บทที่ 3
ขั้นตอนในการดำเนินการวิจัย

เมื่อได้ศึกษาโครงสร้างพยางค์ และคุณลักษณะต่าง ๆของเสียงพูดแล้ว ในบทนี้จะกล่าวถึงขั้นตอนการเก็บตัวอย่างข้อมูลเสียงพูด และรายละเอียดของขั้นตอนการหาขอบเขตพยางค์ โดยอาศัยคุณลักษณะของสัญญาณเสียงพูด

3.1 การเก็บตัวอย่างข้อมูลเสียงพูด

3.1.1 การสร้างรายการคำศัพท์

ในการสร้างรายการคำศัพท์เพื่อนำมาใช้ในการทดสอบหาขอบเขตพยางค์สำหรับคำพูดต่อเนื่องภาษาไทยนั้น ได้กำหนดขึ้นตามคำแนะนำของผู้ช่วยศาสตราจารย์ ดร. สุดาพร ลักษณีย-นาวิน แห่งภาควิชาภาษาศาสตร์ คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ซึ่งเป็นคำศัพท์ที่มีความกำกวมในการกำหนดขอบเขตพยางค์ระหว่างรอยต่อพยางค์ ในโครงสร้างพยางค์ทางสัทวิทยาในกรณีปัญหาดังต่อไปนี้ (อุมาพร ศรีรักษา, 2538)

1. ปัญหาจากเสียงพยัญชนะ

เมื่อพยัญชนะปรากฏอยู่ระหว่างสระ 2 เสียง และเป็นพยัญชนะ [p, t, k, m, n, ŋ, w, j] ซึ่งสามารถทำหน้าที่เป็นพยัญชนะต้นหรือพยัญชนะท้ายก็ได้ ดังโครงสร้าง A และ B

A: CVC#V

B: CV#CV

พยัญชนะ	ประโยคทดสอบ	
	A: CVC#V	*B: CV#CV*
/p/	คนทำบาปอวดตัวว่าเก่ง baap1#uat1	เธอทำฉันบ่าปวดไปหมด baa1#puat1
/t/	เขาเป็นญาติอำภา jaat2#am0	มาลีโดนหญ้าตำที่ขา jaa2#tum0
/k/	ในปากอินทรีย์มีปลาสองตัว paak1#in0	คนกินข้าวแต่ป่ากินน้ำ paa1#kin0
/m/	แม่ไปตามอาที่บ้าน taam0#aa0	แม่บอกว่าตามาอยู่ที่บ้าน taa0#maa0
/n/	ฉันเห็นว่านออกดอกตั้งหลายต้น waan2#ɔɔk1	คำว่าหนอกหมายถึงต้นคอวัว waa2#nɔɔk1

/ŋ/	ฉันจะลองอมเงินคุณแม่ดู lɔɔŋ0#om0	พี่ลองมแหวนขึ้นมาให้ lɔɔ0#ŋom0
/w/	น้องจะเอาว่าวอันนั้น waaw2#an0	อาจารย์บอกว่าวันนี้เป็นวันดี waa2#wan0
/j/	ตากล้องบอกให้เอนตัวไปทางซ้าย haaj2#een0	ปากการาคาห้าเยนเท่านั้น haa2#jeen0

มีประโยคที่มีความกำกวมระหว่างรอยต่อพยางค์จากปัญหาเสียงพยัญชนะ 8 คู่ รวม 16 ประโยค

2. **ปัญหาจากเสียงพยัญชนะเรียง**

เมื่อมีพยัญชนะ 2 เสียงปรากฏเรียงกันอยู่ระหว่างสระ 2 เสียง และพยัญชนะ C_1 ได้แก่ [p, t, k] และพยัญชนะ C_2 ได้แก่ [r, l, w] ในกรณีนี้พยัญชนะ C_1 สามารถทำหน้าที่เป็นพยัญชนะท้ายของพยางค์แรก หรือทำหน้าที่เป็นพยัญชนะต้นควบกล้ำกับพยัญชนะ C_2 ของพยางค์หลัง ดังโครงสร้าง C และ D

C: $CVC_1\#C_2V$

D: $CV\#C_1C_2V$

พยัญชนะ	ประโยคทดสอบ	
	C: $CVC_1\#C_2V$	D: $CV\#C_1C_2V$
/pr/	พวกนั้นโดนปรับรายตัว prap1#raaj0	ฝนตกประปรายเป็นประจำ pra1#praaj0
/tr/	เธอปัดราที่ขนมออก pat1#ra0	มานะปะตราที่หน้ารถยนต์ pa1#traa0
/kr/	ส่วนภูมิภาครองลงมาจากส่วนกลาง phaak2#rɔɔŋ0	ตำรวจใช้ผ้ากรองฝุ่นปิดปาก phaa2#krɔɔŋ0
/pl/	ช่วยกันจับหลีกไปไกล ๆ cap1#liik1	ผมจะปลีกตัวมาทันที ca1#pliik1
/kl/	เขาอยากสักลายเสือที่แขน sak1#laaj0	น้ำในสระกลายเป็นสีดำ sa1#klaaj0
/kw/	เขาแห่นาคเวียนรอบโบสถ์ naak2#wiian0	ข้าวเปลือกอยู่หน้าเกวียนสองกอง naa2#kwiian0

มีประโยคที่มีความกำกวมระหว่างรอยต่อพยางค์จากปัญหาเสียงพยัญชนะเรียง 6 คู่ รวม 12 ประโยค

3. ปัญหาจากเสียงสระเรียง

เมื่อเสียงสระ [i, ii, u, uu, ɯ, ɯɯ] ปรากฏเรียงกับเสียงสระ [a] เสียงสระทั้ง 2 เสียงนี้อาจเป็นสระเดี่ยวทั้งคู่ใน 2 พยางค์ หรือเป็นสระผสมใน 1 พยางค์ ดังโครงสร้าง X และ Y

X : *CVVC#*

Y : *CV#VC#*

สระ	ประโยคทดสอบ	
	X : *CVVC#*	*Y* : *CV#VC#*
/ia/	คำว่าเ<u>เตียบ</u>แปลว่าตะลุ่ม tiap1#	ชาติชายโดนครู<u>ดิอับ</u>อายขายหน้า ti1#ap1
/iia/	เธอจบจากโรงเรียน<u>เตรียม</u>อุดม triiam0#	ดอนเรียนปริญญา<u>ตรีอัม</u>พรเคยได้รับทุน trii0#am0
/ɯɯa/	เขา<u>เคือง</u>ฉันด้วยเรื่องเล็กน้อย khɯɯaŋ0#	เธอ<u>คืออัง</u>คณาทิมดี khɯɯ0#aŋ0
/uua/	ลำดวนเป็นคนใจคอ<u>รวนเร</u> ruuan0#	ปิติปักธงที่<u>รูอันสุด</u>ท้าย ruu0#an0

มีประโยคที่มีความกำกวมระหว่างรอยต่อพยางค์จากปัญหาเสียงสระเรียง 4 คู่ รวม 8 ประโยค

<u>3.1.2 การบันทึกข้อมูล</u>

การบันทึกข้อมูลเพื่อใช้ในการวิจัยนี้ได้จัดสภาพแวดล้อมให้คล้ายกับสถานที่ทำงานทั่วไป ทำการบันทึกข้อมูลเสียงพูดจากผู้พูดจำนวน 28 คน เป็นชาย 15 คน หญิง 13 คน มีอายุระหว่าง 15 - 50 ปี ให้ผู้พูดแต่ละคนอ่านประโยคทดสอบทีละประโยค จำนวน 36 ประโยค โดยพูดในอัตราเร็วที่สม่ำเสมอ เช่นเดียวกับการพูดสนทนาปกติ โดยกำหนดอัตราการสุ่มตัวอย่าง (Sampling Rate) เท่ากับ 11 kHz และจัดเก็บด้วยขนาด 16 บิต ซึ่งอุปกรณ์ที่ใช้ในการบันทึกเสียงมีรายละเอียดดังนี้

1. เครื่องคอมพิวเตอร์ Pentium-S 120 หน่วยความจำ 32 MB Hard disk ขนาด 1.2 GB Floppy Disk 3.5" ขนาด 1.44 MB
2. การ์ดเสียง Sound Blaster 16 ของบริษัท Creative Technology
3. ไมโครโฟน Philips Uni-directional Dynamic Microphone รุ่น SBC-465
4. ระบบปฏิบัติการ MS-DOS Version 6.22 ภาษาไทย

5. ระบบปฏิบัติการ Microsoft Window 95 ภาษาไทย
6. โปรแกรม GoldWave Version 3.24

3.2 รายละเอียดและขั้นตอนในการหาขอบเขตพยางค์

3.2.1 ขั้นตอนของกรรมวิธีการหาขอบเขตพยางค์

ขั้นตอนของกรรมวิธีการหาขอบเขตพยางค์สำหรับคำพูดต่อเนื่องภาษาไทย แสดงดังรูปที่ 3.1 ซึ่งมีรายละเอียดดังนี้

ขั้นตอนที่ 1 รับสัญญาณเสียงพูดที่เป็นประโยคคำพูดต่อเนื่องเข้ามาปรับบรรทัดฐานเชิงขนาดของสัญญาณเสียงพูด โดยเทียบกับจำนวนบิตที่ใช้ในการเก็บข้อมูล

ขั้นตอนที่ 2 ทำการเน้นล่วงหน้าสัญญาณเสียงพูด เพื่อเพิ่มอัตราส่วนของสัญญาณเสียงพูดต่อสัญญาณรบกวน

ขั้นตอนที่ 3 กำหนดขนาดของกรอบเสียงพูดให้แต่ละกรอบเสียงพูดมีขนาด 256 ตัวอย่างสัญญาณเสียงพูด และให้แต่ละกรอบเสียงพูดซ้อนทับกัน $^{3}/_{4}$ ของกรอบเสียงพูด

ขั้นตอนที่ 4 วางฟังก์ชันหน้าต่าง ฟังก์ชันหน้าต่างที่ใช้ในงานวิจัยนี้คือ ฟังก์ชันหน้าต่างแบบ Hamming โดยวางฟังก์ชันก์หน้าต่างบนกรอบเสียงพูดแต่ละกรอบ

ขั้นตอนที่ 5 วิเคราะห์หาค่าคุณลักษณะของสัญญาณเสียงพูด โดยกำหนดคุณลักษณะที่ใช้ในงานวิจัยนี้มีดังนี้คือ

- พลังงาน คุณลักษณะพลังงานในงานวิจัยนี้มี 6 ประเภท คือ
  - พลังงานสัมบูรณ์
  - พลังงานเฉลี่ย
  - พลังงานความถี่และเวลา
  - พลังงานกำลังสอง
  - พลังงาน Teager
  - พลังงานจากการแปลงแบบ Walsh
- อัตราการตัดผ่านระดับกำหนด ในงานวิจัยนี้ได้กำหนดระดับกำหนด (L) เท่ากับ 0.07
- ความถี่มูลฐาน

ขั้นตอนที่ 6 ทำการปรับเรียบคุณลักษณะของสัญญาณเสียงพูด โดยใช้วิธีปรับเรียบโดยค่ากลางในขั้นแรก และใช้วิธีปรับเรียบโดยค่าเฉลี่ยเคลื่อนไหวในลำดับต่อมา

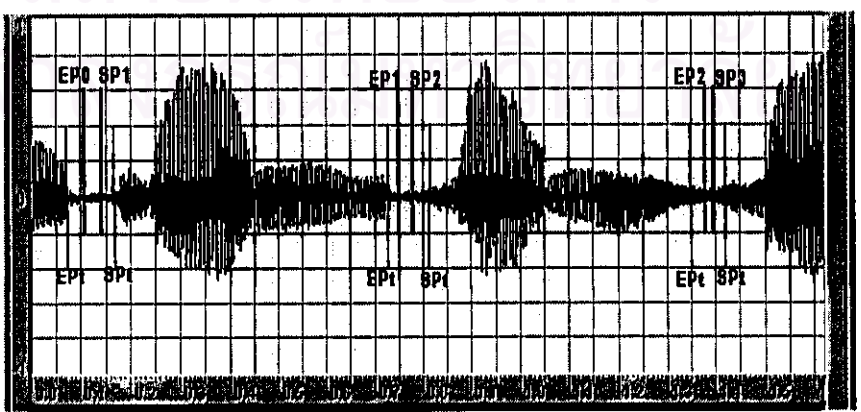
ขั้นตอนที่ 7 ทำการปรับบรรทัดฐานเชิงขนาดคุณลักษณะของเสียงพูด โดยเทียบกับค่าที่มากที่สุดของคุณลักษณะ

รูปที่ 3.1 แผนภาพแสดงขั้นตอนการหาขอบเขตพยางค์

ขั้นตอนที่ 8 ตรวจหาจุดแบ่งพยางค์ วิธีการตรวจหาจุดแบ่งพยางค์ในงานวิจัยนี้มี 4 วิธี การเลือกใช้วิธีการตรวจหาจุดแบ่งพยางค์จะขึ้นอยู่กับค่าคุณลักษณะที่ใช้

1. วิธี EMN (Energy Maximum-Minimum) เมื่อคุณลักษณะที่นำมาใช้ในการหาขอบเขตพยางค์เป็นค่าพลังงาน จะใช้วิธี EMN ในการตรวจหาจุดแบ่งพยางค์
2. วิธี EBR (Energy and Band Crossing Rate) เมื่อคุณลักษณะที่นำมาใช้ในการหาขอบเขตพยางค์เป็นค่าพลังงานและอัตราการตัดผ่านระดับกำหนด จะใช้วิธี EBR ในการตรวจหาจุดแบ่งพยางค์
3. วิธี EBF(Energy, Band Crossing Rate and Fundamental Frequency) เมื่อคุณลักษณะที่นำมาใช้ในการหาขอบเขตพยางค์เป็นค่าพลังงาน อัตราการตัดผ่านระดับกำหนด และความถี่มูลฐาน จะใช้วิธี EBF ในการตรวจหาจุดแบ่งพยางค์
4. วิธีฟังก์ชันการประมาณ เป็นวิธีการตรวจหาจุดแบ่งพยางค์ที่ใช้การประมาณค่าคุณลักษณะพลังงานทางสถิติ

ขั้นตอนที่ 9 ปรับปรุงขอบเขตพยางค์ ณ จุดแบ่งพยางค์หนึ่งจุด หมายถึง จุดปลายพยางค์แรกและจุดต้นพยางค์ที่สอง ในขั้นตอนการตรวจหาจุดแบ่งพยางค์ จะได้จุดต้นและจุดปลายพยางค์ขั้นแรกมา แต่จุดต้นและจุดปลายอาจมีตำแหน่งยังไม่ถูกต้อง จึงต้องนำมาผ่านขั้นตอนการปรับปรุงขอบเขตพยางค์อีกครั้งหนึ่ง ซึ่งตำแหน่งของจุดปลายพยางค์ที่ได้รับการปรับปรุง จะเป็นตำแหน่งที่มีอัตราส่วนพลังงานด้านขวาต่อด้านซ้ายสูงที่สุดในช่วงที่กำหนด (สมการที่ 2.32) และตำแหน่งของจุดต้นพยางค์ที่ได้รับการปรับปรุง จะเป็นตำแหน่งที่มีอัตราส่วนพลังงานด้านซ้ายต่อด้านขวาสูงที่สุดในช่วงที่กำหนด (สมการที่ 2.33) ดังแสดงในรูปที่ 3.2 โดยที่ EPt คือจุดปลายพยางค์แรกดั้งเดิม SPt คือจุดต้นพยางค์ที่สองดั้งเดิม ส่วน EP1 และ SP2 คือจุดปลายพยางค์ที่หนึ่งและจุดต้นพยางค์ที่สองที่ได้รับการปรับตำแหน่งแล้วตามลำดับ



รูปที่ 3.2 การปรับปรุงขอบเขตพยางค์

3.2.2 กรรมวิธีการหาขอบเขตพยางค์ (Syllable Segmentation Algorithm)

จากขั้นตอนการหาขอบเขตพยางค์ ผู้วิจัยได้กำหนดกรรมวิธีการหาขอบเขตพยางค์ตามคุณลักษณะและวิธีการตรวจหาจุดแบ่งพยางค์เป็น 3 ชุด คือ (1) ชุดที่ใช้คุณลักษณะพลังงานและวิธีการตรวจหาจุดแบ่งพยางค์ EMN จำแนกออกเป็น 6 กรรมวิธี (2) ชุดที่ใช้คุณลักษณะพลังงานที่ดีที่สุดในชุดที่ (1) กับวิธีฟังก์ชันการประมาณในการตรวจหาจุดแบ่งพยางค์ มี 1 กรรมวิธี และ (3) ชุดที่ใช้คุณลักษณะพลังงานที่ดีที่สุดในชุดที่ (1) ร่วมกับค่าคุณลักษณะอื่น ๆ จำแนกออกเป็น 2 กรรมวิธี ดังนั้นในงานวิจัยนี้มีกรรมวิธีในการหาขอบเขตพยางค์ทั้งหมด 9 กรรมวิธีดังนี้คือ

กรรมวิธี	คุณลักษณะ	วิธีการตรวจหาจุดแบ่งพยางค์
1. กรรมวิธีพลังงานสัมบูรณ์ (Absolute Energy Algorithm)	พลังงานสัมบูรณ์	วิธี EMN
2. กรรมวิธีพลังงานเฉลี่ย (Root Mean Square Energy Algorithm)	พลังงานเฉลี่ย	วิธี EMN
3. กรรมวิธีพลังงานความถี่และเวลา (Frequency-Time Energy Algorithm)	พลังงานความถี่และเวลา	วิธี EMN
4. กรรมวิธีพลังงานกำลังสอง (Square Energy Algorithm)	พลังงานกำลังสอง	วิธี EMN
5. กรรมวิธีพลังงาน Teager (Teager Energy Algorithm)	พลังงาน Teager	วิธี EMN
6. กรรมวิธีพลังงานจากการแปลงแบบ Walsh (Walsh Transform Energy Algorithm)	พลังงานจากการแปลงแบบ Walsh	วิธี EMN
7. กรรมวิธีฟังก์ชันการประมาณพลังงานเฉลี่ย (Root Mean Square Energy Estimation Function)	พลังงานเฉลี่ย	วิธีฟังก์ชันการประมาณ
8. กรรมวิธีพลังงานเฉลี่ยและอัตราการตัดผ่านระดับกำหนด (Root Mean Square Energy and Band Crossing Rate Algorithm)	พลังงานเฉลี่ย และอัตราการตัดผ่านระดับกำหนด	วิธี EBR
9. กรรมวิธีพลังงานเฉลี่ย อัตราการตัดผ่านระดับกำหนด และความถี่มูลฐาน (Root Mean Square Energy, Band Crossing Rate and Fundamental Frequency Algorithm)	พลังงานเฉลี่ย อัตราการตัดผ่านระดับกำหนด และความถี่มูลฐาน	วิธี EBF

3.2.3 ขั้นตอนและวิธีการตรวจหาจุดแบ่งพยางค์ (Syllable Segmentation Point Detection Method)

วิธีการตรวจหาจุดแบ่งพยางค์ในงานวิจัยนี้มีด้วยกัน 4 วิธี ดังนี้คือ

1. วิธี EMN เป็นวิธีการตรวจหาจุดแบ่งพยางค์ที่ใช้กับคุณลักษณะพลังงานทั้งหมด
2. วิธี EBR เป็นวิธีการตรวจหาจุดแบ่งพยางค์ที่ใช้คุณลักษณะพลังงาน และอัตราการตัดผ่านระดับกำหนด ร่วมกัน
3. วิธี EBF เป็นวิธีการตรวจหาจุดแบ่งพยางค์ที่ใช้คุณลักษณะพลังงาน อัตราการตัดผ่านระดับกำหนด และความถี่มูลฐานร่วมกัน
4. วิธีฟังก์ชันการประมาณ เป็นวิธีการตรวจหาจุดแบ่งพยางค์ที่ใช้การประมาณค่าคุณลักษณะพลังงานทางสถิติ

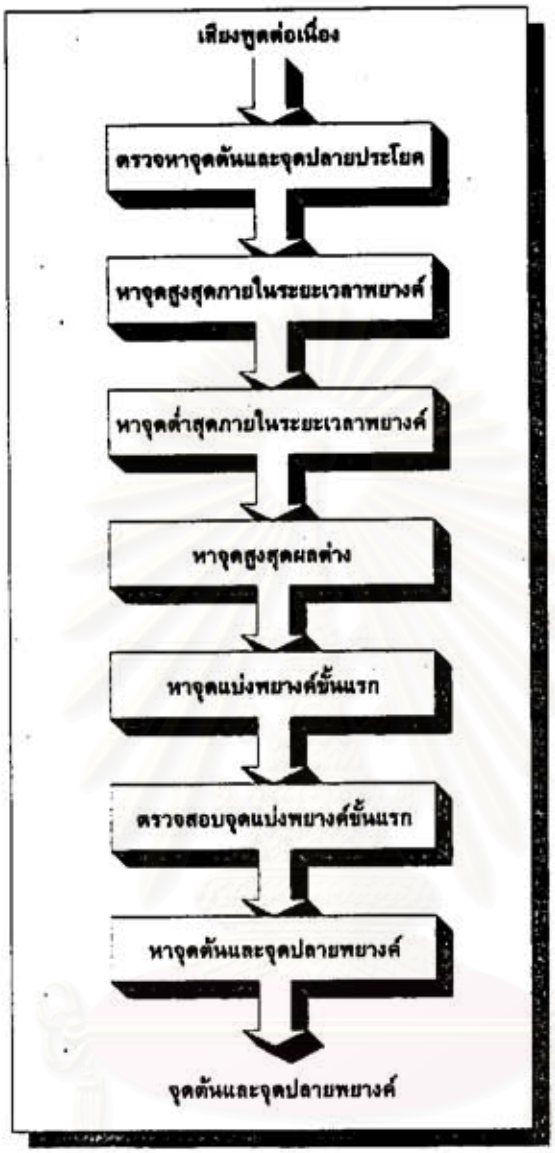
พารามิเตอร์ที่ต้องทำการปรับค่าเพื่อนำไปใช้ในการตรวจหาขอบเขตพยางค์มีดังนี้คือ

- ระยะเวลาพยางค์ (DurTh) คือ การกำหนดจำนวนกรอบเสียงพูดใน 1 พยางค์ (1 กรอบเสียงพูดประมาณ 5.8 มิลลิวินาที)
- ค่าความต่างพลังงาน (DffTh) คือ การกำหนดค่าความต่างระหว่างพลังงานสูงและพลังงานต่ำ

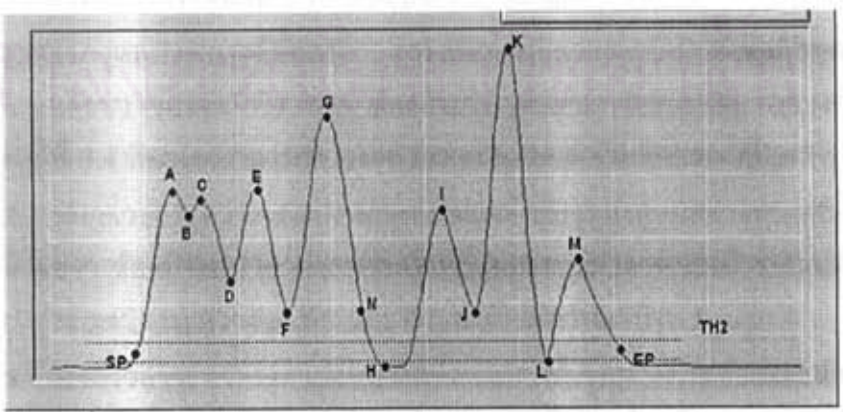
1. วิธี EMN (Energy Maximum-Minimum)

วิธี EMN เป็นวิธีการตรวจหาจุดแบ่งพยางค์ที่ใช้กับคุณลักษณะพลังงานทั้งหมด ขั้นตอนและรายละเอียดแสดงในรูปที่ 3.3 ซึ่งประกอบด้วย 7 ขั้นตอน ดังนี้

ขั้นตอนที่ 1 ทำการตรวจหาจุดต้น (Beginning Point) และจุดปลายของประโยค (Ending Point) การกำหนดระดับพลังงานเริ่มต้นให้กับระบบโดยนำมาใช้ในการกำหนดจุดต้นและจุดปลายของประโยคนั้น มีผลทำให้ตำแหน่งของจุดต้นและจุดปลายของประโยคมีความผิดพลาดมาก เนื่องจากแต่ละประโยค แต่ละคนพูด มีค่าระดับพลังงานไม่เท่ากัน แม้ว่าสัญญาณเสียงจะได้รับการปรับบรรทัดฐานเชิงขนาดแล้วก็ตาม ดังนั้นจึงใช้วิธีการกำหนดจุดต้นและจุดปลายประโยคด้วยวิธีที่ปรับแต่งได้ (Adaptive Threshold) โดยการหาค่าเฉลี่ยของระดับพลังงานภายใน 5 กรอบเสียงพูดที่ต้นประโยคและท้ายประโยค (TH1) เพื่อนำมาใช้เป็นค่าเริ่มต้นในการหาจุดต้นและจุดปลายประโยค จุดต้นประโยคคือ จุดแรกที่มีค่าระดับพลังงานสูงกว่า TH1 และภายในระยะเวลาพยางค์ ที่ต่อจากจุดที่คาดว่าจะเป็นจุดต้นประโยคต้องมีระดับพลังงานสูงกว่า TH2 (TH1+DffTh) จากรูปที่ 3.4 คือ จุด SP และ จุด EP คือจุดปลายประโยค



รูปที่ 3.3 แผนภาพแสดงขั้นตอนและรายละเอียดของวิธี EMN



รูปที่ 3.4 แผนภูมิเส้นระดับพลังงานของสัญญาณเสียงพูด

ขั้นตอนที่ 2 หาจุดสูงสุดภายในระยะเวลาพยางค์ เป็นการหาค่าพลังงานที่มากที่สุดภายในระยะเวลาพยางค์ โดยที่ X_n คือระดับพลังงานที่กรอบเสียงพูด n ซึ่งเป็นกรอบเสียงพูดที่สนใจ จากนั้นพิจารณากรอบเสียงพูด X_n เทียบกับกรอบเสียงพูด $X_{n-(DurTh/2)}$ ถึง $X_{n+(DurTh/2)}$ ถ้ากรอบเสียงพูด X_n มีค่าพลังงานมากที่สุดภายในระยะเวลาที่พิจารณา X_n คือ จุดสูงสุดภายในเวลาที่กำหนด จากนั้นทำการพิจารณาซ้ำเดิม โดยเลื่อนจุดที่สนใจไปทางขวาทีละจุด จากรูปที่ 3.4 จุดสูงสุดภายในระยะเวลาพยางค์คือจุด A, C, E, G, I, K และM

ขั้นตอนที่ 3 หาจุดต่ำสุดภายในระยะเวลาพยางค์ จุดต่ำสุดภายในระยะเวลาพยางค์ คือจุดที่ต่ำที่สุดระหว่างจุดสูงสุดภายในระยะเวลาพยางค์ 2 จุดที่อยู่ต่อเนื่องกัน ทำการหาจุดต่ำสุดภายในระยะเวลาพยางค์จุดอื่นต่อไป โดยเลื่อนจุดสูงสุดภายในระยะเวลาพยางค์ไปทางขวาทีละจุดแล้วพิจารณาซ้ำ จากรูปที่ 3.4 คือ จุดต่ำสุดภายในระยะเวลาพยางค์คือ จุด B, D, F, H, J และ L

ขั้นตอนที่ 4 หาจุดสูงสุดผลต่าง จุดสูงสุดผลต่าง คือ จุดที่มีค่าผลต่างพลังงานระหว่างจุดสูงสุดภายในระยะเวลาพยางค์ กับจุดต่ำสุดภายในระยะเวลาพยางค์ทั้งทางซ้ายและขวามากกว่าค่าความต่างพลังงานที่กำหนด (DffTh) ทำการพิจารณาซ้ำโดยเลื่อนจุดสูงสุดภายในเวลากำหนดไปทางขวาทีละจุด จากรูปที่ 3.4 จุดสูงสุดผลต่างคือจุด A, E, G, I, K และ M จุด C ไม่เป็นจุดสูงสุผลต่าง เนื่องจากค่าความต่างพลังงานระหว่างจุด CB มีค่าน้อยกว่าค่าที่กำหนดกำหนด

ขั้นตอนที่ 5 หาจุดแบ่งพยางค์ขั้นแรก จุดแบ่งพยางค์ คือ จุดที่ที่มีค่าพลังงานต่ำที่สุดระหว่างจุดสูงสุดผลต่าง 2 จุดที่อยู่ต่อเนื่องกัน จากรูปที่ 3.4 จุดแปงพยางค์ขั้นแรกคือจุด D, F, H, J และ L

ขั้นตอนที่ 6 ดรวจสอบจุดแบ่งพยางค์ขั้นแรก จุดแปงพยางค์ที่ได้ในขั้นแรกอาจไม่ใช่จุดแบ่งพยางค์ที่แท้จริง จึงต้องมีการดรวจสอบอีกครั้งหนึ่ง โดยใช้ระยะเวลาเป็นพารามิเตอร์ในการดรวจสอบ การทดสอบมี 3 กรณีคือ

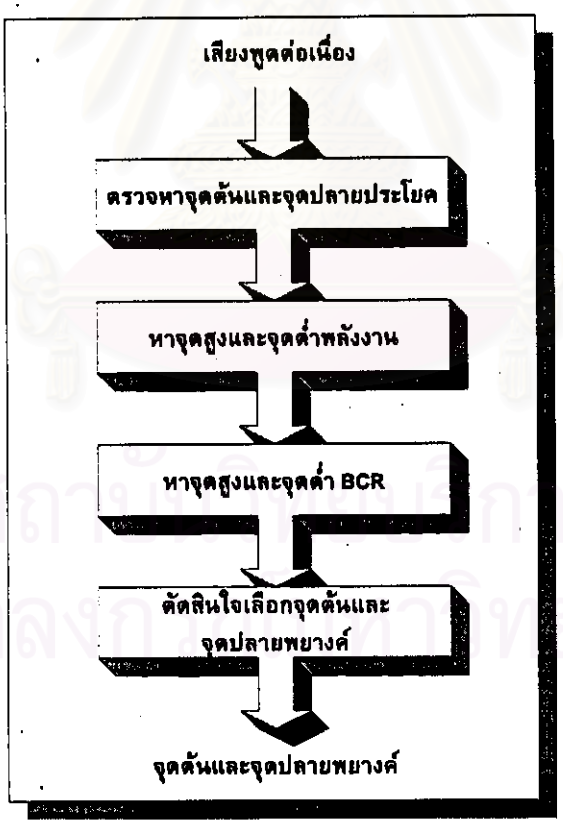
1. จากจุดแบ่งพยางค์ขั้นแรก ระยะเวลาระหว่างจุดที่พิจารณากับจุดด้านซ้ายและขวาของจุดที่พิจารณา ต้องมีระยะเวลามากกว่าระยะเวลาพยางค์ จากรูปที่ 3.4 คือระยะเวลาระหว่างจุด FD และ FH ถ้าเริ่มพิจารณาที่จุด F
2. ระยะเวลาระหว่างจุดแบ่งพยางค์ขั้นแรก กับจุดสูงสุดผลต่างทางด้านซ้ายและขวา ต้องมีระยะเวลามากกว่าครึ่งหนึ่งของระยะเวลาพยางค์ จากรูปที่ 3.4 คือ ระยะเวลาระหว่างจุด FE และ FG ถ้าเริ่มพิจารณาที่จุด F
3. ระยะเวลาระหว่างจุดแปงพยางค์ขั้นแรก กับจุดทางด้านซ้ายและด้านขวาที่มีระดับพลังงานเท่ากันหรือใกล้เคียงกันมากที่สุด ต้องมีระยะเวลามากกว่าครึ่งหนึ่งของระยะเวลาพยางค์ จากรูปที่ 3.4 คือระยะเวลาระหว่างจุด FD และ FN

จุดแปงพยางค์ที่แท้จริง คือมีจำนวนจุดแปงพยางค์เท่ากัน 2 ใน 3 ของการทดสอบ หากจำนวนจุดแปงพยางค์ทั้ง 3 การทดสอบมีจำนวนไม่เท่ากันเลย จะเลือกกรณีที่ให้จำนวนจุดแปงพยางค์ที่มากที่สุดเป็นจุดแปงพยางค์ที่แท้จริง

ขั้นตอนที่ 7 หาจุดต้น (Starting Syllable Point) และจุดปลายพยางค์ (Ending Syllable Point) หรือเรียกรวมกันว่า จุด Endpoint ณ จุดแปงพยางค์ 1 จุด จะเป็นจุดปลายของพยางค์แรก และจุดต้นของพยางค์ที่สอง จุดปลายของพยางค์แรกคือ จุดแรกที่มีค่าพลังงานมากกว่าค่าพลังงานที่จุดแปงพยางค์ไปทางซ้าย และจุดต้นของพยางค์ที่สองคือจุดที่มีค่าพลังงานมากกว่าค่าพลังงานที่จุดแปงพยางค์ไปทางขวา

2. วิธี EBR (Energy and Band Crossing Rate)

วิธี EBR เป็นวิธีการตรวจหาจุดแปงพยางค์ที่ใช้คุณลักษณะพลังงาน และอัตราการตัดผ่านระดับกำหนด ร่วมกัน ซึ่งมีขั้นตอนและรายละเอียดในการหาขอบเขตพยางค์ดังรูปที่ 3.5 โดยมี 4 ขั้นตอนดังนี้

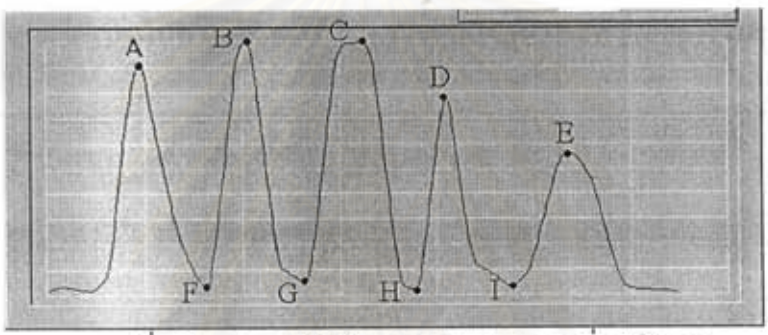


รูปที่ 3.5 แผนภาพแสดงขั้นตอนและรายละเอียดของวิธี EBR

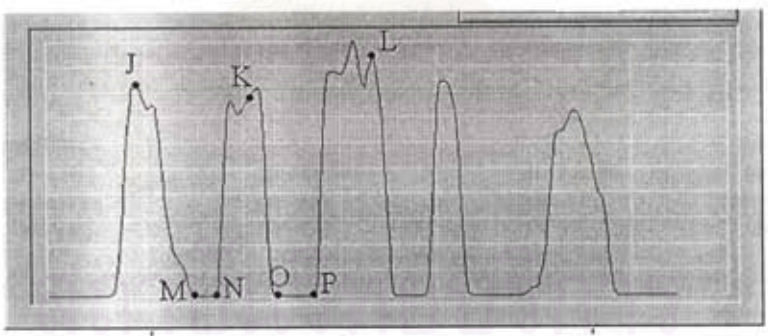
ขั้นตอนที่ 1 ตรวจหาจุดต้นและจุดปลายประโยค เป็นวิธีเดียวกับการตรวจหาจุดต้นและจุดปลายประโยคในวิธี EMN

ขั้นตอนที่ 2 หาจุดสูงและจุดต่ำพลังงาน จุดสูงและจุดต่ำพลังงานคือ จุดสูงสุดและจุดต่ำสุดภายในระยะเวลาพยางค์ที่หาได้จากวิธี EMN จากรูปที่ 3.6a จุดสูงพลังงานคือ จุด A, B, C, D, E และจุดต่ำพลังงานคือจุด F, G, H, I

ขั้นตอนที่ 3 หาจุดสูงและจุดต่ำ BCR จุดต่ำ BCR คือจุดที่อยู่ระหว่างจุดสูงพลังงาน 2 จุดที่อยู่ต่อเนื่องกัน ตัวอย่างแสดงในรูปที่ 3.6(a-b) จุด AB จะให้จุดต่ำ BCR คือจุด M ส่วนจุดสูง BCR คือจุดที่มีค่าอัตราการตัดผ่านระดับกำหนดมากที่สุดระหว่างจุดสูงพลังงานกับจุดต่ำ BCR เช่น จุด J คือจุดสูง BCR ที่อยู่ระหว่างจุด AM



รูปที่ 3.6a แผนภูมิเส้นแสดงจุดสูงและจุดต่ำพลังงาน



รูปที่ 3.6b แผนภูมิเส้นแสดงจุดสูงและจุดต่ำ BCR

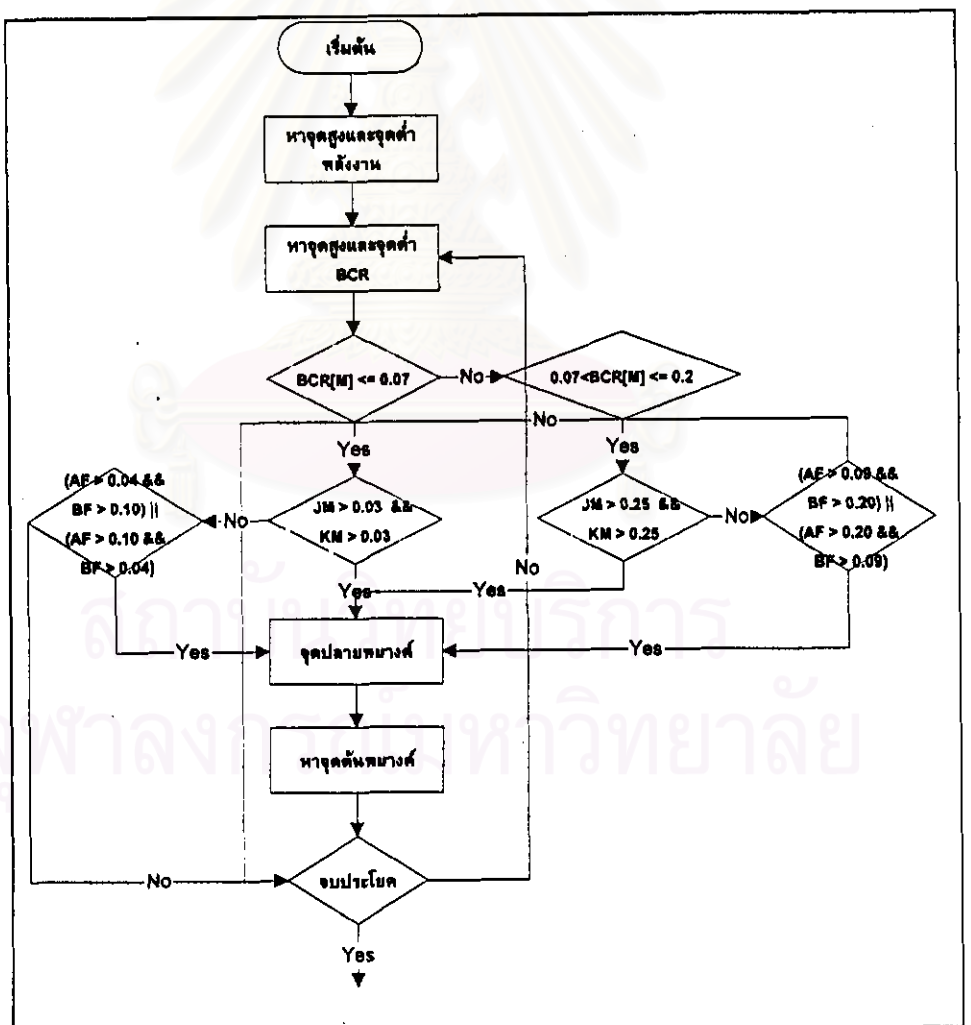
ขั้นตอนที่ 4 ตัดสินใจเลือกจุดต้นและจุดปลายพยางค์ วิธีการหาขอบเขตพยางค์เมื่อใช้คุณลักษณะพลังงานและอัตราการตัดผ่านระดับที่กำหนด มีขั้นตอนการตัดสินใจดังรูปที่ 3.7 ซึ่งแสดงตัวอย่างการตัดสินใจว่าจุด M ในรูปที่ 3.6b เป็นจุดปลายพยางค์หรือไม่ โดยที่ BCR[M] คืออัตราการตัดผ่านระดับกำหนดที่จุด M JM คือผลต่างอัตราการตัดผ่านระดับกำหนดที่จุด J กับจุด M KM คือผลต่างอัตราการตัดผ่านระดับกำหนดที่จุด K กับจุด M AF คือผลต่างของระดับพลังงานที่จุด A กับจุด F BF คือผลต่างของระดับพลังงานที่จุด B และจุด F เมื่อได้จุดปลายพยางค์แรก

แล้ว จุดต้นพยางค์ที่สองคือจุดแรกที่มีค่าอัตราการตัดผ่านระดับกำหนดสูงกว่าจุดปลายพยางค์แรก

ขั้นตอนในการตัดสินใจเลือกจุดต้นและจุดปลายพยางค์ ของวิธี EBR มีดังนี้

ขั้นแรก ตรวจสอบอัตราการตัดผ่านระดับกำหนด อัตราการตัดผ่านระดับกำหนดที่จุดแบ่งพยางค์ควรมีค่าต่ำ แต่อาจไม่เป็นไปตามที่คาดไว้ จึงกำหนดระดับของอัตราการตัดผ่านระดับกำหนดไว้ 2 ระดับ เพื่อกำหนดค่าคงที่ต่าง ๆ ที่ต่างกัน ในการตรวจสอบเพื่อหาจุดแบ่งพยางค์

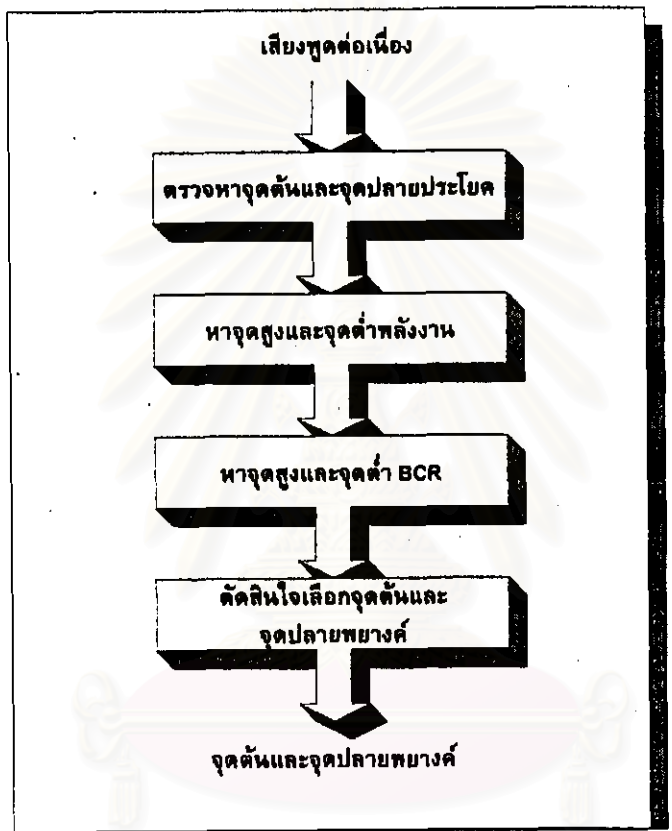
ขั้นที่สอง ตรวจสอบผลต่างอัตราการตัดผ่านระดับกำหนดระหว่างจุดสูงBCR กับจุดต่ำBCR ผลต่างอัตราการตัดผ่านระดับกำหนดระหว่างจุดสูงและจุดแบ่งพยางค์ควรมีค่าสูง หากไม่เป็นไปตามที่คาดไว้ ไปตรวจสอบผลต่างระหว่างพลังงานจุดสูงกับจุดต่ำ



รูปที่ 3.7 แผนภาพกระบวนการตัดสินใจเลือกจุดปลายพยางค์และจุดต้นพยางค์

3. วิธี EBF (Energy, Band Crossing Rate and Fundamental Frequency)

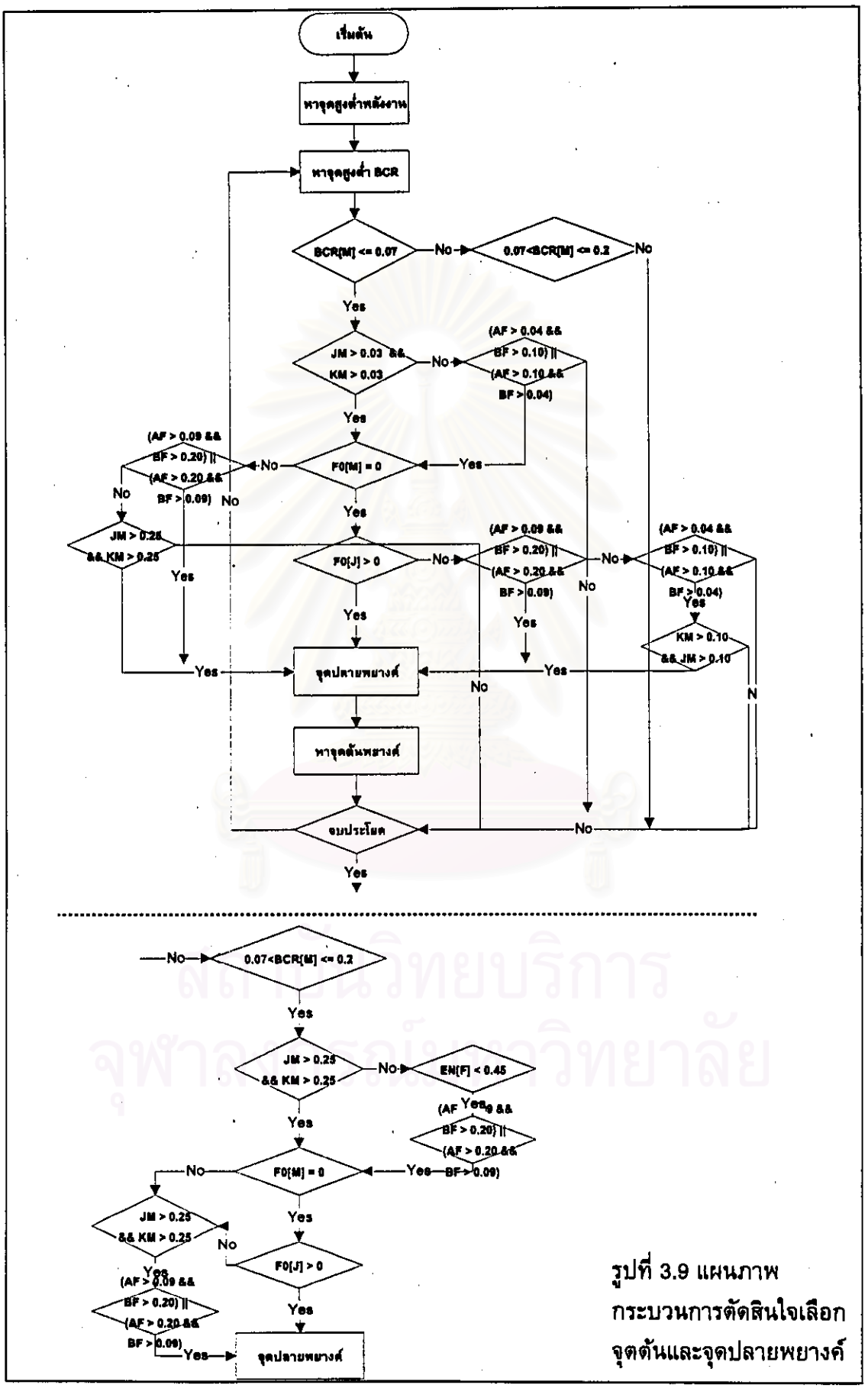
วิธี EBF เป็นวิธีที่นำมาใช้ในการหาขอบเขตพยางค์ เมื่อคุณลักษณะที่ใช้เป็นค่าพลังงาน อัตราการตัดผ่านระดับกำหนด และความถี่มูลฐานของสัญญาณเสียงพูด ซึ่งมีกระบวนการและขั้นตอนในการหาขอบเขตพยางค์ดังรูปที่ 3.8 ขั้นตอนการตัดสินใจเลือกจุดปลายและจุดต้นพยางค์เป็นขั้นตอนเดียวที่ต่างจากวิธี EBR ซึ่งมีวิธีการดังแสดงในรูปที่ 3.9



รูปที่ 3.8 แผนภาพแสดงขั้นตอนและรายละเอียดของวิธี EBF

เมื่อตรวจหาจุดต้นและจุดปลายประโยคได้แล้ว ในขั้นตอนต่อมาทำการหาจุดสูงและจุดต่ำพลังงานซึ่งมีวิธีการเหมือนในวิธี EMN แล้วนำจุดสูงพลังงานมาใช้ในการหาจุดต่ำและจุดสูง BCR

ณ จุดแบ่งพยางค์ค่าพลังงานและอัตราการตัดผ่านระดับกำหนดควรจะมีค่าต่ำ และไม่ควรจะมีค่าความถี่มูลฐาน แต่ในความเป็นจริง ที่คุณลักษณะพลังงาน อัตราการตัดผ่านระดับกำหนดและความถี่มูลฐาน อาจไม่สอดคล้องกันตามที่คาดไว้ ดังนั้นจึงมีกรณีในการตรวจสอบตามกระบวนการในรูปที่ 3.9 ซึ่ง BCR[M] คือ อัตราการตัดผ่านระดับกำหนดที่จุด M JM คือ ผลต่างระหว่างจุด J กับจุดต่ำ M AF คือผลต่างระหว่างจุด A กับจุด F F_0[M] คือ ความถี่มูลฐานที่จุด M และรูปที่ 3.10 (a-c) แสดงแผนภูมิเส้นของระดับพลังงาน อัตราการตัดผ่านระดับกำหนด และความถี่มูลฐาน



รูปที่ 3.9 แผนภาพกระบวนการตัดสินใจเลือกจุดต้นและจุดปลายพยางค์

ขั้นตอนในการตัดสินใจเลือกจุดต้นและจุดปลายพยางค์ ของวิธี EBF มีดังนี้

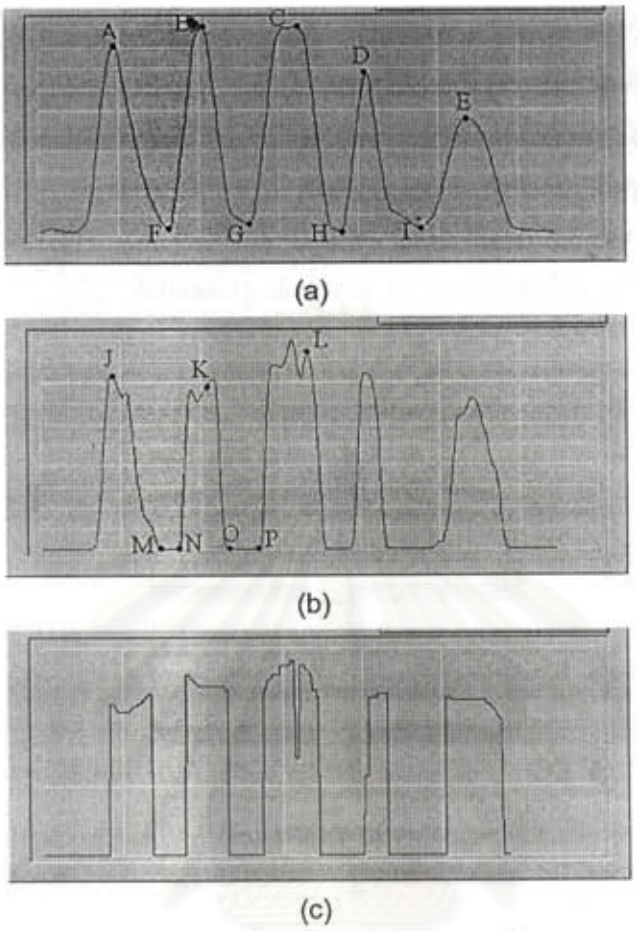
ขั้นแรก ตรวจสอบอัตราการตัดผ่านระดับกำหนด อัตราการตัดผ่านระดับกำหนดที่จุดแบ่งพยางค์ควรมีค่าต่ำ หากไม่เป็นไปตามที่คาดไว้ จึงกำหนดระดับของอัตราการตัดผ่านระดับกำหนดไว้ 2 ระดับ เพื่อกำหนดค่าคงที่ต่าง ๆ ที่ต่างกัน ในการตรวจสอบเพื่อหาจุดแบ่งพยางค์

ขั้นที่สอง ตรวจสอบผลต่างอัตราการตัดผ่านระดับกำหนดระหว่างจุดสูงBCR กับจุดต่ำBCR ผลต่างอัตราการตัดผ่านระดับกำหนดระหว่างจุดสูงและจุดแบ่งพยางค์ควรมีค่าสูง หากผลต่างมีค่าน้อยกว่าค่าที่กำหนด ไปตรวจสอบผลต่างพลังงานที่จุดสูงและจุดต่ำ

ขั้นที่สาม ตรวจสอบความถี่มูลฐานที่จุดแบ่งพยางค์ เพื่อยืนยันความถูกต้องในการตัดสินใจเลือกจุดแบ่งพยางค์จากคุณลักษณะอัตราการตัดผ่านระดับกำหนด ณ จุดแบ่งพยางค์ จะไม่มีค่าความถี่มูลฐาน หากมีความถี่มูลฐานอาจเกิดจากการประมาณค่าความถี่มูลฐานผิดพลาด หรืออาจเกิดจากคุณลักษณะของเสียงที่มีลักษณะเป็นรายคาบ จึงต้องไปตรวจสอบความต่างของระดับพลังงาน หากระดับพลังงานมีค่าต่ำกว่าที่กำหนด อาจเกิดจากคุณสมบัติของเสียงพูดที่ให้พลังงานน้อย จึงไปตรวจสอบที่อัตราการตัดผ่านระดับกำหนดอีกครั้งหนึ่ง

ขั้นที่สี่ ตรวจสอบความถี่มูลฐานที่จุดสูง BCR ก่อนที่จะมีจุดแบ่งพยางค์ จำเป็นต้องมีสัญญาณเสียงมาก่อน ดังนั้น ณ จุดที่มีค่าอัตราการตัดผ่านระดับกำหนดสูง จึงต้องมีค่าความถี่มูลฐาน หากไม่มีความถี่มูลฐาน อาจเกิดจากการประมาณความถี่มูลฐานผิดพลาด จึงต้องไปตรวจสอบที่ค่าความต่างของอัตราการตัดผ่านระดับกำหนด

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 3.10 (a) แผนภูมิเส้นแสดงจุดสูงต่ำพลังงาน (b) จุดสูงต่ำ BCR (c) ความถี่มูลฐาน

4. วิธีฟังก์ชันการประมาณ

วิธีฟังก์ชันการประมาณเป็นวิธีการทางสถิติ นำมาใช้ในการตัดสินใจหาขอบเขตพยางค์ เมื่อคุณลักษณะที่ใช้คือพลังงาน รูปที่ 3.11 แสดงขั้นตอนของวิธีฟังก์ชันการประมาณ ซึ่งมี 5 ขั้นตอน ดังนี้

ขั้นตอนที่ 1 หาจุดสูงจุดต่ำพลังงาน ซึ่งเป็นวิธีเดียวกับการหาจุดสูงต่ำในวิธี EMN นำจุดต่ำพลังงานไปใช้เป็นข้อมูลในการบ่งชี้ตำแหน่งจุดแบ่งพยางค์ ในการคำนวณทางสถิติต่อไป

ขั้นตอนที่ 2 คำนวณค่าการประมาณ (d) ซึ่งเป็นการประมาณเชิงเส้น ตำแหน่งที่มีค่า d ค่ามากที่สุดคือจุดแบ่งพยางค์ ดังแสดงในรูปที่ 3.12

$d = \bar{a}^T \cdot \bar{x}$

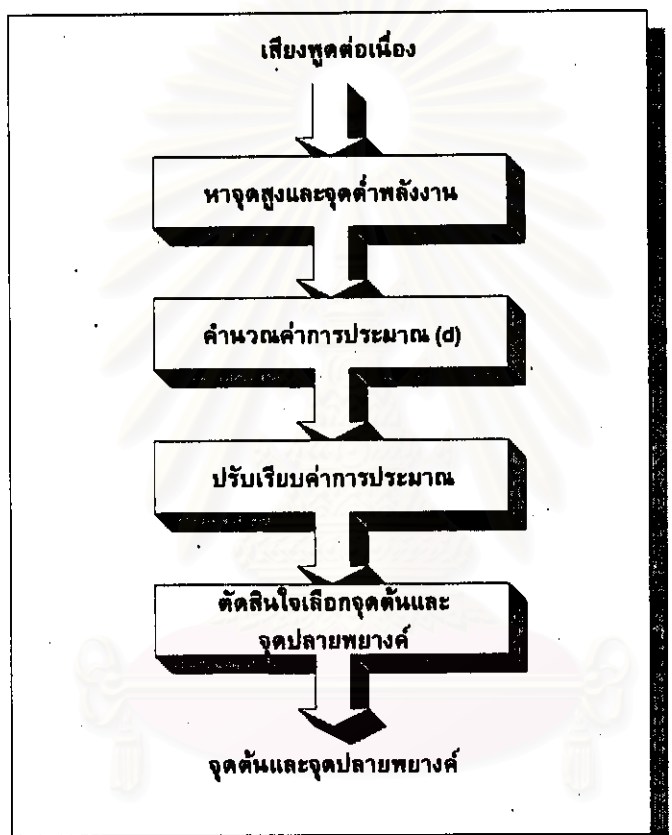
$\bar{x}^T = \{x_1, x_2, .., x_L\}$ $\bar{x}$ คือเวกเตอร์พลังงานเสียงพูด และ L คือจำนวนกรอบเสียงพูด

$\bar{a} = \bar{R}^{-1} \cdot \bar{\rho}$, $\bar{a}$ คือเวกเตอร์สัมประสิทธิ์การประมาณ

$\vec{R} = E[\vec{x} \cdot \vec{x}^T]$ $\vec{R}$ คือการคำนวณค่าคาดหวัง (Expectation Value) เวกเตอร์ของพลังงานเสียงพูด

$\vec{\rho} = E[z \cdot \vec{x}]$ $\vec{\rho}$ คือการคำนวณค่าคาดหวังเวกเตอร์ของพลังงานเสียงพูดที่มีขอบเขตพยางค์อยู่กึ่งกลางกรอบเสียงพูด

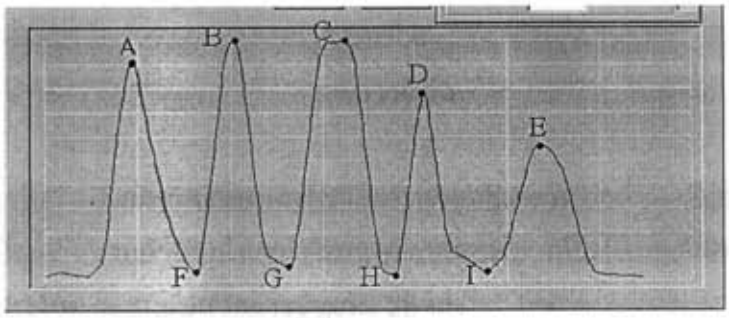
$$z = \begin{cases} 1, & \textit{boundary lie in the middle of the window} \\ 0, & \textit{boundary does not lie in the middle of the window} \end{cases}$$



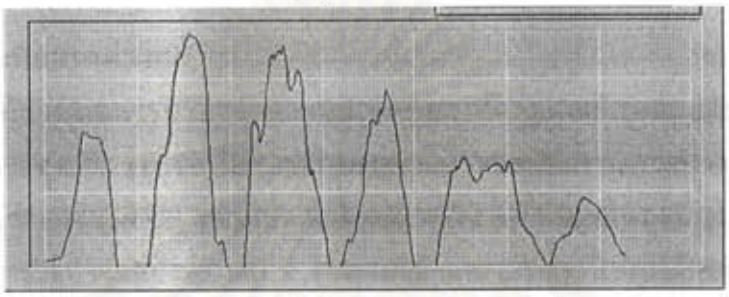
รูปที่ 3.11 แผนภาพแสดงขั้นตอนและรายละเอียดของวิธีฟังก์ชันการประมาณ

ขั้นตอนที่ 3 ปรับเรียบค่าการประมาณ จำเป็นต้องมีการปรับเรียบเพื่อป้องกันการมีค่าสูงจำนวนมาก ซึ่งต้องเลือกขนาดให้มีความเหมาะสม เพื่อจะได้จุดแบ่งพยางค์ที่มีความถูกต้อง

ขั้นตอนที่ 4 ตัดสินใจเลือกจุดต้นและจุดปลายพยางค์ จุดแบ่งพยางค์คือ จุดที่มีค่าการประมาณสูงสุดภายในระยะเวลาพยางค์ และค่าการประมาณที่จุดสูงสุดภายในระยะเวลาพยางค์ต้องมีค่ามากกว่า 0.30 (เป็นค่าที่ได้จากการทดลอง) จุดปลายพยางค์แรกคือจุดที่อยู่ห่างจากจุดแบ่งพยางค์ไปทางซ้ายด้วยระยะเวลา 1/10 ของระยะเวลาพยางค์ ส่วนจุดต้นพยางค์ที่สองคือ จุดที่อยู่ห่างจากจุดแบ่งพยางค์ไปทางขวาด้วยระยะเวลา 1/10 ของระยะเวลาพยางค์



(a)



(b)

รูปที่ 3.12 (a) แผนภูมิเส้นระดับพลังงาน (b) แผนภูมิเส้นค่าการประมาณ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย