

การรู้จำเสียงพูดสระภาษาไทยโดด ๆ ไม่ขึ้นกับผู้พูดโดยการวัดสเปกตรัมคิสแทนซ์  
และใช้ไดนามิกไทม์วาร์บิง



นาย ชีระ ภัทรพรนันท์

ศูนย์วิทยทรัพยากร

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

จุฬาลงกรณ์มหาวิทยาลัย  
ภาควิชาวิศวกรรมไฟฟ้า

บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย

พ.ศ. 2538

ISBN 974-632-703-8

ลิขสิทธิ์ของบัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย

Speaker-Independent Isolated Thai Spoken Vowel Recognition by Using  
Spectrum Distance Measurement and Dynamic Time Warping

Mr. Teera Phatrapornnant



ศูนย์วิทยทรัพยากร

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Engineering

Department of Electrical Engineering

Graduate School

Chulalongkorn University

1995

ISBN 974-632-703-8

หัวข้อวิทยานิพนธ์

การรู้จำเสียงพูดสระภาษาไทยโดด ๆ ไม่ขึ้นกับผู้พูดโดยการวัดสเปกตรัม  
ดิสแตนท์และใช้ไดนามิกไทม์วาร์ปิง

โดย

นาย ชีระ ภัทราพรนันท์

ภาควิชา


วิศวกรรมไฟฟ้า

อาจารย์ที่ปรึกษา

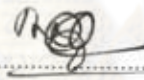
รองศาสตราจารย์ ดร. สมชาย จิตะพันธ์กุล





บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้หัวข้อวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่ง  
ของการศึกษาตามหลักสูตรปริญญาโทบัณฑิต


.....คณบดีบัณฑิตวิทยาลัย  
(รองศาสตราจารย์ ดร.สันติ ฤงสูรณ)

คณะกรรมการสอบวิทยานิพนธ์

.....ประธานกรรมการ  
(รองศาสตราจารย์ ดร.ณรงค์ อยู่ถนอม)

.....อาจารย์ที่ปรึกษา  
(รองศาสตราจารย์ ดร.สมชาย จิตะพันธ์กุล)

.....กรรมการ  
(อาจารย์ ดร.วาทีต เบนจพลงกุล)

.....กรรมการ  
(ดร.บวรกุล จิตต์ประเสริฐ)

พิมพ์ต้นฉบับบทคัดย่อวิทยานิพนธ์ภายในกรอบสี่เหลี่ยมนี้เพียงแผ่นเดียว

ธีระ ภัทพรนันท์ : การรู้จำเสียงพูดสระภาษาไทยโคด ๆ ไม่ขึ้นกับผู้พูดโดยการวัดสเปกตรัมดิสแตนซ์ และใช้ไดนามิกไทม์วาร์ปิง ( SPEAKER-INDEPENDENT ISOLATED THAI SPOKEN VOWEL RECOGNITION BY USING SPECTRUM DISTANCE MEASUREMENT AND DYNAMIC TIME WARPING ) อ.ที่ปรึกษา : รศ. ดร. สมชาย จิตะพันธ์กุล, 151 หน้า. ISBN 974-632-703-8

ระบบรู้จำเสียงสระภาษาไทยคำเดี่ยวแบบต่างบุคคล โดยการวัดสเปกตรัมดิสแตนซ์ที่ผ่านการทรานส์ฟอร์ม ด้วยวิธีฟาสฟูเรียร์ร่วมกับเทคนิคของไดนามิกไทม์วาร์ปิง ทดสอบกับบุคคลในกลุ่มและบุคคลต่างกลุ่ม จำนวนกลุ่มละ 5 คน โดยใช้แบบอ้างอิงเสียงสระภาษาไทยจำนวน 24 ตัวและเป็นแบบอ้างอิง 1 แบบต่อหนึ่งสระ ซึ่งใช้ 2 ชุด แบบอ้างอิงชุดแรกเป็นแบบอ้างอิงไม่ได้จัดกลุ่ม และแบบอ้างอิงชุดที่สองเป็นแบบที่ได้รับการจัดกลุ่มแล้ว ผลการทดสอบกับบุคคล ทั้งสองกลุ่มโดยใช้แบบอ้างอิงจากคน 10 คน แบบไม่จัดกลุ่มได้อัตราความถูกต้องของบุคคลในกลุ่มร้อยละ 84.44 และ ของบุคคลต่างกลุ่มร้อยละ 83.33 ส่วนผลการทดสอบกับแบบอ้างอิงที่จัดกลุ่มแล้ว ได้อัตราความถูกต้องของบุคคลในกลุ่ม ร้อยละ 90.83 และของบุคคลต่างกลุ่มร้อยละ 86.17

นอกจากนี้ยังทดสอบการจำแนกเสียงวรรณยุกต์ภาษาไทยทั้ง 5 เสียงจากสระ 3 ตัวคือ “อา”, “อิ”, และ “อุ” โดยการหาค่าความผิดพลาดยกกำลังสองเทียบกับสมการแทนวรรณยุกต์ ผลการทดสอบกับผู้ชาย และผู้หญิงกลุ่มละ 5 คน ได้อัตราความถูกต้องร้อยละ 81.00



ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ภาควิชา ..... วิศวกรรมไฟฟ้า  
สาขาวิชา ..... วิศวกรรมไฟฟ้า  
ปีการศึกษา ..... 2538

ลายมือชื่อนิสิต ..... ธีระ ภัทพรนันท์  
ลายมือชื่ออาจารย์ที่ปรึกษา .....  
ลายมือชื่ออาจารย์ที่ปรึกษาร่วม .....

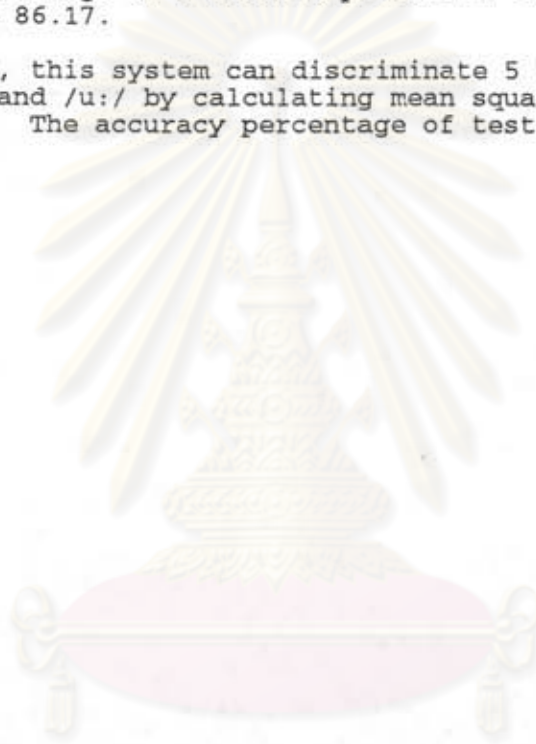


## C515777:MAJOR DIGITAL SIGNAL PROCESSING

KEY WORD: RECOGNITION / SPEAKER-INDEPENDENT / ISOLATED THAI SPOKEN VOWEL.  
TEERA PHATRAPORNANT : SPEAKER-INDEPENDENT ISOLATED THAI SPOKEN  
VOWEL RECOGNITION BY USING SPECTRUM DISTANCE MEASUREMENT AND  
DYNAMIC TIME WARPING. THESIS ADVISOR : ASSO. PROF. SOMCHAI  
JITAPUNKUL, Ph. D. 151 pp. ISBN 974-632-703-8

Speaker-independent isolated Thai vowel recognition by using spectrum distance measurement and Dynamic Time Warping technique are applied with 5 trained persons and 5 non-trained persons. There are 24 Thai vowel reference templates and each reference is represented by one template. Two reference sets are used in this thesis, first reference set is unclassified group and second is classified group. By reference template created from 10 persons, the first reference set's accuracy percentage of 5 trained person is 84.44 and its 5 non-trained persons is 83.33. The second reference set's accuracy percentage of 5 trained person is 90.83 and its 5 non-trained persons is 86.17.

Otherwise, this system can discriminate 5 tone of Thai from 3 Thai vowel /a:/, /i:/, and /u:/ by calculating mean square error with tone's reference formula. The accuracy percentage of testing with 5 males and 5 females is 81.00



ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ภาควิชา.....วิศวกรรมไฟฟ้า

สาขาวิชา.....วิศวกรรมไฟฟ้า

ปีการศึกษา..... 2533

ลายมือชื่อนิสิต..... ธี. ภัทพรินทร์

ลายมือชื่ออาจารย์ที่ปรึกษา.....

ลายมือชื่ออาจารย์ที่ปรึกษาร่วม.....

## กิตติกรรมประกาศ

ผู้เขียนขอขอบพระคุณ รองศาสตราจารย์ ดร. สมชาย จิตะพันธ์กุล อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่ได้ให้คำปรึกษาและแนะนำงานวิทยานิพนธ์นี้สำเร็จลุล่วงไปด้วยดี

ขอขอบพระคุณคณะกรรมการสอบวิทยานิพนธ์ รองศาสตราจารย์ ดร.ณรงค์ อยู่ถนอม, ดร. บวรกุล จิตต์ประเสริฐ, Senior Manager Broadcasting Engineering, IBC Public Company Limited , ดร. วาทิต เบญจพลกุล , ผศ. ดร. สุดาพร ลักษณะนิยานวิน ภาควิชาภาษาศาสตร์ คณะอักษรศาสตร์ ที่ได้กรุณาสละเวลาและให้คำแนะนำต่างๆ ในการทำวิทยานิพนธ์ รวมทั้งบัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัยที่ได้ให้ทุนอุดหนุนในการทำวิจัยในครั้งนี้ด้วย  
ขอขอบคุณเพื่อน ชาวเนคเทค ที่ช่วยเหลือในด้านจัดพิมพ์วิทยานิพนธ์และบันทึกเสียง  
ขอขอบคุณน้อง ๆ ปี 4 และ ห้องปฏิบัติการ DSP ทุกคนที่ให้ความช่วยเหลือต่าง ๆ



ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญ



	หน้า
บทคัดย่อภาษาไทย .....	ง
บทคัดย่อภาษาอังกฤษ .....	จ
กิตติกรรมประกาศ .....	ฉ
บทที่ 1 บทนำ .....	1
ความเป็นมาของปัญหา .....	1
วัตถุประสงค์ของวิทยานิพนธ์ .....	3
ขอบเขตของวิทยานิพนธ์ .....	3
ขั้นตอนและวิธีดำเนินการวิจัย .....	3
ประโยชน์ที่จะได้รับจากการวิจัยนี้ .....	4
บทที่ 2 ลักษณะเชิงสัทศาสตร์ของหน่วยเสียงภาษาไทย .....	5
เสียงพยัญชนะ .....	5
เสียงสระ .....	13
เสียงวรรณยุกต์ .....	15
บทที่ 3 อักขรวิธีใหม่และเทคนิคต่าง ๆ .....	18
ไดนามิกไทม์วาร์ปิง .....	20
ค่าระยะห่าง .....	21
การแบ่งกลุ่มแบบอ้างอิง .....	22
การจำแนกเสียงวรรณยุกต์ .....	26
บทที่ 4 ขั้นตอนการทำงานของระบบ .....	32
ส่วนรับอินพุต .....	33
ส่วนพรีโพรเซสซิง .....	33
ส่วนการแปลงข้อมูล .....	34
ส่วนโพลีโพรเซสซิง .....	34
ส่วนการสร้างแบบอ้างอิง .....	34
ส่วนการสร้างรหัสเสียงวรรณยุกต์ .....	35
ส่วนระบบรู้จำเสียงวรรณยุกต์ .....	38
ส่วนการจัดกลุ่มแบบอ้างอิง .....	40
ส่วนการทดสอบการรู้จำเสียงสระ .....	40
บทที่ 5 การทดสอบและการประเมินผล .....	42
การทดสอบเสียงสระ .....	42
การทดสอบเสียงวรรณยุกต์ .....	46

บทที่ 6	สรุปและข้อเสนอแนะ .....	49
	สรุปผล .....	49
	ข้อเสนอแนะ .....	49
รายการอ้างอิง .....		51
ภาคผนวก .....		53
ภาคผนวก ก.	สเปคโตรแกรมเสียงสระต่าง ๆ ในเพศชาย .....	54
ภาคผนวก ข.	สเปคโตรแกรมเสียงสระต่าง ๆ ในเพศหญิง .....	78
ภาคผนวก ค.	สเปคโตรแกรมแบบอ้างอิงของเสียงสระต่าง ๆ ในเพศชาย .....	103
ภาคผนวก ง.	สเปคโตรแกรมของเสียงวรรณยุกต์ต่าง ๆ ในเสียงสระอา .....	128
ภาคผนวก จ.	ผลการทดลองการรู้จำเสียงสระในเพศหญิงกับเสียงแบบอ้างอิงเพศชาย .....	134
ภาคผนวก ฉ.	ผลการทดลองการรู้จำเสียงสระในเพศชายกับเสียงแบบอ้างอิงเพศชาย .....	137
ภาคผนวก ช.	ค่ากำหนดที่ใช้ในการหาขอบเขตสัญญาณเสียง .....	148
ประวัติผู้เขียน .....		151


  
 ศูนย์วิทยทรัพยากร  
 จุฬาลงกรณ์มหาวิทยาลัย



## สารบัญตาราง

ตารางที่	หน้า
2.1 หน่วยเสียงพยัญชนะภาษาไทย .....	12
2.2 เสียงสระจำแนกตามลักษณะของเสียง และตำแหน่งของลิ้น .....	13
2.3 เปรียบเทียบลักษณะทางสัทศาสตร์ของวรรณยุกต์ตั้งแต่ต้นศตวรรษ จนถึงปัจจุบัน .....	17
3.1 สระที่ถูกจัดอยู่ในกลุ่มต่าง ๆ เมื่อ threshold = 0.5 เท่าของค่าเฉลี่ยพลังงานทั้งหมด .....	24
3.2 สระที่ถูกจัดอยู่ในกลุ่มต่าง ๆ เมื่อ threshold = 1.0 เท่าของค่าเฉลี่ยพลังงานทั้งหมด .....	24
3.3 สระที่ถูกจัดอยู่ในกลุ่มต่าง ๆ เมื่อ threshold = 1.5 เท่าของค่าเฉลี่ยพลังงานทั้งหมด .....	25
3.4 สระที่ถูกจัดอยู่ในกลุ่มต่าง ๆ เมื่อ threshold = 0.6 ถึง 1.4 เท่าของค่าเฉลี่ยพลังงานทั้งหมด .....	25
5.1 แสดงเปอร์เซ็นต์ความถูกต้องรวมของการทดสอบเสียงสระ .....	42
5.2 แสดงเวลาเฉลี่ยต่อคำของการทดสอบเสียงสระ .....	43
5.3 ตัวเลขแสดงผลความถูกต้องของบุคคลกลุ่ม A โดยใช้แบบอ้างอิง ค. แบบแบ่งกลุ่ม .....	44
5.4 ตัวเลขแสดงผลความถูกต้องของบุคคลกลุ่ม B โดยใช้แบบอ้างอิง ค. แบบแบ่งกลุ่ม .....	45
5.5 ตัวเลขแสดงผลความถูกต้องของเสียงต่าง ๆ ที่รู้จำเป็นสระอื่น .....	46
5.6 ตัวเลขแสดงผลความถูกต้องของการแยกวรรณยุกต์เสียง "อา" .....	47
5.7 ตัวเลขแสดงผลความถูกต้องของการแยกวรรณยุกต์เสียง "อิ" .....	47
5.8 ตัวเลขแสดงผลความถูกต้องของการแยกวรรณยุกต์เสียง "อุ" .....	47

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

## สารบัญภาพ

รูปที่	หน้า
1.1 บล็อกไดอะแกรมของระบบรู้จักเสียงพูด .....	1
2.1 แสดงถึงฐานริมฝีปาก .....	6
2.2 แสดงถึงฐานริมฝีปากกับฟัน .....	6
2.3 แสดงถึงฐานฟัน .....	7
2.4 แสดงถึงฐานปุ่มเหงือก .....	7
2.5 แสดงถึงฐานบริเวณหลังปุ่มเหงือก .....	8
2.6 แสดงถึงลิ้นแข็ง .....	8
2.7 แสดงถึงฐานปุ่มเหงือกเพดานแข็ง .....	8
2.8 แสดงถึงฐานเพดานแข็งปุ่มเหงือก .....	9
2.9 แสดงถึงฐานเพดานแข็ง .....	9
2.10 แสดงถึงฐานเพดานอ่อน .....	9
2.11 แสดงถึงฐานลิ้นไก่ .....	10
2.12 แสดงถึงฐานแก้มคอ .....	10
2.13 แสดงถึงฐานเส้นเสียง .....	11
2.14 สระประสม เริ่มที่สระหน้าไปจบที่สระหลัง .....	13
2.15 สระประสมเริ่มออกเสียงที่สระหลัง ไปจบที่สระหน้า .....	14
2.16 สระประสมสองเสียงเริ่มออกเสียงที่สระต่ำ ไปจบที่สระสูง .....	14
2.17 สระประสมสองเสียงเริ่มออกเสียงที่สระสูง ไปจบที่สระต่ำ .....	14
2.18 สระประสมสามเสียง .....	15
2.19 ระดับเสียงหลัก 5 ระดับ .....	15
2.20 สัญลักษณ์ของเสียงวรรณยุกต์โดยใช้เส้นแสดงระดับเสียง .....	15
2.21 สัญลักษณ์ของเสียงวรรณยุกต์โดยใช้อักษรวรรณยุกต์ .....	16
2.22 สัญลักษณ์ของเสียงวรรณยุกต์โดยใช้ตัวเลขแทนระดับเสียง .....	16
2.23 สัญลักษณ์ของเสียงวรรณยุกต์โดยใช้การบรรยายสัทศาสตร์ระดับเสียง .....	16
2.24 ความถี่มูลฐานของเสียงวรรณยุกต์ 5 เสียงในภาษาไทย .....	17
3.1 Dynamic Time Warping ระหว่าง A และ B .....	19
3.4 ความถี่มูลฐานของเสียง "อา" .....	27
3.5 ความถี่มูลฐานของเสียง "อ่า" .....	27
3.6 ความถี่มูลฐานของเสียง "อ๊า" .....	28
3.7 ความถี่มูลฐานของเสียง "อึ๊" .....	28
3.8 ความถี่มูลฐานของเสียง "อึ๊" .....	28

3.9	รูปกราฟของสมการเส้นตรง $y = f_{\max} - ax$ .....	29
3.10	รูปกราฟของสมการพาราโบลา $(x - h)^2 = 4C(y - k)$ .....	30
4.1	แผนผังการทำงานของระบบ .....	32
4.2	ฟอร์มเมตของไฟล์เสียง .....	33
4.3	ขั้นตอนการทำงานของการทำงานที่สวรรณยุคต์ .....	36
4.4	ขั้นตอนการทำงานของการทำงานที่จำแนกเสียงวรรณยุคต์ .....	37
4.5	ขั้นตอนการทำงานของการทำงานที่จัดกลุ่มอ้างอิง .....	39
4.6	ขั้นตอนการตัดสินใจการรู้จำเสียงสระ .....	41



ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย