บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงทฤษฎีทางสัทศาสตร์และภาษาศาสตร์ที่อธิบายถึงการกำเนิดเสียงพูด คุณลักษณะของเสียงพูด และกล่าวถึงกรรมวิธีทางด้านการวิเคราะห์สัญญาณที่จะนำมาวิเคราะห์หาองค์ประกอบและคุณสมบัติที่สำคัญของเสียงพูด เพื่อที่จะนำผลของการวิเคราะห์เหล่านั้นเก็บไว้เป็นฐานข้อมูลหน่วยเสียง เพื่อนำฐานข้อมูลหน่วยเสียงนี้มาสังเคราะห์เสียงต่อไป

งานวิจัยที่เกี่ยวข้อง

งานที่ผ่านมาในอดีตมีงานของหลายคนซึ่งก็ใช้เทคนิคที่ต่างกันและชนิดของหน่วยเสียงที่ต่างกัน แต่ส่วนมากเป็นงานที่เก็บองค์ประกอบของข้อมูลเป็นหน่วยเสียงและนำมาสังเคราะห์เสียงกลับในภายหลัง

Charpentier และ Moulines ได้เสนอกรรมวิธี PSOLA (Pitch-Synchronous Overlap-Add) ซึ่งใช้ในการเปลี่ยนท่วงทำนองของหน่วยเสียงแบบหน่วยคู่เสียงและหน่วยอัฒพยางค์ (Charpentier and Moulines 1989) เป็นกรรมวิธีที่เกี่ยวกับการเปลี่ยนแปลงระยะพิทช์ของเสียงเสียใหม่เพื่อให้เสียงสังเคราะห์มีท่วงทำนองใหม่ โดยหาตำแหน่งของพิทช์ของเสียงต้นแบบเตรียมเอาไว้จากนั้นจะมาเลื่อนระยะห่างของพิทช์เหล่านี้ให้มีท่วงทำนองเป็นไปตามที่ต้องการ

Luksaneeyanawin (1995) ได้นำกรรมวิธี PSOLA มาใช้ในการสังเคราะห์ซ้ำเสียงวรรณยุกต์ภาษาไทยโดยใช้หน่วยพยางค์ต้นแบบที่มีวรรณยุกต์เอก โดยการปรับระยะพิทช์ของเสียงโดยใช้กรรมวิธี PSOLA ให้เป็นไปตามรูปแบบค่าความถี่หลักมูลของแต่ละวรรณยุกต์ และได้ทำการทดสอบการรับฟังสองแบบคือการทดสอบการระบุเสียงวรรณยุกต์ของกลุ่มตัวอย่างกับการทดสอบทัศนคติของกลุ่มตัวอย่างต่อคุณภาพเสียง ซึ่งผลที่ได้จากการทดสอบจะให้ข้อเสนอแนะว่าหน่วยต้นแบบที่เป็นวรรณยุกต์ของเสียงสระหน่วยใดที่เหมาะจะนำมาสังเคราะห์เสียงวรรณยุกต์ ทำให้ฐานข้อมูลหน่วยเสียงลดจำนวนลงเนื่องจากไม่ต้องเก็บหน่วยเสียงจากทุกวรรณยุกต์

2.1 ทฤษฎีทางด้านภาษาศาสตร์

เสียงที่มนุษย์ใช้ในภาษาพูดนั้นจะมีลักษณะที่แตกต่างกันมากบ้างน้อยบ้าง อย่างไรก็ตามจะมีลักษณะที่สำคัญบางประการ ซึ่งปรากฏในเสียงพูดหลายเสียงซึ่งอาจจะเรียกได้ว่าเป็นลักษณะร่วมของเสียงพูด ลักษณะร่วมของเสียงพูดมีหลายอย่างดังนี้ (กาญจนา นาคสกุล 2524)

หอสมุดกลาง สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

2.1.1 ความก้องหรือไม่ก้องของเสียง

เกิดจากเส้นเสียงซึ่งมีลักษณะเป็นแผ่นกล้ามเนื้อ 2 แผ่นวางขวางอยู่ตรงกล่องเสียงตรงปลายท่อหลอดลม ในเวลาที่ไม่พูดเส้นเสียงจะอยู่ห่างจากกันเปิดช่องระหว่างเส้นเสียงหรือช่องคอหอย (Glottis) ให้ลมหายใจผ่านเข้าออกได้โดยสะดวก เรียกเส้นเสียงในตำแหน่งนี้ว่าเส้นเสียงเปิด ในขณะที่ออกเสียงพูดบางครั้งเส้นเสียงอาจเปิดกว้างกว่าเวลาหายใจตามปกติก็ได้ มีเสียงพูดของคนหลายเสียงซึ่งเกิดขึ้นในขณะที่เส้นเสียงเปิด เรียกเสียงเหล่านี้ว่าเสียงไม่ก้อง (Voiceless sounds) เช่นเสียง พ ท ป ส ในภาษาไทย

ในการออกเสียงพูดบางเสียง เส้นเสียงจะถูกดึงเข้ามาใกล้กันจนเกือบจะปิดช่องทางลมจนสนิท ตำแหน่งนี้เรียกว่าเส้นเสียงปิด ถ้าออกเสียงพูดในขณะที่เส้นเสียงปิด ลมที่ดันขึ้นมาจากปอดจะทำให้เส้นเสียงทั้งสองสั่น ลมที่ออกมาไม่สะดวกเพราะต้องบีบตัวผ่านช่องแคบ ๆ เป็นจังหวะจึงทำให้เกิดเป็นเสียงขึ้น เรียกว่าเสียงก้อง (Voiced sounds) เช่นพยัญชนะ ม น ง

ลักษณะของเสียงก้องและเสียงไม่ก้องนี้สามารถสังเกตได้จากรูปคลื่นเสียง ถ้าเป็นเสียงก้องเช่นเสียงสระต่าง ๆ ในภาษาไทยคลื่นเสียงจะมีลักษณะเป็นคลื่นเสียงแบบกึ่งรายคาบซึ่งจะเห็นรูปคลื่นที่มีรูปร่างคล้าย ๆ กันเกิดซ้ำกัน ในกรณีของเสียงไม่ก้องหรือเสียงพยัญชนะส่วนมากในภาษาไทยคลื่นเสียงที่ได้จะมีรูปแบบไม่ซ้ำเดิม เช่นในเสียง ส คลื่นเสียงที่เห็นจะมีรูปร่างไม่เป็นระเบียบ

2.1.2 ความยาวของเสียง

เสียงพูดบางเสียงอาจจะเปล่งออกมาติดต่อกันได้นาน เช่นเสียงสระ เสียงพยัญชนะนาสิก พยัญชนะเสียดแทรก เป็นต้น แต่บางเสียงก็ไม่สามารถจะออกได้นาน เช่น เสียงพยัญชนะระเบิดไม่ก้องเป็นต้น ความยาวของเสียงอาจมีได้หลายขนาดแต่โดยทั่ว ๆ ไป จะพูดถึงเพียง 2 ขนาดคือ สั้นและยาว ในภาษาไทยเสียงพูดซึ่งจะต้องพูดถึงความสั้นยาวก็มีเพียงเสียงสระเท่านั้น

ถ้าดูจากคลื่นเสียง คลื่นเสียงที่มีความยาวมากจะมีคาบเวลาในการเกิดนานกว่าคลื่นเสียงที่มีความยาวน้อย ตัวอย่างที่เห็นได้ชัดคือสระเสียงสั้นและสระเสียงยาวซึ่งสระเสียงยาวจะมีคาบเวลาในการเกิดนานกว่าสระเสียงสั้น

2.1.3 ระดับเสียงสูงต่ำ (Pitch)

เสียงจะมีระดับเสียงสูงต่ำอยู่ที่ความถี่เสียงถ้าความถี่ต่ำเสียงที่ออกมาจะต่ำ ความถี่ของเสียงหมายถึงอัตราการสั่นสะเทือนของอุปกรณ์กำเนิดของเสียง ในกรณีของเสียงพูดอวัยวะส่วนที่ทำให้เสียงมีระดับสูงหรือต่ำคือเส้นเสียง การที่เส้นเสียงเกิดการสั่นจะทำให้เกิดคลื่นเสียงแบบเป็นรายคาบเกิดขึ้นซึ่งระดับเสียงสูงต่ำนี้ขึ้นอยู่กับความถี่ของคลื่นเสียงที่เกิดจากการสั่นของเส้นเสียง ค่าความถี่

ของเส้นเสียงนี้สามารถวิเคราะห์หาค่าออกมาได้โดยวิธีการทางคณิตศาสตร์ หรือถ้าดูจากรูปคลื่นจะสังเกตได้จากคลื่นเสียงแบบเป็นรายคาบที่เกิดซ้ำกันของเสียงสระ เสียงใดที่มีคาบเวลาของการครบรอบน้อยเสียงนั้นจะมีระดับเสียงสูง และตรงกันข้ามเสียงใดที่มีคาบเวลาของการครบรอบมากเสียงนั้นจะมีระดับต่ำ ถ้าดูรูปคลื่นเสียงของผู้หญิงและคลื่นเสียงของผู้ชายเปรียบเทียบกันในเสียงสระเดียวกันจะเห็นได้ชัดว่า ในเสียงของผู้ชายจะมีคาบเวลามากกว่าเสียงของผู้หญิง

2.1.4 ความดัง

ความดัง (Loudness) ขึ้นอยู่กับปริมาณของลมที่ผู้พูดเปล่งออกมาในช่วงหนึ่ง ๆ รวมกับคุณลักษณะประจำตัวของเสียงและการลงน้ำหนักเสียง ค่าความดังนี้คือค่าแอมปลิจูดของคลื่นเสียงค่าแอมปลิจูดของคลื่นเสียงที่ได้ขึ้นอยู่กับแรงลมที่เปล่งออกมา คลื่นเสียงช่วงที่มีแอมปลิจูดมากจะมีความดังมาก คลื่นเสียงช่วงที่มีแอมปลิจูดน้อยจะมีความดังน้อยเพราะเนื่องจากเสียงมีการเปลี่ยนแปลงไปตามเวลาและการเน้นเสียง ดังนั้นค่าแอมปลิจูดของคลื่นเสียงจะมีค่าไม่เท่ากันและจะเปลี่ยนแปลงไปตลอดเสียง

2.1.5 การลงน้ำหนัก

การลงน้ำหนัก (Stress) หมายถึง การออกเสียงพยางค์ใดพยางค์หนึ่งให้ดังเน้นมากหรือน้อยกว่าพยางค์อื่นที่อยู่ข้างเคียง การลงน้ำหนักพยางค์จึงขึ้นอยู่กับกำลังแรงที่ใช้ในการเปล่งเสียงของพยางค์แต่ละพยางค์ การลงน้ำหนักนี้ก็คือการเพิ่มหรือลดค่าแอมปลิจูดในคลื่นเสียงนั่นเอง เพราะถ้าใช้แรงลมในการเปล่งเสียงมากจะทำให้แอมปลิจูดของคลื่นเสียงมีค่ามาก ถ้าลงน้ำหนักเบาหรือไม่ค่อยเน้นเสียงซึ่งใช้แรงลมน้อยทำให้คลื่นเสียงที่ได้มีค่าแอมปลิจูดไม่มาก

2.1.6 เสียงสระ

หน่วยเสียงที่สำคัญในภาษาทุกภาษาคือหน่วยเสียงสระ เสียงสระเป็นเสียงที่เกิดจากลมซึ่งผ่านเส้นเสียงในตำแหน่งปิดเกือบสนิท และลมที่ต้องดันตัวออกมาทำให้เส้นเสียงสั่นเกิดเป็นเสียงดังขึ้นเรียกว่า เสียงก้อง เสียงที่เรียกว่าเป็นเสียงสระจะต้องถูกเปล่งออกมาทางปากโดยที่ไม่มีอวัยวะส่วนใดในปากมาปิดกั้นทางลมไว้ แต่อวัยวะในปากอาจจะอยู่ในท่าและตำแหน่งต่าง ๆ ที่ทำให้โพรงปากมีลักษณะต่างกันได้หลายแบบ ลมที่ผ่านออกมาจึงเกิดเป็นเสียงต่าง ๆ กัน การห่อริมฝีปากหรือไม่ห่อริมฝีปาก การยกลิ้นส่วนหนึ่งส่วนใด การยกลิ้นสูงต่ำต่าง ๆ ย่อมมีส่วนในการทำให้เกิดเสียงสระต่าง ๆ กันทั้งนั้น หน่วยเสียงสระในภาษาไทยมีทั้งหมด 21 หน่วย เป็นหน่วยเสียงสระเดี่ยวเสียงสั้น 9 หน่วย หน่วยเสียงสระเดียวยาว 9 หน่วย และหน่วยเสียงสระประสม 3 หน่วย

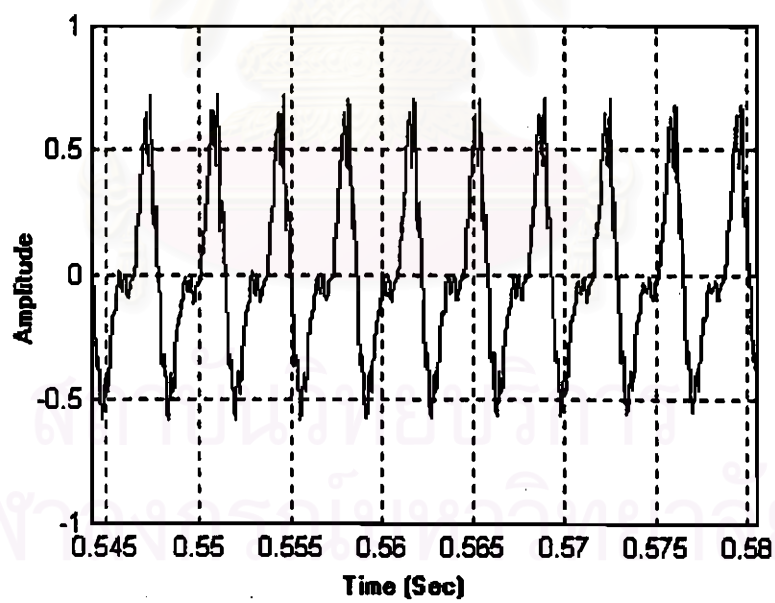
หน่วยเสียงสระเดี่ยวสั้น 9 หน่วยมี ิ, เ-ะ, แ-ะ, ึ, เ-อะ, -ะ, ุ, โ-ะ, เ-าะ

หน่วยเสียงสระเดี่ยวยาว 9 หน่วยมี ี, เ-, แ-, ื, เ-อ, -า, ู, โ-, -อ

หน่วยเสียงสระประสม 3 หน่วยมี มีเสียงย่อยหน่วยละ 2 เสียง เป็นสระสั้นและสระยาวคือ เ-ียะ, เ-ีย, เ-อะ, เ-อ, -ัวะ, -ัว

เสียงสระเป็นเสียงก้องที่ดังกระจายไปได้ไกลกว่าเสียงพยัญชนะ เสียงสระจึงเป็นแกนของพยางค์ซึ่งทำให้เสียงอื่น ๆ ในพยางค์นั้นได้ยินได้ หน่วยเสียงสระในภาษาไทยสามารถเกิดร่วมกับวรรณยุกต์ได้ทุกหน่วย แต่เกิดตามหลังหน่วยเสียงพยัญชนะต้นและนำหน่วยเสียงพยัญชนะสะกดไม่ได้ทุกหน่วย

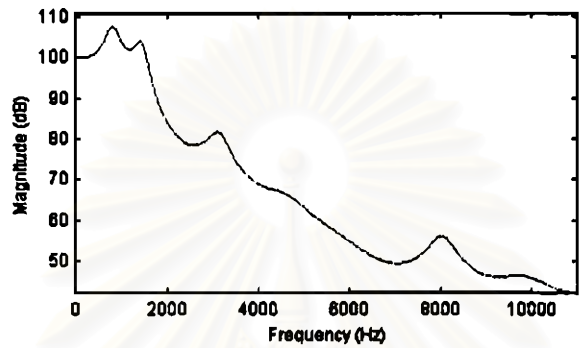
เนื่องจากเสียงสระเป็นเสียงก้อง เกิดจากการสั่นของเส้นเสียงผ่านอวัยวะกล่อมเกลาเสียงในช่องทางเดินเสียงแล้วออกจากปากโดยไม่มีอวัยวะอื่นใดมาขวางทางออกของเสียง ถ้าดูรูปคลื่นของเสียงสระจะพบว่าเป็นรูปคลื่นที่เป็นรายคาบ โดยมีคลื่นเสียงที่มีความถี่หลักมูลเท่ากับอัตราที่เส้นเสียงสั่น และเนื่องจากเกิดการกำทอนจากช่องทางเดินเสียงคือช่องปากและลำคอทำให้มีค่าความถี่ฮาร์มอนิกส์อื่นผ่านออกมาได้มากน้อยต่าง ๆ กันปนเข้ากับคลื่นเสียงจากเส้นเสียงด้วย ทำให้เกิดเป็นคลื่นเสียงที่ซับซ้อนดังรูปที่ 2.1



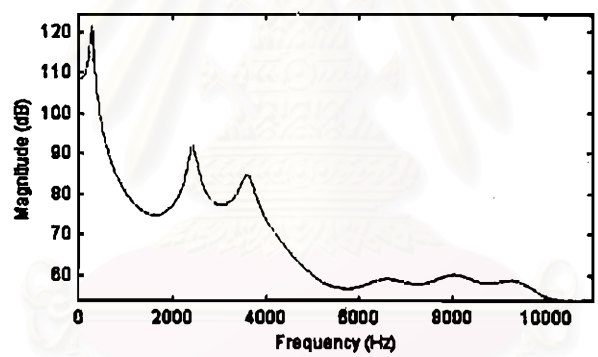
รูปที่ 2.1 แสดงรูปคลื่นเสียงของสระ "อี"

เสียงสระแต่ละตัวจะมีลักษณะเฉพาะตัวของความถี่เสียงที่ต่างกันออกไป ถ้านำสระต่าง ๆ มาวิเคราะห์หาองค์ประกอบทางความถี่ดู จะพบว่าสระต่าง ๆ จะมีรูปแบบของความถี่ฟอร์แมนท์ (Formant Frequency) ที่ต่างกันออกไป เนื่องจากรูปร่างของช่องทางเดินเสียงนี้จะเป็นตัวกำหนด

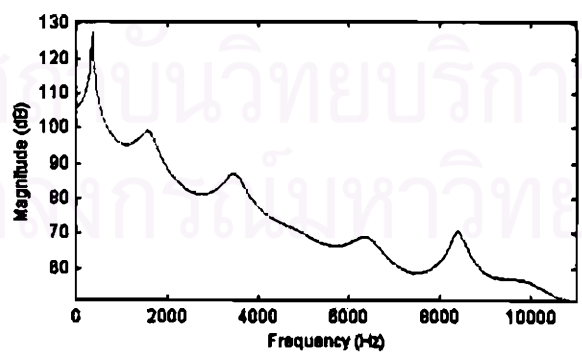
ความถี่ที่จะเกิดการกำทอน และรูปร่างของช่องทางเดินเสียงนี้จะเปลี่ยนไปตามแต่ละสระ นั่นคือจะทำให้ค่าความถี่ฟอร์แมนท์ของสระต่าง ๆ มีค่าต่างกัน รูปที่ 2.2 ถึงรูปที่ 2.10 จะแสดงสเปกตรัมของสระเสียงยาวในภาษาไทยซึ่งแต่ละยอดของสเปคตรัมจะบอกถึงค่าความถี่ฟอร์แมนท์ต่าง ๆ ของแต่ละสระ



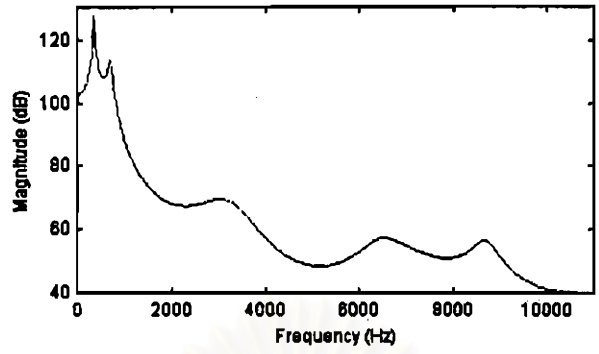
รูปที่ 2.2 แสดงสเปคตรัมของสระอา



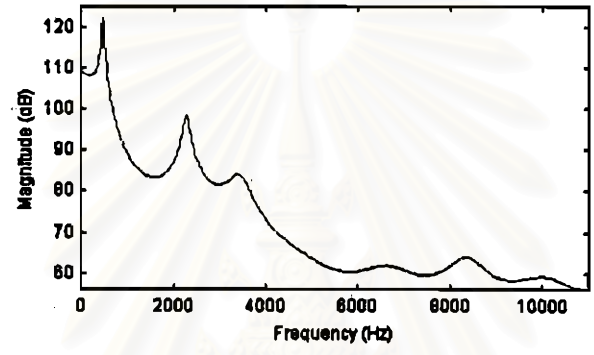
รูปที่ 2.3 แสดงสเปคตรัมของสระอี



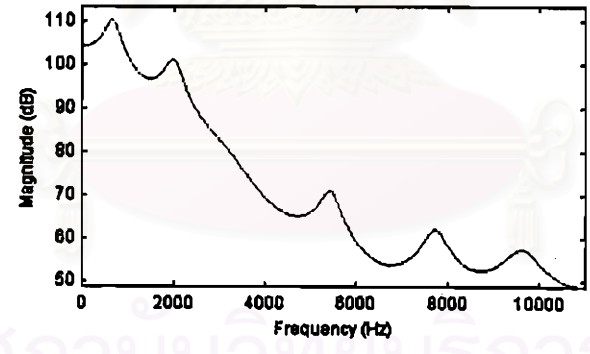
รูปที่ 2.4 แสดงสเปคตรัมของสระอือ



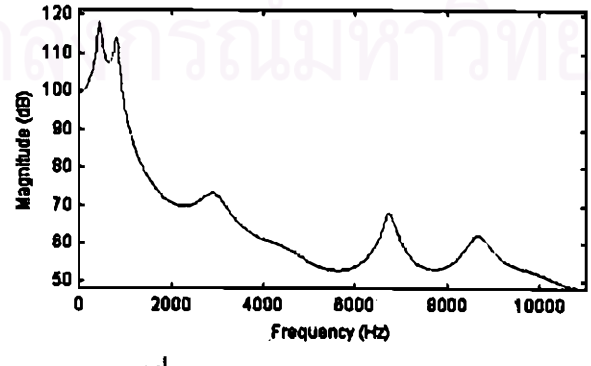
รูปที่ 2.5 แสดงสเปคตรัมของสระอู



รูปที่ 2.6 แสดงสเปคตรัมของสระเอ

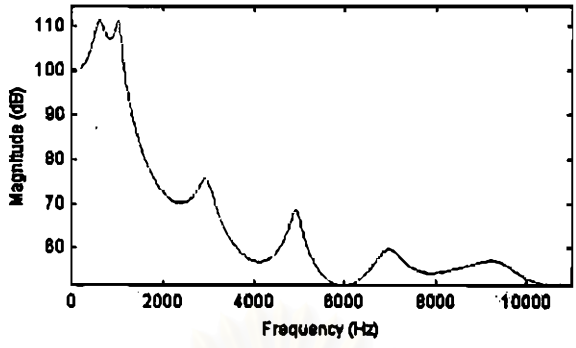


รูปที่ 2.7 แสดงสเปคตรัมของสระแอ

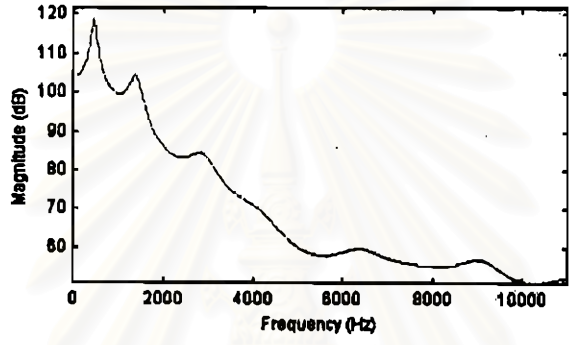


รูปที่ 2.8 แสดงสเปคตรัมของสระโอ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 2.9 แสดงสเปคตรัมของสระออ



รูปที่ 2.10 แสดงสเปคตรัมของสระเออ

2.1.7 เสียงพยัญชนะ

เสียงสำคัญที่ใช้ในภาษาไทยอีกพวกหนึ่งคือเสียงพยัญชนะ เป็นเสียงที่เกิดจากลมซึ่งผ่านเส้นเสียง แล้วมาถูกดัดแปลงด้วยอวัยวะออกเสียงส่วนต่าง ๆ ในปาก ทำให้เกิดเสียงขึ้น เสียงพยัญชนะมีหลายประเภทมีลักษณะการออกเสียงแตกต่างกันหลายแบบ ความแตกต่างของเสียงจะมีความสำคัญยิ่งขึ้นเมื่อความแตกต่างนั้นทำให้ความหมายของคำในภาษาต่างกันหรือเรียกได้ว่า เป็นความแตกต่างที่ทำให้เสียงสองเสียงทำหน้าที่เป็นสองหน่วยเสียงต่างกันในภาษานั้นเพื่อให้ทราบว่าในภาษามีหน่วยเสียงกี่หน่วย หน่วยเสียงพยัญชนะในภาษาไทยมี 21 หน่วยเสียง

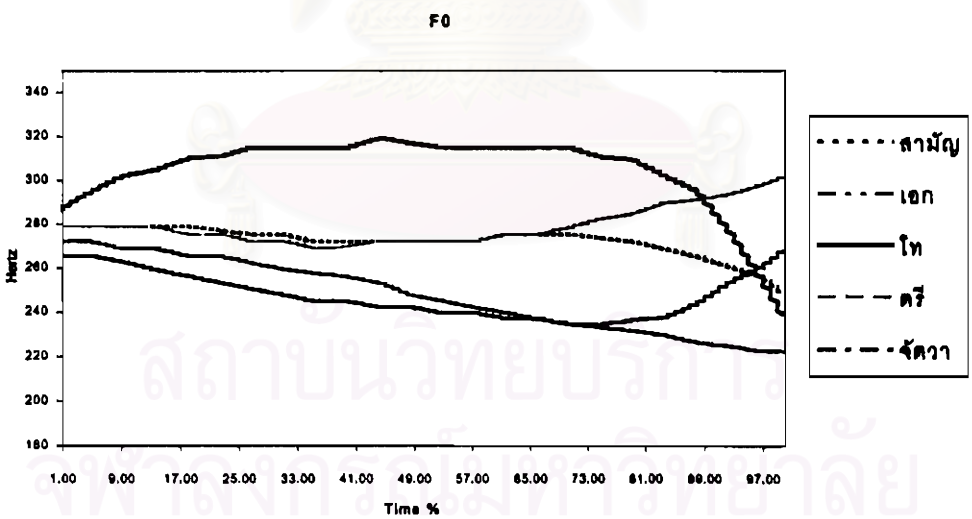
หน่วยเสียงพยัญชนะในภาษาไทยจะปรากฏในตำแหน่งต่าง ๆ ได้ 4 ตำแหน่งคือ

1. เกิดต้นคำ นำหน้าเสียงสระในพยางค์หนึ่ง ๆ เป็นหน่วยเสียงพยัญชนะต้น
2. เกิดนำเสียงพยัญชนะอื่นอีกเสียงหนึ่งในตำแหน่งต้นคำเรียกว่า พยัญชนะควบ
3. เกิดตามเสียงพยัญชนะอื่นในต้นคำ คือเกิดควบกับพวกที่ 2
4. เกิดตามหลังเสียงสระ เป็นเสียงพยัญชนะสะกด

2.1.8 เสียงวรรณยุกต์

ในการออกเสียงพูดในภาษาหนึ่ง ๆ นั้น ไม่ว่าผู้พูดจะเปล่งเสียงสระพยัญชนะเรียงติดต่อกันออกมาเป็นคำวลีหรือเป็นประโยคก็ตาม ในการออกเสียงนั้นผู้พูดอาจจะดัดแปลงเสียงสระเสียงพยัญชนะนั้น ๆ ให้อยู่ในระดับเสียง สูง ต่ำ หรือเปลี่ยนแปลงเสียงขึ้นลงต่าง ๆ ได้อีก เสียงสูงต่ำในภาษาพูดเกิดได้ด้วยการสั่นสะเทือนของเส้นเสียงในอัตราต่าง ๆ กัน แต่เสียงที่เปล่งออกมาในขณะที่เส้นเสียงสั่นนั้นต้องเป็นเสียงก้อง พยางค์ในคำพูดทั่วไปจะประกอบด้วยเสียงสระและพยัญชนะ สระเป็นเสียงก้องจึงเป็นเสียงที่ช่วยทำให้เกิดระดับสูงต่ำได้ โดยมองว่าระดับเสียงสูงต่ำของสระในพยางค์หนึ่ง ๆ ก็เป็นระดับเสียงสูงต่ำของพยางค์นั้นด้วย เพราะเมื่อเสียงสระซึ่งเป็นแกนของพยางค์มีระดับใดเสียงอื่น ๆ ในพยางค์นั้นก็ย่อมมีระดับเสียงนั้นตามไปด้วยเช่นกัน หรืออีกนัยหนึ่งคือเสียงสูงต่ำในภาษาคือการที่ค่าความถี่หลักมูลของเสียงสระมีการเปลี่ยนแปลงไปเวลา

ในภาษาไทยระดับเสียงสูงต่ำของคำเรียกว่า วรรณยุกต์ เป็นส่วนสำคัญที่ทำให้คำตั้งแต่ 2 คำขึ้นไปซึ่งมีส่วนประกอบอื่น ๆ คือพยัญชนะ สระ และตัวสะกดอย่างเดียวกันมีความหมายต่างกันได้ วรรณยุกต์ในภาษาไทยจึงจัดเป็นหน่วยเสียงเรียกว่าหน่วยเสียงวรรณยุกต์ หน่วยเสียงวรรณยุกต์จัดว่าเป็นหน่วยเสียงซ้อน เสียงวรรณยุกต์ในภาษาไทยมีลักษณะเค้าโครงและความถี่ของเสียงโดยประมาณดังรูป 2.11



รูปที่ 2.11 แสดงความถี่หลักมูลของวรรณยุกต์ทั้ง 5 เสียง

รูปที่ 2.11 นี้เป็นความถี่หลักมูลของสระเสียงยาว ในสระเสียงสั้นจะมีทางเดินของความถี่หลักมูลคล้าย ๆ เช่นนี้แต่คาบเวลาจะสั้นกว่า หน่วยเสียงวรรณยุกต์ทั้ง 5 หน่วยนี้สามารถจัดกลุ่มได้เป็น 2 กลุ่ม

2.1.8.1 **วรรณยุกต์คงระดับ**

วรรณยุกต์ซึ่งมีระดับความถี่ของเสียงค่อนข้างคงที่ตลอดพยางค์ จัดได้ว่ามีวรรณยุกต์คงระดับ ในการออกเสียงพูดนั้นโดยปกติต้นพยางค์และท้ายพยางค์มักจะไม่อยู่ในระดับเดียวกันทีเดียว ต้นพยางค์จะมีเสียงสูงกว่าดังกว่าเสมอ ในทางสัทศาสตร์ถือว่าระดับเสียงที่แตกต่างกันไปนี้เล็กน้อยเมื่อเทียบกับวรรณยุกต์อีกพวกหนึ่ง เสียงวรรณยุกต์คงระดับในภาษาไทยมี 3 หน่วยคือ

- หน่วยเสียงวรรณยุกต์ระดับต่ำ

  หน่วยเสียงวรรณยุกต์ระดับต่ำ (Low tone) คือวรรณยุกต์เอก วรรณยุกต์ระดับต่ำจะปรากฏในพยางค์ได้ทุกแบบ

- หน่วยเสียงวรรณยุกต์ระดับกลาง

  หน่วยเสียงวรรณยุกต์กลาง (Mid tone) คือวรรณยุกต์สามัญ วรรณยุกต์นี้มีระดับเสียงกลาง ๆ ในปลายพยางค์จะมีความถี่ลดลง วรรณยุกต์สามัญไม่ปรากฏในพยางค์ที่สะกดด้วยพยัญชนะกัก

- หน่วยเสียงวรรณยุกต์ระดับสูง

  หน่วยเสียงวรรณยุกต์ระดับสูง (High tone) คือวรรณยุกต์ตรี เสียงวรรณยุกต์นี้มีลักษณะเด่นที่มีระดับเสียงสูง โดยจะค่อย ๆ สูงขึ้นทีละน้อยจากต้นเสียง ในภาษาไทยหน่วยเสียงวรรณยุกต์ตรีไม่ปรากฏในพยางค์ที่ประสมด้วยสระเสียงยาวที่มีตัวสะกดเป็นเสียงหลัก

2.1.8.2 **วรรณยุกต์เปลี่ยนระดับ**

วรรณยุกต์เปลี่ยนระดับเป็นเสียงซึ่งมีระดับความถี่ของการออกเสียงเปลี่ยนแปลงมากในช่วงพยางค์หนึ่ง ๆ เช่น ต้นพยางค์ออกเสียงให้มีระดับสูงแล้วลดระดับเสียงลดอย่างเร็วไปสู่ระดับต่ำที่ปลายพยางค์ หรือต้นพยางค์เสียงมีระดับต่ำแล้วเสียงเพิ่มระดับเสียงอย่างรวดเร็วเป็นระดับสูงที่ปลายพยางค์ หรือเปลี่ยนสูงแล้วต่ำแล้วสูงอีกก็ได้ ในภาษาไทยวรรณยุกต์เปลี่ยนระดับมี 2 หน่วยคือ

- หน่วยเสียงวรรณยุกต์เปลี่ยนตก

  หน่วยเสียงวรรณยุกต์เปลี่ยนตก (Falling tone) คือเสียงวรรณยุกต์โท วรรณยุกต์โทจะไม่ปรากฏในพยางค์ที่มีสระเสียงสั้นและะสะกดด้วยพยัญชนะกัก

- หน่วยเสียงวรรณยุกต์เปลี่ยนขึ้น

  หน่วยเสียงวรรณยุกต์เปลี่ยนขึ้น (Rising tone) คือเสียงวรรณยุกต์จัตวา วรรณยุกต์จัตวาจะไม่ปรากฏในพยางค์ที่มีพยัญชนะสะกดเป็นเสียงกัก

2.1.9 พยางค์ในภาษาไทย

ในการพูดของคนมิได้เปล่งเสียงออกมาเป็นห้วง ๆ ทีละเสียง แต่จะเปล่งเสียงหลายเสียงติดต่อกันมาเป็นกลุ่ม ๆ เรียงกันเป็นคำพูด คำพูดแต่ละกลุ่มอาจจะมีจำนวนเสียงมากบ้างน้อยบ้างแล้วแต่ลักษณะของแต่ละภาษา ในแต่ละกลุ่มเสียงนั้นจะมีเสียงบางเสียงที่ดังเด่นกว่าเสียงอื่น ในการพูดประโยคหนึ่งหรือวรรคหนึ่งถ้ามีเสียงที่ดังเด่นกว่าเสียงอื่นเป็นจำนวนเท่าใด ก็จะฟังได้ว่าประโยคนั้นหรือวรรคนั้นมีจำนวนพยางค์เท่ากับจำนวนเสียงที่ดังเด่นนั้น ดังนั้นพยางค์จึงหมายถึงจำนวนเสียงที่ดังเด่นซึ่งปรากฏในกลุ่มเสียงที่เรียงกันเป็นคำพูด เสียงอื่น ๆ ที่อยู่ข้างเคียงก็จะประกอบเข้าเป็นส่วนของพยางค์

โดยทั่ว ๆ ไปคนเรามีความสามารถที่จะแยกพยางค์ในภาษาของตนได้โดยอาศัยความเคยชินเสียงที่เราได้ยินว่าออกมาครั้งหนึ่งาะมีกี่เสียงรวมกันก็ตามก็จะเป็นพยางค์หนึ่ง

พยางค์ที่เปล่งออกมาพยางค์หนึ่ง ๆ ไม่จำเป็นจะต้องมีความหมายกำกับเสมอไป แต่เมื่อพยางค์หรือกลุ่มของพยางค์ที่ประกอบกันขึ้นมาอย่างน้อยที่สุดมีความหมายและปรากฏได้โดยลำพังเมื่อนั้นพยางค์หรือกลุ่มของพยางค์นั้นก็จะเป็นคำของภาษา ดังนั้นคำในภาษาจึงหมายถึงกลุ่มเสียงที่เล็กที่สุดที่สามารถปรากฏได้โดยลำพังและมีความหมาย

กลุ่มเสียงที่ปรากฏว่าเป็นคำมีความหมายในภาษาไทย มีลักษณะต่าง ๆ กันหลายแบบ คำในภาษาไทยจึงจัดตามลักษณะของพยางค์เป็นได้หลายแบบดังนี้คือ

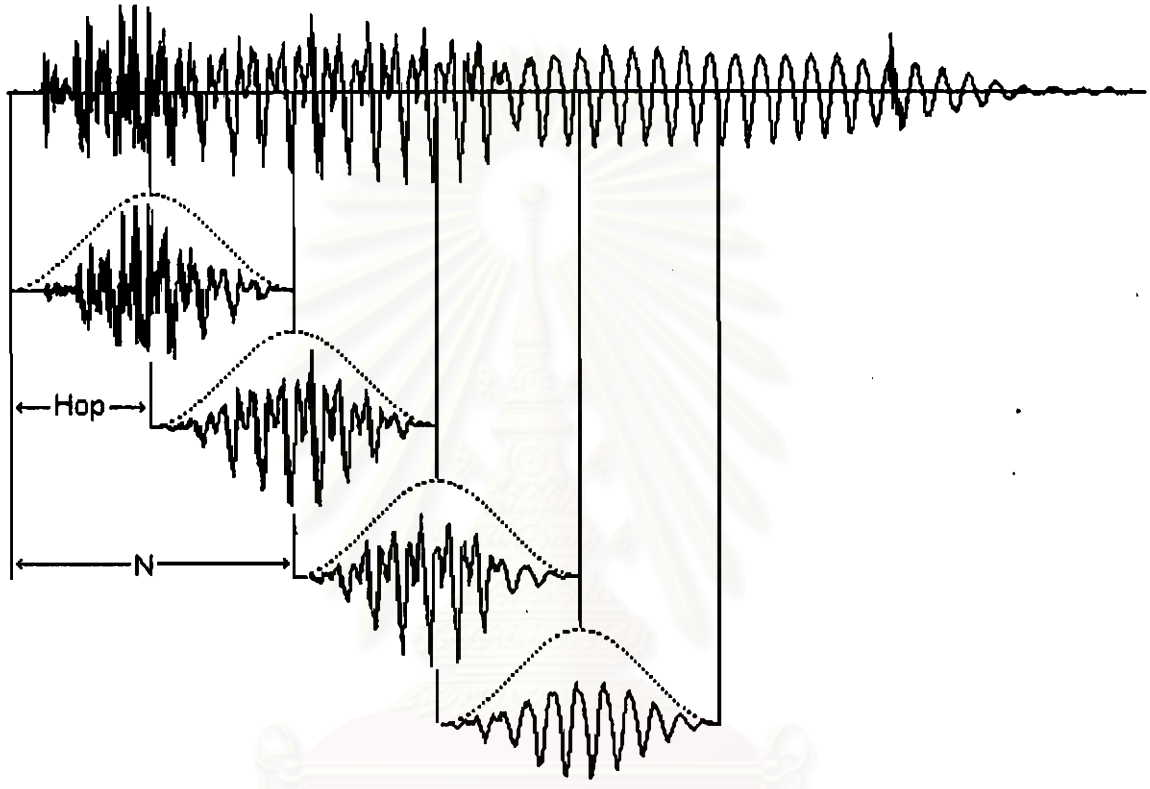
- คำพยางค์เดียว
- คำสองพยางค์
- คำหลายพยางค์

ในภาษาไทย พยางค์หนึ่ง ๆ จะต้องประกอบด้วยหน่วยเสียงพยัญชนะต้นอย่างน้อย 1 หน่วยเสียง ประกอบด้วยหน่วยเสียงสระ 1 หน่วยเสียง และหน่วยเสียงวรรณยุกต์ 1 หน่วยเสียง พยางค์จึงมีส่วนประกอบอย่างน้อย 3 หน่วยเสียง

2.2 ทฤษฎีการวิเคราะห์เสียง

เนื่องจากสัญญาณเสียงพูดจะมีคุณสมบัติแปรเปลี่ยนไปตามเวลา (Non Stationary) ดังนั้นในการประยุกต์ใช้งานกรรมวิธีประมวลผลสัญญาณเชิงเลขกับเสียงพูดจึงต้องแบ่งสัญญาณเสียงพูดออกเป็นส่วนย่อย ๆ เรียกว่ากรอบเสียงพูด (Speech Frame) โดยแต่ละกรอบเสียงพูดจะมีความยาวประมาณ 10-30 มิลลิวินาที ซึ่งถือได้ว่าสัญญาณเสียงพูดในแต่ละกรอบเสียงพูดไม่แปรเปลี่ยนตามเวลา จากนั้นจึงจะทำการประมวลผลสัญญาณเสียงพูดในแต่ละกรอบเสียงพูดได้

ในการแบ่งกรอบเสียงพูดถ้านำกรอบเสียงพูดมาวิเคราะห์โดยเรียงลำดับต่อ ๆ กัน จะเห็นว่าแอมปลิจูดบริเวณช่วงต่อระหว่างกรอบมีค่าน้อยอาจทำให้ข้อมูลสำคัญบริเวณขอบสูญหายไปได้ ดังนั้นในการแบ่งกรอบเสียงพูดจึงนิยมแบ่งให้กรอบเสียงพูดเหลื่อมกันเป็นบางส่วน เช่นในรูปข้างล่างให้กรอบเสียงพูดเหลื่อมกัน 50%



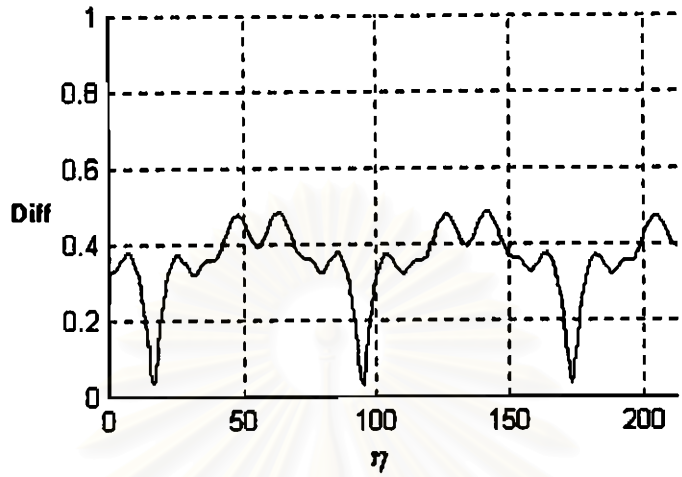
รูปที่ 2.12 แสดงการแบ่งกรอบเสียงพูดมาวิเคราะห์เป็นส่วนย่อย

2.2.1 กรรมวิธี AMDF (Average Magnitude Difference Function)

กรรมวิธี AMDF เป็นกรรมวิธีที่จะหาค่าความแตกต่างเฉลี่ยของแอมปลิจูดของสัญญาณซึ่งสามารถนำมาใช้หาค่าความถี่หลักมูลได้ (Deller, Proakis and Hansen, 1993) โดยนำสัญญาณมาสองสัญญาณโดยสัญญาณที่สองเป็นสัญญาณที่เกิดจากการเลื่อนทางเวลาของสัญญาณตัวแรก ในการที่เลื่อนสัญญาณไปแต่ละครั้ง จะทำการหาค่าเฉลี่ยของผลต่างระหว่างสัญญาณทั้งสองแล้ววาดออกมาเป็นกราฟดังรูปที่ 2.13

ขนาดกรอบสัญญาณที่จะนำมาหาค่าความถี่หลักมูลด้วยกรรมวิธีนี้จะต้องครอบคลุมสัญญาณมากกว่าสองคาบเสียงพูดขึ้นไป โดยที่ $s(n)$ คือสัญญาณเสียงพูดจากกรอบที่เลือกมา และ $s(n+\eta)$ คือสัญญาณตัวแรกที่ถูกเลื่อนทางเวลา

$$AMDF(\eta;m) = \frac{1}{N}\sum_{n=m-N+1}^{m}\left|s(n)-s(n+\eta)\right| \qquad (2.1)$$

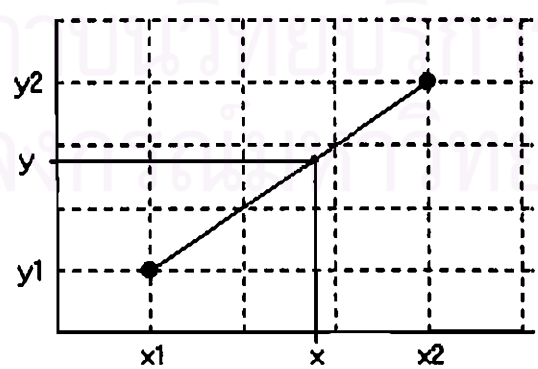


รูปที่ 2.13 แสดงผลลัพธ์จากกรรมวิธี AMDF

จุดต่ำสุดจุดแรกที่ปรากฏในกราฟคือเวลาที่ล้าหลังซึ่งสามารถนำมาคำนวณหาค่าความถี่หลักมูลได้ ค่า η ในกรรมวิธีนี้ไม่นิยมเริ่มจากค่า 0 แต่จะเริ่มที่ค่า ๆ หนึ่งเพราะเหตุว่าจะทำให้ง่ายต่อการหาจุดต่ำสุดจุดแรก

2.2.2 การประมาณค่าในช่วงแบบเชิงเส้น (Linear Interpolation)

การประมาณค่าในช่วงเชิงเส้นเป็นการประมาณค่าที่ไม่รู้ระหว่างจุด 2 จุดที่รู้ค่า โดยการลากเส้นตรงระหว่างจุด 2 จุดที่รู้ค่าจะสามารถคำนวณค่าที่อยู่ระหว่างจุดสองจุดบนเส้นตรงนี้ได้ (Kreyszig, 1988)



รูปที่ 2.14 แสดงการหาค่า y จากจุดสองจุด

จากรูปที่ 2.14 ถ้ารู้ค่า x1, x2, y1, y2 และต้องการรู้ว่าที่จุด x จะมีค่า y เท่าไร สามารถหาได้จากความสัมพันธ์

$$y = \left(\frac{y_2 - y_1}{x_2 - x_1}\right)(x - x_1) + y_1 \qquad (2.2)$$

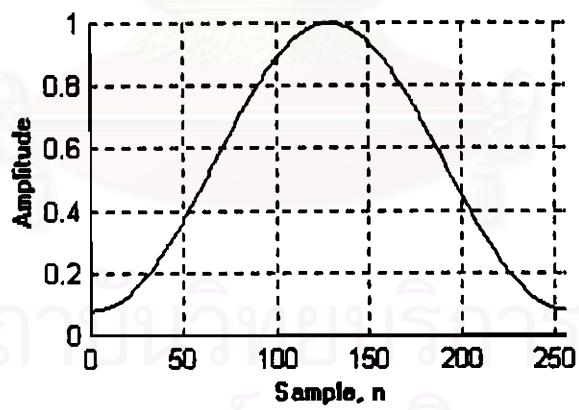
2.2.3 ฟังก์ชั่นหน้าต่าง (Window function)

ในการวิเคราะห์สัญญาณในช่วงเวลาสั้น ๆ หรือการแบ่งกรอบเสียงพูดจะต้องมีการตัดสัญญาณในช่วงเวลาสั้น ๆ จากสัญญาณทั้งหมดมาทำการวิเคราะห์ทีละส่วน การตัดสัญญาณช่วงสั้น ๆ ออกมานี้จะนำสัญญาณมาคูณกับฟังก์ชั่นหน้าต่างก่อนที่จะนำสัญญาณไปวิเคราะห์ เพื่อจะเฉลี่ยน้ำหนักให้ตรงขอบปลายทั้งสองข้างของสัญญาณมีการเปลี่ยนแปลงอย่างไม่ทันทีทันใด

ฟังก์ชั่นหน้าต่างมีหลายแบบ แต่ที่นิยมใช้กันมากในการวิเคราะห์เสียงพูดคือฟังก์ชั่นหน้าต่างแบบแฮมมิ่ง (Hamming window) ซึ่งมีนิยามดังนี้ (Rabiner and Schafer, 1978)

$$w(n) = 0.54 - 0.46\cos\left(\frac{2\pi n}{N-1}\right) \qquad (2.3)$$

เมื่อ $n = 0,1,\ldots,N-1$



รูปที่ 2.15 แสดงรูปหน้าต่างแบบแฮมมิ่ง

2.2.4 การแปลงฟูเรียร์แบบเต็มหน่วย (Discrete Fourier Transform)

การแปลงฟูเรียร์เป็นการวิเคราะห์สัญญาณในโดเมนความถี่ เพื่อใช้หาองค์ประกอบเชิงความถี่หรือที่เรียกว่าสเปคตรัมของสัญญาณ โดยสัญญาณที่จะนำมาวิเคราะห์จะต้องเป็นสัญญาณที่ถูกชักตัวอย่างมาและเป็นสัญญาณรายคาบ โดยมีนิยามดังนี้ (O'Shaughnessy, 1987)

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{\frac{-j2\pi kn}{N}} \qquad (2.4.1)$$

สมการ 2.4.1 เป็นการแปลงฟูเรียร์แบบเต็มหน่วยจำนวน N จุดกับสัญญาณรายคาบ การแปลงฟูเรียร์จะมีการแปลงกลับฟูเรียร์ซึ่งเป็นคู่การแปลงดังสมการ 2.4.2

$$x(n) = \frac{1}{N} \sum_{n=0}^{N-1} X(k) e^{\frac{j2\pi kn}{N}} \qquad (2.4.2)$$

การแปลงฟูเรียร์แบบเต็มหน่วยถูกใช้ในการคำนวณค่าสเปคตรัม เนื่องจากเสียงพูดเป็นสัญญาณที่ซับซ้อนเป็นสัญญาณที่ความถี่ปนกันอยู่มากมาย ในการวิเคราะห์เสียงพูดเราสามารถใช้การแปลงฟูเรียร์แบบเร็วในการหาองค์ประกอบความถี่ของเสียงพูดได้

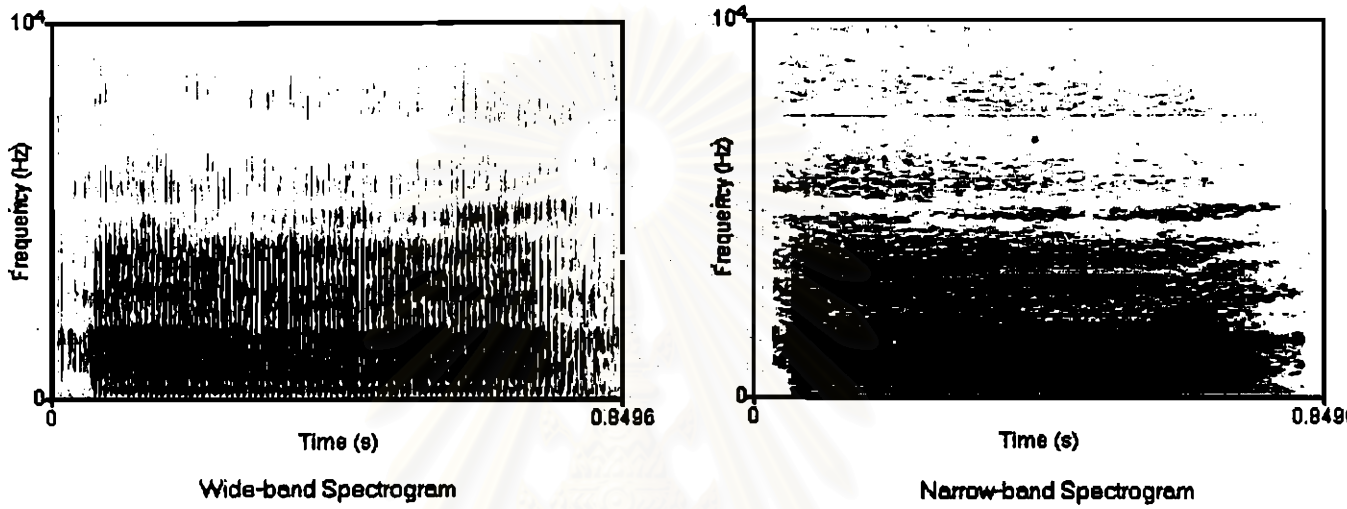
2.2.5 แผ่นภาพคลื่นเสียง (Spectrogram)

แผ่นภาพคลื่นเสียงเป็นเครื่องมืออย่างหนึ่งที่ใช้ในการเรียนรู้เรื่องเสียง ซึ่งจะแสดงความถี่ในแกนตั้งแสดงเวลาในแกนนอนและแสดงแอมปลิจูดด้วยความเข้มของสีในรูปเดียวกัน นักสัทศาสตร์จะใช้รูปภาพนี้ในการวิเคราะห์เสียงสูงต่ำ รูปภาพนี้ยังใช้ในการวิเคราะห์เสียงพยัญชนะและเสียงสระทั้งในการพูดเป็นคำและการพูดเป็นประโยค (สุดาพร ลักษณียนาวิน, 2529) แผ่นภาพคลื่นเสียงนี้ได้มาจากการทำการแปลงฟูเรียร์กับสัญญาณเสียงพูดแล้วนำสเปคตรัมที่ได้มาพล็อตโดยให้แกนตั้งเป็นแกนความถี่ และแกนนอนเป็นแกนเวลาแสดงแอมปลิจูดของสัญญาณด้วยความเข้ม ซึ่งแผ่นภาพคลื่นเสียงมีทั้งแบบช่วงกรองกว้าง (Wide-band spectrogram) และแบบช่วงกรองแคบ (Narrow-band spectrogram)

แผ่นภาพคลื่นเสียงแบบช่วงกรองกว้าง จะแสดงความละเอียดทางแกนเวลาได้ดีแต่ในแกนความถี่ช่วงจะกว้างแสดงสัญญาณที่เกิดจากการกระตุ้นของเส้นเสียงได้ไม่ชัดเจน แต่จะเห็นถึงคาบพิทช์ออกมาในแนวตั้งซึ่งเหมาะสำหรับการพิจารณาความถี่ฟอร์แมนท์ การกำทอนของฮาร์มอนิกส์ภายใต้ฟอร์แมนท์แต่ละฟอร์แมนท์จะอยู่ในช่วงแถบความถี่ประมาณ 300 เฮิรทซ์ แสดงให้เห็นเป็นแถบสีดำครอบคลุมในแต่ละฟอร์แมนท์ทางแนวตั้ง ความถี่ตรงกึ่งกลางของแต่ละแถบสีดำก็คือความถี่ฟอร์แมนท์โดยประมาณ

แผ่นภาพคลื่นเสียงแบบช่วงกรองแคบ จะแสดงความละเอียดของฮาร์มอนิกส์ทางแกนความถี่ได้ดีแต่ในแกนเวลาจะไม่ละเอียดนัก แผ่นภาพคลื่นเสียงชนิดนี้จะแสดงฮาร์มอนิกส์ที่เกิดจากการกระตุ้นของเส้นเสียงได้ละเอียดชัดเจนกว่าแผ่นภาพคลื่นเสียงแบบช่วงกรองกว้าง เหมาะสำหรับดูการ

เปลี่ยนแปลงฮาร์มอนิกส์อย่างละเอียดในการเชื่อมต่อระหว่างเสียง สามารถนำมาใช้ในการตัดสินปัญหาในการวิเคราะห์เสียงวรรณยุกต์ได้ เช่นในบางกรณีที่ไม่แน่ใจว่าเสียงวรรณยุกต์ในคำหนึ่งต่างหรือเหมือนกับเสียงวรรณยุกต์ในคำอีกคำหนึ่งหรือไม่และไม่สามารถตัดสินจากการฟังได้ จะสามารถใช้แผ่นภาพคลื่นเสียงแบบช่วงกรองแคบแสดงเสียงสูงต่ำของวรรณยุกต์มาเปรียบเทียบกันได้



รูปที่ 2.16 แสดงแผ่นภาพคลื่นเสียงแบบช่วงกรองกว้างและช่วงกรองแคบ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย