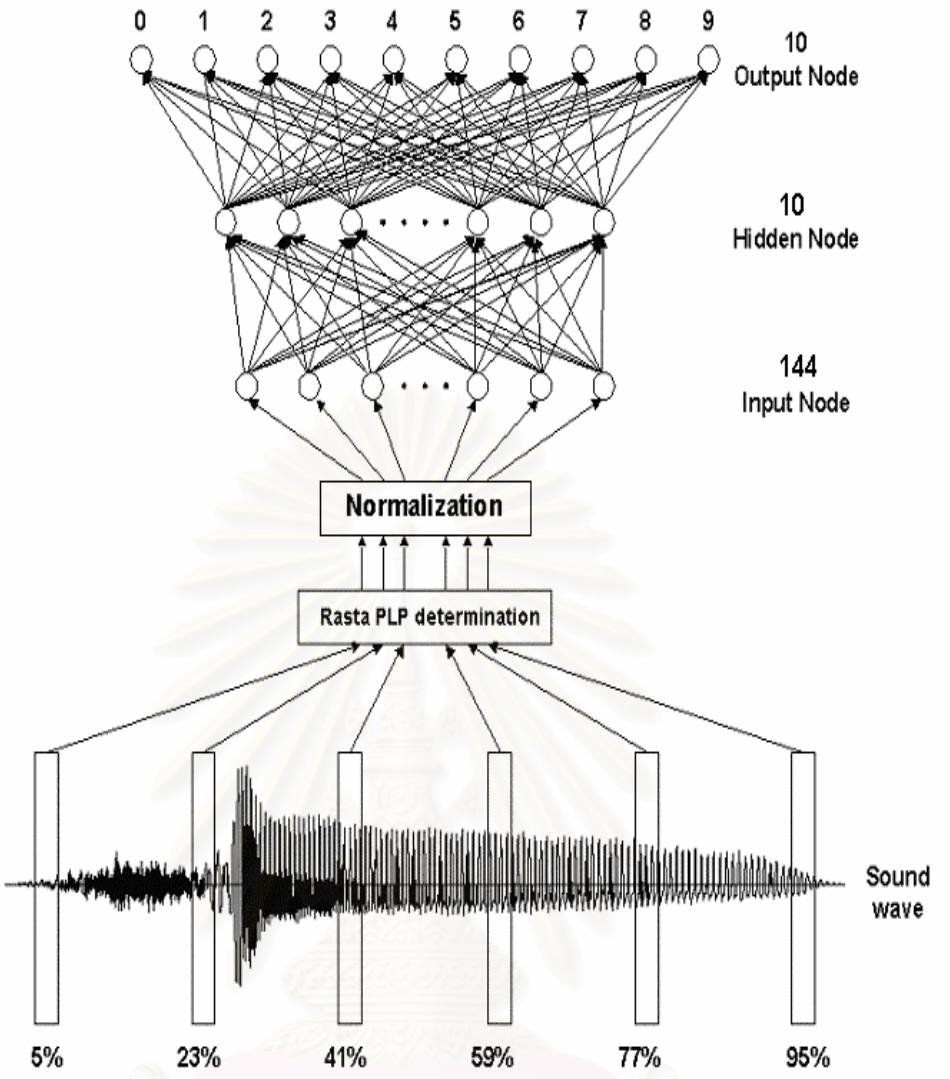
จำนวนกรอบ 15 กรอบ เลือกกรอบจากตำแหน่งที่ 5, 11, 18, 24, 31, 37, 44, 50, 56, 63, 69, 76, 82, 89, 95 เปอร์เซ็นต์ ตามแกนเวลาของข้อมูลเสียง

3.3 ข่ายงานระบบประสาทเทียม

ระบบข่ายงานประสาทเทียมที่ใช้ในงานวิจัยนี้ใช้ NICO's Toolkit (Beta Version for Windows 32 bits) [21] ที่พัฒนาโดย Nikko Strom, Department of Speech,Music and Hearing,KTH, Sweden ในส่วนของคำสั่งการนำข้อมูลเข้าและการกำหนดค่าพารามิเตอร์ของข่ายงานประสาทเทียม โดยนายณัฐกร ทับทอง ภาควิชาฟิสิกส์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ประเทศไทย ในการฝึกจะใช้วิธีความผิดพลาดแบบแพร่กระจายย้อนกลับ (Error back-propagation) และกำหนดค่าโมเมนตัมเท่ากับ 0.9 และค่าอัตราการเรียนรู้ที่ใช้เท่ากับ 0.0001 และ 0.00001 ตามลำดับ

ในการฝึกฝนข่ายงานระบบประสาทเทียมมีหลักการทำงานของโปรแกรมก็คือ สร้างเครือข่ายที่ประกอบด้วยระดับชั้นข้อมูลเข้า (input layer), ระดับชั้นข้อมูลซ่อนตัว (hidden layer) และ ระดับชั้นข้อมูลออก (output layer) ซึ่งแต่ละชั้นติดต่อกันหมด จากนั้นทำการฝึกฝนให้เครือข่าย โดยการอ่านข้อมูลเข้าของชุดฝึก กำหนดค่าน้ำหนักเริ่มต้นโดยการสุ่ม แล้วนำค่าน้ำหนักที่ได้มาหาค่าของข้อมูลขอออก แล้วมาเปรียบเทียบกับค่าเป้าหมายเพื่อหาเปอร์เซ็นต์ความถูกต้องและค่าความผิดพลาด เพื่อนำมาใช้ในการปรับค่าน้ำหนักสำหรับรอบถัดไป ค่าน้ำหนักที่ได้ในแต่ละรอบจะถูกนำไปใช้ในการหาข้อมูลขาออกของชุดทดสอบ แล้วนำมาเทียบกับค่าเป้าหมาย หาเปอร์เซ็นต์ความถูกต้อง นั่นคือ หนึ่งรอบของการฝึก จากนั้นจะทำการฝึกเรื่อยๆ ซึ่งเป็นการปรับค่าน้ำหนักเพื่อให้ได้เปอร์เซ็นต์ความถูกต้องเมื่อเทียบกับชุดทดสอบให้มากที่สุด

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

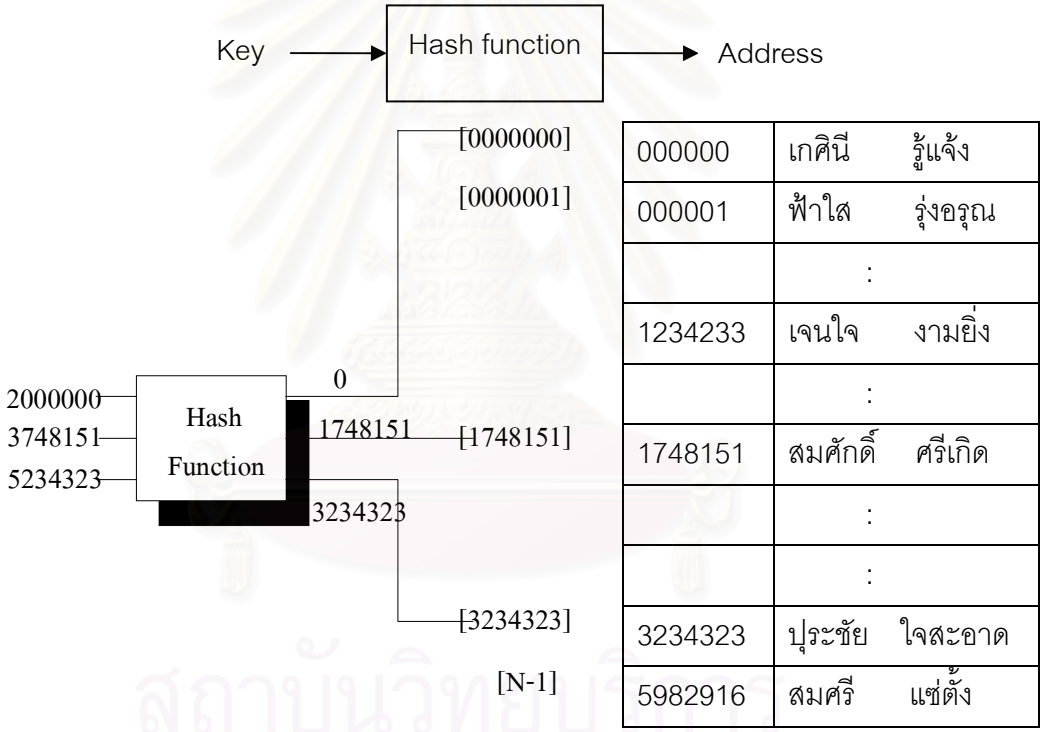


รูปที่ 3.8 ขั้นตอนการรู้จำเสียงเมื่อใช้ข้อมูล 6 กรอบ

จากรูปที่ 3.8 เป็นขั้นตอนในการรู้จำเสียงของตัวเลขที่ทำการแบ่งพยางค์แล้ว โดยใช้จำนวนกรอบ 6 กรอบ เลือกกรอบจากตำแหน่งที่ 5, 23, 41, 59, 77, 95 เปอร์เซ็นต์ ตามแกนเวลาของข้อมูลเสียง กรอบละ 25 มิลลิวินาที แต่ละกรอบจะผ่านขั้นตอนการหาลักษณะสำคัญของเสียงต่างๆ และ กรรมวิธีปรับบรรทัดฐาน จากนั้นจะนำข้อมูลที่ได้เข้าสู่ข่ายงานระบบประสาทเทียม ซึ่งใช้จำนวนบัพของระดับชั้นข้อมูลออก (output layer) เท่ากับ 10 บัพ เนื่องจากผลลัพธ์ที่ต้องการคือ เลข 0-9 จำนวนสิบตัวเลข ส่วนจำนวนบัพของระดับชั้นข้อมูลเข้า (input layer) คือ 144 บัพ ได้จากจำนวนกรอบคูณกับจำนวนอันดับของค่าคุณลักษณะสำคัญแล้วนำมาคูณด้วยสอง เนื่องจากการเพิ่มค่าอนุพันธ์อันดับหนึ่งเข้าไป และระดับชั้นข้อมูลซ่อนตัว (hidden layer) เท่ากับ 10 บัพ

3.4 การสืบค้นข้อมูลของรายนามผู้ใช้โทรศัพท์

การสืบค้นข้อมูลที่มีข้อมูลเป็นจำนวนมาก สิ่งที่ต้องคำนึงถึงที่เป็นส่วนสำคัญคือระยะเวลาที่ใช้ในการค้นคืนของข้อมูล ซึ่งในปัจจุบันมีผู้คิดค้นเทคนิคที่ใช้ในการค้นหาข้อมูลอยู่มากมายขึ้นอยู่กับความเหมาะสมและการนำไปประยุกต์ใช้งาน เทคนิคที่ใช้ในการค้นคืนและการเก็บข้อมูลของรายนามผู้ใช้โทรศัพท์ได้แก่ เทคนิคแบบแฮชชิ่ง (Hashing) เนื่องจากมีความเหมาะสมกับการค้นหาข้อมูลที่มีปริมาณมากและใช้เวลาในการทำงานเป็น O(1) โดย นำค่าคีย์ (key)ซึ่งได้แก่เลขหมายโทรศัพท์มาทำการหาค่าผลต่างกับค่า 2000000 ซึ่งเป็นเลขหมายโทรศัพท์เริ่มต้นในส่วนของโทรศัพท์พื้นฐานในเขตนครหลวงและปริมณฑลเนื่องจากงานวิจัยนี้ใช้ข้อมูลในส่วนของพื้นที่ในเขตนครหลวงและปริมณฑลเท่านั้น ผลลัพธ์ที่ได้จะเป็นค่า address ที่ใช้ในการเก็บของข้อมูลนั้น ซึ่งวิธีนี้จะสามารถทำให้ไม่เกิดการชนกันของข้อมูล ดังแสดงตัวอย่างในรูปที่ 3.9



รูปที่ 3.9 การใช้เทคนิคแบบแฮชชิ่งในการค้นหาข้อมูลรายนามผู้ใช้โทรศัพท์

3.5 การสังเคราะห์เสียงพูดรายนามผู้ใช้โทรศัพท์

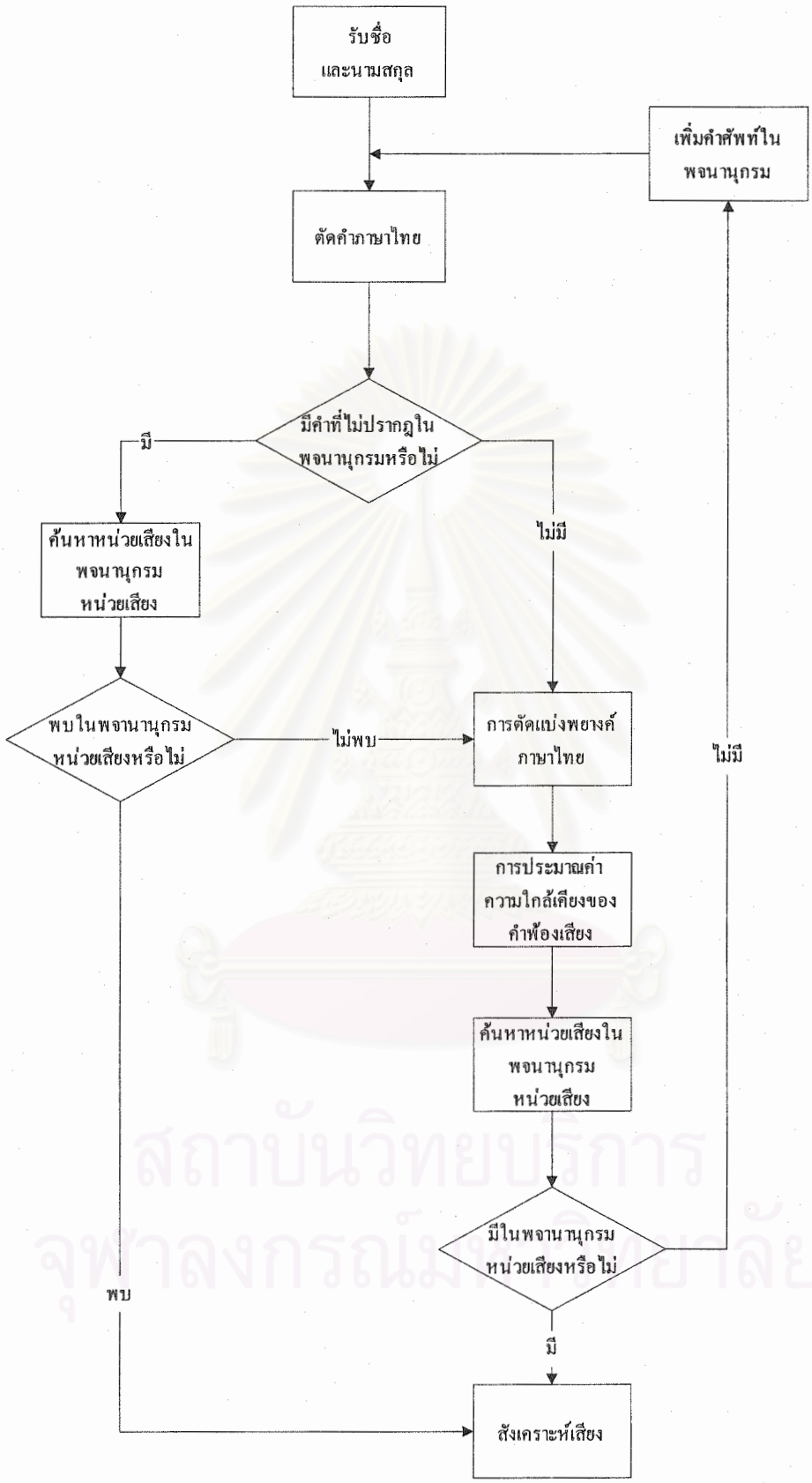
ในระบบสอบถามรายนามผู้ใช้โทรศัพท์แบบอัตโนมัติได้นำเอาหลักการสังเคราะห์เสียงพูดมาประยุกต์ใช้เพื่อใช้ในการแสดงผลออกมาทางเสียงที่ประกอบด้วยชื่อและนามสกุลของผู้ใช้โทรศัพท์ สำหรับในงานวิจัยนี้ได้นำโปรแกรม การสังเคราะห์เสียงพูดภาษาไทยสำหรับคำทับ

ศัพท์ภาษาอังกฤษและคำนามเฉพาะ ของอัจจิมา ตันสกุล[11] มาประยุกต์ใช้งานร่วมกับระบบ โดยโปรแกรมจะมีหลักการทำงานคือ นำเสียงที่เก็บไว้ก่อนแล้วนำมาต่อกันเป็นเสียงพูดที่ต้องการ และมีหลักการตัดคำโดยใช้พจนานุกรม และวิธีตัดคำให้ยาวที่สุด(Longest Matching) ซึ่งการตัดคำโดยใช้พจนานุกรมที่ทำการหาขอบเขตของหน่วยคำที่เป็นวิสามานยนาม(ชื่อเฉพาะ)นั้น ถ้าหากเก็บทุกชื่อหรือนามสกุลที่มีอยู่ในฐานข้อมูลลงในพจนานุกรมทั้งหมด จากนั้นก็ค้นหาและเปรียบเทียบหาคำศัพท์นั้นๆ ว่ามีอยู่ในพจนานุกรมหรือไม่ เพียงเท่านั้นก็จะสามารถหาขอบเขตของคำแต่ละคำได้ทั้งหมด แต่ในความเป็นจริงแล้วถ้ามีข้อมูลจำนวนมากหรือเพิ่มข้อมูลให้มากขึ้นจะทำให้เกิดความยุ่งยากในการจัดเก็บลงในพจนานุกรมเป็นอย่างมาก ดังนั้นการบรรจุคำไว้ในพจนานุกรมจะมีทั้งหน่วยคำที่ย่อยที่สุดที่มีหรือไม่มีความหมาย อาจเป็นคำประสม หรือเป็นคำวิสามานยนามนั้นทั้งคำ ถ้าหากมีชื่อนั้นปรากฏในฐานข้อมูลเป็นจำนวนมาก

ในงานวิจัยนี้ได้เพิ่มขั้นตอนของการสังเคราะห์ชื่อและนามสกุลเสียงพูดภาษาไทยเพื่อให้มีความถูกต้องมากยิ่งขึ้นจึงทำการเพิ่มเขตของข้อมูล(field)ของคำอ่านออกเสียงภาษาไทยของชื่อและนามสกุลจากฐานข้อมูลเดิมที่มีเพียงแค่ชื่อและนามสกุลเท่านั้น เพื่อช่วยทำให้การอ่านของชื่อและนามสกุลมีความถูกต้องใกล้เคียงความเป็นจริงมากยิ่งขึ้น เนื่องจากผู้ดูแลระบบจะสามารถแก้ไขคำอ่านให้ถูกต้องได้ โดยก่อนที่จะนำเอาฐานข้อมูลของชื่อและนามสกุลเข้านั้นผู้ดูแลระบบจะต้องนำเอาชื่อและนามสกุลเหล่านั้นมาให้โปรแกรมสังเคราะห์เสียงพูดทำการอ่านออกเสียงแล้วผู้ดูแลระบบจะเป็นผู้ตัดสินใจว่าโปรแกรมอ่านออกเสียงได้ถูกต้องหรือไม่ ถ้าไม่ถูกต้องผู้ดูแลระบบจะทำการแก้ให้เป็นเสียงของคำอ่านที่ถูกต้องแล้วจึงนำเข้าในฐานข้อมูล

นอกจากนี้ยังได้เพิ่มส่วนของการอ่านคำพ้องเสียงที่ไม่พบในพจนานุกรม เนื่องจากในส่วนของการอ่านชื่อและนามสกุลนั้นบางครั้งพบคำที่ไม่ได้บรรจุไว้ในพจนานุกรมแต่เป็นคำพ้องเสียงของคำที่บรรจุไว้ในพจนานุกรม จึงทำให้ผู้ดูแลระบบต้องทำการเพิ่มคำเข้าไปในพจนานุกรมหรืออาจจะต้องไปแก้ในส่วนของคำอ่านให้เป็นคำที่บรรจุไว้ในพจนานุกรม ดังนั้นในงานวิจัยนี้จึงนำทฤษฎีเซตวิภัชนัย (Fuzzy Set Theory) มาใช้เพื่อคำนวณหาค่าสัดส่วนของความเป็นสมาชิกของเซตวิภัชนัย (Fuzzy set) เพื่อเปรียบเทียบหาพยางค์หรือคำที่พ้องเสียงที่ให้เสียงได้เหมือนกันที่สุด

ขั้นตอนของการสังเคราะห์เสียงพูดของรายนามผู้ใช้โทรศัพท์ แสดงดังรูปที่ 3.10



รูปที่ 3.10 ขั้นตอนของการสังเคราะห์เสียงพูด

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

3.5.1 การตัดคำภาษาไทย

ลักษณะทางโครงสร้างของภาษาอังกฤษจะมีประโยคซึ่งประกอบด้วยคำหลายๆคำเรียงต่อกันไปโดยมีช่องว่าง เป็นตัวคั่นคำ การตัดคำจึงทำค่อนข้างสะดวก แต่สำหรับโครงสร้างทางภาษาไทยและภาษาอื่นๆ ในภูมิภาคเอเชียซึ่งมีความใกล้เคียงทางโครงสร้างภาษา เช่น ภาษาลาว ภาษาจีน ภาษาญี่ปุ่นจะมีความซับซ้อนมากกว่า โดยการเขียนประโยคภาษาไทยเป็นการเขียนติดต่อกันเป็นส่วนใหญ่ จึงต้องมีวิธีการให้คอมพิวเตอร์มีความสามารถที่จะเรียนรู้ขอบเขตของคำได้

หลักการตัดคำในภาษาไทยสามารถแบ่งออกได้เป็น 2 หลักการใหญ่คือ

ก. หลักการตัดคำโดยใช้กฎเกณฑ์ ซึ่งขั้นตอนการตัดคำในยุคแรกๆจะใช้วิธีการตรวจสอบกฎเกณฑ์ของคำภาษาไทย เช่น กฎเกณฑ์ของตัวอักษรที่อยู่ติดกัน หรือกฎเกณฑ์ที่กำหนดโดยราชบัณฑิตยสถาน วิธีการนี้มีข้อจำกัดมากนั่นคือผลของการตัดคำอาจได้กลุ่มของคำ ซึ่งในความเป็นจริงยังสามารถตัดคำแยกย่อยออกไปได้อีก นั่นคือความถูกต้องของคำหลังการตัดต่ำ

ข. หลักการตัดคำโดยใช้พจนานุกรม ในยุคต่อมาขั้นตอนวิธีการตัดคำภาษาไทยโดยส่วนใหญ่จะใช้พจนานุกรมเข้ามาช่วย วิธีการนี้ถึงแม้จะใช้เนื้อที่ของหน่วยความจำหลักมาก แต่ก็เป็นวิธีที่ให้ความถูกต้องในการตัดคำสูงวิธีหนึ่ง ซึ่งในงานวิจัยนี้ได้เลือกขั้นตอนการตัดคำโดยใช้พจนานุกรมนี้

3.5.2 **พจนานุกรมหน่วยเสียง** (Dictionary Based Pronunciation)

คำศัพท์ที่เก็บอยู่ในพจนานุกรมหน่วยเสียงนั้นแตกต่างจากคำศัพท์ในพจนานุกรมทั่วไป โดยจะเก็บคำศัพท์ทั้งระดับคำและระดับพยางค์ เพราะเก็บเป็นพยางค์ตามการอ่านออกเสียง เช่น คำว่า สิญชัย นำสิน พจนานุกรมการอ่านออกเสียงจะแยกเก็บเป็นคำว่า สิน นำ และ ชัย ทั้งนี้พยางค์ที่พ้องเสียงกันคือ สิน กับ สิญ พจนานุกรมจะทำการเก็บพยางค์ สิน เพียงหนึ่งพยางค์ เนื่องจากออกเสียงเหมือนกันสามารถใช้ข้อมูลเสียงร่วมกันได้ ซึ่งข้อมูลเสียงที่ใช้ในงานวิจัยนี้ได้นำบางส่วนมาจากงานวิจัยเรื่อง การสังเคราะห์ข้อความเสียงพูดภาษาไทยสำหรับคำทับศัพท์ภาษาอังกฤษและคำนามเฉพาะ ของอัจจิมา ตันสกุล[11] และได้ทำการแก้ไขเพิ่มเติมในส่วนที่เสียงไม่ชัดเจนคุณภาพไม่ดี และเพิ่มความเป็นธรรมชาติมากยิ่งขึ้น นอกจากนี้ยังได้ทำการเพิ่มคำศัพท์บางคำที่ใช้ในส่วนของชื่อและนามสกุลของรายนามผู้ใช้โทรศัพท์เป็นจำนวนหลายครั้งลงในพจนานุกรมหน่วยเสียง รวมทั้งสิ้น 4,682 เสียง โดยบันทึกเสียงบนเครื่องคอมพิวเตอร์ รุ่น Pentium4 การ์ดเสียงของบริษัท Creative Technology รุ่น Creative SB Audio PCI ไมโครโฟน AKG รุ่น D50S ใช้โปรแกรม Cool Edit2000 Version1.1 ในการเก็บแฟ้มข้อมูลเสียงแบบโมโน 16 บิต ที่อัตราการชักตัวอย่าง 11,025 เฮิรตซ์

3.6 การประมาณค่าความใกล้เคียงของคำพ้องเสียง

การเปรียบเทียบในการประมาณค่าความใกล้เคียงของคำพ้องเสียงเพื่อคำนวณหาค่าสัดส่วนของความเป็นสมาชิกของเซตวิภัชนัย (Fuzzy Set) ทำได้โดยการกำหนดให้มีการจัดกลุ่มอักขระตามการพ้องเสียง ทั้งนี้จะมีสัดส่วนของเซตวิภัชนัยที่ไม่เท่ากัน โดยกำหนดกลุ่มของอักขระดังต่อไปนี้

กลุ่มของพยัญชนะต้น

การจัดกลุ่มของพยัญชนะไทยที่ใช้ในงานวิจัยนี้ใช้วิธีการจัดกลุ่มโดยนำพยัญชนะ 16 ตัวที่มิได้เกี่ยวข้องกับการผันมา นำมาจากพยัญชนะเดิม 13 ตัวซึ่งไม่มีส่วนในการผันวรรณยุกต์ได้แก่ ฆ ฌ ญ ฏ ฐ ฑ ฒ ณ ธ ภ ฬ ศ ษ (มีเพียง 15 คำที่ปรากฏในคำไทยที่มีการผัน) และ ฃ ฅ ฎ แม้จะเป็นกลุ่มพยัญชนะเดิม ฎ ก็ใช้เขียนคำไทยแต่โบราณคำเดียวคือ กฎ นอกนั้นใช้เขียนคำแผลงจาก ฎ ในภาษาบาลี-สันสกฤต และมักใช้เป็นพยัญชนะต้น ส่วน ฃ ฅ แม้จะผันวรรณยุกต์ได้ แต่เมื่อเป็นพยัญชนะต้นที่เลิกใช้แล้ว ก็เท่ากับไม่มีส่วนในการผันวรรณยุกต์ โดยกลุ่มคำเหล่านี้มาพิจารณาถ้าคำใดมีพยัญชนะเสียงซ้ำกับตัวใดใน 28 ตัว ก็จัดเข้าหมู่เดียวกับตัวนั้นๆไป เป็นการกำหนดพื้นเสียงของคำ [5] (เสียงวรรณยุกต์ของคำที่ยังไม่ได้ผันจะเป็นเสียงวรรณยุกต์ใดขึ้นอยู่กับชนิดตัวอักษร ลักษณะพยางค์ (คำเป็น-คำตาย) และเสียงสระสั้น-ยาว) เพื่อใช้ในการอ่านและการพูดเท่านั้น มิใช่เพื่อการผัน ซึ่งแสดงการจำแนกกุล่มของพยัญชนะไทยได้ในตารางที่ 3.1

ตารางที่ 3.1 การจัดกลุ่มเสียงพยัญชนะไทยเพื่อใช้ในการอ่านและพูด

ฎ ด	ฌ ช	
ฏ ต	ญ ย	
ฐ ถ	ฑ ฒ ธ ท	
ศ ษ ส	ณ น	
ฆ ค	ภ พ	
ฬ ล	ฅ ค	ฃ ข

กลุ่มของตัวสะกดไทย

มีหลักเกณฑ์การจัดกลุ่มตามมาตรา คือ แม่บทแจกลูกอักษรตามหมวดคำที่มีตัวสะกดหรือออกเสียงอย่างเดียวกัน แบ่งเป็น 8 มาตรา โดยนำมาประยุกต์ใช้เพื่อเปรียบเทียบตัวสะกด ถ้าอยู่ในกลุ่มนี้จะได้ค่าคะแนนความใกล้เคียงของคำพ้องเสียง ดังแสดงในตารางที่ 3.2

ตารางที่ 3.2 การจัดกลุ่มของเสียงตัวสะกดไทยที่สอดคล้องกัน 8 กลุ่ม

ก ข ค ฆ	น ญ ณ ร ล ฬ	ง
ด จ ช ซ ฌ ฎ ฎ ฐ ฑ ฒ ต ถ ท ธ ศ ษ ส	ย	ม
บ ป พ ฟ ภ	ว	

3.6.1 การกำหนดค่าตัววัดความใกล้เคียงของคำพ้องเสียง

จากหลักไวยากรณ์ภาษาไทยที่ใช้ในการประสมพยางค์หรือคำ ซึ่งพยางค์หนึ่งมีส่วนประสมต่างๆเรียงตามลำดับดังนี้

1.)สระนำ

2.)พยัญชนะต้น

3.)พยัญชนะควบกล้ำ

4.)สระตาม

5.)ตัวสะกด

6.)วรรณยุกต์

การกำหนดค่าความใกล้เคียงของคำพ้องเสียงนั้น ผู้วิจัยได้กำหนดตามความเหมาะให้กับคำพ้องเสียงของพยัญชนะต้นและสระไว้ดังนี้

ค่าตัววัดความใกล้เคียงของคำพ้องเสียงของพยัญชนะต้น = 6

ค่าตัววัดความใกล้เคียงของคำพ้องเสียงของตัวสะกด = 4

ส่วนประสมของพยางค์หรือคำที่เหลือได้แก่ สระนำ พยัญชนะควบกล้ำ สระตาม วรรณยุกต์ ไม่ได้นำมาคำนวณหรือนำมาเป็นตัวเปรียบเทียบ ดังนั้นรูปของ สระนำ พยัญชนะควบกล้ำ สระตาม และวรรณยุกต์ จะยังคงรูปเดิมไม่เปลี่ยนแปลง

3.6.2 ความสัมพันธ์แบบวิภัชนัยของการประมาณค่าความใกล้เคียงของคำพ้องเสียง

การคำนวณหาค่าสัดส่วนของความเป็นสมาชิกของเซตวิภัชนัย เพื่อเปรียบเทียบหาพยางค์หรือคำที่มีค่าความใกล้เคียงของคำพ้องเสียงสูงที่สุด โดยได้กำหนดสมการได้ดังนี้

$$F = \frac{\left(\left(\sum C[i] \times M_i\right) \times 100\right)}{\left(\sum M_i \times n\right)} \qquad (3.1)$$

โดยที่ F = ค่าสัดส่วนของความเป็นสมาชิกของเซตวิภัชนัย

$C[i]$ = สัดส่วนของเซตวิภัชนัยตามกลุ่มของตัวอักขระ

M_i = ค่าตัววัดความใกล้เคียงของคำพ้องเสียง

n = จำนวนครั้งของการใช้ค่าตัววัดความใกล้เคียงของคำพ้องเสียง

สัดส่วนของเซตวิภัชนัยที่ทำให้เกิดเป็นคำพ้องเสียง ตามกลุ่มของตัวอักขระต่างๆจะอยู่ในช่วง [0,1] ดังตัวอย่างต่อไปนี้

คำที่ให้อ่านคือ สุญ คำที่ค้นพบในพจนานุกรม คือ สุน ซึ่งแสดงความสัมพันธ์ของเซตวิภัชนัยได้ตามรูปที่ 3.11

M_i	M_1	M_2
M_1	1	0
M_2	0	0.9

รูปที่ 3.11 ความสัมพันธ์แบบวิภัชนัยบน { M_1 , M_2}

จากรูป M_1 เป็นค่าตัววัดความใกล้เคียงของพยัญชนะที่พ้องเสียงจากการจำแนกกลุ่มพยัญชนะภาษาไทย และ M_2 เป็นค่าตัววัดความใกล้เคียงของพยัญชนะที่พ้องเสียงจากการจำแนกกลุ่มตัวสะกดภาษาไทย ดังนั้นค่าสัดส่วนของเซตวิภัชนัยตามกลุ่มของตัวอักขระ = [1, 0.9] โดยที่ส่วนประกอบที่เหลือของคำเช่นสระ วรรณยุกต์ จะต้องไม่เปลี่ยนแปลงหรือผันเสียงในกรณีที่เป็นวรรณยุกต์ แล้วนำมาคำนวณหาค่าสัดส่วนของความเป็นสมาชิกของเซตวิภัชนัย

$$F = \frac{\left((1\times 6)+(0.9\times 4)\right)\times 100}{\left((6\times 1)+(4\times 1)\right)} \qquad (3.2)$$

$$= 96.6$$

และนำผลที่ได้จากการหาค่าสัดส่วนของความเป็นสมาชิกของเซตวิภัชนัยนั้นมาพิจารณา โดยค่าของคำที่เข้าใกล้ค่า 100 มากที่สุดจะเป็นคำที่มีค่าความใกล้เคียงของคำพ้องเสียงสูงที่สุด แต่ในกรณีที่มีค่าเท่ากันสองคำ จะเลือกคำแรกที่ค้นคืนได้ก่อนเป็นอันดับแรก

3.7 การหาค่าอัตราความถูกต้องของระบบ

เนื่องจากในระบบสอบถามรายนามผู้ใช้โทรศัพท์แบบอัตโนมัติแบ่งส่วนออกเป็น 2 ส่วนใหญ่ๆ คือ การรู้จำเสียงพูด และสังเคราะห์เสียงพูด ดังนั้นในการคำนวณหาอัตราความถูกต้องของทั้งระบบก็สามารถหาได้โดยพูดหมายเลขที่ต้องการเข้าไปในระบบแล้วให้ระบบพูดชื่อของเจ้าของหมายเลขนั้นออกมาแล้วมาหาค่าเปอร์เซ็นต์ความถูกต้องที่ระบบสามารถไปหาชื่อของรายนามผู้ใช้โทรศัพท์ออกมาแล้วออกเสียงชื่อและนามสกุลได้ถูกต้อง ซึ่งโดยความเป็นจริงแล้วในงานวิจัยนี้ส่วนของการสังเคราะห์เสียงผู้วิจัยมีเจตนาที่จะให้ผลของการอ่านออกเสียงชื่อและนามสกุลมีค่าความถูกต้อง 100 เปอร์เซ็นต์ จึงได้เพิ่มเติมในส่วนของคำอ่านของชื่อและนามสกุล ถ้าชื่อหรือนามสกุลใดที่ระบบสังเคราะห์เสียงพูด พูดผิดหรือบางคำไม่มีในพจนานุกรมก็ให้ผู้ดูแลระบบเป็นผู้แก้ไขหรือทำการเพิ่มคำศัพท์เข้าไปในพจนานุกรม ซึ่งผลของค่าอัตราความถูกต้องจะขึ้นกับค่าอัตราความถูกต้องของการรู้จำเสียงพูดนั่นเอง

ดังนั้นอัตราความถูกต้องของระบบจะเท่ากับอัตราความถูกต้องของการรู้จำเสียงพูดในระดับคำต่อเนื่องซึ่งสามารถคำนวณได้จากวิธีการดังต่อไปนี้

กำหนดให้ A_i เป็นคำตอบที่ได้จากข่ายงานระบบประสาทเทียมตามลำดับของชุดหมายเลข

หลังจากที่หาค่า A_i ได้แล้ว ก็คำนวณหาความถูกต้อง โดยให้ค่าคะแนนความถูกต้อง(Score)กับชุดหมายเลขที่ถูกต้อง คือถ้าคำตอบที่ได้ตรงกับคำตอบที่ถูกต้องทุกตัวก็ให้คะแนนเท่ากับ 1 แต่ถ้ามีตัวหนึ่งตัวใดไม่ตรงกับคำตอบที่ถูกต้องจะให้คะแนนเท่ากับ 0 ดังสมการต่อไปนี้

$$score_j = \begin{cases} 1; & A_i = C_j \\ 0; & A_i \neq C_j \end{cases} \qquad (3.3)$$

เมื่อ j คือ ลำดับของชุดหมายเลข

i คือ ลำดับของตัวเลขในชุดหมายเลข

C_j คือ คำตอบที่ถูกต้องทั้งลำดับและตัวเลขของชุดหมายเลข

ดังนั้นเมื่อทำการให้ค่าคะแนนความถูกต้องไปจนครบทุกชุดหมายเลขแล้ว ก็นำค่าคะแนนรวมของชุดหมายเลขทั้งหมด แล้วนำมาคำนวณหาค่าอัตราความถูกต้องของระบบได้ดังสมการที่ 3.4

$$\text{อัตราความถูกต้องของระบบ} = \frac{\sum score_j \times 100}{\text{จำนวนชุดหมายเลขทั้งหมด}} \qquad (3.4)$$

บทที่ 4

การทดลองและผลการทดลอง

ในบทนี้จะกล่าวถึงวิธีการทดลองและผลการทดลองของระบบสอบถามรายนามผู้ใช้โทรศัพท์อัตโนมัติซึ่งเป็นการทดลองหาผลของอัตราความถูกต้องของระบบทั้งหมด โดยในส่วนของการรับเสียงที่ทำหน้าที่รับเสียงพูดเข้ามาในระบบนั้น ในการทดลองเสียงที่ได้มานั้นมาจากสองแหล่ง คือ ไมโครโฟน และจากแฟ้มข้อมูลเสียงซึ่งเก็บเสียงพูดไว้เพื่อความสะดวกในการตรวจสอบความถูกต้อง

4.1 วิธีการทดลอง

1. นำแฟ้มข้อมูลเสียงของชุดทดสอบ (Test Set) ซึ่งประกอบด้วยข้อมูลจากผู้พูดทั้งหมด 10 คน แบ่งเป็นชาย 5 คน หญิง 5 คน จำนวนชุดหมายเลข 100 ชุด รวมทั้งสิ้น 700 ตัวเลข มาทำการทดสอบโดยนำเข้าโปรแกรมแล้วให้โปรแกรมทำการรู้จำเสียงแล้วแสดงผลการรู้จำที่ได้ออกมาทางจอภาพแล้วตรวจสอบอัตราความถูกต้องที่ได้
2. ทำการปรับค่าพารามิเตอร์ต่างๆเพื่อหาค่าที่เหมาะสมเพื่อให้ได้ค่าอัตราความถูกต้องที่สูงเพียงพอ ค่าพารามิเตอร์ที่ใช้ในการรู้จำเสียงพูดที่ใช้ในงานวิจัยมีดังต่อไปนี้

- เสียงที่ใช้บันทึกที่อัตราการชักข้อมูล 11,025 เฮิรตซ์ 16 บิต บันทึกแบบช่องสัญญาณเดี่ยว (Mono channels)

- ลักษณะสำคัญของเสียงใช้ราสต้า-พีแอลพีอันดับที่ 12 จำนวนกรอบ 6, 9, 12, และ15 กรอบ ตามลำดับ โดยแต่ละกรอบจะมีความยาว 25 มิลลิวินาที

- ข่ายงานระบบประสาทเทียมที่ใช้ในการเรียนรู้เป็นแบบเพอร์เซปตรอนหลายชั้น และวิธีการเรียนรู้แบบแพร่กระจายย้อนกลับ
- ค่าโมเมนตัม 0.9 ค่าอัตราการเรียนรู้ 0.0001 และ 0.00001
- จำนวนรอบที่นำมาทำการทดลอง 500, 2,000, และ 5,000 รอบ
- จำนวนบัพของชั้นข้อมูลขาเข้าที่ใช้ 144, 216, 288, 360 บัพ
- จำนวนบัพชั้นข้อมูลซ่อนตัวที่ใช้ 10, 20, 50, 100 บัพ
- จำนวนบัพของชั้นข้อมูลขาออก 10 บัพ

3. ทำการหาค่าความผิดพลาดซึ่งเกิดจากการตัดแบ่งพยางค์ โดยนำแฟ้มข้อมูลเสียงทั้งหมดมาผ่านขั้นตอนการตัดแบ่งพยางค์แล้วนำมาตรวจสอบโดยใช้โปรแกรม SpeechView [13] เป็นเครื่องมือช่วยแล้ววิเคราะห์ดูว่า พยางค์ใดที่ตัดออกมาแล้วมีส่วนประกอบของพยางค์ไม่ครบซึ่ง

ได้แก่ ส่วนของพยัญชนะ ส่วนสระ และส่วนตัวสะกด ถ้าพยางค์ใดขาดส่วนหนึ่งส่วนใดไปถือว่าเกิดความผิดพลาดของการตัดแบ่งพยางค์

ในงานวิจัยนี้ผู้วิจัยได้แบ่งการทดลองออกเป็น 9 ชุดการทดลองดังตารางที่ 4.1 โดยแต่ละชุดการทดลองจะมีการปรับเปลี่ยนค่าพารามิเตอร์ตามความสนใจและความเหมาะสม ถ้าหากค่าใดที่ทำให้ค่าอัตราความถูกต้องสูงที่สุด จะนำค่านั้นมาใช้ในงานทดลองต่อไป

ตารางที่ 4.1 ค่าพารามิเตอร์ที่ปรับเปลี่ยนของแต่ละชุดการทดลอง

การทดลองชุดที่	1	2	3	4	5	6	7	8	9
จำนวนกรอบ	6	9	12	15	12	12	12	12	12
จำนวนชั้นข้อมูลซ่อนตัว	1	1	1	1	1	1	2	2	3
ค่าอัตราการเรียนรู้	10^{-4}	10^{-4}	10^{-4}	10^{-4}	10^{-4}	10^{-4}	10^{-5}	10^{-5}	10^{-5}
จำนวนบัพของชั้นข้อมูลขาเข้า	144	216	288	360	288	288	288	288	288
จำนวนบัพของชั้นข้อมูลซ่อนตัวชั้นที่หนึ่ง	10	10	10	10	50	100	10	50	10
จำนวนบัพของชั้นข้อมูลซ่อนตัวชั้นที่สอง	-	-	-	-	-	-	20	100	50
จำนวนบัพของชั้นข้อมูลซ่อนตัวชั้นที่สาม	-	-	-	-	-	-	-	-	100
จำนวนบัพของชั้นข้อมูลขาออก	10	10	10	10	10	10	10	10	10
จำนวนรอบที่ใช้ในการฝึก	500	2000	2000	2000	2000	2000	5000	5000	5000

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

4.2 ผลการทดลอง

จากการทดลองดังกล่าวข้างต้นได้ผลการทดลองดังตารางที่ 4.2

ตารางที่ 4.2 อัตราความถูกต้องของแต่ละชุดการทดลอง

การทดลองชุดที่	อัตราความถูกต้องของระบบ(%)	อัตราความถูกต้องระดับพยางค์(%)
1	69	92.9
2	71	92.9
3	73	93
4	73	92.9
5	73	93.6
6	73	93.6
7	75	94
8	75	94
9	73	93.6

การคำนวณหาค่าอัตราความถูกต้องของระบบคำนวณได้จากนำค่าคะแนนรวมของชุดหมายเลขทั้งหมด ตามสมการที่ 3.3 แล้วนำมาคำนวณโดยใช้สมการที่ 3.4 ส่วนค่าอัตราความถูกต้องระดับพยางค์ คำนวณจากผลรวมของหมายเลขที่ถูกต้องคูณด้วยหนึ่งร้อยแล้วนำมาหารด้วยจำนวนหมายเลขเดี่ยวทั้งหมดที่ทำการทดลอง

ค่าความถูกต้องซึ่งเกิดจากการตัดแบ่งพยางค์มีค่าเท่ากับ 86 เปอร์เซนต์ ซึ่งได้จากการนำแฟ้มข้อมูลเสียงจำนวน 500 ชุดหมายเลข ที่บันทึกเมื่อนำเข้าขั้นตอนการตัดแบ่งพยางค์แล้วปรากฏว่าตัดแบ่งพยางค์ผิดพลาดคือมีส่วนประกอบของพยางค์ไม่ครบและต้องนำมาทำการแก้ไขก่อนนำไปสู่ในขั้นตอนฝึกของข่ายงานระบบประสาทเทียม จำนวน 70 ชุดหมายเลข

4.3 วิเคราะห์ผลการทดลอง

จากการทดลองชุดที่ 1,2,3 และ 4 เป็นการเปลี่ยนจำนวนกรอบจาก 6 เป็น 9 ,12 และ15 กรอบตามลำดับ ซึ่งผลปรากฏว่าจำนวนกรอบที่เท่ากับ 12 กรอบให้ค่าอัตราความถูกต้องของระบบและอัตราความถูกต้องระดับพยางค์สูงที่สุดคือ 73 และ 93 เปอร์เซ็นต์ ตามลำดับ ดังนั้นจึงใช้จำนวนกรอบ 12 กรอบในชุดการทดลองต่อไป ในการทดลองชุดที่ 5 และ 6 เป็นการทดลองเพิ่มจำนวนบัพของชั้นข้อมูลซ่อนตัวของข่ายงานประสาทเทียมเป็น 50 บัพ และ 100 บัพ ตามลำดับ

ผลการทดลองปรากฏว่าให้ค่าอัตราความถูกต้องของระบบและอัตราความถูกต้องระดับพยางค์เพิ่มขึ้น แต่ทำเวลาในการประมวลผลเมื่อนำไปใช้งานจริงเพิ่มขึ้นเป็น 2 เท่า คือจากเดิมใช้เวลาในส่วนของการรู้จำเสียงประมาณ 1 วินาที เพิ่มขึ้นเป็นประมาณเกือบ 2 วินาที ซึ่งก็เป็นเวลาที่ยอมรับได้เมื่อนำไปใช้ในระบบจริง ส่วนในการทดลองที่ 7 และ 8 จึงเป็นการเพิ่มจำนวนชั้นของชั้นข้อมูลซ่อนตัวและทดลองปรับจำนวนบัพในแต่ละชั้นที่ต่างกัน นอกจากนั้นยังทำการปรับค่าอัตราการเรียนรู้กับจำนวนรอบที่ใช้ในฝึกเพื่อให้ได้ค่าอัตราการรู้จำที่ยอมรับได้โดยใช้เวลาไม่นาน ซึ่งผลปรากฏว่ามีค่าอัตราความถูกต้องของระบบและค่าอัตราความถูกต้องระดับพยางค์เพิ่มขึ้นมาเป็น 75 และ 94 เปอร์เซ็นต์ตามลำดับ หลังจากนั้นได้ทำการทดลองชุดที่ 9 โดยการเพิ่มจำนวนชั้นของชั้นข้อมูลซ่อนตัวเป็น 3 ชั้นผลปรากฏว่าให้ค่าอัตราความถูกต้องลดลงเล็กน้อย ซึ่งเมื่อทำการทดลองทั้ง 9 ชุดแล้วได้ ค่าพารามิเตอร์ที่เหมาะสมที่สุด ได้แก่ ใช้ลักษณะทางสวนศาสตร์คือ ราสต้า-พีแอลพี และอนุพันธ์อันดับหนึ่ง โดยใช้จำนวนกรอบ 12 กรอบ ส่วนการเรียนรู้ของข่ายงานระบบประสาทเทียมใช้จำนวนชั้นของข้อมูลขาเข้า 288 บัพ จำนวนชั้นข้อมูลซ่อนตัว 2 ชั้น ชั้นที่หนึ่ง 50 บัพ และชั้นที่สอง 100 บัพ และจำนวนชั้นข้อมูลขาออก 10 บัพ

จากผลการทดลองพบว่าการปรับค่าทั้งส่วนของลักษณะทางสวนศาสตร์และส่วนของการเรียนรู้ของข่ายงานระบบประสาทเทียมเพิ่มค่าอัตราความถูกต้องของระบบไม่มากนัก ดังนั้นถ้าต้องการที่จะเพิ่มค่าอัตราความถูกต้องให้มากขึ้นจึงควรปรับปรุงในส่วนของการตัดพยางค์ให้มีค่าความผิดพลาดให้น้อยที่สุด โดยอาจใช้ค่าต่างๆร่วมประกอบในการพิจารณาหาจุดตัดของพยางค์ เช่น ค่าสัมประสิทธิ์อัตสหสัมพันธ์ (Autocorrelation Coefficients) [24] และปรับปรุงส่วนของการตัดพยางค์ให้มีความละเอียดขึ้นเนื่องจากในงานวิจัยนี้จุดปลายของพยางค์หนึ่งเป็นจุดต้นของพยางค์หนึ่ง ซึ่งในกรณีที่ผู้พูดเว้นระยะเวลาในการพูดแต่ละพยางค์นานก็จะทำให้การตัดพยางค์รับเอาเสียงเงียบที่อยู่ระหว่างพยางค์เข้าไปด้วย

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

ในงานวิจัยนี้เป็นการนำเอาเทคโนโลยีทางด้านการรู้จำเสียงพูดและการสังเคราะห์เสียงพูดภาษาไทยมาพัฒนาแล้วนำมาประยุกต์ใช้กับระบบสอบถามรายนามผู้ใช้โทรศัพท์แบบอัตโนมัติ โดยจำลองระบบลงบนเครื่องไมโครคอมพิวเตอร์รับเสียงข้อมูลขาเข้าโดยผ่านทางไมโครโฟนและส่งเสียงข้อมูลขาออกโดยผ่านทางลำโพง

5.1 สรุปผลการวิจัย

ระบบสอบถามรายนามผู้ใช้โทรศัพท์แบบอัตโนมัตินั้นจะประกอบด้วย 2 ขั้นตอนใหญ่ๆ ได้แก่ การรู้จำเสียงพูดภาษาไทย และการสังเคราะห์เสียงพูดภาษาไทย

ในขั้นตอนของการรู้จำเสียงพูดซึ่งเป็นเสียงพูดของตัวเลขต่อเนื่องนั้น สามารถแบ่งเป็นขั้นตอนย่อยได้แก่ ขั้นตอนการหาจุดตัดหัวท้ายหน่วยและการตัดแบ่งพยางค์ ซึ่งผู้วิจัยได้เสนอวิธีในการตัดแบ่งโดยใช้ค่าพลังงานและค่าอัตราตัดผ่านแกนศูนย์ร่วมกันในการพิจารณาหาจุดแบ่ง ซึ่งให้ผลของอัตราความถูกต้องในการตัดแบ่งพยางค์เท่ากับ 86 เปอร์เซ็นต์ ส่วนขั้นตอนการรู้จำเสียงพูดได้ใช้คุณลักษณะสำคัญของเสียงคือราสต้า-พีแอลพี อันดับที่ 12 และอนุพันธ์อันดับที่ 1 และใช้ข่ายงานระบบประสาทเทียมช่วยในการฝึกฝนและการรู้จำ ซึ่งได้ค่าพารามิเตอร์ที่เหมาะสมสำหรับการเรียนรู้จำเสียงพูดตัวเลขต่อเนื่องคือ จำนวนกรอบ 12 กรอบ จำนวนชั้นข้อมูลซ่อนตัว 2 ชั้น จำนวนบัพของชั้นข้อมูลขาเข้า 288 บัพ จำนวนบัพของชั้นข้อมูลซ่อนตัวที่หนึ่ง 50 บัพ จำนวนบัพของชั้นข้อมูลซ่อนตัวที่สอง 100 บัพ จำนวนบัพของชั้นข้อมูลขาออก 10 บัพ ซึ่งได้ค่าอัตราความถูกต้องของระบบเท่ากับ 75 เปอร์เซ็นต์และได้ค่าอัตราความถูกต้องระดับพยางค์ 94 เปอร์เซ็นต์

ส่วนในขั้นตอนการสังเคราะห์เสียงพูดใช้วิธีการสังเคราะห์เสียงโดยใช้วิธีการตัดคำโดยใช้เปรียบเทียบจากพจนานุกรมแล้วนำหน่วยเสียงที่ได้จากพจนานุกรมหน่วยเสียงมาต่อกันเป็นเสียงพูด ในงานวิจัยนี้ได้เพิ่มเติมส่วนที่เป็นคำอ่านของชื่อและนามสกุลเพื่อให้มีค่าความถูกต้องมากที่สุดคือ 100 เปอร์เซ็นต์ และได้เสนอวิธีการประมาณค่าความใกล้เคียงของคำพ้องเสียงในกลุ่มของพยัญชนะต้นและตัวสะกด โดยใช้ทฤษฎีเซตวิภัชนัยเข้ามาช่วยเพื่อลดจำนวนการเพิ่มคำศัพท์ในพจนานุกรม ซึ่งในชื่อและนามสกุลจะพบคำพ้องเสียงอยู่ 20 เปอร์เซ็นต์โดยประมาณ

5.2 ข้อเสนอแนะ

1. ผู้บอกภาษาที่ใช้ในการทดลองไม่ได้จำกัดสำเนียง และความชัดเจนในการพูดของผู้พูด ซึ่งมีบางคนที่เสียงคุณภาพไม่ดีเท่าที่ควร และมีสำเนียงที่ต่างกัน ความยาวนานในการการออกเสียงแต่ละพยางค์ไม่เท่ากันดังนั้นควรมีการคัดเลือกผู้บอกภาษาเพื่อให้ได้เสียงที่มีคุณภาพที่ดี ซึ่งจะทำให้ผลการรู้จำดีขึ้น
2. ในการทดลอง ใช้ข้อมูลจากผู้บอกภาษาจำนวน 50 คนเท่านั้นซึ่งน้อยเกินไป ทำให้เกิดความแปรปรวนมาก และเมื่อนำมาฝึกโดยใช้ข่ายงานระบบประสาทเทียมแล้วไม่สามารถครอบคลุมความแปรปรวนทั้งหมดได้ ดังนั้นควรที่จะเพิ่มจำนวนผู้บอกภาษาให้มากกว่านี้ โดยนำเสียงที่บันทึกจากการทดสอบแบบทันกาลมาทำการฝึกโดยใช้ข่ายงานระบบประสาทเทียมไปเรื่อยๆ เพื่อเพิ่มจำนวนผู้บอกภาษา เพื่อให้อัตราการรู้จำน่าเชื่อถือมากขึ้น
3. อัตราความถูกต้องของระบบขึ้นอยู่กับค่าอัตราความถูกต้องของการตัดแบ่งพยางค์มากกว่าการปรับค่าพารามิเตอร์ของข่ายงานระบบประสาทเทียม ดังนั้นถ้าเพิ่มค่าอัตราความถูกต้องของการตัดแบ่งพยางค์หรือค่าความผิดพลาดที่เกิดจากการตัดแบ่งพยางค์ก็จะทำอัตราความถูกต้องของระบบเพิ่มขึ้นมาก
4. ขั้นตอนในการตัดแบ่งพยางค์ควรจะตัดแบ่งพยางค์ให้ละเอียดขึ้นเพื่อให้ได้พยางค์นั้นๆจริง เนื่องจากในการพูดตัวเลขต่อเนื่องในการพูดของแต่ละคนจะมีลักษณะและทำนองของการพูดที่ไม่เหมือนกันทำให้เมื่อทำการตัดแบ่งพยางค์แล้วเกิดช่องว่างขึ้นในพยางค์นั้นๆไม่เท่ากันทำให้อัตราในการรู้จำไม่มากเท่าที่ควร
5. ในส่วนของการสังเคราะห์เสียงที่เป็นคำวิสามานยนามในภาษาไทยมีการออกเสียงที่เป็นเสียงอะ กึ่งเสียงที่เป็นตัวสะกดที่ไม่ประวิสรรชนีย์ค่อนข้างมาก ซึ่งทำให้ต้องแก้ไขในส่วนนี้มาก ดังนั้นจึงควรมีการศึกษาเพิ่มเติมในส่วนของการสังเคราะห์เสียงที่เป็นคำวิสามานยนามในภาษาไทยในมีความถูกต้องมากยิ่งขึ้น
6. โปรแกรมควรจะพัฒนาให้สามารถรับจำนวนหลักของชุดตัวเลขที่ไม่จำกัด แต่ต้องมีอัตราความถูกต้องใกล้เคียงกับการจำกัดจำนวนหลักของชุดตัวเลข
7. สามารถนำเอาโปรแกรมไปประยุกต์ใช้กับระบบงานต่างๆได้อาทิเช่น ระบบสอบถามผลสอบหรือการลงทะเบียนเรียน ระบบการซื้อขายฉลากกินแบ่งรัฐบาล หรืออาจนำเอาไปประยุกต์ใช้ในการรู้จำชื่อหรือคำที่ต้องการ เช่น ระบบโอนสายโทรศัพท์อัตโนมัติ ระบบสอบถามการพยากรณ์อากาศ ระบบสอบถามการซื้อขายหลักทรัพย์ เป็นต้น

รายการอ้างอิง

1. กาญจนา นาคสกุล. ระบบเสียงภาษาไทย. หนังสือชุด บรมราชกุมารีอักษรานุรักษ์ คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิยาลัย ลำดับที่ 4. กรุงเทพมหานคร : โครงการเผยแพร่ผลงานวิชาการ, 2545.
2. ชัย วุฒิวิวัฒน์ชัย. การรู้จำเสียงคำไทยหลายพยางค์แบบไม่ขึ้นกับผู้พูดโดยใช้เทคนิคแบบฟัซซีและนิวรอลเน็ตเวิร์ก. วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาวิศวกรรมไฟฟ้า วิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2540.
3. ธีระพันธ์ ล.ทองคำ. ภาษาไทยหลากหลายสำเนียง. กรุงเทพมหานคร : โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย, 2545.
4. ปิสันธ ปิยศิริเวช และภิรตา ตันติวีรสุด. การรู้จำคำพูดต่อเนื่องภาษาไทยระดับพยางค์โดยใช้ข่ายงานระบบประสาท. โครงงานทางวิศวกรรมระดับปริญญาตรี ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2542.
5. พิศศรี กมลเวชช. ครบครันเรื่องวรรณยุกต์. พิมพ์ครั้งที่ 2. กรุงเทพมหานคร : หจก.หอรัตนชัยการพิมพ์, 2544.
6. วิรัช ศรเลิศล้ำวาณิช. การตัดคำไทยในระบบแปลภาษา. ห้องปฏิบัติการวิจัยภาษาและวิทยาการความรู้, การแปลภาษาด้วยคอมพิวเตอร์,หน้า 50-56. ศูนย์เทคโนโลยีอิเลคทรอนิกส์และคอมพิวเตอร์แห่งชาติ.
7. สิริขวัญ เลิศวิไล และสิริวงศ์ วงศ์ทองเสริม. การรู้จำคำพูดตัวเลขเดี่ยวภาษาไทยแบบทันกาล. โครงงานทางวิศวกรรมระดับปริญญาตรี ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2545.
8. สุทธิชัย ปทุมล่องทอง. อ่านไทยอย่างไรให้ถูกต้อง. โรงพิมพ์ O.N.G : สำนักพิมพ์บุ๊คแบงค์, 2544.
9. อภิศักดิ์ สุพจน์เฉลิมขวัญ และปราการ อุณจักร. ขั้นตอนการหาขอบเขตพยางค์สำหรับคำพูดต่อเนื่องภาษาไทย. โครงงานทางวิศวกรรมระดับปริญญาตรี ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2542.
10. อมร ทวีศักดิ์. สัทศาสตร์. พิมพ์ครั้งที่ 3. บริษัท สหธรรมิก จำกัด : สถาบันวิจัยภาษาและวัฒนธรรมเพื่อพัฒนาชนบท มหาวิทยาลัยมหิดล, 2542.
11. อัจจิมา ตันสกุล. การสังเคราะห์ข้อความเสียงพูดภาษาไทยสำหรับคำทับศัพท์ภาษาอังกฤษและคำนามเฉพาะ. วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2544.

12. Bandemer, H. and Gottwald, S. Fuzzy sets Fuzzy logic Fuzzy Methods with Appication. United States of America : John Wiley &Sons,Inc, 1995.

13. Cool Edit Pro, Available from: http://www.adobe.com[2003,May 1].

14. CSLU (Center for Spoken Language Understanding), CSLU Toolkit, Available from: http://cslu.cse.ogi.edu/toolkit/ [2003,May 1]

15. Furui, S. Digital Speech Processing,Synthesis and Recognition. United States of America : Marcel Dekker,Inc, 2001.

16. Gold, B., and Morgan, N. Speech and Audio Signal Processing: Processing and Perception of Speech and Music. United States of America:John Wiley &Sons,Inc, 2000.

17. Hermansky, H.and Morgan, N., RASTA Processing of Speech. IEEE Transaction on Speech and Audio Processing, Vol.2 No. 4 , pp. 587-589, 1994.

18. Huang, X., Acero, A. and Hon, H. , Spoken Language Processing : A guide to Theory ,Algorithm, and System Development . United States of America : Prentice Hall PTR, 2001.

19. Jelinek, F. Statistical Methods for Speech Recognition. United States of America : The MIT Press, 1999.

20. Rabiner, L. and Juang, B. Fundamentals of Speech Recognition. New Jersey : Prentice-Hall,Inc, 1993.

21. Strom, N., The NICO (Neural Inference Computation) Toolkit, Available from: http://www.speech.kth.se/NICO/index.html[2003,Apirl 3]

22. Thomas A. Standish. Data Structures, Algorithms & Software Principles in C. United States of America : Addison-Wesley,1995.

23. Thomas F. Quatieri. Discrete-Time Speech Signal Processing Principles and Practice. Prentice-Hall signal processing series, 2001.

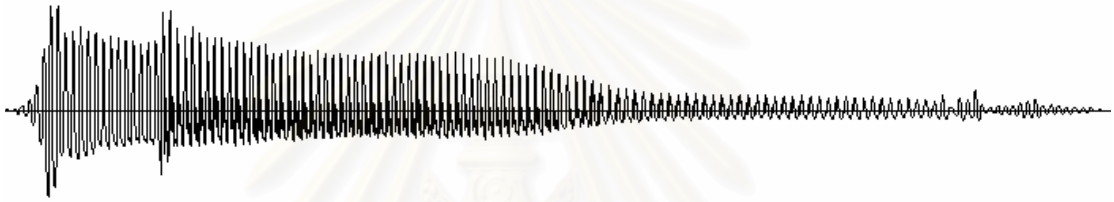
24. Thubthong, N., and Kijsirikul, B. A Syllable-Based Connected Thai Digit Speech Recognition Using Neural Network and Duration Modeling: The 1999 IEEE International Symposium on Intelligent Signal Processing and Communication Systems, pp. 785-788, 1999.

ภาคผนวก

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

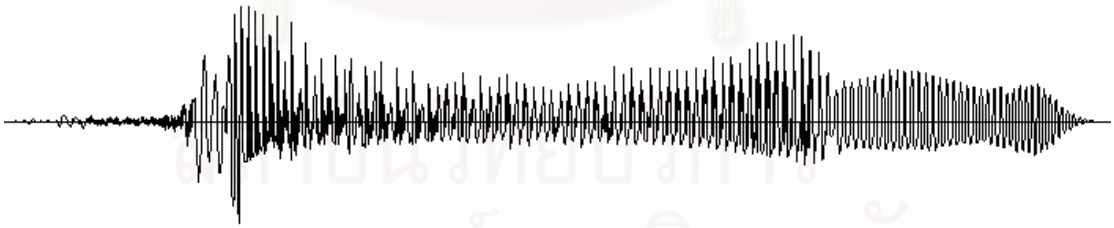
ภาคผนวก ก
ลักษณะคลื่นเสียงพูดตัวเลขภาษาไทย

เสียงศูนย์



เสียงหนึ่ง

เสียงสอง



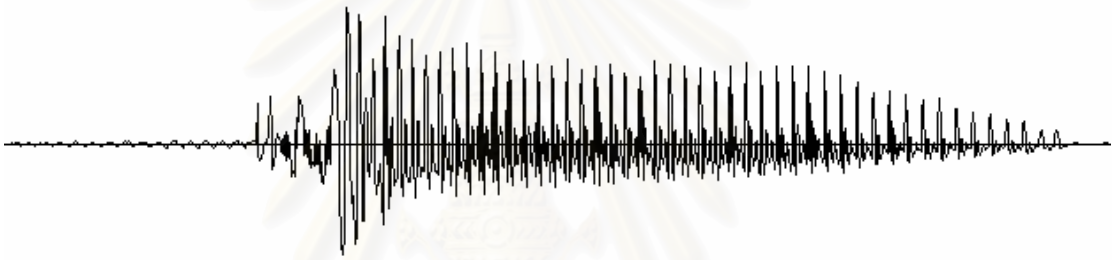
เสียงสาม

เสียงสี่

รูปที่ ก-1 แสดงคลื่นเสียงพูดตัวเลขภาษาไทยเลข ศูนย์ถึงสี่

เสียงห้า

เสียงหก



เสียงเจ็ด

เสียงแปด

เสียงเก้า

รูปที่ ก-2 แสดงคลื่นเสียงพูดตัวเลขภาษาไทยเลข ห้าถึงเก้า

สถาบันวิทยบริการ

ภาคผนวก ข
การแจกแจงผลการรู้จำของแต่ละชุดการทดลอง

ตารางที่ ข-1 ผลของการทดลองชุดที่ 1

	0	1	2	3	4	5	6	7	8	9	อัตราการรู้จำ
0	97.8	0	0	0	2.2	0	0	0	0	0	97.83
1	0	97.2	1.4	0	0	0	0	1.4	0	0	97.22
2	0	0	96.5	2.3	0	0	1.2	0	0	0	96.51
3	0	0	1.2	92.9	0	1.2	0	0	1.2	3.6	92.86
4	0	0	0	0	100	0	0	0	0	0	100
5	0	0	0	0	0	89.3	0	2.7	8	0	89.33
6	0	0	3	0	0	1.5	95.5	0	0	0	95.52
7	0	5.6	0	0	3.7	0	0	90.7	0	0	90.74
8	0	0	0	0	0	1.4	0	1.4	97.2	0	97.22
9	0	0	0	1.6	0	3.2	0	0	0	95.2	95.16

อัตราการรู้จำโดยเฉลี่ยเท่ากับ 95.29 เปอร์เซ็นต์

ตารางที่ ข-2 ผลของการทดลองชุดที่ 2

	0	1	2	3	4	5	6	7	8	9	อัตราการรู้จำ
0	95.7	0	0	0	2.2	0	2.2	0	0	0	95.65
1	0	95.8	0	0	2.8	0	0	1.4	0	0	95.83
2	0	0	96.5	2.3	0	0	1.2	0	0	0	96.51
3	0	0	1.2	96.4	0	0	0	0	1.2	1.2	96.43
4	0	0	0	0	95.1	0	0	4.9	0	0	95.12
5	0	1.3	0	0	0	92	0	0	6.7	0	92
6	0	0	1.5	0	1.5	1.5	95.5	0	0	0	95.52
7	0	1.9	0	0	3.7	0	0	92.6	1.9	0	92.59
8	0	0	0	0	0	1.4	0	2.8	95.8	0	95.83
9	0	0	0	1.6	0	3.2	0	0	1.6	93.5	93.55

อัตราการรู้จำโดยเฉลี่ยเท่ากับ 95 เปอร์เซ็นต์

ตารางที่ ข-3 ผลของการทดลองชุดที่ 3

	0	1	2	3	4	5	6	7	8	9	อัตราการรู้จำ
0	100	0	0	0	0	0	0	0	0	0	100
1	0	93.1	0	0	5.6	0	0	1.4	0	0	93.06
2	0	0	96.5	1.2	0	0	0	0	1.2	1.2	96.51
3	0	0	2.4	92.9	0	0	0	0	0	4.8	92.86
4	0	0	0	0	98.8	0	0	1.2	0	0	98.78
5	0	0	0	0	0	93.3	0	0	6.7	0	93.33
6	0	0	1.5	0	0	1.5	97	0	0	0	97.01
7	0	0	1.9	0	0	0	1.9	96.3	0	0	96.3
8	0	0	0	0	0	0	0	2.8	97.2	0	97.22
9	0	0	0	1.6	0	3.2	0	0	0	95.2	95.16

อัตราการรู้จำโดยเฉลี่ยเท่ากับ 95.86 เปอร์เซ็นต์

ตารางที่ ข-4 ผลของการทดลองชุดที่ 4

	0	1	2	3	4	5	6	7	8	9	อัตราการรู้จำ
0	98	0	0	0	0	0	2	0	0	0	98
1	0	93.8	2.1	0	2.1	0	0	2.1	0	0	93.75
2	2.3	1.1	83	13.6	0	0	0	0	0	0	82.95
3	0	0	2.6	91	0	0	0	0	0	6.4	91.03
4	0	0	0	0	98.5	1.5	0	0	0	0	98.53
5	0	0	0	5	0	93.3	0	0	0	1.7	93.33
6	0	1.7	0	0	0	0	96.6	0	1.7	0	96.55
7	0	1.7	0	0	0	0	3.3	93.3	0	1.7	93.33
8	0	0	0	2.8	0	4.2	0	0	91.7	1.4	91.67
9	0	0	8.6	4.3	0	5.7	0	0	2.9	78.6	78.57

อัตราการรู้จำโดยเฉลี่ยเท่ากับ 91.29 เปอร์เซ็นต์

ตารางที่ ข-5 ผลของการทดลองชุดที่ 5

	0	1	2	3	4	5	6	7	8	9	อัตราการรู้จำ
0	100	0	0	0	0	0	0	0	0	0	100
1	0	95.8	0	0	1.4	1.4	0	0	0	1.4	95.83
2	0	0	98.8	0	0	0	1.2	0	0	0	98.84
3	0	0	1.2	94	0	2.4	0	0	1.2	1.2	94.05
4	0	0	0	0	98.8	0	0	1.2	0	0	98.78
5	0	0	0	0	0	98.7	0	0	1.3	0	98.67
6	0	0	1.5	0	0	1.5	97	0	0	0	97.01
7	0	0	0	0	1.9	0	0	98.1	0	0	98.15
8	0	0	1.4	0	0	1.4	0	4.2	93.1	0	93.06
9	0	1.6	0	0	0	3.2	0	0	0	95.2	95.16

อัตราการรู้จำโดยเฉลี่ยเท่ากับ 96.86 เปอร์เซ็นต์

ตารางที่ ข-6 ผลของการทดลองชุดที่ 6

	0	1	2	3	4	5	6	7	8	9	อัตราการรู้จำ
0	100	0	0	0	0	0	0	0	0	0	100
1	0	94.4	0	0	2.8	2.8	0	0	0	0	94.44
2	0	0	100	0	0	0	0	0	0	0	100
3	0	0	0	96.4	0	1.2	0	0	0	2.4	96.43
4	0	0	0	0	98.8	0	0	1.2	0	0	98.78
5	0	0	0	0	0	97.3	0	0	2.7	0	97.33
6	0	0	1.5	0	0	1.5	97	0	0	0	97.01
7	0	0	0	0	0	0	0	98.1	1.9	0	98.15
8	0	0	1.4	0	0	0	0	1.4	97.2	0	97.22
9	0	0	0	1.6	0	3.2	0	0	0	95.2	95.16

อัตราการรู้จำโดยเฉลี่ยเท่ากับ 97.43 เปอร์เซ็นต์

ตารางที่ ข-7 ผลของการทดลองชุดที่ 7

	0	1	2	3	4	5	6	7	8	9	อัตราการรู้จำ
0	97.8	0	0	0	2.2	0	0	0	0	0	97.83
1	2.8	97.2	0	0	0	0	0	0	0	0	97.22
2	0	0	98.8	1.2	0	0	0	0	0	0	98.84
3	0	0	2.4	91.7	0	1.2	0	0	1.2	3.6	91.67
4	0	0	0	0	98.8	0	0	1.2	0	0	98.78
5	0	0	0	0	0	96	0	0	4	0	96
6	0	0	1.5	0	0	1.5	97	0	0	0	97.01
7	0	1.9	0	0	3.7	0	0	94.4	0	0	94.44
8	0	0	0	0	0	0	0	2.8	95.8	1.4	95.83
9	0	0	0	1.6	0	1.6	0	0	0	96.8	96.77

อัตราการรู้จำโดยเฉลี่ยเท่ากับ 96.43 เปอร์เซ็นต์

ตารางที่ ข-8 ผลของการทดลองชุดที่ 8

	0	1	2	3	4	5	6	7	8	9	อัตราการรู้จำ
0	100	0	0	0	0	0	0	0	0	0	100
1	0	97.2	0	0	1.4	1.4	0	0	0	0	97.22
2	0	0	98.8	0	0	0	1.2	0	0	0	98.84
3	0	0	1.2	96.4	0	0	0	0	0	2.4	96.43
4	0	0	0	0	98.8	0	0	1.2	0	0	98.78
5	0	1.3	0	0	0	96	0	0	2.7	0	96
6	0	0	1.5	0	0	1.5	97	0	0	0	97.01
7	0	0	0	0	0	0	0	98.1	1.9	0	98.15
8	0	0	0	0	0	0	0	4.2	94.4	1.4	94.44
9	0	0	0	1.6	0	1.6	0	0	0	96.8	96.77

อัตราการรู้จำโดยเฉลี่ยเท่ากับ 97.29 เปอร์เซ็นต์

ตารางที่ ข-9 ผลของการทดลองชุดที่ 9

	0	1	2	3	4	5	6	7	8	9	อัตราการรู้จำ
0	97.8	0	0	0	2.2	0	0	0	0	0	97.83
1	2.8	95.8	0	0	0	0	0	1.4	0	0	95.83
2	0	0	100	0	0	0	0	0	0	0	100
3	0	0	0	90.5	0	1.2	0	0	2.4	6	90.48
4	1.2	0	0	0	97.6	0	0	1.2	0	0	97.56
5	0	1.3	0	0	0	92	0	0	4	2.7	92
6	0	1.5	1.5	0	0	1.5	95.5	0	0	0	95.52
7	0	0	0	1.9	1.9	0	3.7	92.6	0	0	92.59
8	0	0	0	0	0	0	0	2.8	97.2	0	97.22
9	0	0	0	1.6	0	1.6	0	0	6.5	90.3	90.32

อัตราการรู้จำโดยเฉลี่ยเท่ากับ 95 เปอร์เซ็นต์

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ค

พจนานุกรมหน่วยเสียง

ก	ควอน	ชุม	ตุง	เน้น	ผัว	มัม	ลาม	โสม
ก็	ควัก	ชุ่ม	ตุ้ง	เน้น	ผัส	มัย	ล่าม	โสมนัส
กก	ควัน	ชู	ตุน	เนป	ผ้า	มัล	ลาย	ใส
กง	ควา	ชู้ด	ตุ้น	เนฟ	ผา	มัว	ลาว	ใส่
กฎ	คว้า	ชูต	ตุ๋น	เนม	ผ่า	มัส	ลำ	ไส้
กด	ควาก	เช	ตุร	เนย	ผาง	มัสการ	ล่ำ	ห
กน	ควาน	เชค	ตุส	เนล	ผาด	มัสยิด	ล้ำ	หก
กบ	ความ	เช็ค	ตู	เนส	ผ่าน	ม่า	ลิ	หง
กม	ควาย	เชง	ตู้	เนอ	ผาม	ม้า	ลิก	หงิก
กร	ควิ	เชด	ตูน	เนอส	ผิ	มา	ลิ่ง	หญ้า
กรก	ควิก	เช็ด	ตูม	เนา	ผิง	มาก	ลิง	หญิง
กรกฎา	ควิด	เช่น	เต	เน่า	ผิด	มาค	ลิช	หด
กรง	ควิท	เช่น	เตก	เนาะ	ผิน	มาช	ลิตร	หน
กรณี	ควิน	เชฟ	เต็ก	เนิด	ผิว	มาด	ลิน	หนด
กรด	ควิป	เชย	เตจ	เนิน	ผี	มาต	ลิ้น	หนวก
กรน	ควิส	เชล	เตด	เนีย	ผึ่ง	มาตรการ	ลิบ	หน่วง
กรบ	คอ	เชษ	เต็ด	เนียง	ผึ้ง	มาตรฐาน	ลิม	หน่วย
กรม	คอก	เชส	เตต	เนียน	ผุ	มาตรา	ลิล	หนอ
กรมหมื่น	ค็อก	เชอ	เต็ต	เนียบ	ผู้	มาท	ลิล	หนอง
กรมหลวง	คอง	เช่า	เตท	เนียม	ผูก	มาน	ลิว	หน่อย
กรร	คองเกรส	เช้า	เตน	เนียล	เผ	ม่าน	ลิส	หนัก
กรรม	คอด	ชำ	เต็น	เนียว	เผชิญ	มาบ	ลี่	หนัง
กรรมการ	คอต	เชิง	เต้น	เนียส	เผด็จ	มาย	ลี	หน้า
กรรมา	ค็อต	เชิญ	เตน	เนื้อ	เผย	มารค	ลี้	หนา
กรอ	คอท	เชิด	เต็ป	เนื่อง	เผลอ	มารยาท	ลีก	หน่า
กร้อ	คอน	เชิ้ต	เตม	เนือง	เผ่า	มาล	ลีด	หนาด
กรอง	ค่อน	เชีย	เต็ม	แน	เผา	มาส	ลีท	หนาม
กร่อน	คอบ	เชียง	เตย	แน่	เผื่อ	ม่ำ	ลีน	หน่าย
กรอบ	คอม	เชียน	เตร	แนก	แผ่	มิ	ลีบ	หน่ำ
กรอย	คอย	เชียม	เตรต	แนค	แผน	มิก	ลีส	หนำ
กร่อย	ค่อย	เชียร	เตรท	แนช	แผ่น	มิค	ลึก	หนิ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

กรอส คอล เชียล เตรส แนน แผนก มิง ลืม หนี้
กระ คอส เชี่ยล เตรีย แน่น แผล มิ้ง ลื่อ หนี
กรัง คะ เชี่ยว เตรียม แนฟ แผ้ว มิ่ง ลื้อ หนึ่ง
กรัม คั่ง เชียส เตล แนล โผน มิจ ลือ หนุ
กรัล คัญ เชื้อ เต็ล แนว โผล่ มิช ลุ หนุน
กรา คัด เชื่อ เตส แน่ว ไผ่ มิท ลุก หนุ่ม
กราน คัน เชื่อม เตอ แนส ฝ มิน ลุง หนู
กราบ คั้น แช เต่า แนะ ฝน มิป ลุด หมด
กราฟ คับ แช่ เต้า โน ฝรั่ง มิร ลุม หมวก
กราม คัม แชง เติก โน่ ฝอย มิล ลุ่ม หมวด
กราว คัล แช่ง เติบ โนด ฝะ มิว ลุย หมอ
กราส คัว แชน เติม โน้ต ฝัก มิส ลุส หม้อ
กริ คัส แชม เติล โนน ฝัง มี ลู หมอก
กริด คา แช่ม เติ้ล โน่น ฝั่ง มี่ ลู่ หมอน
กริล ค่า แชล เตีย โน้น ฝัน มีด ลูก หม่อน
กรี ค้า แชส เตียง โน้ม ฝา มีท ลูด หมอบ
กรีด ค่าง โช เตียน โนส ฝ่า มีน ลูน หม่อม
กรีน ค้าง โชก เตือน ใน ฝ้า มีล ลูน หมัก
กรีม คาด โช้ค แต ไน ฝาก มืด ลูป หมัน
กรุ คาน โชค แต่ ไนท ฝ่าย มือ เล่ หมั้น
กรุง ค้าน โชติ แต้ บ ฝ้าย มุ เล หมา
กรุณา คาบ ใช่ แตก บก ฝาย มุก เลก หมาก
กรุ่น คาพ ใช้ แต่ง บง ฝั่น มุ่ง เล็ก หมาย
กรุ๊ป คาม ไช แตง บ่ง ฝี มุง เล็ก หมิ่น
กรุย คาย ซ แตน บด ฝึก มุติ เลข หมี
กรู ค่าย ซอ แต้น บท ฝุ่น มุน เลค หมี่
กรูด คาร ซ็อก แตน บน ฝูง มุม เลง หมึก
กรู๊ป คารวะ ซอง แตป บ่น เฝ้า มุส เล็ง หมื่น
กฤติ คาล ซ่อง แตร บบ เฝือ มุสลิม เลจ หมุน
กฤษ คาว ซอน แตะ บม เฝือก มู เลช หมู
กฤษฎี ค่าว ซ่อน โต บ่ม แฝก มู้ด เลิน หมู่
กล ค้าว ซอบ โต้ บรม แฝง มูน เลด หยด
กลม คาส ซอฟ โตก บรร แฝด มูฟ เลต หยวน
กลยุท คำ ซอม โตน บรรพ ใฝ่ เม เล็ต หย่อน

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

กล้วย
กลวย
กลวิ
กลอง
กล่อง
กล้อง
กลอน
กล่อม
กลั่น
กลับ
กลับ
กลัว
กลา
กล้า
กลาง
กล้าม
กลาย
กล่าว
กลาส
กลาโหม
กลี
กลึง
กลืน
กลุ่ม
กลุ้ม
กลู
กวด
ก๊วน
กวม
กวย
กวัย
กวา
กว่า
กวาง

คิ
คิง
คิด
คิน
คิว
คิ้ว
คิส
คี
คีช
คีด
คีต
คีน
คึก
คืน
คืบ
คือ
คุ
คุก
คุณ
คุณธรรม
คุณประโยช
คุณภาพ
คุณลักษณะ
คุณวุฒิ
คุณสม
คุณูป
คุต
คุน
คุ้น
คุม
คุ้ม
คุย
คุ้ย
คู

ซ้อม
ซ่อม
ซอย
ซอล
ซอส
ซะ
ซัก
ซัค
ซัต
ซัน
ซับ
ซัม
ซั่ม
ซัล
ซัส
ซ่า
ซา
ซาก
ซาด
ซาต
ซาน
ซาบ
ซ้าย
ซาล
ซาว
ซ้ำ
ซ้ำ
ซิ
ซิก
ซิง
ซิ่ง
ซิ้ง
ซิท
ซิน

โตฟ
โตร
โตร
โตรก
โตส
โต๊ะ
ใต้
ไต
ไต้
ไต่
ไตค
ไตย
ไตร
ถ
ถก
ถด
ถน
ถนน
ถนอม
ถนัด
ถม
ถล่ม
ถลุง
ถ่วง
ถ้วน
ถ้วย
ถวาย
ถอด
ถอน
ถอย
ถ้อย
ถะ
ถัก
ถัง

บรอค
บรอด
บรอน
บรัช
บรั่น
บรัม
บรัส
บรา
บราว
บริ
บริการ
บริก
บริง
บริจ
บริด
บริว
บรี
บรุก
บรู
บรูม
บล็อก
บลัง
บลัฟ
บลาส
บลิ
บลิก
บลิว
บลี
บลู
บวก
บวง
บ่วง
บวช
บวน

พ
พก
พง
พด
พ่น
พน
พ้น
พบ
พร
พรม
พรรค
พรรณนา
พรรษา
พรหม
พรอ
พร็อก
พร่อง
พร่อง
พร็อบ
พร้อม
พร้อม
พระ
พรัก
พรา
พร่า
พร้า
พราก
พราน
พร้าว
พราหม
พริ
พริน
พริ้น
พริม

เม็ก
เมขลา
เมค
เมฆ
เม่ง
เมจ
เมด
เม็ด
เมต
เมตร
เมท
เม้น
เมน
เมม
เมล
เมล็ด
เมส
เมอ
เมา
เม้า
เมาะ
เมิฆ
เมิด
เมิน
เมี่ย
เมียน
เมี้ยน
เมียม
เมื่อ
เมือ
เมือง
เมื่อย
แม
แม้

เลต
เล็ท
เลท
เลน
เล่น
เล็บ
เลป
เลฟ
เลม
เล่ม
เลย
เลว
เลส
เลห
เลอ
เละ
เล่า
เลา
เลาะ
เลิก
เลิฟ
เลิศ
เลี่ย
เลี่ยง
เลียด
เลียน
เลี่ยน
เลียม
เลี่ยม
เลียว
เลี้ยว
เลียส
เลือก

หย่อม
หยัด
หยา
หยา
หย่า
หยาม
หยุ
หยุด
หยุ่น
หร
หรอ
หรับ
หรา
หราน
หร่าน
หริ
หรี่
หรีด
หรือ
หล
หลง
หลบ
หลวง
หลวม
หล่อ
หลอก
หลอน
หลอม
หลัก
หลัง
หลั่ง
หลับ
หล้า
หล้า

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

กว้าง คู่ ซิ่น ถัด บวร พริส แม็ก เลือด หลาก
กวาด คู้ ซิบ ถัน บอ พรี แมก เลื่อน หลาก
กว่าน คูน ซิม ถัม บ่อ พรีม แม็กซ เลือน หลาด
กว๊าน คูป ซิล ถั่ว บ็อก พรุ่ง แมค เลื่อม หลาด
กวิน คูล ซิส ถา บอก พรู แมง แล หลาน
กวี คูส ซี ถ้า บอง พรู้ฟ แมซ แล็ก หลาน
กษ เค ซี่ ถาก บอด พฤ แมด แลก หลาย
กห เค็ก ซี้ ถาง บอต พฤก แมต แลค หลาย
กอ เค้ก ซีด ถ่าง บอน พฤติ แม่น แลง หลำ
ก่อ เคท ซีน ถาน บ่อน พฤษภ แมน แล้ง หลิบ
ก้อ เคน ซีพ ถ่าน บ๊อบ พฤหัสบดี แม้น แลต หลี
กอก เค็น ซีล ถาบ บอบ พล แม็ป แลน หลีก
ก๊อก เค้น ซีส ถาม บอม พลบ แมล แล่น หลุด
กอง เคป ซี้ง ถ่าย บอย พลเมือง แมลง แลบ ห่วง
กอซ เค็ม ซี่ง ถ้ำ บ่อย พลเรือน แม้ว แล็ป ห้วง
ก็อด เคย ซึม ถิ บอล พลวัต แมส แลม หวง
กอต เคร ซื่อ ถิต บอส พลอ โม แลล หวน
ก็อต เคร่ง ซื้อ ถิ่น บะ พลอด โม่ แล้ว ห้วย
กอน เคร็ท ซุ ถี บัก พล็อต โมง และ หวะ
ก่อน เครน ซุง ถี่ บัง พล้อต โมช โล หวัง
ก้อน เครา ซุน ถึก บั้ง พลอย โมท โล่ หวัด
กอบ เคราะ ซุบ ถึง บัด พลัง โมทนา โลก หวัน
ก๊อป เครียด ซุ้ม ถือ บัตร พลัด โมน โลง หว่า
กอฟ เครือ ซุส ถุ บัท พลัน โมย โล่ง หว้า
กอย เครื่อง ซู ถุง บัน พลับ โมล โลน หว่าง
กอส เคล ซูน ถุน บั่น พลัส โมส โลป หวาด
กะ เคล็ด ซูป ถูก บับ พลา ไม่ โลม หวาน
กัก เคลม ซูม ถูป บัพ พลาด ไม โลส หวาย
กัง เคลา ซูล เถร บัฟ พลาย ไม้ โละ หวีด
กั้ง เคลิบ เซ เถอะ บัม พลาส ย ไล ห่อ
กัด เคลิ้ม เซก เถา บั้ม พลิ ยก ไล่ หอ
กัต เคลีย เซ็ก เถาะ บั๊ม พลิก ยง ว หอก
กัน เคลียส เซค เถิด บัล พลี ยน วก ห้อง
กั้น เคลื่อน เซ่ง เถียง บัว พลีส ยม วง หอบ

กับ เคลือบ เซด เถือ บัส พลู ยล วจ หอม
กัม เคว เซต เถื่อน บา พวก ยว วน หอย
กัล เควน เซ็ต แถบ บ่า พ่วง ยวง วร ห้อม
กัลยา เควล เซ็น แถม บ้า พวง ยวด วอ หะ
กั๊ว เควส เซ่น แถลง บาก พวน ยวน วอก หัก
กั่ว เคส เซ้น แถว บาง พ่อ ยศ วอด หัส
กัส เคอ เซน โถ บ้าง พอ ยอ วอต หา
กา เคา เซป โถง บาด พอก ยอก วอน ห้า
ก่า เค้า เซ็พ ท บ้าน พ้อง ย่อง วอฟ หาก
ก้า เคาะ เซฟ ทก บาน พอง ยอง วอม หาด
กาก เคิฟ เซม ทง บาบ พอต ยอช วอย หาน
กาง เคิล เซล ทด บ่าย พ็อต ยอด วอล ห้าม
กาช เคียง เซส ทน บาย พ็อป ย้อน วอลต หาย
ก๊าช เคียน เซอ ทบ บาร พอพ ย่อม วะ หาร
กาญจ เคียฟ เซา ทม บารมี พอย ย้อม วัก หิ
กาด เคียว เซาะ ทร บาล พอล ยอม วัง หิต
กาน เคี้ยว เซิน ทรง บ่าว พอส ย่อย วัฒนะ หุ
ก้าน แค เซิฟ ทรรศ บาส พะ ยะ วัฒนา หุ่ง
กาพ แค่ เซิล ทรรศการ บำ พัก ยัก วัด เหตุ
กาม แคก เซีย ทรรศนะ บิ พัง ยัง วัต เห็น
ก้าม แคช เซียท ทรวง บิก พัชร ยั้ง วัน เหน็จ
กาย แคด เซียน ทรอ บิ๊ก พัฒนา ยั่ง วัย เหมา
ก่าย แค็ต เซียบ ทร็อต บิง พัด ยัด วัล เหมาะ
กายภาพ แคน เซียม ทรอน บิด พัน ยัน วัลลภ เหมือง
การ แค่น เซียส ทรอม บิต พั้น ยับ วัว เหมือน
การบูร แค้น แซ ทรอย บิ้น พัพ ยัล วัสดุ เห็ด
การเล่น แคบ แซ่ ทรอส บิน พัม ยัว วา เหย
กาว แคป แซก ทระ บิล พัล ยั่ว ว่า เหยิง
ก้าว แคม แซ็ก ทรัป บิว พัสดุ ยา ว้า เหยียด
กาศ แค้ม แซง ทรัพ บิส พา ย่า วาก เหร่
กาศนียบัตร แคร็ก แซด ทรัม บี พาก ยาก วาง เหรด
กาส แครง แซท ทรัล บี้ พาด ยาง ว่าง เหรี่ยง
กำ แคล แซน ทรัส บี่ พาต ย่าง วาด เหรียญ
กำเนิด แคล็ก แซฟ ทรา บี้ พาท ยาด วาน เหล็ก

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

กิ แคลน แซม ทราน บีต พาธ ยาตรา ว่าน เหลว
กิ่ แคล้ว แซล ทราบ บีท พาน ยาน วาม เหล่า
กิ้ แคว โซ ทราม บีบ พาม ย่าน วาย เหลี่ยม
กิจ แคว้น โซ่ ทราย บึก พาย ยาม ว่าย เหลียว
กิณกะ แคส โซ่ง ทราส บึ้ง พ่าย ยาย วาล เหลือ
กิต โค โซน ทริ บึง พาส ยาท วาว เหลือก
กิน โคก โซม ทริก บึน พาหนะ ย้าย ว่าว เหลือง
กิ้น โค้ก โซล ทริน บือ พัว ยาล วาส เหลือบ
กิบ โค้ง โซส ทรี บุ พำ ยาว วิ เหลือม
กิ๊บ โคช ไซ ทรีท บุ๊ก พิ ย่ำ วิก เหลื่อม
กิล โค้ช ไซ้ ทรีส บุก พิก ย้ำ วิกรม เหว
กิว โคด ฌ ทรุด บุคลา พิง ยำ วิกฤต เหวก
กิส โค้ด ญ ทรู บุคลิก พิจารณา ยิ วิค เหวี่ยง
กี โคต ญวน ทฤษฎี บุ้ง พิณ ยิ่ง วิง เห่า
กี่ โคตร ญอล ทล บุ่ง พิด ยิด วิ่ง เหียน
กี้ โคน ญัก ทลาย บุช พิทยา ยิน วิช แห
กีด โค่น ญัต ท่วง บุด พิธ ยิป วิญ แห่
กีบ โคป ญัตติ ท้วง บุตร พิน ยิม วิด แหง
กีส โคม ญา ทวน บุตรี พิม ยิ้ม วิทยา แห่ง
กึง โคร ญาต ท่วม บุน พิว ยี่ วิทยุ แห้ง
กึ่ง โครก ญาน ทว่า บุฟ พิษ ยี วิน แหน
กึ๋น โครง ญี่ ทวาร บุร พี่ ยีน วิป แหน่ง
กุ โครน ญู ทวิ บุรุษ พี ยีส วิล แหม่ม
กุ๊ก โครม ฎ ทวิน บุล พี้ ยืด วิว แหล่
กุ้ง โครส ฎก ทวี บู่ พีช ยืน วิส แหล่ง
กุด โคล ฎา ทวีป บู้ พึง ยื่น วิสกี้ แหลม
กุน โคลง ฎาน ทไว บู พึ่ง ยืม วี่ แหละ
กุบ โคลน ฎี ทอ บูช พืช ยึด วี แหวน
กุม โคลส ฎ ท่อ บูม พื้น ยุ วีค โห
กุย โคว ฎิ ท้อ บูร พุ ยุก วีต โหน
กุล โควตา ฎิญ ทอก บูรณา พุง ยุค วีน โหนด
กุส โคส ฐ ท็อก บูล พุ่ง ยุคลบาท วุ โหม
กู โคร ฐา ทอค เบ พุช ยุ่ง วุฒิ โหลด
กู้ โคร่ ฐาน ทอง เบก พุด ยุง วุด โหว่

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

กู่	ไค	ฐิน	ท่อง	เบ็ก	พุ่ม	ยุด	วุธ	โหวก
กูด	ไคท	ฑ	ท้อง	เบค	พุม	ยุบ	วุ่น	โหวต
กู๊ด	ไคร	ฑิต	ทอช	เบด	พุล	ยุ่ย	วุล	ให้
กู๊ป	ไคร้	ฑุ	ทอด	เบ็ด	พู	ยู	วู	ใหญ่
กูร	ไคล	ฒ	ทอน	เบต	พู่	ยู่	วู้ด	ใหม่
กูล	ไคล้	ณ	ท่อน	เบท	พูด	ยูร	วูด	ใหม
เก	ไคว	ณรง	ท้อน	เบน	พูน	ยูส	เว	ไห
เก้	ศ	ณา	ทอป	เบร	พูล	เย	เวก	ไหน
เก๋	ฒ	ณี	ท็อป	เบรก	เพ	เย้	เวค	ไหม้
เกง	โฆษ	ณีต	ทอฟ	เบรน	เพก	เยน	เวช	ไหม
เก็ง	โฆษณา	ด	ทอม	เบรา	เพ็ก	เย่น	เวด	ไหล่
เก่ง	ง	ดก	ทอย	เบรียล	เพจ	เย็น	เวต	ไหล
เก๋ง	งก	ดง	ท็อย	เบล	เพชร	เยล	เวท	ไหว
เกจ	งง	ดน	ทอล	เบล็ม	เพ็ญ	เยส	เว็ท	ฟ
เกณ	งด	ดม	ทอ	เบลอ	เพด	เยอ	เว่น	อ
เกต	งบ	ดรรช	ทะ	เบลา	เพ็ด	เยอะ	เวน	อก
เก็ต	งม	ดรรชนี	ทัก	เบส	เพท	เยา	เว้น	อคติ
เกท	ง่วง	ดรอ	ทักษ	เบอ	เพ็ท	เย้า	เวบ	อง
เก็ท	งวด	ดรอก	ทั้ง	เบ้อ	เพน	เยิ่น	เว็บ	อด
เกน	งอ	ดร็อป	ทัง	เบอก	เพ่น	เยีย	เวฟ	อน
เก็น	งอย	ดรัก	ทั่ง	เบา	เพ้น	เยี่ยง	เวร	อเนก
เก็บ	งะ	ดรัม	ทัช	เบ้า	เพนียง	เยี่ยม	เวล	อบ
เกม	งัก	ดรา	ทัด	เบาะ	เพป	เยียว	เวส	อพยพ
เกร	งัด	ดราฟ	ทัน	เบิก	เพร	เยื้อ	เวอ	อม
เกรง	งัน	ดริ	ทับ	เบิค	เพรส	เยื่อ	เวอค	อย่า
เกรด	งับ	ดริก	ทัม	เบิด	เพร่า	เยือก	เวอส	อยาก
เกร็ด	งา	ดริค	ทัล	เบิต	เพร่า	เยือน	เวิค	อย่าง
เกรน	ง่า	ดริง	ทัว	เบิธ	เพราะ	แย่	เวิด	อยู่
เกรส	งาน	ดริด	ทั่ว	เบิน	เพรียง	แย	เวิต	อร
เกรอะ	งาม	ดรี	ทัศนะ	เบิ้ล	เพรียง	แย่ง	เวิท	อรรถ
เกราะ	งาย	ดรู	ทัส	เบิล	เพล	แย้ง	เวิบ	อรัญ
เกรียง	ง่าย	ดวก	ทา	เบีย	เพล็ก	แย็บ	เวิล	อริ
เกรียบ	งำ	ด้วง	ท่า	เบี้ย	เพลง	แยม	เวิส	อรุณ
เกล	งิ	ดวง	ท้า	เบี่ยง	เพลท	โย	เวีย	อวด

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

เกลอ	งุ้ม	ด่วน	ทาก	เบียด	เพลน	โยก	เวียง	อ้วน
เกลา	งู	ด้วย	ทาง	เบียน	เพลส	โยง	เวียด	อวน
เกล้า	เงย	ดวล	ทาน	เบียบ	เพลา	โยช	เวียน	อวย
เกลี่ย	เงาะ	ดวิก	ทาน	เบียม	เพลิง	โยน	เวียม	ออ
เกลือ	เงิน	ดอ	ท่าน	เบื่อ	เพลิด	ใย	แว	ออก
เกวียน	เงือน	ดอก	ทาย	เบื้อ	เพลิน	ร	แวะ	ออง
เกศ	เงื่อน	ด็อก	ท้าย	เบื้อง	เพลีย	รก	แวก	ออต
เกษตร	เงือน	ดอง	ทาล	เบือน	เพศ	รง	แวค	ออน
เกษตรกร	แง	ดอต	ทาว	แบ	เพส	รด	แว้ง	อ่อน
เกษตรกรรม	แง่	ด็อท	ท่าว	แบก	เพอ	รน	แวง	อ้อน
เกษตรศาสต	แงะ	ดอน	ท้าว	แบ็ก	เพา	รบ	แวด	ออพ
เกษม	ไง	ดอฟ	ทาส	แบค	เพิก	รม	แวต	ออม
เกษียณ	จ	ดอม	ทำ	แบ็ค	เพิ่ง	รวง	แวน	อ้อม
เกส	จก	ดอย	ทิ	แบง	เพิต	ร่วง	แว่น	ออย
เกอ	จง	ด้อย	ทิง	แบ่ง	เพิ่ม	รวจ	แวม	อ้อย
เกะ	จด	ดอล	ทิ้ง	แบด	เพิล	รวด	โว	ออส
เกา	จน	ดอส	ทิจ	แบ็ท	เพีย	รวน	โว่	อะ
เก่า	จบ	ดะ	ทิน	แบน	เพียง	ร่วน	โวค	อัก
เก้า	จม	ดัก	ทิป	แบบ	เพี้ยน	รวบ	โวฟ	อัง
เก่า	จร	ดัค	ทิม	แบ็พ	เพียบ	ร่วม	โวล	อั้ง
เกาะ	จรกา	ดัง	ทิล	แบร	เพียม	รวม	ไว	อัด
เกิด	จรส	ดั้ง	ทิว	แบร็ก	เพียว	รวย	ไว้	อัตรา
เกิต	จรัส	ดัช	ทิส	แบล	เพื่อ	รอ	ศ	อัธยา
เกิน	จรา	ดัชนี	ที	แบ็ลก	เพื่อน	ร่อ	ศรัท	อั้น
เกิล	จราย	ดัด	ที่	แบล็ค	แพ้	รอค	ศรี	อัน
เกีย	จริ	ดัต	ทีน	โบ้	แพ	รอง	ศว	อันตราย
เกียน	จริง	ดัน	ทีป	โบ	แพก	ร่อง	ศัก	อับ
เกียรติ	จ้วง	ดับ	ทีพ	โบ๊ก	แพ็ก	ร้อง	ศักดิ์	อัพ
เกียว	จวน	ดัม	ทีฟ	โบ๊ต	แพค	รอด	ศักราช	อัม
เกี่ยว	จวบ	ดัล	ทีม	โบน	แพ็ค	ร็อต	ศัตรู	อัย
เกี้ยว	จอ	ดัส	ทึก	โบร	แพง	รอน	ศาส	อัล
เกือ	จอก	ดา	ทึม	โบรก	แพ่ง	ร่อน	ศาสตร์	อัว
เกื้อ	จ็อก	ด้า	ทุ	โบล	แพด	ร้อน	ศาสดา	อัศจรร
เกือบ	จ๊อก	ด่า	ทุก	โบส	แพ็ท	รอบ	ศาสนา	อัส

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

แก	จอง	ดาก	ทุ่ง	ใบ	แพท	รอม	ศิ	อา
แก่	จอจ	ด่าง	ทุจริต	ไบ	แพน	รอย	ศิล	อ่า
แก้	จอด	ดาด	ทุน	ไบน	แพม	ร่อย	ศิล	อาก
แก่ง	จอน	ดาต	ทุ่น	ไบร	แพร	ร้อย	ศิลป	อาค
แก๊ง	จ็อบ	ดาน	ทุบ	ป	แพร่	รอล	ศิษ	อาคเน
แก๊ซ	จอม	ด้าน	ทุ้ม	ปก	แพรค	ระ	ศีล	อาง
แกต	จอย	ด่าน	ทุม	ปด	แพร่ง	รัก	ศูน	อ่าง
แกน	จะ	ดาบ	ทุ่ม	ปน	แพล	รัง	เศวต	อ้าง
แก่น	จัก	ดาม	ทุษ	ป่น	แพลง	รั้ง	เศษ	อาช
แก๊ป	จักรี	ดาล	ทู	ปบ	แพลน	รัด	โศก	อาชญา
แกม	จัง	ดาว	ทูต	ปร	แพ้ว	รัต	โศภ	อาด
แก้ม	จัด	ด้าว	ทูธ	ปรน	แพส	รัน	โศลก	อาต
แกร	จั๊ต	ดาษ	ทูน	ปรอง	โพ	รับ	ไศล	อ่าน
แกร่ง	จัตตุรัส	ดาส	ทูม	ประ	โพด	รัม	ษ	อาน
แกรม	จัตวา	ดำ	ทูล	ปรัก	โพน	รั่ม	ษะ	อาบ
แกล	จัน	ดำรัส	เท	ปรัง	โพ้น	รัล	ษัตรี	อาฟ
แก้ว	จั่น	ดำริ	เทก	ปรัชญา	โพย	รั่ว	ษา	อาม
แกส	จับ	ดิ	เท็ก	ปรับ	โพร	รั้ว	ษิต	อาย
แก๊ส	จัม	ดิก	เทค	ปรัส	โพรง	รัส	ส	อ้าย
แกะ	จัส	ดิ้ง	เทง	ปรา	โพรช	รา	สก	อารยะ
โก	จา	ดิ่ง	เทจ	ปราก	โพล	ร่า	สกปรก	อาว
โก้	จ้า	ดิช	เท็จ	ปราง	โพส	راก	สกฤต	อ่าว
โก๋	จาก	ดิน	เทต	ปราช	ไพ	ราง	สกล	อ้าว
โก่ง	จ่าง	ดิ้น	เทน	ปราณ	ไพน	ร่าง	สเคป	อำ
โกง	จ้าง	ดิบ	เท็น	ปราบ	ไพร	ร้าง	สง	อิ
โกณ	จาง	ดิพ	เทป	ปราม	ไพร่	راจ	ส่ง	อิก
โกด	จาน	ดิว	เทพ	ปราย	ไพร้	راด	สงบ	อิง
โกป	จาม	ดิษ	เทม	ปรารถนา	ฟ	ران	สงวน	อิจ
โกฟ	จาย	ดิษฐาน	เท็ม	ปราศ	ฟก	ร้าน	สงัด	อิน
โกร	จ่าย	ดิส	เทร	ปราศจาก	ฟรอซ	راบ	สด	อิป
โกล	จาร	ดิต	เทรด	ปริ	ฟรอน	ราฟ	สตรี	อิม
โกศ	จำ	ดี	เทรน	ปริง	ฟร้อน	ราม	สถาปนา	อิ่ม
โกส	จ้ำ	ดี้	เทรียม	ปริเฉท	ฟรัก	รามเกียรต	สถาปนิก	อิล
โก๊ะ	จิ	ดีด	เทรียล	ปริญ	ฟรัง	ร้าย	สถาปัตยกรร	อิว

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ใกล้ จิก ดีน เทล ปริน ฟรา ราย สน อิส
ไก จิ๊ก ดีพ เทว ปริ้น ฟราน ราล สนอง อิสระ
ไก่ จิค ดีม เทวดา ปริ๊น ฟราย ราว สนั่น อี้
ไก๊ จิง ดีล เทศ ปริบ ฟริ ร้าว สนับ อี
ไกร จิต ดีส เทศนา ปริบท ฟริซ ราษฎ สนาน อีก
ไกล จิน ดึก เทส ปริมาณ ฟริส ราษฎร สนาม อีน
ไกล่ จิ้น ดึง เทอ ปรี ฟรี ราส สนิก อีส
ข จิฟ ดื่น เทอน ปรึก ฟรีด รำ สนิก อึ
ขง จิม ดื่ม เทอม ปรุ ฟรีส ร่ำ สนุก อึ้ง
ขณะ จิ้ม ดื้อ เทา ปรุง ฟรุต ริ สนุน อึด
ขด จิว ดุ เท่า ปรู๊ฟ ฟลอ ริก สบ อื่น
ขน จิ๋ว ดุก เท้า ปลด ฟลัก ริง สเปน อุ
ข้น จิส ดุง เทิง ปลั้น ฟลิค ริ่ง สม อุด
ขนบ จี ดุล เทิน ปลอด ฟลิน ริช ส้ม อุทยาน
ขนม จี้ ดุส เทีย ปลอท ฟลุก ริน สมการ อุ่น
ขนอน จี๋ ดู เทียน ปลอม ฟลุ๊ค ริบ สมร อุป
ขนาด จีน ดูก เทียบ ปล่อย ฟลู ริม สมอ อุปกร
ขนาน จีฟ ดูด เทียม ปละ ฟอ ริมทรัพย สมอง อุปถัม
ขบ จึง ดูล เที่ยว ปลั๊ก ฟอก ริล สมัคร อุปโภค
ขม จืด เด เทือก ปลัด ฟอง ริ้ว สมัย อุปสรรค
ข่ม จุ เดก เทือน ปลา ฟอซ ริส สมาน อุ้ม
ขมุ จุก เด็ก แท ปลาก ฟอด รี สม่ำ อุส
ขยะ จุด เดค แท้ ปลาด ฟอต รี่ สมุด อู
ขยับ จุน เด็จ แทก ปลาบ ฟ็อต รีช สมุน อู่
ขยาย จุบ เดช แท็ก ปลาย ฟอท รีด สยาม อู้
ขระ จุ๊บ เดด แทค ปลาส ฟอน รีน สรรเสริญ เอ
ขรุ จุ่ม เดต แท็ค ปลิ ฟ้อน รีบ สรวง เอ็ก
ขรุขระ จุ้ย เดท แท็ง ปลิง ฟอม รีฟ สรวล เอก
ขลา จุลจอม เดน แท้ง ปลิต ฟอย รีม สร้อย เอกฉัน
ขวบ จู เด้น แท่ง ปลีก ฟอล รีล สระ เอกชน
ขวัญ จู๋ เด็น แทน ปลื้ม ฟอส รีส สร้าง เอกเทศ
ขวาก จูง เด่น แท่น ปลุก ฟะ รึก สราญ เอกภาพ
ขวาง จูด เดนมาก แทป ปวง ฟัก รื่น สรุป เอกราช

ขอ	จูน	เดม	แท็ป	ปวด	ฟัง	รื้อ	สล็อต	เอกลัก
ข้อ	จูล	เด็ด	แทร	ป่วน	ฟัน	รุ	สละ	เอกสาร
ข็อก	เจ	เด็ม	แทร็ก	ป่วย	ฟัล	รุก	สลัก	เอกสิทธ
ของ	เจก	เดร	แทรก	ปอ	ฟ่า	รุง	สลัด	เอง
ข้อง	เจ็ก	ดรีม	แทรค	ปอก	ฟ้า	รุ่ง	สลับ	เอจ
ขอน	เจ็ก	เดรต	แทรน	ปอง	ฟ่า	รุณกรรม	สลัม	เอซ
ข้อน	เจค	เดรน	แทล	ป้อง	ฟ่าง	รุด	สลาก	เอด
ขอบ	เจ็ค	เดรส	แทส	ป๋อง	ฟ่าง	รุ่น	สลาม	เอ็ด
ขอม	เจ็ด	เดรีย	โท	ปอต	ฟ่าน	รุน	สลาย	เอท
ข่อย	เจต	เดรียน	โทด	ปอท	ฟ่าม	รุ่ม	สลิง	เอ็ท
ขะ	เจตนา	เดรียส	โทน	ปอน	ฟ่าย	รุส	สลิด	เอ็น
ขัง	เจ็ท	เดล	โทป	ปอน	ฟ่าว	รู้	สลิป	เอน
ขัด	เจน	เดส	โทร	ป้อน	ฟ่าส	รู	สลิม	เอป
ขัน	เจ็บ	เดอ	โทรม	ป็อป	ฟิ	รูจ	สลุต	เอม
ขั้น	เจม	เดอะ	โทรล	ปอม	ฟิก	รูน	สวด	เอ็ม
ขับ	เจร	เดิน	โทรส	ปอย	ฟิง	รูป	สวน	เอ่ย
ขัย	เจรจา	เดิม	โทล	ปะ	ฟิช	รูม	ส่วน	เอล
ขัว	เจริญ	เดิล	โทษ	ป๊ะ	ฟิต	รูส	สวบ	เอว
ขา	เจล	เดีย	โทส	ปัก	ฟิน	เร	สวม	เอส
ข่า	เจส	เดียท	ไท	ปัง	ฟิบ	เร่	ส่วย	เออ
ข้า	เจอ	เดียน	ไทย	ปัด	ฟิม	เรก	สวย	เอ่อ
ขาด	เจ๋อ	เดียม	ไทร	ปัต	ฟิล	เร็ก	สวรรค	เออะ
ขาน	เจา	เดียล	ธ	ปัน	ฟิว	เรค	สวัน	เอ๊ะ
ขาม	เจ้า	เดี๋ยว	ธง	ปั่น	ฟิศ	เร็ค	สวัส	เอะ
ข้าม	เจาะ	เดียว	ธน	ปั้น	ฟี	เรง	สวา	เอา
ขาย	เจิน	เดี่ยว	ธม	ปับ	ฟี่	เร็ง	สว่าง	เอ้า
ข่าย	เจิม	เดื่อ	ธร	ป๊บ	ฟีช	เร่ง	สวาย	เอ๊า
ขาว	เจีย	เดื่อง	ธรณี	ปัป	ฟีด	เรจ	สวิง	เอิก
ข่าว	เจี๊ยก	เดือด	ธรรม	ปั๊ม	ฟีน	เรด	สวิช	เอิด
ข้าว	เจียน	เดือน	ธรรมชาติ	ปัว	ฟีล	เรต	สวี	เอิท
ขำ	เจือ	แด่	ธอ	ปัส	ฟืน	เรท	สห	เอิน
ขิง	แจ	แดง	ธอด	ป่า	ฟื้น	เรล	สหรัฐ	เอีย
ขิต	แจก	แดด	ธอน	ป๋า	ฟุ	เรน	สอ	เอียง
ขิม	แจ็ก	แดน	ธะ	ป้า	ฟุ้ง	เร็น	สอง	เอียด

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ขี่	แจ๊ค	แดป	ธัน	ป่า	ฟุต	เร้น	ส่อง	เอียน
ขี้	แจง	แดม	ธา	ปาก	ฟุ่ม	เร็นต	สอด	เอียว
ขีด	แจ้ง	แดล	ธาน	ปาค	ฟุล	เรฟ	สอต	เอื้อ
ขี้น	แจ่ม	โด	ธนาคาร	ปาง	ฟุส	เร็ม	สอน	เอื่อย
ขี้น	แจ๋ว	โดด	ธานินทร์	ปาด	ฟู	เรล	สอบ	แอ
ขืน	โจ	โดน	ธาร	ปาน	ฟูก	เร็ว	สอย	แอก
ขึ่น	โจ๊ก	โดป	ธารณะ	ป่าน	ฟูด	เรส	สะ	แอค
ขื่อ	โจท	โดม	ธิ	ปาม	ฟู้ด	เรอ	สัก	แอง
ขุด	โจน	โดย	ธิค	ปาย	ฟูม	เร่อ	สักวา	แอ็ง
ขุน	โจม	โดร	ธิบ	ป่าย	ฟูล	เรา	สั่ง	แอด
ขุม	ใจ	โดรม	ธิม	ป้าย	เฟ	เร้า	สัง	แอต
ขุย	ไจ	โดว	ธี	ปาล	เฟ่	เราะ	สังคายนา	แอท
ขู่	ฉ	โดส	ธุรกิจ	ปาว	เฟก	เริง	สัชนา	แอน
เข	ฉก	ใด	เธ	ปาส	เฟ็ก	เริญ	สัญลักษ	แอบ
เข้	ฉบับ	ได	เธค	ปิ	เฟค	เริน	สัด	แอป
เขต	ฉลอง	ได้	เธอ	ปิก	เฟซ	เริบ	สัน	แอฟ
เขน	ฉลาก	ไดฟ	เธีย	ปิง	เฟด	เริ่ม	สั้น	แอม
เข็น	ฉวย	ไดร	น	ปิ้ง	เฟต	เรีย	สับ	แอล
เขป	ฉ้อ	ไดส	นก	ปิด	เฟท	เรียก	สัม	แอส
เข็ม	ฉ่อง	ต	นข	ปิน	เฟน	เรียง	สัมมนา	แอะ
เข้ม	ฉะ	ตก	นครบาล	ปี	เฟม	เรียด	สา	โอ
เขมร	ฉัตร	ตง	นครหลวง	ปี่	เฟร	เรียน	ส่า	โอ้
เขมือบ	ฉัน	ตด	นครินท	ปี้	เฟรช	เรียบ	สาก	โอ๊ก
เขา	ฉับ	ตน	นง	ปีก	เฟร็น	เรียม	สาง	โอก
เข้า	ฉัย	ต๋น	นด	ปีช	เฟรม	เรียล	สาด	โอง
เขิน	ฉา	ต้น	นบ	ปีด	เฟรส	เรียว	สาท	โอ่ง
เขียน	ฉาก	ตบ	นม	ปีน	เฟรา	เรียส	สาธารณรัฐ	โอด
เขียว	ฉาง	ต้ม	นรก	ปีบ	เฟล	เรื่อ	สาน	โอ๊ต
เขี้ยว	ฉาน	ตม	นรา	ปืก	เฟว	เรื้อ	สาบ	โอน
เขือ	ฉาย	ตร	นรินท	ปืน	เฟส	เรือ	สาม	โอม
เขื่อง	ฉิ่	ตรง	นวด	ปุ	เฟอ	เรือก	สามเณร	โอล
เขื่อน	ฉีด	ตรวจ	นวน	ปุจ	เฟ้อ	เรื่อง	สามารถ	ไอ
แข	ฉุก	ตรอ	น่วน	ปุ่น	เฟิท	เรือง	สาย	ไอซ
แขก	ฉุด	ตรอน	นวย	ปุ้ย	เฟิน	เรื้อน	สารคดี	ไอศ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

แข็ง เฉ ตรอม นวล ปุย เฟิม เรือน สารบัญ ฮ
แข่ง เฉก ตระ นส ปุ้ย เฟิล แร สารพัด ฮวน
แขนง เฉพาะ ตรัค นอ ปุส เฟิส แร่ สารภี ฮอ
แขม เฉย ตรัง นอก ปู เฟีย แร็ก สารรูป ฮ่อ
แขวง เฉลิม ตรัต น็อก ปู่ เฟื่อ แรก สาว ฮอก
แขวน เฉลี่ย ตรัม น็อค ปูน เฟื้อ แรค สำ ฮอค
โข เฉียง ตรัย น่อง ปูม เฟือย แรง สำรวจ ฮ็อค
โขนง เฉี่ยว ตรัส นอง เป แฟ แร้ง สำเร็จ ฮ่อง
โขลน โฉด ตรา น้อง เป้ แฟก แร็ด สิ ฮอด
ไข โฉนด ตร้า นอต เป็ก แฟ็ก แรต สิก ฮ็อต
ไข่ โฉม ตราย น็อต เปก แฟค แรน สิค ฮอท
ไข้ ช ตราส นอท เปช แฟต แรฟ สิ่ง ฮอน
ฃ ชง ตราบ นอน เปญ แฟน แรม สิง ฮอฟ
ค ชด ตริ นอฟ เป็ด แฟ้น แรล สิต ฮอย
คก ชน ตริก นอม เปต แฟร แรส สิน ฮอยซ
คง ชนนี ตริง นอย เป็ต แฟรก แระ สิ้น ฮอล
คด ชนบท ตริป น้อย เปน แฟรง โร สิบ ฮะ
คน ชม ตรี นอล เป็น แฟรน โร่ สี ฮัก
ค้น ชร ตรีท นะ เป็ป แฟรม โร้ค สี่ ฮัง
คบ ชรา ตรีม นัก เปร แฟล โรค สีด ฮัท
คม ช่วง ตรึง นักษัตร เปรม แฟลช โรง สึก ฮัน
คมนา ชวน ตลก นัง เปรส แฟลต โรช สืบ ฮับ
คร ช่วย ตลอด นั่ง เปรอะ โฟ โรน สือ ฮัม
ครก ชวล ตลาด นัด เปรียญ โฟค โรป สื่อ ฮัล
ครบ ชวัน ตลิ่ง นัน เปรียบ โฟน โรม สุ ฮ่า
ครร ชวา ตอ นั่น เปรี้ยว โฟม โรย สุก ฮา
ครอง ชวา ต่อ นั้น เปรียว โฟร โรล สุจริต ฮาด
ครอน ชอ ต๊อก นับ เปล โฟล โรส สุด ฮาท
ครอบ ช่อ ต๊อก นั่ม เปล่ง ไฟ ไร สุน ฮาน
ครอฟ ชอก ตอก นัม เปล่า ไฟล ไร่ สุนทรี ฮาพ
ครอส ช็อก ต็อค นัล เปลี่ยน ภ ไร้ สุ่ม ฮิ
คระ ช็อค ตอง นัส เปลือก ภา ฤ สุม ฮิท
ครั้ง ชอง ต้อง นา เปลือง ภรรยา ฤก สุล ฮิน
ครั่ง ช่อง ตอน น่า เปลื้อง ภะ ฤกษ์ สุว ฮิบ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ครัน	ช้อง	ต้อน	น้า	เปลือย	ภัย	ฤทธิ์	สู่	ฮิล
ครับ	ชอต	ตอบ	นาก	เปส	ภัณฑ์	ล	สู้	ฮิว
ครัม	ชอท	ต่อม	นาง	เปอ	ภัทร	ลก	สู	ฮิส
ครัว	ช็อท	ต่อย	นาฏศิล	เปอม	ภาชนะ	ลง	สูง	ฮี
ครา	ช้อน	ตอล	นาด	เป้า	ภาพ	ลด	สูต	ฮี้
คร่า	ชอบ	ตะ	นาน	เปิด	ภิก	ลน	สูตร	ฮีท
คราบ	ชอป	ตั๊ก	น่าน	เปิท	ภู	ล้น	สูติ	ฮีล
คราฟ	ช็อป	ตั๊กแตน	นาบ	เปิม	ภูม	ลบ	สูท	ฮุ
คราม	ช็อย	ตัง	นาฟ	เปิล	เภอ	ล้ม	สูบ	ฮุก
คราว	ช้อย	ตั่ง	นาม	เปิ้ล	เภา	ลม	เส	ฮุน
คร่าว	ชะ	ตั้ง	นาย	เปิส	โภค	ล่ม	เสก	ฮู
คริก	ชัก	ตัด	น่าย	เปีย	โภคภัน	ล่วง	เสจ	ฮูด
คริป	ชัง	ตั๊น	นาว	เปียง	โภชนา	ลวง	เสด	ฮูพ
คริส	ช้ชวาล	ตัน	น้าว	เปียน	ม	ล้วน	เสด็จ	เฮ
ครี	ชัด	ตับ	นาส	เปี้ยน	มก	ลอ	เสถียร	เฮก
ครีบ	ชั่น	ตัฟ	น้ำ	เปี่ยม	มกราคม	ล่อ	เส็ท	เฮด
ครีม	ชัน	ตั๊ฟ	นำ	แป	มง	ล้อ	เสน	เฮน
ครึ	ชั้น	ตัม	นิ	แป้ง	มงคลสมัย	ล็อก	เสน่ห์	เฮฟ
ครึก	ชันษา	ตัว	นิก	แปด	มด	ลอก	เส้น	เฮม
ครึ่ง	ชับ	ตั๋ว	นิค	แปน	มน	ล็อค	เสนอ	เฮล
ครื้ม	ชัย	ตัส	นิง	แป๊บ	ม่น	ลอง	เสบ	เฮส
ครื้น	ชัยภูมิ	ตา	นิ่ง	แปร	มรกต	ลอจ	เสม็ด	เฮอ
ครุ	ชัยมง	ต้า	นิช	แปล	มรดก	ลอด	เสมอ	เฮา
ครู	ชัว	ตาก	นิด	แปลก	มรรถ	ลอต	เสมียน	เฮ้า
ครู่	ชั่ว	ต่าง	นิน	แปลง	มรสุม	ล็อต	เสมือน	เฮิซ
ครูก	ชา	ตาด	นิป	แปลน	มฤดก	ลอน	เสร็จ	เฮิร
คล	ช่า	ตาต	นิม	แปลช	มลา	ล็อบ	เสริม	เฮิรต
คลอ	ช้า	ตาท	นิล	โป	มวก	ลอบ	เสล	เฮีย
คลอง	ช้าง	ตาน	นิว	โป่ง	มวก	ลอฟ	เสลา	เฮือน
คล่อง	ชาง	ต่าน	นิ้ว	โป๊ป	ม่วง	ลอม	เสวนา	แฮ
คล้อง	ช่าง	ต้าน	นิส	โปร	มวย	ล้อม	เสวย	แฮ็ง
คลอด	ชาจ	ตาบ	นี	โปร่ง	มวล	ลอย	เสา	แฮน
คลอย	ชาด	ตาฟ	นี่	โปรด	มหกรรม	ลอส	เส้า	แฮฟ
คลัง	ชาต	ตาม	นี้	โปล	มหรสพ	ลออ	เสิจ	แฮม

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

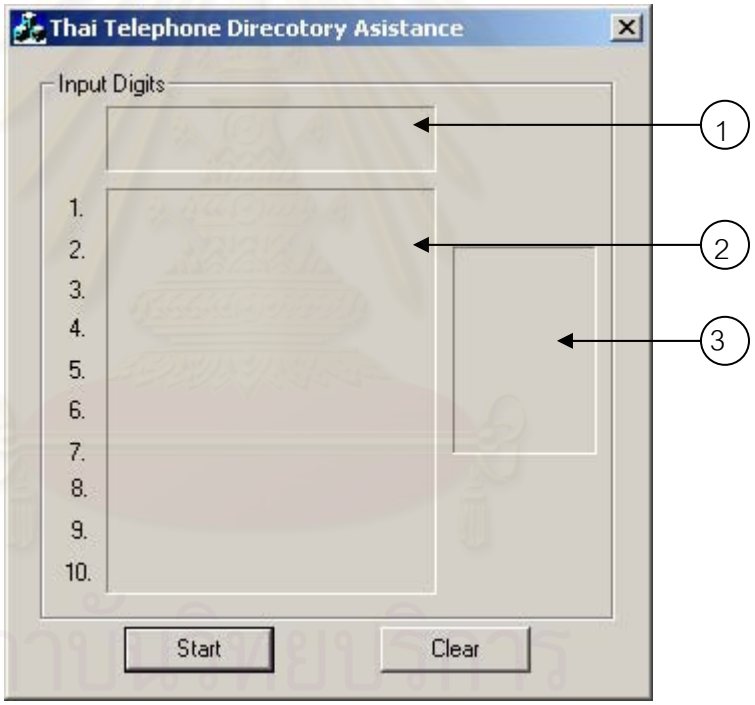
คลัต	ชาน	ตาย	นีด	โปส	มหัศจรร	ละ	เสิช	แฮล
คลับ	ชาป	ต่าย	นีย	ไป	มหาตมะ	ละคร	เสิด	แฮส
คลา	ชาย	ตาล	นีล	ไปร	มหาสารคราม	ล่ะ	เสิต	โฮ
คล้า	ชาว	ตาว	นีส	ไปรษณี	มหาสมุทร	ลัก	เสิฟ	โฮม
คลาก	ชิ	ตำ	นึก	ผ	มอ	ลักษณะ	เสิรฟ	โฮล
คลาด	ชิก	ต่ำ	นึ่ง	ผก	มอค	ลัง	เสี่ย	โฮลต
คลาย	ชิ่ง	ตำรวจ	นึง	ผง	มอง	ลัด	เสีย	โฮส
คล้าย	ชิง	ติ	นุ	ผนวก	มอญ	ลัท	เสียง	โฮป
คลาส	ชิด	ติก	นุก	ผนวช	มอท	ลัน	เสี่ยง	ไฮ
คลิ	ชิน	ติง	นุ่ง	ผนัง	มอน	ลับ	เสียด	ไฮส
คลิก	ชิ้น	ติ้ง	นุช	ผนึก	ม็อบ	ลัม	เสียม	
คลิง	ชิป	ติณ	นุ่น	ผม	มอบ	ลัย	เสียว	
คลิด	ชิม	ติด	นุ่ม	ผล	มอย	ล้วะ	เสี่ยว	
คลิน	ชิล	ติน	นุม	ผลง	ม่อย	ลัส	เสี้ยว	
คลิป	ชี	ติม	นู	ผลอ	ม้อย	ล่า	เสื่อ	
คลิม	ชี้	ติล	นูบ	ผลัก	มอล	ลา	เสื้อ	
คลี	ชีค	ติว	นูล	ผลัด	มอส	ล้า	เสือ	
คลี่	ชีต	ติส	เน	ผลาญ	มะ	ลาก	เสื่อม	
คลีก	ชีน	ตี้	เนก	ผลิต	มะ	ลาง	แส	
คลีน	ชีพ	ตี	เน็ก	ผลู	มัก	ล่าง	แสก	
คลึง	ชีฟ	ตีน	เนค	ผอง	มัค	ล้าง	แสง	
คลื่น	ชีส	ตึก	เนจ	ผ่อน	มัง	ลาจ	แสดง	
คลุม	ชื่น	ตึง	เนต	ผอม	มั่ง	ลาด	แสตม	
ควง	ชื้น	ตื่น	เน็ต	ผะ	มัช	ลาน	แสน	
ควบ	ชื่อ	ตื้น	เนตร	ผัก	มัด	ล่าน	แสลง	
ควร	ชุ	ตุ	เนท	ผัง	มัธยม	ล้าน	แสวง	
ควอ	ชุด	ตุก	เน็ท	ผัน	มัน	ลาบ	โส	
ควอต	ชุน	ตุ๊กตา	เนน	ผับ	มั่น	ลาฟ	โสด	

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ง
ส่วนอธิบายโปรแกรม

ระบบสอบถามรายนามผู้ใช้โทรศัพท์แบบอัตโนมัติที่นำเสนอในงานวิจัยนี้เป็นต้นแบบ (Prototype) ของระบบ โดยทำการจำลองระบบลงบนเครื่องไมโครคอมพิวเตอร์ รับเสียงข้อมูลขาเข้าโดยผ่านทางไมโครโฟนและส่งเสียงข้อมูลขาออกโดยผ่านทางลำโพง โดยลักษณะการทำงานจริงจะเป็นรูปแบบที่ผ่านทางสายโทรศัพท์ผู้ใช้จะติดต่อผ่านระบบโดยใช้เสียงเท่านั้น แต่เนื่องจากต้องการให้สามารถตรวจสอบความถูกต้องได้ง่ายขึ้นจึงออกแบบให้สามารถแสดงผลออกทางหน้าจอโดยแสดงเป็นหมายเลขโทรศัพท์

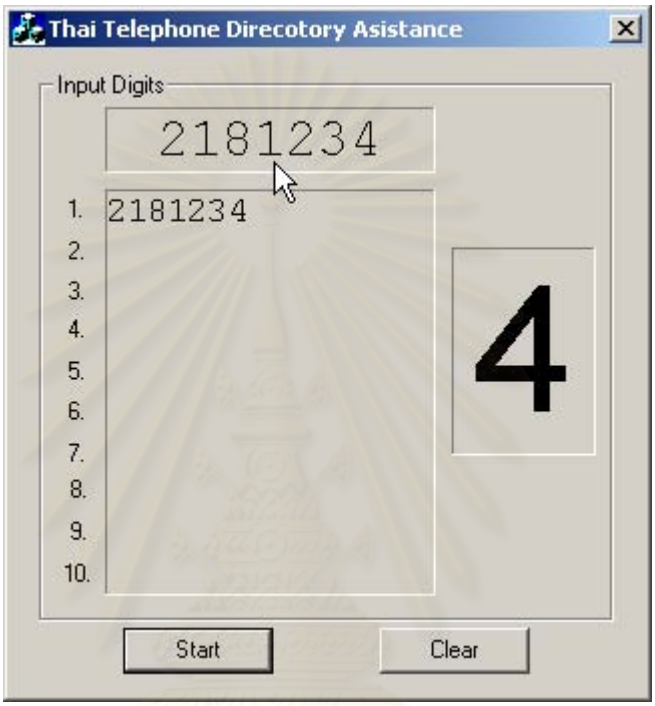
เริ่มต้นโปรแกรมโดยสั่งงานจากโปรแกรมชื่อ TTDA.exe จะปรากฏหน้าต่างโปรแกรมดังรูปที่ ง-1



รูปที่ ง-1 โปรแกรมระบบสอบถามรายนามผู้ใช้โทรศัพท์แบบอัตโนมัติ

จากรูป หมายเลข 1 แสดงผลของชุดหมายเลขที่พูดแต่ละครั้ง หมายเลข 2 แสดงผลของชุดหมายเลขทั้งหมดจำนวน 10 ชุด หมายเลข 3 แสดงส่วนของหมายเลขเดี่ยวของชุดหมายเลขตัวสุดท้าย ปุ่ม Start เป็นปุ่มเริ่มต้นการทำงานของระบบ ปุ่ม Clear เป็นปุ่มล้างชุดหมายเลขทั้งหมดในกรณีที่ชุดหมายเลขมีจำนวนเกิน 10 ชุด

เมื่อเริ่มต้นโปรแกรมโดยกดปุ่ม Start ระบบจะกล่าวต้อนรับและบอกให้พูดหมายเลขโทรศัพท์ที่ต้องการทราบชื่อผู้ใช้โทรศัพท์หลังได้ยินเสียงสัญญาณ เมื่อผู้พูดบอกเลขหมายโทรศัพท์ที่ต้องการทราบ ระบบจะทำการประมวลผลและแจ้งกลับมาในส่วนของตัวเลขแสดงทางจอภาพ และพร้อมกันนั้นระบบจะไปทำการค้นหารายชื่อผู้ใช้โทรศัพท์จากฐานข้อมูล นำกลับมาโดยพูดชื่อและนามสกุลนั้นออกมา ดังแสดงในรูปที่ ง -2



รูปที่ ง-2 โปรแกรมแสดงผลของการรู้จำหมายเลขโทรศัพท์

ในส่วนของฐานข้อมูลผู้ใช้สามารถเข้าไปเรียกดูและสามารถเพิ่มหรือลบรายนามของผู้ใช้โทรศัพท์ได้โดยเรียกจากโปรแกรมชื่อ TTDADB.exe จะแสดงรายการเมนูหลักของคำสั่งดังแสดงในรูปที่ ง-3 ถึง ง-5 ซึ่งประกอบด้วย

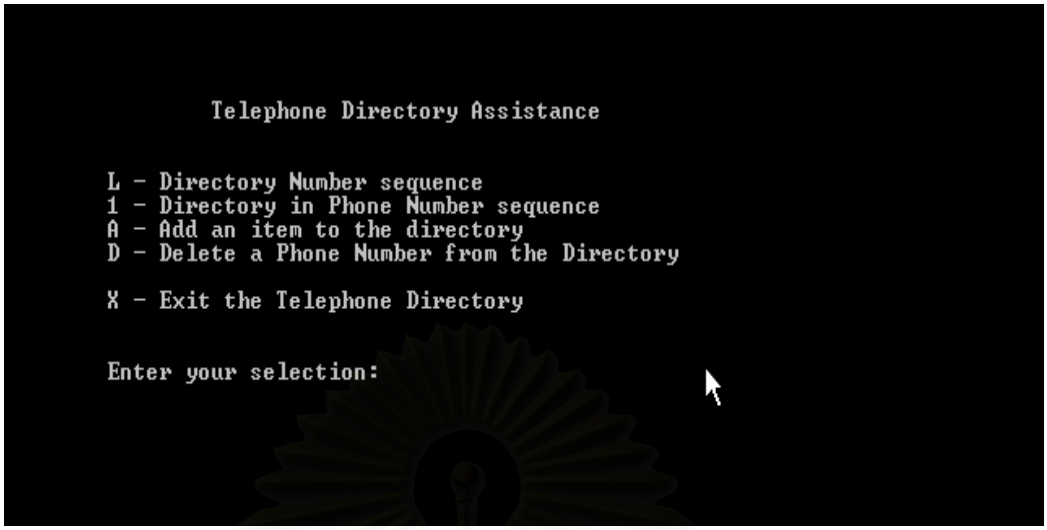
L แสดงรายการของหมายเลขโทรศัพท์และรายนามผู้ใช้โทรศัพท์ตามลำดับการนำข้อมูลเข้า

1 แสดงรายการของหมายเลขโทรศัพท์และรายนามผู้ใช้โทรศัพท์ตามลำดับเลขหมาย

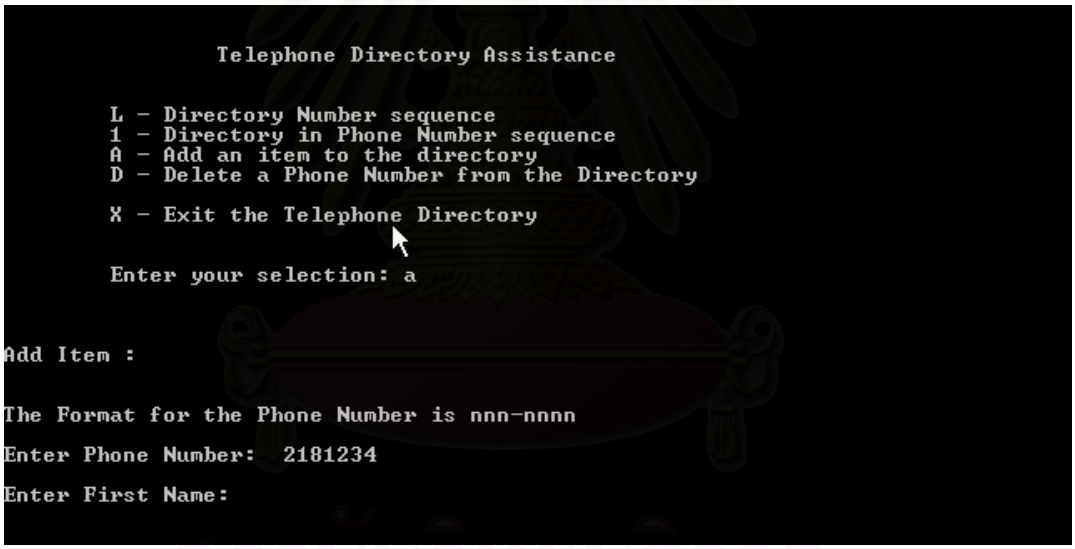
A คำสั่งในการเพิ่มหมายเลขโทรศัพท์ลงในฐานข้อมูล

D คำสั่งในการลบหมายเลขโทรศัพท์ออกจากฐานข้อมูล

X ออกจากระบบ

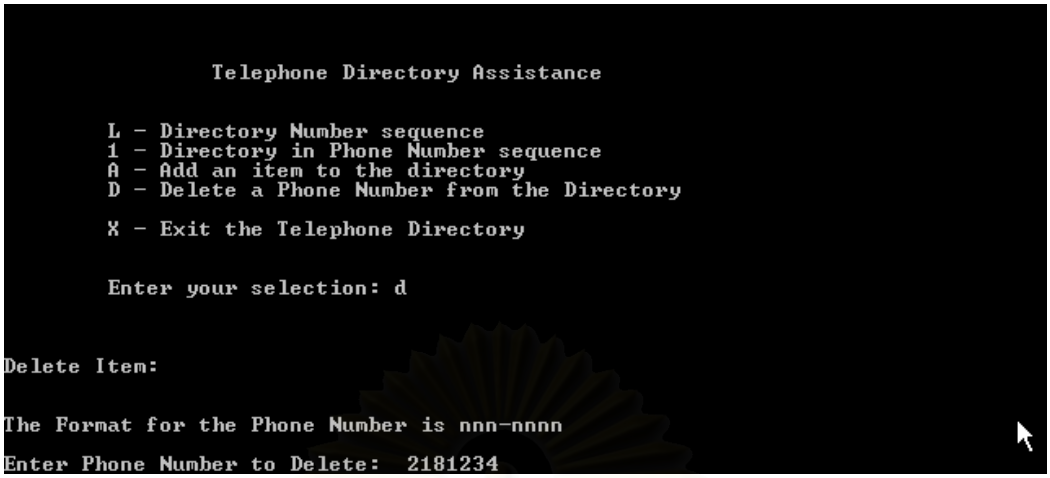


รูปที่ ง-3 เมนูแสดงคำสั่งที่ใช้กับฐานข้อมูล



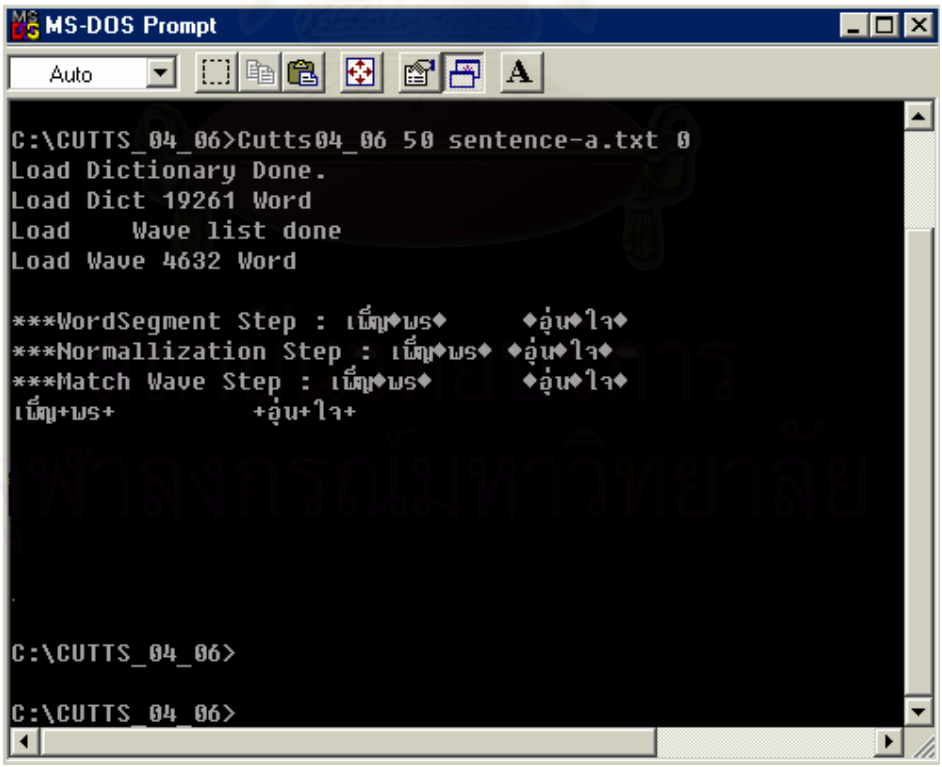
รูปที่ ง-4 การเพิ่มหมายเลขโทรศัพท์ลงในฐานข้อมูล

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ ง-5 การลบหมายเลขโทรศัพท์ออกจากฐานข้อมูล

ในส่วนของการสังเคราะห์เสียงผู้ดูแลระบบจะต้องนำชื่อและนามสกุลมาผ่านขั้นตอนการสังเคราะห์เสียงก่อนเพื่อตรวจสอบความถูกต้องในการอ่านออกเสียง โดยการฟังจากลำโพง ดังแสดงในรูปที่ ง-6 แต่ถ้าการอ่านออกเสียงนั้นผิดพลาดก็ให้ผู้ดูแลระบบแก้ไขโดยใส่คำอ่านที่ถูกต้อง



รูปที่ ง-6 การตรวจสอบการอ่านออกเสียงชื่อและนามสกุล

ประวัติผู้เขียนวิทยานิพนธ์

นายทักษพล พึ่งยนต์ เกิดวันที่ 10 พฤศจิกายน พ.ศ. 2518 เกิดที่ จังหวัดกรุงเทพมหานคร สำเร็จการศึกษาปริญญาวิทยาศาสตรบัณฑิต สาขาวิทยาการคอมพิวเตอร์ประยุกต์ คณะวิทยาศาตร์ประยุกต์ สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ ในปีการศึกษา 2540 และเข้าศึกษาต่อในหลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย เมื่อปีการศึกษา 2544 ปัจจุบันทำงานที่ บริษัท ทศท คอร์ปอเรชั่น จำกัด(มหาชน)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย