บทที่ 2

ทฤษฎีและหลักการ

ในบทนี้จะกล่าวถึงทฤษฎี แนวความคิดและหลักการที่เกี่ยวข้องกับการเกิดเสียงพูด โดยเริ่มตั้งแต่เทคนิคการสังเคราะห์เสียงพูดในรูปแบบต่าง ๆ ลักษณะทางกายภาพของการเปล่งเสียงซึ่งจะกล่าวถึงหน้าที่ของอวัยวะที่ใช้ในการเปล่งเสียงเริ่มตั้งแต่ปอด (lungs) จนถึงริมฝีปาก (lips) รวมถึงปรากฏการณ์การเกิดเสียง ต่อมาจะเป็นทฤษฎีด้านภาษาศาสตร์ซึ่งเกี่ยวข้องกับลักษณะของหน่วยเสียงของคำในภาษาไทยซึ่งการก่อเกิดของหน่วยเสียงมีความสัมพันธ์กับลักษณะทางกายภาพโดยตรงรวมถึงการวิเคราะห์เสียงในงานด้านการประมวลผลของสัญญาณเสียงเพื่อหาค่าความถี่หลักมูลของเสียงแต่ละวรรณยุกต์ หลังจากนั้นจะเป็นทฤษฎีเกี่ยวกับเสียงเพื่อนำไปวิเคราะห์และสร้างแบบจำลองที่ใช้แทนกระบวนการทำให้เกิดเสียงพูด

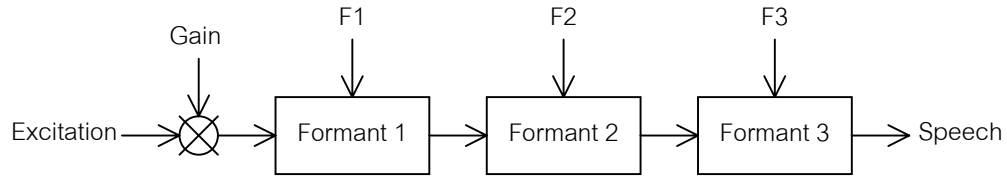
2.1 บทนำ

เทคนิคการสังเคราะห์เสียงพูดที่ใช้กันในปัจจุบันนี้มี 3 วิธีคือ

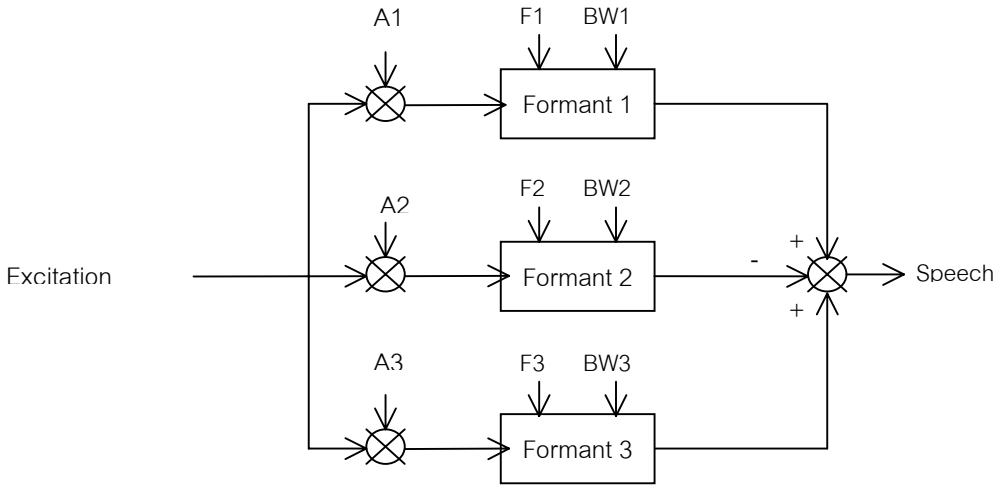
1. Concatenative synthesis เป็นการนำหน่วยเสียงที่ทำการบันทึกเอาไว้มาเล่นใหม่โดยใช้วิธีการนำหน่วยเสียงที่บันทึกมาต่อกัน ตัวอย่างของหน่วยเสียงที่ใช้ เช่น คำ (words) พยางค์ (syllables) หน่วยคู่เสียง (diphones) และหน่วยเสียงอนุภาค (microphonemes) นอกจากนี้อาจจะเก็บแต่ละหน่วยเสียงให้อยู่ในรูปของค่าพารามิเตอร์ ด้วยวิธีการวิเคราะห์และสังเคราะห์ออกมา (analysis-synthesis method) เมื่อต้องการสังเคราะห์เสียงออกมาก็นำค่าพารามิเตอร์เหล่านี้มาต่อกัน

2. Formant synthesis เป็นการวิเคราะห์หาค่าฟอร์แมนต์เพื่อเป็นตัวแทนของหน่วยเสียงในภาษาเพื่อนำค่าที่ได้นี้ไปเป็นอินพุทให้กับ ทรานเฟอร์ฟังก์ชันของช่องทางเดินเสียง (vocal tract) เพื่อสังเคราะห์เป็นเสียงออกมา สามารถแบ่งแบบจำลองได้เป็น 3 แบบคือ แบบอนุกรม แบบขนาน และแบบอนุกรมขนานดังแสดงในภาพประกอบ 2-1

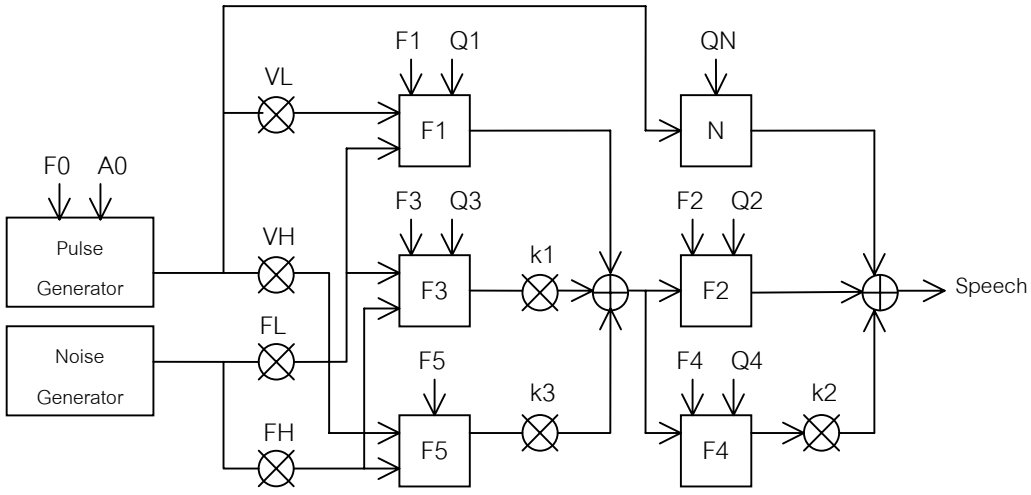
3. Articulatory synthesis เป็นวิธีจำลองระบบของการเกิดเสียงพูดของมนุษย์ประกอบด้วยแบบจำลองของ 3 ส่วนหลักคือ แหล่งกำเนิดเสียง (sound source production), รูปร่างช่องทางเดินเสียง (vocal tract), การแพร่เสียง (radiation) ซึ่งในวิทยานิพนธ์ชิ้นนี้จะใช้วิธีการนี้ดังจะได้กล่าวรายละเอียดในหัวข้อต่อไป



(ก) แบบอนุกรม



(ข) แบบขนาน



(ค) แบบอนุกรมขนาน

ภาพประกอบ 2-1 แบบจำลองสำหรับการสังเคราะห์เสียงแบบ Formant synthesis (ก) แบบอนุกรม (ข) แบบขนาน (ค) แบบอนุกรมขนาน (Lemmetty, 1999)

2.2 เสียงและอวัยวะที่ใช้ในการเปล่งเสียง

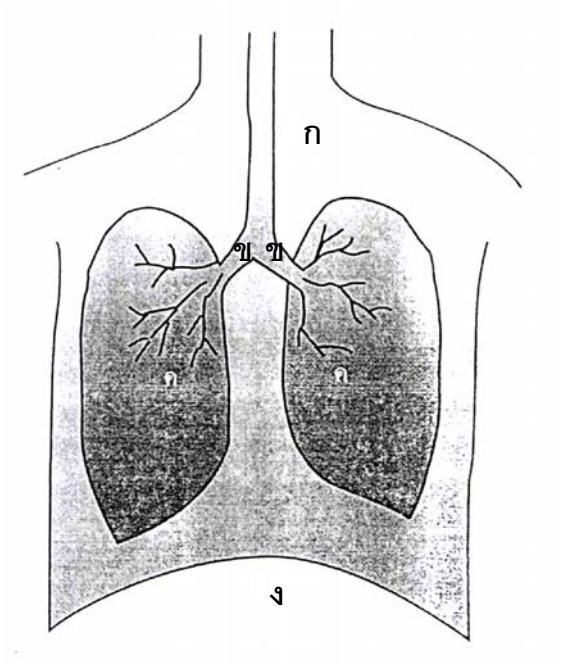
ในการศึกษาเรื่องเสียงในภาษา เราจำเป็นต้องเรียนรู้เกี่ยวกับลักษณะหน้าที่และการทำงานของอวัยวะต่าง ๆ ที่ใช้ในการเปล่งเสียง เนื่องจากความรู้ดังกล่าวจะช่วยให้เราเข้าใจกระบวนการของการเปล่งเสียง ซึ่งความเข้าใจนี้จะช่วยให้เราสามารถฟัง และจำแนกเสียงในภาษาได้ดี

ลมเป็นองค์ประกอบที่สำคัญในการเปล่งเสียง เสียงในภาษาต่าง ๆ ส่วนใหญ่จะใช้ลมที่เกิดจากปอดซึ่งเป็นลมหายใจออก อย่างไรก็ตามมีบางภาษาที่เสียงพยัญชนะบางเสียงใช้ลมที่เกิดจากคอและเป็นลมเข้า (Implosives) เสียงพยัญชนะบางเสียงใช้ลมที่เกิดจากคอเป็นลมออก (ejectives) พยัญชนะบางเสียงใช้ลมที่เกิดจากปากและเป็นลมเข้า (clicks)

อวัยวะที่ใช้ในการออกเสียงพูดแบ่งออกเป็น 3 กลุ่ม ได้แก่

1. ปอดและหลอดลม

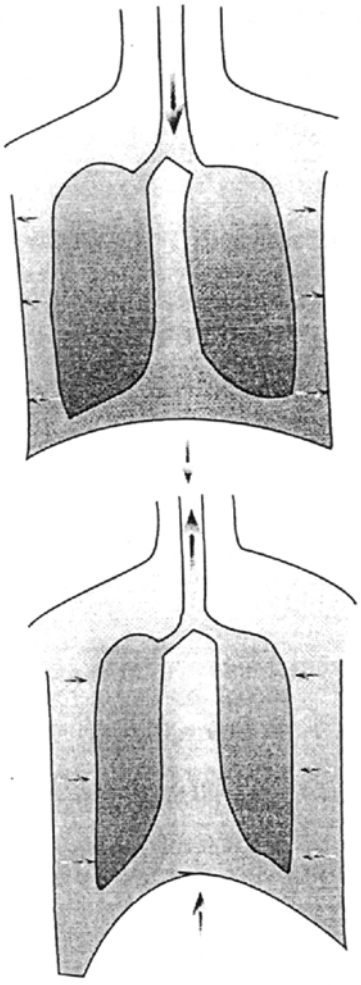
ปอดทั้งสองข้างประกอบด้วยถุงลมขนาดเล็กจำนวนมาก ปอดมีลักษณะเหมือนฟองน้ำไม่มีกล้ามเนื้อประกอบอยู่ ภายในปอดมีหลอดลมฝอยอยู่มากมาย หลอดลมฝอยนี้จะมารวมกันเป็นหลอดลมแขนง และหลอดลมแขนงของปอดทั้งสองข้าง จะมารวมกันเป็นหลอดลม ดังภาพประกอบ 2-2



ก. หลอดลม
ข. หลอดลมแขนง
ค. ปอด
ง. กะบังลม

ภาพประกอบ 2-2 บริเวณช่องอกและคอ (จินดา เฮงสมบูรณ์, 2542)

ด้านล่างของปอดมีกะบังลมซึ่งเป็นแผ่นกล้ามเนื้อรูปโค้ง แผ่นกะบังลมนี้จะเป็นตัวช่วยกดดันให้ช่องอกใหญ่ขึ้นหรือเล็กลงเวลาหายใจ กล่าวคือถ้าแผ่นกะบังลมลดตัวลง กระดูกซี่โครงจะถูกกล้ามเนื้อดึงให้ขยายตัวออก ขนาดของช่องอกจะใหญ่ขึ้น ความดันของอากาศในช่องอกจะต่ำกว่าความดันของอากาศภายนอกช่องอก อากาศก็จะไหลจากภายนอกร่างกายเข้าไปในปอดเป็นลมหายใจเข้า เมื่อสูดลมหายใจเข้าแล้ว ช่องอกจะอยู่ในสภาพพัก แผ่นกะบังลมจะดันตัวโค้งเล็กน้อย กระดูกซี่โครงจะถูกกล้ามเนื้อระหว่างซี่โครงดึงให้เข้ามาหากันขนาดของช่องอกจะเล็กลง ความดันของอากาศในช่องอกจะสูงกว่าความดันภายนอกช่องอกอากาศก็จะไหลจากปอดไปนอกร่างกายเป็นลมหายใจออกดังภาพประกอบ 2-3



ลมเข้าสู่ปอด

(ข)

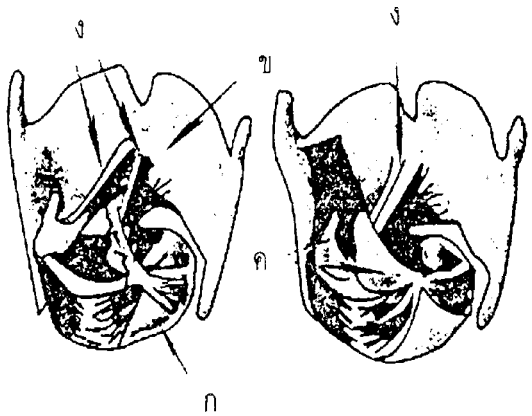
ลมออกจากปอด

ภาพประกอบ 2-3 การทำงานของกล้ามเนื้อกะบังลมเวลาหายใจ (จินดา เฮงสมบูรณ์, 2542)

เราใช้ลมหายใจออกเป็นสิ่งสำคัญที่ช่วยให้เกิดเสียงพูด โดยที่ลมจะไหลผ่านหลอดลมเข้าสู่กล่องเสียง ผ่านสู่อวัยวะในช่องปากหรือช่องจมูก อวัยวะเหล่านี้จะดัดแปลงลมให้เป็นเสียงพูดต่าง ๆ กัน

2. กล่องเสียงและเส้นเสียง

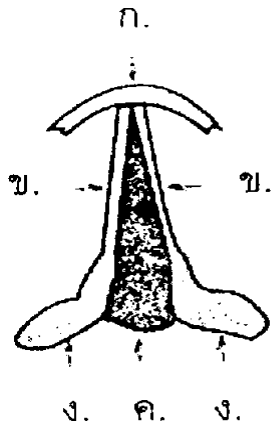
กล่องเสียงประกอบขึ้นด้วยกระดูก กระดูกอ่อน และกล้ามเนื้อหลายชิ้น ส่วนหนึ่งของกล่องเสียงคือ สิ่งที่เราเรียกกันว่า ลูกกระเดือก กล่องเสียงอยู่เหนือหลอดลม ด้านล่างของกล่องเสียงประกอบด้วยกระดูกอ่อนชื่อ ไครคอยด์ (cricoid) เป็นกระดูกอ่อนรูปคล้ายแหวน ด้านกว้างหันมาทางด้านหลังของร่างกาย ทางด้านหน้าของกล่องเสียงประกอบด้วย กระดูกอ่อนชื่อ ไทรอยด์ (thyroid) เป็นกระดูกอ่อนแผ่นเดียวรูปโค้งครึ่งวงกลม กระดูกชิ้นนี้เป็นส่วนของกล่องเสียงที่เราสัมผัสได้และเห็นได้ที่บริเวณคอ นอกจากนี้กล่องเสียงยังประกอบด้วยกระดูกอ่อนรูปปิรามิดอีก 1 คู่ ชื่อกระดูกอ่อน อาริตินอยด์ (arytenoid) กระดูกอ่อนคู่นี้อยู่ด้านหลังของกล่องเสียงบนกระดูกอ่อนรูปวงแหวน ส่วนประกอบที่สำคัญที่สุดของกล่องเสียง ได้แก่ เส้นเสียง ซึ่งมีอยู่ 2 เส้น มีลักษณะเป็นแผ่นประกอบด้วยเส้นเอ็นและกล้ามเนื้อ เส้นเสียงทั้งคู่พาดตัวตามแนวนอนอยู่ตรงกลางของกล่องเสียงปลายด้านหนึ่งของเส้นเสียงทั้งสองเส้นอยู่ติดกันและเชื่อมอยู่กับกระดูกอ่อนไทรอยด์ทางด้านหน้าของกล่องเสียงปลายอีกด้านหนึ่งของเส้นเสียงแต่ละเส้นเชื่อมอยู่กับกระดูกอ่อนอาริตินอยด์แต่ละอัน ดังภาพประกอบ 2-4, 2-5, 2-6



ก. กระดูกอ่อนไครคอยด์
ข. กระดูกอ่อนไทรอยด์
ค. กระดูกอ่อนอาริตินอยด์
ง. เส้นเสียง

ภาพประกอบ 2-4 ส่วนประกอบของกล่องเสียง (จินดา เฮงสมบูรณ์, 2542)

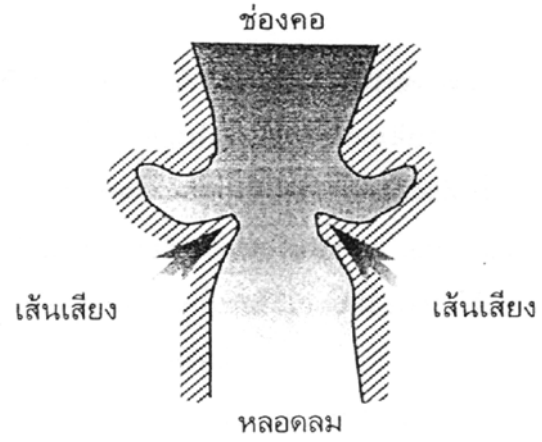
ด้านหน้าของกล่องเสียง



ก. กระดูกอ่อนไทรอยด์
ข. เส้นเสียง
ค. ช่องระหว่างเส้นเสียง
ง. กระดูกอ่อนอาริตินอยด์

ด้านหลังของกล่องเสียง

ภาพประกอบ 2-5 บริเวณเส้นเสียงขณะหายใจปกติ (จินดา เฮงสมบูรณ์, 2542)

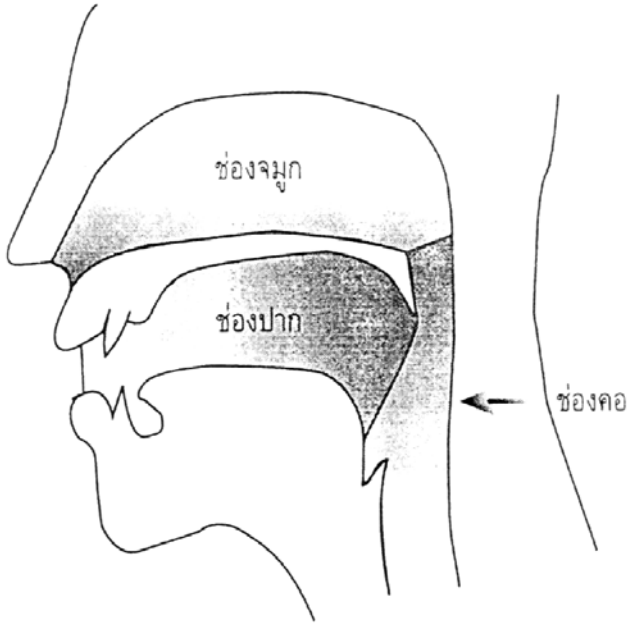


ภาพประกอบ 2-6 ภาพตัดของเส้นเสียงตามแนวตั้งขณะหายใจ (จินดา เฮงสมบูรณ์, 2542)

เส้นเสียงมีหน้าที่หลักในการปิดกั้นไม่ให้อาหารตกลงไปในหลอดอาหาร ในการเปล่งเสียงหน้าที่สำคัญของเส้นเสียงคือ เปลี่ยนลมจากปอดให้เป็นคลื่นเสียง

3. อวัยวะที่อยู่เหนือเส้นเสียง

ลมจากปอดเมื่อผ่านเส้นเสียงแล้วจะผ่านขึ้นมายังช่องคอ และจะออกสู่ภายนอกทางช่องปาก หรือ ช่องจมูก ช่องใดช่องหนึ่งหรือทั้งสองช่องพร้อม ๆ กันดังแสดงในภาพประกอบ 2-7 อาณาเขตของ ช่องคอ ช่องปากและ ช่องจมูกมีดังนี้

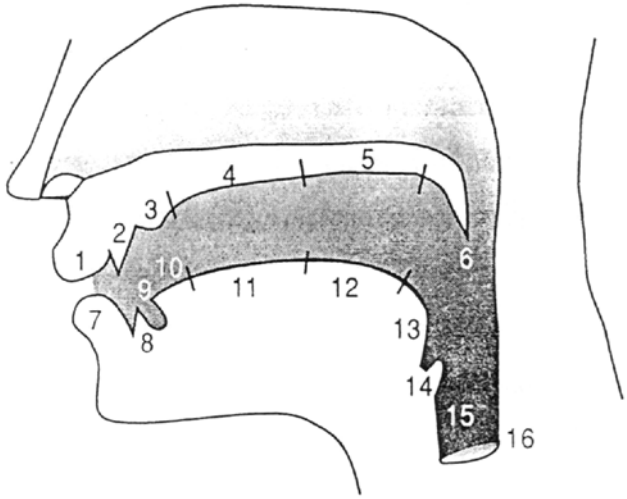


ภาพประกอบ 2-7 อวัยวะที่อยู่เหนือเส้นเสียง (จินดา เฮงสมบูรณ์, 2542)

อวัยวะที่อยู่เหนือเส้นเสียงซึ่งมีองค์ประกอบของอวัยวะต่าง ๆ ดังแสดงในภาพประกอบ 2-8 เหล่านี้อาจจำแนกตามหน้าที่ในการช่วยให้เกิดเสียงเป็น 2 ประเภท คือ

1) อวัยวะที่เป็นตำแหน่งที่เกิดเสียงต่าง ๆ (passive articulator) เรียกอีกอย่างหนึ่งว่า ฐานที่เกิดเสียง เป็นอวัยวะที่ไม่เคลื่อนที่ในขณะที่ออกเสียงอยู่ทางด้านบนของช่องปาก ได้แก่ ริมฝีปากบน ฟันบน ปุ่มเหงือก เพดานแข็ง เพดานอ่อน และ ลิ้นไก่

2) อวัยวะที่เป็นส่วนกระทำอาการ (active articulator) เรียกอีกอย่างหนึ่งว่า กรณ์ เป็นอวัยวะที่เคลื่อนที่ขณะออกเสียง อยู่ด้านล่างของช่องปาก ได้แก่ ริมฝีปากล่าง ฟันล่าง ปลายลิ้น ลิ้นส่วนปลาย ลิ้นส่วนหน้าและลิ้นส่วนหลัง



ภาพประกอบ 2-8 บริเวณต่าง ๆ ของอวัยวะในช่องปาก (จินดา เฮงสมบูรณ์, 2542)

1. ริมฝีปากบน (Upper lip)
2. ฟันบน (Upper teeth)
3. ปุ่มเหงือก (Alveolar ridge)
4. เพดานแข็ง (Hard palate)
5. เพดานอ่อน (Soft palate)
6. ลิ้นไก่ (Uvula)
7. ริมฝีปากล่าง (Lower lip)
8. ฟันล่าง (Lower teeth)
9. ปลายลิ้น (The tip of the tongue)
10. ลิ้นส่วนปลาย (The blade of the tonge)
11. ลิ้นส่วนหน้า (The front of the tonge)
12. ลิ้นส่วนหลัง (The back of the tonge)
13. โคนลิ้น (The root of the tonge)
14. แผ่นเนื้อปากหลอดลม (Epiglottis)
15. กรวยคอ (Pharynx)
16. เส้นเสียง (Vocal cord)

คำอธิบายภาพ

1,7 ริมฝีปากบนและริมฝีปากล่าง เป็นอวัยวะที่เคลื่อนไหวได้มาก ริมฝีปากทั้งคู่นี้จะดัดแปลงลมให้เป็นเป็นเสียงต่าง ๆ กัน โดยปิดริมฝีปาก เปิดเล็กน้อย เปิดกว้าง รูปปากแผ่ รูปปากห่อ ฯลฯ

2,8 ฟัน ช่วยกักลมทำให้เกิดเสียงหลายประเภท เช่น เมื่อฟันบนกดลงบนริมฝีปากล่างและปล่อยให้ลมเสียดแทรกช่องฟันออกมาเป็นเสียง ฟ [f] หรือใช้ลิ้นส่วนปลายแตะหลังฟันบน ทำให้เกิดเสียงต่าง ๆ เช่น ด [d] ต [t] ฯลฯ

3. ปุ่มเหงือก คือส่วนนูนออกมาจากแผ่นเพดานบน อยู่หลังฟันบน เป็นส่วนสำคัญในการเปล่งเสียงพยัญชนะอีกส่วนหนึ่ง เช่น ร [r] ล [l] ฯลฯ

4. เพดานแข็ง คือส่วนโค้งที่เป็นกระดูกแข็งอยู่ถัดจากปุ่มเหงือกเข้าไป เป็นบริเวณที่เกิดเสียงพยัญชนะอีกส่วนหนึ่ง เช่น จ [c] ช [ch] ฯลฯ

5. เพดานอ่อน คือ กระดูกที่อยู่ต่อกับเพดานแข็ง สามารถเคลื่อนขยับขึ้นลงได้เล็กน้อย เป็นบริเวณที่เกิดเสียงพยัญชนะอีกส่วนหนึ่ง เช่น ก [k] ค [kh] ฯลฯ

6. ลิ้นไก่ เป็นก้อนเนื้อเล็ก ๆ อยู่ปลายเพดานอ่อนตรงกลางปาก สามารถขยับขึ้นลงได้เป็นอวัยวะสำคัญที่ช่วยในการเปล่งเสียงทางจมูก และเสียงที่ไม่ใช่เสียงทางจมูก

9-13 ลิ้น เป็นอวัยวะที่เคลื่อนไหวได้มากในการเปล่งเสียงพูด แบ่งออกเป็น 5 ส่วนตามตำแหน่งที่ช่วยให้เกิดเสียงต่าง ๆ กัน ได้แก่

9. ปลายลิ้น คือส่วนปลายสุดของลิ้น สามารถยกขึ้นไปแตะอวัยวะต่าง ๆ ในปากส่วนบนได้ง่าย ช่วยในการเกิดเสียง พยัญชนะ ร [r] ล [l] ฯลฯ

10. ลิ้นส่วนปลาย คือ ส่วนที่ติดกับปลายลิ้น สามารถยกขึ้นไปแตะอวัยวะในปากส่วนบนได้ ช่วยในการเกิดเสียง พยัญชนะ ต [t] ด [d] ฯลฯ

11. ลิ้นส่วนหน้า คือส่วนที่อยู่ตรงกับเพดานแข็ง เมื่อวางลิ้นราบ ช่วยในการเกิดเสียง พยัญชนะ จ [c] ช [ch] และสระหน้า ฯลฯ

12. ลิ้นส่วนหลัง คือส่วนของลิ้นที่อยู่ตรงข้ามกับเพดานอ่อน เมื่อวางลิ้นราบ ช่วยในการเกิดเสียง พยัญชนะ ก [k] ค [kh] และสระหลัง ฯลฯ

13. โคนลิ้น คือส่วนที่อยู่ต่อจากลิ้นส่วนหลังลงไปในช่องคอ

14. แผ่นเนื้อปากหลอดลม เป็นก้อนเนื้อเล็ก ๆ คล้ายลิ้นไก่อยู่ต่อจากโคนลิ้นลงไปในช่องคอ มีหน้าที่ปิดช่องลมเมื่อกลืนอาหาร และเปิดช่องลมเมื่อพูด

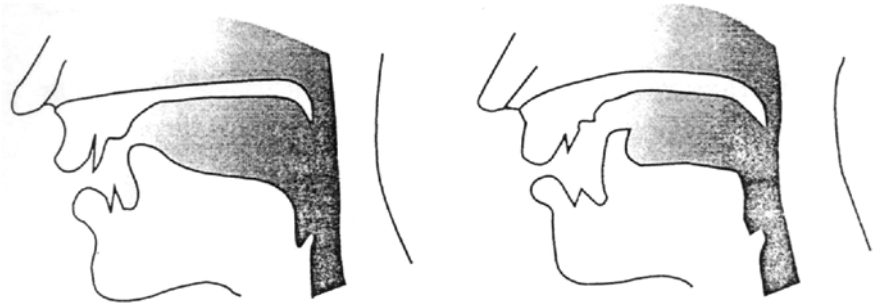
15. กรวยคอ คือ โพรงคอที่อยู่ถัดจากช่องปากลงไปถึงเส้นเสียง

16. เส้นเสียง เป็นส่วนที่สำคัญที่ช่วยให้เกิดเสียงพูด

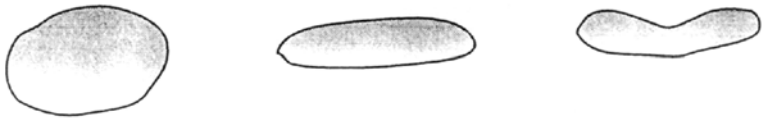
อวัยวะต่าง ๆ ภายในบริเวณช่องปากนี้ มีหน้าที่หลักเกี่ยวกับการกินอาหาร ในการเปล่งเสียงนั้นอวัยวะสำคัญที่มีหน้าที่ดัดแปลงลมให้เกิดเป็นเสียงต่าง ๆ ได้แก่ ลิ้น ริมฝีปาก ลิ้นไก่ และเพดานอ่อน

ก. ลิ้น เป็นอวัยวะที่เคลื่อนไหวได้อย่างคล่องแคล่วในหลายทิศทาง ทั้งกระดกขึ้น กระดกลง แลบไปข้างหน้า ดึงไปข้างหลัง นอกจากนี้ปลายลิ้นยังเคลื่อนตัวได้มากกว่าส่วนอื่น อาจกระดกขึ้นไปในทิศทางของฟันบน และปุ่มเหงือกหรือ งอปลายลิ้นกลับเข้าไปด้านใน และด้านล่าง

ของลิ้น กระดกขึ้นไปในทิศทางของฟันบน ปุ่มเหงือก เพดานแข็งแสดงในภาพประกอบ 2-9 (ก) - (ข) ทำให้เกิดเสียงต่าง ๆ กันได้มาก นอกจากนี้ลักษณะของลิ้นยังแบ่งออกเป็น 3 ลักษณะดังแสดงในภาพประกอบ 2-9 (ค) - (จ)



(ก) ปลายลิ้นกระดกขึ้นโดยไม่งอกลับ (ข) ปลายลิ้นกระดกขึ้นและงอกลับ



(ค) กลม (ง) แบนราบ (จ) มีร่องตรงกลาง

ภาพประกอบ 2-9 ลิ้นลักษณะต่าง ๆ (จินดา เฮงสมบูรณ์, 2542)

ข. ริมฝีปาก มีส่วนในการดัดแปลงลมให้เป็นเสียงต่าง ๆ กัน โดยริมฝีปากอาจปิดสนิท เปิดเล็กน้อย เปิดกว้าง ห่อกลม หรือทำเป็นรูปรีก็ได้ดังภาพประกอบ 2-10

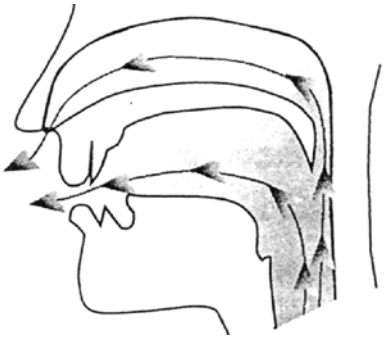
(ก) ห่อ (ข) เปิดธรรมดา

(ค) เปิดกว้าง

ภาพประกอบ 2-10 ริมฝีปากลักษณะต่าง ๆ (จินดา เฮงสมบูรณ์, 2542)

ค. ลิ้นไก่และเพดานอ่อน เป็นส่วนที่เคลื่อนไหวได้เล็กน้อย โดยอาจยกขึ้นไปติดกับผนังคอทำให้ลมจากปอดเคลื่อนไปสู่ช่องจมูกไม่ได้ หรืออาจลดตัวลงปล่อยให้ลมจากปอดเข้าสู่ช่องจมูก

(ก) ลมออกทางช่องปาก (ข) ลมออกทางช่องจมูก



(ค) ลมออกทางช่องปากและช่องจมูก

ภาพประกอบ 2-11 การเคลื่อนตัวของลิ้นไก่และเพดานเพื่อควบคุมทิศทางของลม
(จินดา เฮงสมบูรณ์, 2542)

สามารถอธิบายภาพประกอบ 2-11 ได้ว่าหากเพดานอ่อนและลิ้นไก่เคลื่อนไปติดกับผนังคอ ลมจากปอดจะผ่านไปทางช่องจมูกไม่ได้ต้องผ่านทางช่องปากเท่านั้น (ภาพประกอบ 2-11 ก) เสียงที่เกิดขึ้นก็เป็นเสียงที่เกิดจากช่องปาก แต่ถ้าเพดานอ่อนและลิ้นไก่ลดลงอยู่ในลักษณะพักธรรมดา ลมจากปอดจะผ่านขึ้นไปทางช่องจมูกได้ เสียงที่เกิดขึ้นจะมี 2 ประเภทคือ

- ถ้ามีการปิดกักลม ณ ที่ใดที่หนึ่งในช่องปากลมจะผ่านออกได้เฉพาะทางจมูก (แสดงในภาพประกอบ 2-11 ข) เสียงที่เกิดขึ้นเป็นเสียงทางจมูก (nasal sounds) ซึ่งเกิดขึ้นได้ที่เสียงท้าย

พยางค์ เช่น คำว่า ลม เสียง "ม" จะปรากฏที่พยางค์ท้าย โดยลักษณะของเพดานอ่อนและลิ้นไก่จะลดระดับลงมาปิดกั้นไม่ให้ลมผ่านออกทางปาก

- ถ้าไม่มีการปิดกักลมในช่องปาก ลมจะผ่านออกได้ทั้งทางปากทั้งทางจมูก เสียงที่เกิดขึ้นเป็น เสียงขึ้นจมูก (ภาพประกอบ 2-11 ค) ซึ่งจะเกิดกับเสียงสระที่ชิดกับเสียงทางจมูกและในบางภาษาจะมีสระทั้งที่เป็นเสียงสระขึ้นจมูก และเสียงสระที่ออกทางช่องปากโดยการเปล่งเสียงริมฝีปากทั้งคู่อยู่ชิดกันกักลมไว้ ในขณะเดียวกันเพดานอ่อนและลิ้นไก่ลดระดับลงอยู่ในลักษณะพัก ทำให้ลมผ่านออกทางช่องจมูกเมื่อเปิดที่กักลมที่ริมผีปาก ลมจะออกทางปากด้วย

2.3 ทฤษฎีด้านภาษาศาสตร์

หลักการที่ใช้ในการสังเคราะห์เสียงพูดภาษาไทยในวิทยานิพนธ์ชิ้นนี้เกี่ยวข้องกับทฤษฎีด้านภาษาศาสตร์ และทฤษฎีที่เกี่ยวข้องกับเสียงและอวัยวะที่ใช้ในการเปล่งเสียงรวมถึงการวิเคราะห์แบบจำลองที่ใช้ร่วมในการสังเคราะห์เสียง ซึ่งทฤษฎีด้านภาษาศาสตร์และทฤษฎีที่เกี่ยวข้องกับเสียงและอวัยวะที่ใช้ในการเปล่งเสียงใช้ในการทำความเข้าใจลักษณะของการเปล่งเสียงภาษาไทย เช่น อวัยวะที่เกี่ยวข้องกับการเปล่งเสียงนั้น ๆ ลักษณะของลิ้น และริมฝีปาก เป็นต้น เมื่อเข้าใจหลักการเปล่งเสียงแล้วก็สามารถนำความรู้ส่วนนี้ไปหาค่าพารามิเตอร์ที่เหมาะสมป้อนให้กับแบบจำลองที่ใช้ในการสังเคราะห์เสียงเพื่อให้ออกมาเป็นเสียงภาษาไทยได้

นักภาษาศาสตร์ได้ศึกษาและบรรยายการเปล่งเสียงอย่างเป็นระบบ และกำหนดสัญลักษณ์ทางภาษาเพื่อแสดงคุณสมบัติของเสียงแต่ละเสียง สัญลักษณ์ที่ใช้เรียกว่าสัทอักษร (phonetic symbol) สัทอักษรที่ใช้เป็นภาษาสากลเรียกว่า IPA (International Pronunciation Alphabet) สัทอักษรแต่ละตัวทำหน้าที่แทนหน่วยเสียงซึ่งแสดงให้เห็นว่าแต่ละเสียงทำหน้าที่ต่างกัน

หน่วยเล็กที่สุดของการพูดภาษาไทยคือพยางค์ ในการเปล่งเสียงแต่ละพยางค์ประกอบด้วยส่วนสำคัญได้แก่ เสียงสระ 1 เสียง เสียงพยัญชนะ 1 เสียงหรือเสียงพยัญชนะควบกล้ำ และเสียงวรรณยุกต์ 1 เสียง ส่วนเสียงพยัญชนะท้ายอาจมีหรือไม่มีก็ได้ เช่น "กิน" ประกอบด้วยเสียงพยัญชนะต้นคือ "ก" มีเสียงสระคือ "อิ" เสียงพยัญชนะท้ายคือ "น" และเสียงวรรณยุกต์เป็นเสียงสามัญ

ภาษาไทยมีตัวพยัญชนะทั้งหมด 44 รูป แต่เปล่งเสียงได้ต่างกัน 21 เสียง ทั้ง 21 เสียงสามารถเกิดต้นพยางค์หรือเป็นพยัญชนะต้นได้ แต่เสียงพยัญชนะสะกดมีเพียง 9 หน่วยเสียง ส่วนเสียงสระประสม 3 เสียง เสียงสระเดี่ยวมีความแตกต่างกันในเรื่องความสั้น-ยาวของเสียง ดังนั้นจึงสามารถแบ่งได้เป็นสระเสียงสั้น 9 เสียง และสระเสียงยาว 9 เสียง ส่วนวรรณยุกต์ในภาษาไทยมี 4 รูป แต่มี 5 เสียง (กำชัย ทองหล่อ, 2540) ดังรายละเอียดต่อไปนี้

2.3.1 เสียงสระ

หน่วยเสียงที่สำคัญในภาษาทุกภาษาคือหน่วยเสียงสระ เสียงสระเป็นเสียงที่เกิดจากลมซึ่งผ่านเส้นเสียงในตำแหน่งปิดเกือบสนิท และลมที่ต้องดันตัวออกมาทำให้เส้นเสียงสั่นเกิดเป็นเสียงดังขึ้นเรียกว่าเสียงก้อง เสียงที่เรียกว่าเป็นเสียงสระจะต้องถูกเปล่งออกมาทางปากโดยที่ไม่มีอวัยวะส่วนใดในปากมาปิดกั้นทางลมไว้ แต่อวัยวะในปากอาจจะอยู่ในท่าและตำแหน่งต่าง ๆ ที่ทำให้โพรงปากมีลักษณะต่างกันได้หลายแบบ ลมที่ผ่านออกมาจึงเกิดเป็นเสียงต่าง ๆ กัน การห่อริมฝีปากหรือไม่ห่อริมฝีปาก การยกลิ้นส่วนหนึ่งส่วนใด การยกลิ้นสูงต่ำต่าง ๆ ย่อมมีส่วนในการทำให้เกิดเสียงสระต่าง ๆ กันทั้งสิ้น หน่วยเสียงสระในภาษาไทยมีทั้งหมด 21 หน่วย แบ่งเป็น 2 ประเภทคือ สระเดี่ยว และสระเลื่อนหรือสระประสม เป็นหน่วยเสียงสระเดี่ยวเสียงสั้น 9 หน่วย หน่วยเสียงสระเดี่ยวเสียงยาว 9 หน่วย และหน่วยเสียงสระประสม 3 หน่วย

1. สระเดี่ยว (monophthongs) ได้แก่สระที่ขณะเปล่งเสียงลักษณะของลิ้นและริมฝีปากไม่มีการเปลี่ยนแปลงหรือมีการเปลี่ยนแปลงน้อยมาก ตำแหน่งจุดสูงสุดของลิ้นจะมีตำแหน่งเดียว แบ่งเป็น 2 ชนิดดังนี้

ก. หน่วยเสียงสระเดี่ยวเสียงสั้น 9 หน่วยมี อิ, เอะ, แอะ, อึ, เอะ, อะ, อุ, โอะ, เอาะ

ข. หน่วยเสียงสระเดี่ยวเสียงยาว 9 หน่วยมี อี, เอ, แอ, อือ, เอ, อา, อู ,โอะ, ออ

2. สระเลื่อนหรือสระประสม ได้แก่สระที่ขณะเปล่งเสียง ลักษณะของลิ้นมีการเปลี่ยนแปลงอย่างชัดเจน ส่วนลักษณะริมฝีปากอาจจะเปลี่ยนแปลงหรือไม่ก็ได้ สระประเภทนี้ตำแหน่งจุดสูงสุดของลิ้นจะมีมากกว่า 1 จุด แบ่งเป็น 2 ประเภทคือ

ก. สระเลื่อนสองส่วน (diphthongs) ได้แก่เสียงสระซึ่งเลื่อนไปในทิศทางเดียว ซึ่งเสียงสระประสมภาษาไทยอยู่ในประเภทนี้ได้แก่ หน่วยเสียงสระประสม 3 หน่วยมี เสียงย่อยหน่วยละ 2 เสียง เป็นสระเสียงสั้นและสระเสียงยาวคือ เอียะ, เอีย, เอือะ, เอือ, อัวะ, อัว

ข. สระเลื่อนสามส่วน (triphthongs) ได้แก่เสียงสระซึ่งเลื่อนไปในทิศทางหนึ่งแล้วเลื่อนต่อไปยังอีกทิศทางหนึ่ง

เสียงสระเป็นเสียงก้องทุกเสียงที่ดังกระจายไปได้ไกลกว่าเสียงพยัญชนะ เสียงสระจึงเป็นแกนของพยางค์ซึ่งทำให้เสียงอื่น ๆ ในพยางค์นั้นได้ยินได้ หน่วยเสียงสระในภาษาไทยสามารถเกิดร่วมกับวรรณยุกต์ได้ทุกหน่วย แต่เกิดตามหลังหน่วยเสียงพยัญชนะต้นและนำหน่วยเสียงพยัญชนะสะกดไม่ได้ทุกหน่วย

เนื่องจากเสียงสระเป็นเสียงก้อง เกิดจากการสั่นของเส้นเสียงผ่านอวัยวะกล่อมเกลาเสียงในช่องทางเดินเสียงแล้วออกจากปากโดยไม่มีอวัยวะอื่นใดมาขวางทางออกของเสียง ถ้าดูรูปคลื่น

ของเสียงสระจะพบว่าเป็นรูปคลื่นที่เป็นรายคาบ โดยมีคลื่นเสียงที่มีความถี่หลักมูลเท่ากับอัตราที่เส้นเสียงสั่น และเนื่องจากเกิดการกำทอนจากช่องทางเดินเสียงคือช่องปากและลำคอทำให้มีค่าความถี่ฮาร์มอนิกส์อื่นผ่านออกมาได้มากน้อยต่าง ๆ กันปนเข้ากับคลื่นเสียงจากเส้นเสียงด้วย ทำให้เกิดเป็นคลื่นเสียงที่ซับซ้อน

เสียงสระแต่ละตัวจะมีลักษณะเฉพาะตัวของความถี่เสียงที่ต่างกันออกไป ถ้านำสระต่าง ๆ มาวิเคราะห์หาองค์ประกอบทางความถี่ดู จะพบว่าสระต่างๆ จะมีรูปแบบความถี่ฟอร์แมนท์ (formant frequency) หรือความถี่การกำทอนที่ต่างกันออกไป เนื่องจากรูปร่างของช่องทางเดินเสียงนี้จะเป็นตัวกำหนดความถี่การกำทอน และรูปร่างของช่องทางเดินเสียงนี้จะเปลี่ยนไปตามแต่ละสระนั่นคือจะทำให้ค่าความถี่ฟอร์แมนท์ของสระต่าง ๆ มีค่าแตกต่างกัน

การศึกษาเรื่องสระมีสิ่งที่ควรพิจารณาคือ อวัยวะในช่องปากที่เกี่ยวข้องกับการดัดแปลงลมจากปอดวางตัวห่างจากกัน ลมจากปอดจะผ่านบริเวณกลางของช่องปากออกมาได้โดยสะดวก ไม่ถูกปิดกั้น อวัยวะที่มีบทบาทในการดัดแปลงลมจากปอดมาเป็นเสียงสระต่าง ๆ คือ ลิ้นและริมฝีปาก

1. ลักษณะของลิ้น นักสัทศาสตร์ได้ใช้วิธีถ่ายภาพเอกซเรย์ และภาพยนตร์เอกซเรย์ช่องปาก ในขณะที่ผู้พูดเปล่งเสียงสระต่าง ๆ พบว่าจุดที่สูงสุดของลิ้นจะอยู่ต่างกันเสมอ ดังนั้นสิ่งที่ต้องพิจารณาในการเปล่งเสียงสระคือตำแหน่งของจุดที่สูงสุดของลิ้นขณะที่เปล่งเสียงสระแต่ละเสียง ซึ่งก็คือตำแหน่งของเสียงสระ หรือที่เกิดของเสียงสระ การเคลื่อนไหวของลิ้นในการเปล่งเสียงสระเป็นดังนี้

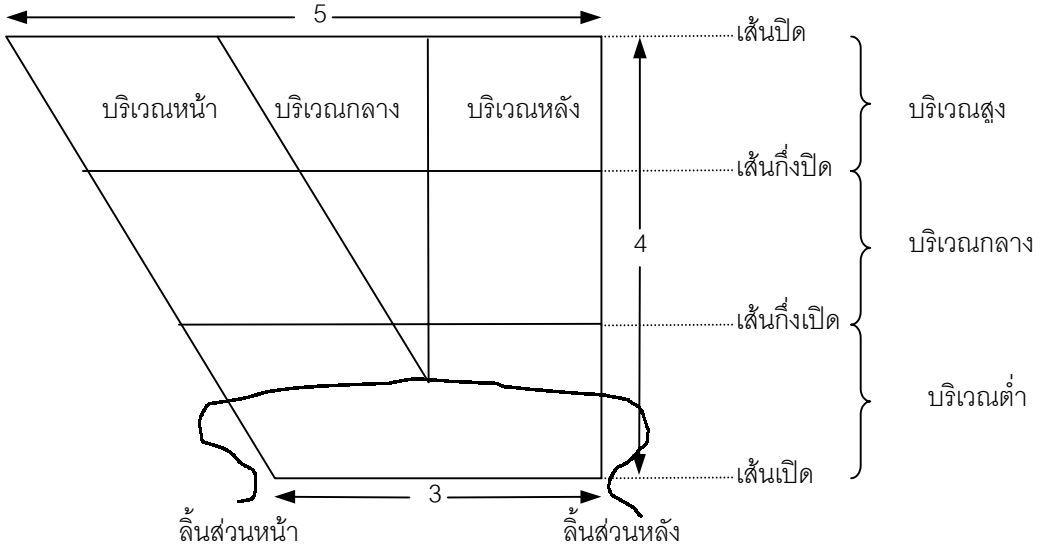
1) ส่วนของลิ้น ขณะที่เปล่งเสียงเราจะใช้ลิ้นส่วนต่าง ๆ คือใช้ลิ้นส่วนหน้า ส่วนกลาง และส่วนหลัง ยกขึ้นในระดับที่ต่าง ๆ กัน เราจึงเรียกสระตามส่วนต่าง ๆ ของลิ้นที่ยกขึ้น ถ้าเปล่งเสียงสระโดยยกลิ้นส่วนหน้าขึ้นเรียกว่า สระหน้า ถ้ายกลิ้นส่วนหลังขึ้นก็เรียกว่า สระหลัง

2) ระดับของลิ้น ความสูงต่ำของลิ้นที่ยกขึ้นใกล้เพดานมีส่วนสำคัญในการทำให้เสียงสระต่างกัน ถ้าลิ้นยกในระดับสูงช่องระหว่างลิ้นกับเพดานก็จะแคบสระนั้นก็เรียกว่า สระปิด (close vowel) ถ้าลิ้นยกในระดับต่ำ ช่องระหว่างลิ้นกับเพดานก็จะกว้าง สระนั้นก็เรียกว่า สระเปิด (open vowel)

2. ลักษณะริมฝีปาก ในการเปล่งเสียงสระริมฝีปากจะอยู่ในลักษณะต่าง ๆ กัน คือริมฝีปากห่อกลม แผ่ หรืออยู่ในลักษณะปกติ

นักสัทศาสตร์ได้สร้างแผนภูมิสระขึ้นดังแสดงในภาพประกอบ 2-12 โดยดัดแปลงรูปมาจากบริเวณที่เกิดเสียงสระในปากประกอบด้วยเส้นนอนขนาน 4 เส้น เส้นตั้งที่อยู่ขวามือสุดจะตั้งฉากกับเส้นทั้ง 4 สัดส่วนระหว่างเส้นนอนเส้นบนสุดกับเส้นนอนเส้นล่างสุด และเส้นที่ตั้งฉากทาง

ขวามือคือ 5:4:3 เรียกเส้นนอนทั้ง 4 จากบนลงล่างว่าเส้นปิด เส้นกึ่งปิด เส้นกึ่งเปิด และเส้นเปิด เรียกบริเวณที่อยู่ระหว่างเส้นทั้ง 4 จากบนลงมาล่างว่า บริเวณสูง บริเวณกลาง และบริเวณต่ำ และเรียกบริเวณของแผนภูมิสระจากซ้ายไปขวาว่า บริเวณหน้า บริเวณกลาง และบริเวณหลังตามลำดับ



ภาพประกอบ 2-12 บริเวณการเกิดเสียงสระ (จินดา เฮงสมบูรณ์, 2542)

ก่อนที่จะพิจารณาการแทนตำแหน่งเสียงสระภาษาไทยในแผนภาพสระ จำเป็นที่จะต้องเข้าใจสัญลักษณ์ที่ใช้แทนสัทอักษรในภาษาสากลให้ตรงกันเสียก่อน ซึ่งสามารถแสดงได้ด้วยตาราง 2-1 โดยเลือกแสดงเฉพาะเสียงสระเท่านั้น สำหรับรายละเอียดของตารางเป็นการแสดงทั้งสระเดี่ยวและสระประสม ซึ่งสระเสียงสั้นสามารถแทนได้ด้วยสัทอักษรตัวเดียวส่วนสระเสียงยาวที่มีบริเวณการเกิดเสียงตำแหน่งเดียวกันแทนได้ด้วยสัทอักษรของคู่สระเสียงสั้นนั้น ๆ 2 ตัวติดกัน หรืออาจจะแทนได้ด้วยสัทอักษรของสระเสียงสั้นตามด้วยเครื่องหมาย " ː "

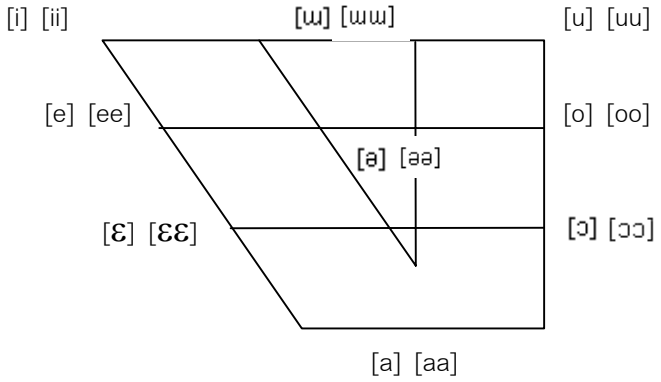
ตาราง 2-1 เปรียบเทียบรูปสระในภาษาไทยกับสัทอักษร

สระ	สัทอักษร
อิ - อี	[i] [ii]
เอะ - เอ	[e] [ee]
แอะ - แอ	[ɛ] [ɛɛ]
อึ - อือ	[ɯ] [ɯɯ]
เออะ - เออ	[ə] [əə]

ตาราง 2-1 (ต่อ)

สระ	สัทอักษร
อะ - อา	[a] [aa]
อุ - อู	[u] [uu]
โอะ - โอ	[o] [oo]
เอาะ - ออ	[ɔ] [ɔɔ]
เอียะ - เอีย	[ia] [iia]
เอือะ - เอือ	[ɯa] [ɯɯa]
อัวะ - อัว	[ua] [uua]

เสียงสระในภาษาไทยสามารถแสดงตำแหน่งการเปล่งเสียงของสระภาษาไทยลงบนแผนภาพสระได้ดังภาพประกอบ 2-13



ภาพประกอบ 2-13 สระในภาษาไทยในแบบของแผนภาพสระ

คำอธิบาย

1. หน้า กลาง หลัง เป็นชื่อเรียกสระตามส่วนของลิ้นที่ยกขึ้นใกล้เพดานดังนี้

1) สระหน้า เกิดจากการที่เปล่งเสียงสระโดยปล่อยออกจากปอดให้ผ่านขึ้นมาในช่องปาก ในขณะที่ยกลิ้นส่วนหน้าขึ้นใกล้เพดานแข็ง สระหน้าในภาษาไทยมี 3 คู่คือ [i] [ii], [e] [ee], [ɛ] [ɛɛ]

2) สระกลาง เกิดจากการที่เปล่งเสียงสระโดยปล่อยออกจากปอดให้ผ่านขึ้นมาในช่องปาก ในขณะที่ยกลิ้นระหว่างลิ้นส่วนหน้ากับลิ้นส่วนหลังขึ้นใกล้บริเวณเพดานแข็งต่อกับบริเวณเพดานอ่อน สระกลางในภาษาไทยมี 3 คู่คือ [∝] [∝∝], [↔] [↔↔], [a] [aa]

3) สระหลัง เกิดจากการที่เปล่งเสียงสระโดยปล่อยออกจากปอดให้ผ่านขึ้นมาในช่องปาก ในขณะที่ยกลิ้นส่วนหลังขึ้นใกล้เพดานอ่อน สระหลังมี 3 คู่คือ [u] [uu], [o] [oo], [□] [□□]

2. สูง กลาง ต่ำ เป็นชื่อเรียกสระตามระดับความสูง-ต่ำของลิ้นดังนี้

1) สระสูง เป็นสระที่เปล่งเสียงโดยยกลิ้นส่วนต่าง ๆ ขึ้นในระดับสูงสุดเท่าที่จะเกิดเสียงสระได้ สระสูงในภาษาไทยมี 3 คู่คือ [i] [ii], [∝] [∝∝], [u] [uu]

2) สระกลาง เป็นสระที่เปล่งเสียงโดยยกลิ้นส่วนต่าง ๆ ขึ้นในระดับกึ่งกลางระหว่างการเปล่งเสียงสระสูงกับสระต่ำ สระกลางในภาษาไทยมี 3 คู่คือ [e] [ee], [↔] [↔↔], [o] [oo]

3) สระต่ำ เป็นสระที่เปล่งเสียงโดยยกลิ้นส่วนต่าง ๆ ขึ้นในระดับต่ำสุดเท่าที่จะเกิดเสียงสระได้ สระต่ำภาษาไทยมี 3 คู่คือ [ɛ] [ɛɛ], [a] [aa], [□][□□]

สระเดี่ยวดังกล่าวอธิบายเสียงแต่ละเสียงได้ด้วยตาราง 2-2 ซึ่งให้ข้อมูลเกี่ยวกับระดับลิ้นในขณะที่เปล่งเสียงสระนั้น ๆ รวมถึงตำแหน่งของลิ้นที่ยกขึ้นสูงสุดระหว่างการเปล่งเสียงและลักษณะของริมฝีปากแบบห่อริมฝีปากหรือไม่ห่อริมฝีปาก

ตาราง 2-2 เสียงสระตามระดับลิ้น ตำแหน่งลิ้นและลักษณะริมฝีปาก

ริมฝีปาก	แผ่ - รี	แผ่ - รี	ห่อ - กลม
ระดับลิ้น ตำแหน่งลิ้น	หน้า	กลาง	หลัง
สูง	[i] [ii]	[ɯ] [ɯɯ]	[u] [uu]
กลาง	[e] [ee]	[ə] [əə]	[o] [oo]
ต่ำ	[ɛ] [ɛɛ]	[a] [aa]	[ɔ] [ɔɔ]

สระเลื่อนในภาษาไทยเป็นสระเลื่อนสองส่วนและเป็นสระเลื่อนลง (opening diphthong) มี 6 เสียงคือ เอียะ [ia] เอีย [iia] เอือะ [∝a] เอือ [∝∝a] อัวะ [ua] และอัว [uua]

ลักษณะการเลื่อนของทั้ง 6 สระเป็นดังนี้

1. เอียะ เป็นสระที่เริ่มต้นที่สระสูง หน้า ริมฝีปากไม่ห่อ แล้วเลื่อนลิ้นไปยังตำแหน่งสระต่ำ กลาง ริมฝีปากไม่ห่อ ระยะเวลาเปล่งเสียงสั้น

2. เอีย มีลักษณะเหมือนเอียะ แต่ระยะเวลาเปล่งเสียงนานกว่า เป็นสระเสียงยาวกว่า

3. เอือะ เป็นสระที่เริ่มต้นที่สระสูง กลาง ริมฝีปากไม่ห่อ แล้วเลื่อนลิ้นไปยังตำแหน่งสระต่ำ กลาง ริมฝีปากไม่ห่อ ระยะเวลาเปล่งเสียงนานพอ ๆ กับเอียะ

4. เอือ มีลักษณะเหมือนเอือะ แต่ระยะเวลาเปล่งเสียงนานกว่า เป็นสระเสียงยาวกว่า

5. อัวะ เป็นสระที่เริ่มต้นที่สระสูง หลัง ริมฝีปากไม่ห่อ แล้วเลื่อนลิ้นไปยังตำแหน่งสระต่ำ กลาง ริมฝีปากไม่ห่อ ระยะเวลาเปล่งเสียงสั้น

6. อัว มีลักษณะเหมือนเอียะ แต่ระยะเวลาเปล่งเสียงนานกว่า เป็นสระเสียงยาวกว่า

2.3.2 เสียงพยัญชนะ

เสียงสำคัญที่ใช้ในภาษาไทยอีกพวกหนึ่งคือเสียงพยัญชนะ เป็นเสียงที่เกิดจากลมซึ่งผ่านเส้นเสียง แล้วมาถูกดัดแปลงด้วยอวัยวะเปล่งเสียงส่วนต่าง ๆ ในปาก ทำให้เกิดเสียงขึ้น เสียงพยัญชนะมีหลายประเภทมีลักษณะการเปล่งเสียงแตกต่างกันหลายแบบ ความแตกต่างของเสียงจะมีความสำคัญยิ่งขึ้นเมื่อความแตกต่างนั้นทำให้ความหมายของคำในภาษาต่างกันหรือเรียกได้ว่า เป็นความแตกต่างที่ทำให้เสียงสองเสียงทำหน้าที่เป็นสองหน่วยเสียงต่างกันในภาษานั้นเพื่อให้ทราบว่าในภาษามีหน่วยเสียงกี่หน่วย หน่วยเสียงพยัญชนะในภาษาไทยมี 21 หน่วย

หน่วยเสียงพยัญชนะในภาษาไทยจะปรากฏในตำแหน่งต่าง ๆ ได้ 4 ตำแหน่งคือ

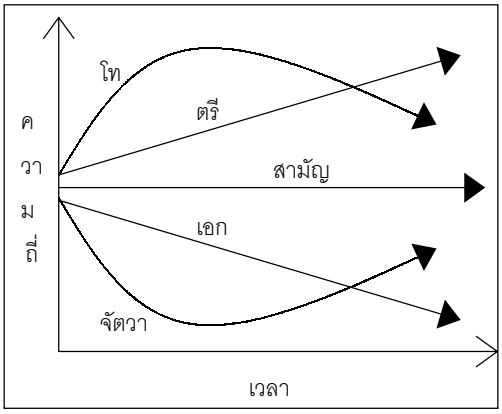
1. เกิดต้นคำ นำหน้าเสียงสระในพยางค์หนึ่ง ๆ เป็นหน่วยเสียงพยัญชนะต้น
2. เกิดนำเสียงพยัญชนะอื่นอีกเสียงหนึ่งในตำแหน่งต้นคำเรียกว่า พยัญชนะควบ
3. เกิดตามเสียงพยัญชนะอื่นในต้นคำ คือเกิดควบกับพวกที่ 2
4. เกิดตามหลังเสียงสระ เป็นเสียงพยัญชนะสะกด

2.3.3 เสียงวรรณยุกต์

ในการเปล่งเสียงพูดในภาษาหนึ่ง ๆ นั้น ไม่ว่าผู้พูดจะเปล่งเสียงสระพยัญชนะเรียงติดต่อกันออกมาเป็นคำวลีหรือเป็นประโยคก็ตาม ในการเปล่งเสียงนั้นผู้พูดอาจจะดัดแปลงเสียงสระเสียงพยัญชนะนั้น ๆ ให้อยู่ในระดับเสียง สูง ต่ำ หรือ เปลี่ยนแปลงเสียงขึ้นลงต่าง ๆ ได้อีก เสียงสูงต่ำในภาษาพูดเกิดได้ด้วยการสั่นสะเทือนของเส้นเสียงในอัตราต่าง ๆ กัน แต่เสียงที่เปล่งออกมาในขณะที่เส้นเสียงสั่นนั้นต้องเป็นเสียงก้อง พยางค์ในคำพูดทั่วไปจะประกอบด้วยเสียงสระและพยัญชนะ สระเป็นเสียงก้องจึงเป็นเสียงที่ช่วยทำให้เกิดระดับสูงต่ำได้ โดยมองว่าระดับเสียงสูงต่ำของสระในพยางค์หนึ่ง ๆ ก็เป็นระดับเสียงสูงต่ำของพยางค์นั้นด้วย เพราะเมื่อเสียงสระซึ่งเป็นแกนของพยางค์มี

ระดับใดเสียงอื่น ๆ ในพยางค์นั้นก็ย่อมมีระดับเสียงนั้นตามไปด้วยเช่นกัน หรืออีกนัยหนึ่งคือเสียงสูงต่ำในภาษาคือการที่ค่าความถี่หลักมูลของเสียงสระมีการเปลี่ยนแปลงไปตามเวลา

ในภาษาไทยระดับเสียงสูงต่ำของคำเรียกว่า วรรณยุกต์ เป็นส่วนสำคัญที่ทำให้คำตั้งแต่ 2 คำขึ้นไปซึ่งมีส่วนประกอบอื่น ๆ คือพยัญชนะ สระ และตัวสะกดอย่างเดียวกันมีความหมายต่างกันได้วรรณยุกต์ในภาษาไทยจึงจัดเป็นหน่วยเสียงเรียกว่าหน่วยเสียงวรรณยุกต์ หน่วยเสียงวรรณยุกต์จัดว่าเป็นหน่วยเสียงซ้อน เสียงวรรณยุกต์ในภาษาไทยมีลักษณะเค้าโครงและความถี่ของเสียงโดยประมาณดังภาพประกอบ 2-14 (Abramson, 1975 : อ้างอิงโดย อาทร นันทิยกุล, 2533)



ภาพประกอบ 2-14 ความถี่ของเสียงวรรณยุกต์ที่เปลี่ยนแปลงเทียบกับเวลา

เสียงสามัญแทนด้วยการเปลี่ยนแปลงที่ค่อนข้างที่จะคงที่ สำหรับเสียง ตรี และ เอก แทนด้วยเส้นตรงความชันเป็นบวกและลบ เส้นโค้งตก และโค้งขึ้นแทนเสียง โท และ เสียงจัตวาตามลำดับ

หน่วยเสียงวรรณยุกต์ทั้ง 5 หน่วยนี้สามารถจัดกลุ่มได้เป็น 2 กลุ่ม

2.3.3.1 วรรณยุกต์คงระดับ

วรรณยุกต์คงระดับซึ่งมีระดับความถี่ของเสียงค่อนข้างคงที่ตลอดพยางค์ จัดได้ว่ามีวรรณยุกต์คงระดับในการออกเสียงพูดนั้นโดยปกติต้นพยางค์และท้ายพยางค์มักจะไม่อยู่ในระดับเดียวกัน ในทางสัทศาสตร์ถือว่าระดับเสียงที่แตกต่างกันไปนี้เล็กน้อยเมื่อเทียบกับวรรณยุกต์อีกพวกหนึ่ง เสียงวรรณยุกต์คงระดับในภาษาไทยมี 3 หน่วยคือ

1. หน่วยเสียงวรรณยุกต์ระดับต่ำ

หน่วยเสียงวรรณยุกต์ระดับต่ำ (low tone) คือวรรณยุกต์เอก วรรณยุกต์ระดับต่ำจะปรากฏในพยางค์ได้ทุกแบบ

2. หน่วยเสียงวรรณยุกต์ระดับกลาง

หน่วยเสียงวรรณยุกต์ระดับกลาง (mid tone) คือวรรณยุกต์สามัญ วรรณยุกต์นี้มีระดับเสียงกลาง ๆ ในปลายพยางค์จะมีความถี่ลดลง วรรณยุกต์สามัญไม่ปรากฏในพยางค์ที่สะกดด้วยพยัญชนะกัก

3. หน่วยเสียงวรรณยุกต์ระดับสูง

หน่วยเสียงวรรณยุกต์ระดับสูง (high tone) คือวรรณยุกต์ตรี เสียงวรรณยุกต์นี้มีลักษณะเด่นที่มีระดับเสียงสูง โดยจะค่อย ๆ สูงขึ้นที่ละน้อยจากต้นเสียง ในภาษาไทยหน่วยเสียงวรรณยุกต์ตรีไม่ปรากฏในพยางค์ที่ประสมด้วยสระเสียงยาวที่ตัวสะกดเป็นเสียงหลัก

2.3.3.2 วรรณยุกต์เปลี่ยนระดับ

วรรณยุกต์เปลี่ยนระดับเป็นเสียงซึ่งมีระดับความถี่ของการเปล่งเสียงเปลี่ยนแปลงมากในช่วงพยางหนึ่ง ๆ เช่น ต้นพยางค์เปล่งเสียงให้มีระดับสูงแล้วลดระดับเสียงอย่างรวดเร็วไปสู่ระดับต่ำที่ปลายพยางค์ หรือต้นพยางค์เสียงมีระดับต่ำแล้วเสียงเพิ่มระดับเสียงอย่างรวดเร็วเป็นระดับสูงที่ปลายพยางค์ หรือเปลี่ยนสูงแล้วต่ำแล้วสูงอีกก็ได้ ในภาษาไทยวรรณยุกต์เปลี่ยนระดับมี 2 หน่วยคือ

1. หน่วยเสียงวรรณยุกต์เปลี่ยนตก

หน่วยเสียงวรรณยุกต์เปลี่ยนตก (falling tone) คือเสียงวรรณยุกต์โท วรรณยุกต์โทจะไม่ปรากฏในพยางค์ที่มีสระเสียงสั้นและสะกดด้วยพยัญชนะกัก

2. หน่วยเสียงวรรณยุกต์เปลี่ยนขึ้น

หน่วยเสียงวรรณยุกต์เปลี่ยนขึ้น (rising tone) คือเสียงวรรณยุกต์จัตวา วรรณยุกต์จัตวาจะไม่ปรากฏในพยางค์ที่มีพยัญชนะสะกดเป็นเสียงกัก

2.4 ทฤษฎีการวิเคราะห์เสียงพูด

เนื่องจากสัญญาณเสียงพูดจะมีคุณสมบัติไม่นิ่ง (non stationary) ดังนั้นในการประยุกต์ใช้งานกรรมวิธีประมวลผลสัญญาณเต็มหน่วยกับเสียงพูด จึงต้องแบ่งสัญญาณเสียงพูดออกเป็นส่วนย่อย ๆ เรียกว่ากรอบเสียงพูด (speech Frame) โดยแต่ละกรอบเสียงพูดจะมีความยาวประมาณ 10-30 มิลลิวินาที (Furui, 2001) ซึ่งถือได้ว่าสัญญาณเสียงพูดในแต่ละกรอบเสียงพูดนิ่ง จากนั้นจึงจะทำการประมวลผลสัญญาณเสียงพูดในแต่ละกรอบเสียงพูดได้

2.4.1 การวิเคราะห์ออโตคอรีเลชันในช่วงเวลาสั้น ๆ

ออโตคอรีเลชันของสัญญาณเสียงคือค่าของสเปคตรัมแมกนิจูด (spectrum magnitude) ยกกำลังสอง ซึ่งมีความสัมพันธ์กับส่วนกลับของการแปลงฟูเรียร์ (fourier transform) ของพลังงานในรูปของสเปคตรัม ออโตคอรีเลชันจะให้ข้อมูลเกี่ยวกับฮาร์โมนิคของสัญญาณเสียง แอมปลิจูดของฟอร์แมนท์ ตลอดจนแสดงถึงความเป็นคาบของสัญญาณโดยไม่เกี่ยวข้องกับเรื่องของเฟส ดังนั้นจึงสามารถจะนำไปประยุกต์ใช้ในด้านการหาคาบของพิตช์ได้ (อาทร นันทิยกุล, 2533) ออโตคอรีเลชันฟังก์ชันนิยามได้ตามสมการ 2.1 ดังนี้

$$\Phi(m) = \frac{1}{N} \sum_{n=0}^{N-m} s(n)s(n+m) \qquad 0 \le m \le M_0 - 1 \qquad (2.1)$$

นั่นคือการรวมผลของการคูณสัญญาณตัวอย่างเข้ากับสัญญาณตัวอย่างตัวก่อนหน้า ภายในจำนวนตัวอย่างที่จำกัดจาก 0 ถึง N-m และทำให้เป็นบรรทัดฐาน (normalized) โดยการหารด้วยจำนวนตัวอย่าง N โดยที่ M_0 เป็นค่าของการนำ (lag) ที่มากที่สุดที่จะให้มีได้สำหรับตัวอย่างที่กำลังพิจารณา ตัวอย่างเช่นถ้าต้องการใช้กับสัญญาณที่เป็นคาบ P ค่า M_0 ก็ควรที่จะมากกว่าค่า P แต่ทั้งนี้ค่า m จะมีค่าอยู่ในช่วงเท่าไรนั้นอยู่กับการนำไปประยุกต์ใช้ในด้านต่าง ๆ

ฟังก์ชันออโตคอรีเลชันจะมีค่าสูงสุดเมื่อ m = 0 และค่า $\Phi(m)$ จะมีค่าเท่ากับค่าพลังงานของสัญญาณ ในกรณีที่สัญญาณมีลักษณะเป็นคาบ ที่มีคาบเท่ากับจำนวน P ตัวอย่างแล้วจะได้ว่าฟังก์ชันออโตคอรีเลชันของสัญญาณนั้นจะมีคาบเท่ากับ P

$$\Phi(m) = \Phi(m+P)$$

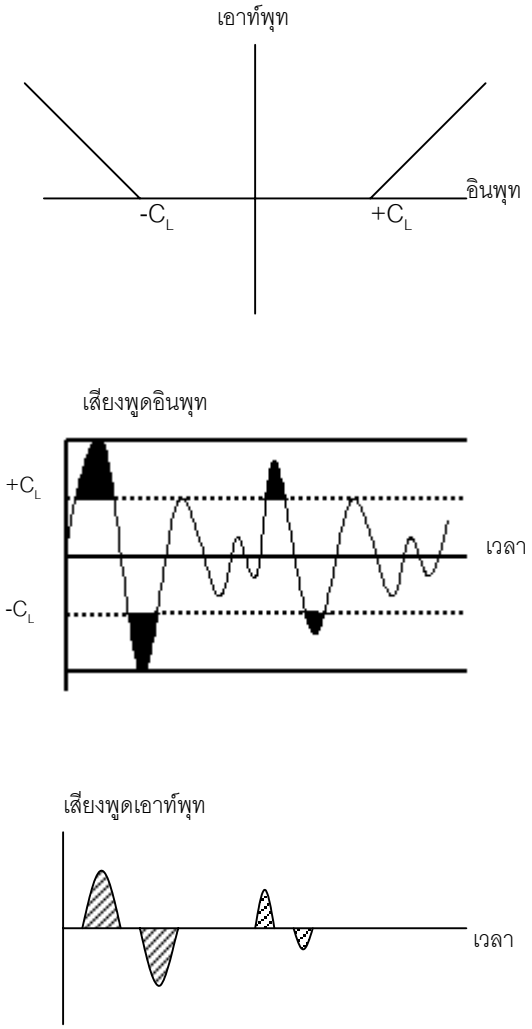
หรืออาจกล่าวได้ว่าค่าสูงสุดของ $\Phi(m)$ จะเกิดขึ้นเมื่อ m = 0, $\pm$P, $\pm$2P, $\pm$3P…..

2.4.2 เทคนิคเซนเตอร์คลิปปิง (center clipping technique)

ปัญหาสำคัญประการหนึ่งในการหาความถูกต้องของพิตช์ของเสียงในช่วงแถบความถี่ที่กว้างของสัญญาณก็คือ ผลของโครงสร้างฟอร์แมนท์ที่มีการวัดสัมพันธ์กับคาบของรูปคลื่นสัญญาณ ดังนั้นเพื่อเพิ่มความถูกต้องในการหาพิตช์ของเสียงจะต้องทำการลดผลของฟอร์แมนท์นี้ลงให้มากที่สุดเท่าที่จะทำได้ซึ่งเทคนิคเซนเตอร์คลิปปิงเป็นวิธีที่นิยมเพราะนำมาประยุกต์ใช้งานได้สะดวก

วิธีการนี้คือการตรวจจับระดับของสัญญาณว่า ถ้ามีค่าต่ำกว่าระดับคลิปปิงที่กำหนด ก็จะกำหนดให้สัญญาณตรงนั้นมีค่าเป็นศูนย์ ส่วนสัญญาณที่ยังมีค่าสูงกว่าระดับคลิปปิงก็จะเอาค่ามาหักลบด้วยค่าระดับคลิปปิงเหลือแต่ส่วนที่เกินไปจากระดับคลิปปิงเท่านั้น ภาพประกอบ 2-15 แสดงถึง

ลักษณะสมบัติของสัญญาณอินพุต-เอาท์พุตที่นำมาทำกระบวนการเซนเตอร์คลิปปิง ซึ่งจะเห็นได้ว่าถ้ามีการกำหนดระดับคลิปปิงที่เหมาะสมแล้วโครงสร้างของรูปคลื่นที่เป็นส่วนของฟอร์แมนท์ต่าง ๆ จะถูกกำจัดออกไปได้เกือบทั้งหมด ซึ่งจะมีผลทำให้สัญญาณที่มีสเปคตรัมราบเรียบที่ได้ออกมามองเห็นคาบของมันได้ง่ายกว่าสัญญาณที่ยังไม่ได้ผ่านกระบวนการทำให้สเปคตรัมราบเรียบ



(ก) ระดับคลิปปิง

(ข) เสียงพูดอินพุท

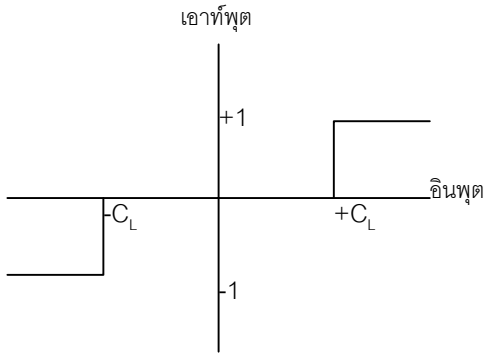
(ค) เสียงพูดเอาท์พุท

ภาพประกอบ 2-15 ลักษณะของสัญญาณเมื่อผ่านกระบวนการเซนเตอร์คลิปปิง

ขั้นตอนแรกของการประมวลผลสำหรับแต่ละเฟรมนั้นก็จะทำการคำนวณระดับคลิปปิงของแต่ละเฟรม เนื่องจากเสียงเป็นสัญญาณที่มีช่วงการเปลี่ยนแปลงกว้างมากดังนั้นการเลือกระดับคลิปปิงจะต้องทำด้วยความระมัดระวัง เพื่อป้องกันการสูญหายของข้อมูลที่สำคัญบางส่วนของรูป

คลื่นในขณะที่รูปคลื่นมีแอมปลิจูดที่กำลังมีขนาดเพิ่มขึ้นหรือกำลังลดลงในระหว่างเฟรมนั้น ๆ กรณีแบบนี้จะเกิดขึ้นในช่วงที่เสียงกำลังจะเริ่มขึ้นหรือกำลังจะจบลง หรืออาจจะเป็นตรงรอยต่อของช่วงการเปลี่ยนแปลงของเสียง สมมติว่าเราเลือกให้ระดับคลิปปิงมีค่า C_L สัญญาณเสียงแต่ละเฟรมมีขนาด 30 มิลลิวินาที จะถูกแบ่งออกเป็น 3 ส่วน ๆ ละ 10 มิลลิวินาที ในส่วนแรกและส่วนสุดท้ายอัลกอริธึมจะทำการหาค่าสมบูรณ์ที่จุดยอดสูงสุดเก็บเอาไว้สำหรับแต่ละส่วน ค่าระดับคลิปปิงก็จะกำหนดเอาจากค่าเปอร์เซนต์ที่แน่นอนของค่าสมบูรณ์ที่จุดยอดสูงสุดที่น้อยที่สุดระหว่างสองค่านั้น ตามวิธีการของ Sondhi (อ้างโดย อาทร นันทิยกุล, 2533) จะกำหนดค่าระดับคลิปปิงไว้ที่ 30 เปอร์เซ็นต์ของค่าสมบูรณ์ที่จุดยอด ซึ่งจะใช้เป็นระดับอ้างอิงกับข้อมูลของเสียงแต่ละเฟรม

หลังจากการหาระดับคลิปปิงได้แล้ว จะนำสัญญาณมาตัดช่วงที่อยู่ระหว่างระดับคลิปปิงนั้นทิ้งไป ซึ่งลักษณะนี้จึงเรียกว่า เซนเตอร์คลิปปิง (center clipping) ดังนั้นผลที่ได้ออกมากำหนดให้เป็น 3 ค่าคือ +1, -1 สำหรับสัญญาณที่มีค่าเกินกว่าระดับคลิปปิงด้านบวกและต่ำกว่าด้านลบตามลำดับ และค่า 0 สำหรับสัญญาณที่อยู่ในช่วงดังกล่าว จะเห็นได้ว่าไม่ว่าค่าของสัญญาณตัวอย่างจะมีค่ามากหรือน้อยแค่ไหน ก็จะถูกทำให้มีค่าเท่ากับ +1 และ -1 เหมือนกันหมด ภาพประกอบ 2-16 แสดงให้เห็นถึงลักษณะของสัญญาณเสียงขาเข้าและขาออกเมื่อผ่านการเซนเตอร์คลิปปิง เพื่อนำผลมาทำการออโตคอรีเลชันต่อไป ผลของการเซนเตอร์คลิปปิงนี้จะเป็นการลดความยุ่งยากในการคำนวณออโตคอรีเลชันลงอย่างมาก เหมาะสำหรับการนำไปปฏิบัติในขั้นตอนการเขียนโปรแกรมคอมพิวเตอร์



ภาพประกอบ 2-16 ลักษณะของสัญญาณเสียงขาเข้าและขาออกเมื่อผ่านการเซนเตอร์คลิปปิง

2.5 แบบจำลองที่ใช้ในการสังเคราะห์เสียง

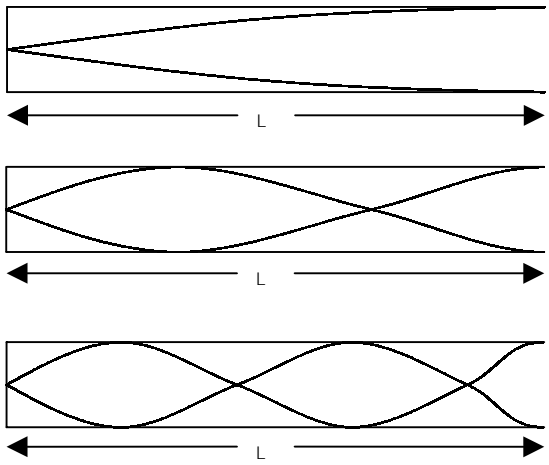
มนุษย์สร้างเสียงพูดขึ้นมาได้โดยมนุษย์จะมีอวัยวะสำคัญที่ทำให้เกิดเสียงหรือที่เรียกว่าเส้นเสียง (vocal cords) ซึ่งอยู่ในกล่องเสียง (larynx) และจากการสั่นของเส้นเสียงจะทำให้เกิดคลื่นเสียงความถี่ต่าง ๆ ผ่านเข้าสู่ลำคอ ผ่านจากลำคอเข้าสู่ช่องปากหรือช่องจมูกออกไปภายนอก ซึ่งขนาดและรูปร่างของอวัยวะภายในช่องปากนี้จะเป็นสิ่งกำหนดว่าคลื่นเสียงความถี่ไหนปรากฏออกมาให้ได้ยิน หรือคลื่นความถี่ไหนจะถูกดูดซับไว้ไม่ยอมให้ปรากฏออกมา ปรากฏการณ์ที่ทำให้เกิดเสียงพูดออกมาได้นี้เกิดขึ้นได้เนื่องจากรูปร่างของช่องคอและช่องปากของมนุษย์ ที่สามารถเปลี่ยนแปลงไปมาได้อย่างอิสระทำให้ผู้พูดมีความสามารถในการกำหนดคลื่นเสียงความถี่ต่าง ๆ ให้ปรากฏออกมาหรือดูดซับเอาไว้ได้ตามความต้องการ การเรียงลำดับชั้นของการคัดเลือกความถี่ของคลื่นเสียงระดับต่าง ๆ เหล่านี้นั่นเองที่ทำให้เกิดเสียงพูดออกมา

อวัยวะที่สำคัญในการสร้างความถี่คือเส้นเสียง มีลักษณะเป็นกล้ามเนื้อ 2 แผ่นอยู่ภายในกล่องเสียงและปิดขวางปากช่องหลอดลมจากด้านหลังมาด้านหน้า ระหว่างเส้นเสียงจะมีช่องว่างซึ่งเป็นทางให้ลมผ่านออกจากปอดได้ ช่องนี้เรียกว่า ช่องว่างระหว่างเส้นเสียง (glottis) เส้นเสียงทั้งสองสามารถดึงให้ห่างกัน หรือให้ชิดกันได้

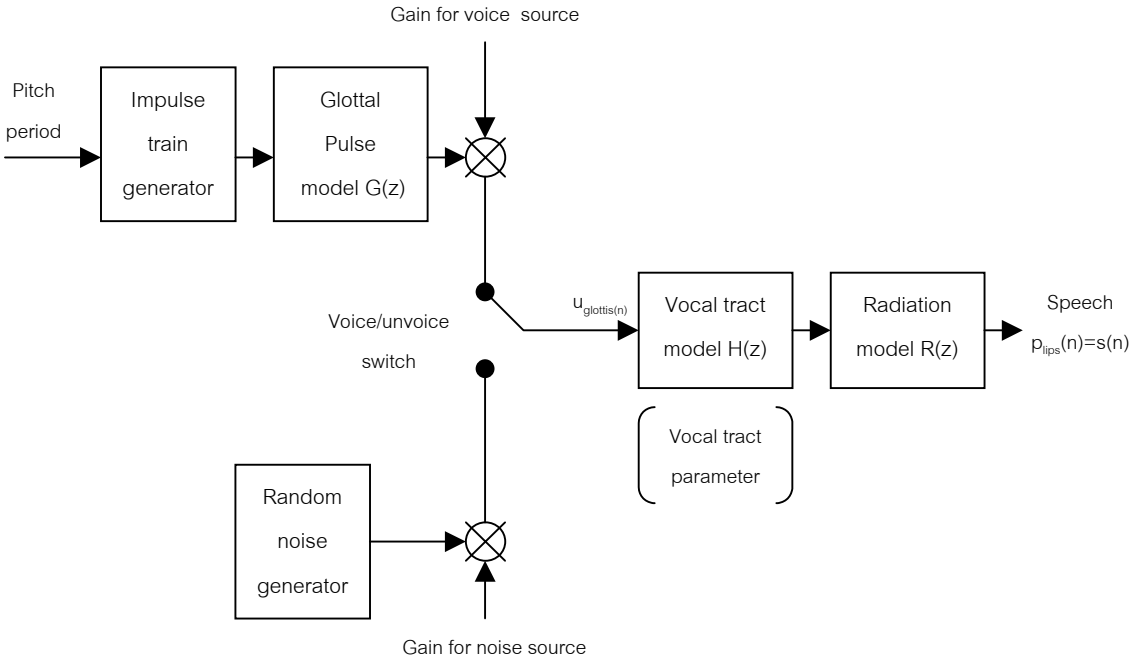
ระบบของเสียงพูดสามารถที่จะพิจารณาได้ว่าประกอบด้วยลำดับของท่อที่ต่อออกมาจากปอดซึ่งมีความยาวโดยประมาณ 17 เซนติเมตร (Parsons, 1972) รูปแบบจำลองของช่องเสียงอย่างง่ายอาจมองได้เป็นลักษณะของท่อทรงกลมที่มีต้นกำเนิดของเสียงอยู่ที่ปลายข้างหนึ่ง (ส่วนของกล่องเสียง) และปลายอีกข้างหนึ่งจะเปิด (ส่วนของช่องปาก) ดังแสดงในภาพประกอบ 2-17 ดังนั้นตามหลักการทางฟิสิกส์จะเกิดการกำทอนภายในท่อได้ที่ความยาวคลื่นเท่ากับ 4L, 4L/3, 4L/5, …. โดยที่ L คือความยาวท่อที่ใช้แทนช่องทางเดินเสียง ซึ่งถ้าคิดเป็นความถี่ที่เกิดการกำทอนก็จะได้ความถี่ที่ c/4L, 3c/4L, 5c/4L, …. Hz โดยที่ c คือความเร็วของเสียงในอากาศ ถ้าแทนความเร็วของเสียงในอากาศด้วยค่า 340 เมตรต่อวินาที และ L เท่ากับ 17 เซนติเมตร ก็จะมีการกำทอนที่ความถี่ประมาณ 500 Hz, 1500 Hz, 2500 Hz, …. ตามลำดับ

ในการสังเคราะห์เสียงด้วยแบบจำลองของอวัยวะที่ใช้ในการเปล่งเสียง ส่วนที่สำคัญที่สุดได้แก่ลักษณะของช่องทางเดินเสียงซึ่งจะมีลักษณะที่ซับซ้อน ในทางปฏิบัติการวิเคราะห์แบบจำลองที่ใช้สำหรับการสังเคราะห์เสียงจะมีการประมาณค่าของช่องทางเดินเสียงเพื่อให้วิเคราะห์แบบจำลองให้ง่ายขึ้นโดยเฉพาะในงานวิจัยชิ้นนี้ซึ่งเป็นกาสังเคราะห์เสียงสระ ในส่วนของการสังเคราะห์เสียงในวิทยานิพนธ์นี้แบ่งเป็น 3 ส่วน สามารถแทนแบบจำลองทั้งหมดได้ด้วยภาพประกอบ 2-18 ส่วนแรกเป็นปอดและกล่องเสียงได้ผลลัพธ์เป็นคลื่นรูปสามเหลี่ยมไม่สมมาตร แสดง

ความเป็นคาบด้วย อิมพัลส์เทรน ส่วนลักษณะรูปคลื่นสามเหลี่ยมแสดงด้วย การผ่านตัวกรองกลอททัล (glottal pulse filter) โดยมีตัวกำหนดเป็นคาบพิตช์ (pitch period) ส่วนที่ 2 เป็นตัวกรองของช่องทางเดินเสียง (vocal tract filter) ซึ่งจะคัดเลือกความถี่เพื่อให้ได้คลื่นเสียงที่ต้องการ ส่วนสุดท้ายเป็นส่วนของการแพร่เสียงเพื่อชดเชยด้านพลังงานและสัญญาณรบกวน



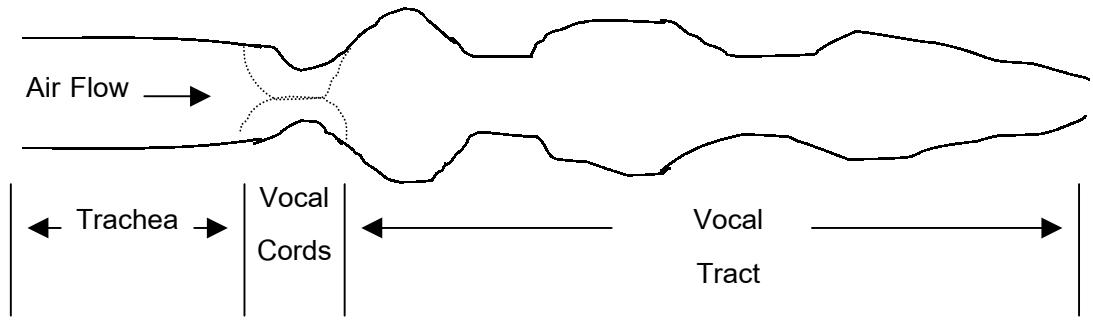
ภาพประกอบ 2-17 แสดงการกำทอนของช่องทางเดินเสียงซึ่งแทนด้วยท่อตรงความยาว L ที่ความยาวคลื่น 4L,4L/3,4L/5 ตามลำดับ



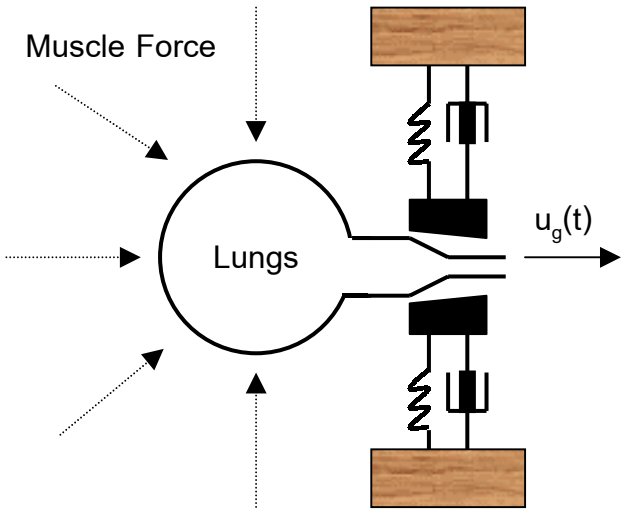
ภาพประกอบ 2-18 แสดงแบบจำลองที่ใช้ในการสังเคราะห์เสียง

2.5.1 แบบจำลองของปอดและกล่องเสียง (lung and larynx model)

จุดประสงค์คือต้องการได้ค่าของพลังงานและความถี่ที่ผลิตขึ้นเพื่อให้เกิดการกำทอนออกมาเป็นเสียงที่ต้องการ ตามภาพประกอบ 2-18 จะเห็นได้ว่าส่วนของการผลิตพลังงานคือส่วนของแบบจำลองกลอททัลพัลส์ (glottal pulse model) ซึ่งเป็นตัวกำหนดแอมปลิจูดให้กับคลื่นเสียงส่วนคาบพิตช์ (pitch period) ใช้กำหนดความถี่ในการผลิตเสียงซึ่งการสร้างรูปคลื่นในส่วนของแบบจำลองกลอททัลพัลส์นั้นมีอยู่หลายวิธีแต่ก็เป็นผลมาจากการศึกษาแบบจำลองของมวลสองชิ้น (two-mass model of vocal cords) ซึ่งในงานชิ้นนี้ได้ศึกษาการทำงานของอวัยวะเส้นเสียงของมนุษย์ในภาพประกอบ 2-19 ในส่วนของกล้ามเนื้อเส้นเสียง (vocal cords) จะมีการเปิด-ปิดตามความถี่ของเสียงพูดของคนนั้น (ตามเส้นประคือ ปิด) ทำให้สามารถแทนการทำงานส่วนนี้ได้ดังภาพประกอบ 2-20



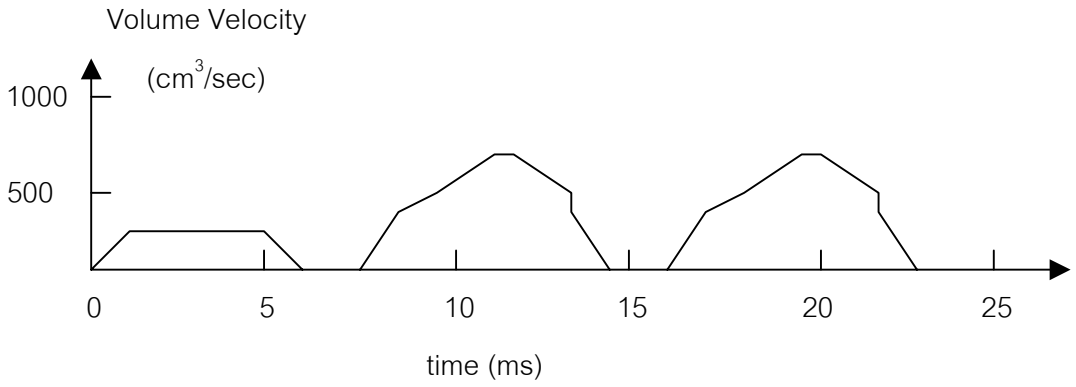
ภาพประกอบ 2-19 การทำงานของเส้นเสียง



ภาพประกอบ 2-20 แบบจำลองที่ใช้แทนเส้นเสียง

(Ishizaka and Flanagan, 1972)

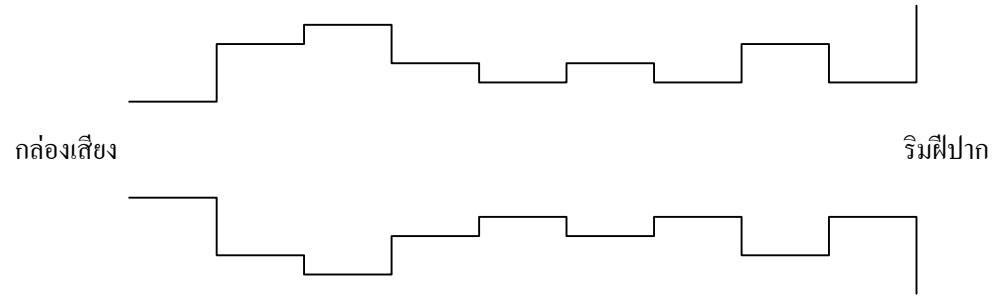
ซึ่งแบบจำลองมวลสองชิ้นนี้ให้ผลลัพธ์ออกมาเป็นความเร็วของปริมาตรของอากาศที่ไหลผ่านออกมาจากกล่องเสียงเป็นรูปสามเหลี่ยมไม่สมมาตรแสดงได้ด้วยภาพประกอบ 2-21



ภาพประกอบ 2-21 แสดงความเร็วของปริมาตรของอากาศที่ไหลผ่านออกมาจากกล่องเสียงกับเวลา

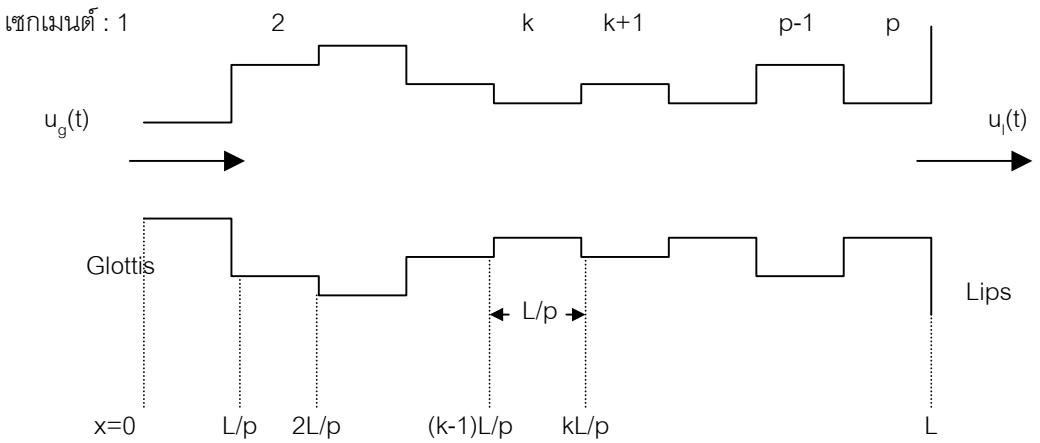
2.5.2 แบบจำลองของช่องทางเดินเสียง (vocal tract model)

ดังที่ได้กล่าวไว้แล้วในเบื้องต้นว่าเราจะแทนช่องทางเดินเสียงด้วยลำดับของท่อขนาดต่าง ๆ กันดังภาพประกอบ 2-22 ด้านซ้ายคือส่วนที่ต่อกับกล่องเสียงด้านขวาแทนริมฝีปาก



ภาพประกอบ 2-22 ช่องทางเดินเสียงแทนด้วยท่อตรงพื้นที่หน้าตัดต่าง ๆ กัน

เราจะแบ่งความยาวของท่อเป็น p เซกเมนต์ ขนาดเท่า ๆ กันซึ่งจะคิดว่าช่องทางเดินเสียงมีความยาวโดยตลอดเป็น L และใช้ทฤษฎีในการชักตัวอย่าง ในการพิจารณาขนาดความยาวของแต่ละเซกเมนต์ โดยใช้ความถี่ในการชักตัวอย่างที่ 10 kHz ซึ่งเพียงพอสำหรับครอบคลุมแบนด์วิดท์ (band width) ของเสียงพูดทั่วไปได้ (ปฏิมากร กิ้มสวัสดิ์, 2542) สามารถแบ่งเซกเมนต์ได้ดังภาพประกอบ 2-23



ภาพประกอบ 2-23 การแบ่งเซกเมนต์ในช่องทางเดินเสียง

การพิจารณาขนาดความยาวของท่อ

กำหนด

- f คือความถี่ในการชักตัวอย่าง (sampling) มีค่าเป็น 10 kHz
- T = 1/ f เป็นคาบการชักตัวอย่าง
- c = ความเร็วของเสียงมีค่าเป็น 340 เมตรต่อวินาที

กำหนดให้เวลาที่เสียงใช้ในการเดินทางในแต่ละเซกเมนต์เป็นเวลาครึ่งหนึ่งของความถี่การชักตัวอย่าง เพราะฉะนั้นความยาวของแต่ละเซกเมนต์ (L_k) คือ

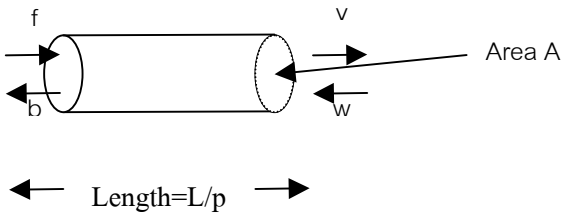
$$L_k = 0.5 * c * T \qquad (2.2)$$

$$= 0.017 \text{ เมตร}$$

จะได้จำนวนของท่อเป็น $\frac{0.17m}{0.017m} = 10$ ท่อ ถ้าคิดว่าช่องทางเดินเสียงยาว 0.17 เมตร

สำหรับการวิเคราะห์คลื่นเสียงในท่อเพื่อที่จะกำหนดทรานเฟอร์ฟังก์ชันให้กับแบบจำลองจะพิจารณาเป็นลำดับดังต่อไปนี้

2.5.2.1 คลื่นเสียงในท่อ



ภาพประกอบ 2-24 ท่อขนาดความยาว L/p พื้นที่หน้าตัด A

ในการวิเคราะห์สมการการเคลื่อนที่ของอนุภาคเราจะคิดว่าอนุภาคของอากาศเคลื่อนที่ในทิศทางเดียวคือทางแกน x และไม่มีการสูญเสียพลังงานเนื่องจากความร้อน ความหนืดและการสูญเสียในรูปแบบอื่น ๆ (leaky lossess) เนื่องจากมีค่าน้อยมาก (Deller, 2000) โดยพิจารณาจาก สมการอนุพันธ์ย่อย 2 สมการคือสมการของโมเมนตัม (momentum equation) และสมการของการคงมวล (mass conservation equation) ดังนี้ (Furui, 2001)

$$-\frac{\partial U}{\partial t} = \frac{A(x)}{\rho}\frac{\partial P}{\partial x} \qquad (2.3)$$

$$-\frac{\partial P}{\partial t} = \frac{\rho\, c^2}{A(x)}\frac{\partial U}{\partial x} \qquad (2.4)$$

U ความเร็วเชิงปริมาตร (volume velocity)

P ความดันเสียง (sound pressure)

c ความเร็วเสียง (sound velocity)

ρ ความหนาแน่นของอากาศ (air density)

A(x) ฟังก์ชันพื้นที่หน้าตัดของทางเดินเสียง (area function)

พิจารณาในภาพประกอบ 2-24 พบว่าที่บริเวณรอยต่อจะมีการแพร่กลับของความเร็วเชิงปริมาตร เพราะฉะนั้นการพิจารณาภาพประกอบ 2-24 คือตัดจากภาพประกอบ 2-23 มาวิเคราะห์ที่เวลา t ใด ๆ ความเร็วเชิงปริมาตรทางด้านซ้ายคือความเร็วด้านไป (f) ลบด้วยความเร็วด้านกลับ (b)

$$U = f\text{-}b \tag{2.5}$$

แทน (2.5) ใน (2.4) จะได้ความดัน (Furui, 2001)

$$P = \rho c/A\ (f+b) \tag{2.6}$$

พิจารณาภาพประกอบ 2-24 จะได้เวลาที่เสียงเดินทางในระยะ L/p เป็น L/cp เพราะฉะนั้น

$$f(t) = v\left(t + L/cp\right)$$

$$b(t) = w\left(t - L/cp\right) \tag{2.7}$$

เนื่องจากความยาวของท่อที่ออกแบบเพื่อให้เสียงเดินทางครึ่งหนึ่งของคาบการชักตัวอย่าง (ดูสมการที่ 2.2 ประกอบ) ถ้าพิจารณาเวลาเป็นดิสครีต จะได้

$$f(n) = v(n+1/2)$$
$$b(n) = w(n-1/2) \tag{2.8}$$

ดังนั้นเขียนอยู่ในรูปของการแปลง z (z-transform) ได้เป็น

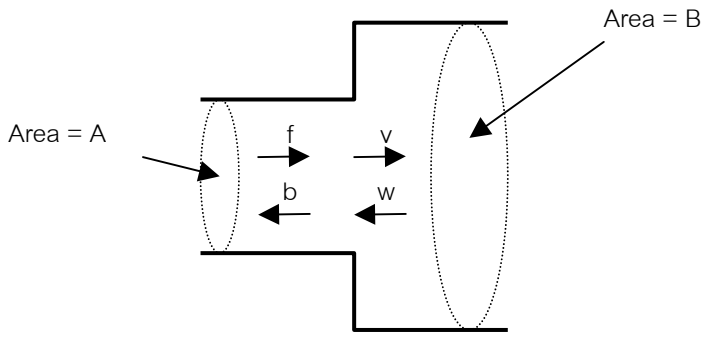
$$F(z) = z^{+1/2}\ V(z)$$
$$B(z) = z^{-1/2}\ W(z) \tag{2.9}$$

เขียนเป็นรูป เมทริกซ์

$$\begin{pmatrix} F \\ B \end{pmatrix} = \begin{pmatrix} z^{1/2} & 0 \\ 0 & z^{-1/2} \end{pmatrix} \begin{pmatrix} V \\ W \end{pmatrix} = z^{1/2} \begin{pmatrix} 1 & 0 \\ 0 & z^{-1} \end{pmatrix} \begin{pmatrix} V \\ W \end{pmatrix} \tag{2.10}$$

2.5.2.2 บริเวณรอยต่อ

พิจารณาบริเวณรอยต่อในภาพประกอบ 2-25 โดยคิดว่าไม่มีการสูญเสียพลังงาน



ภาพประกอบ 2-25 บริเวณรอยต่อของท่อพื้นที่หน้าตัด A และ B

จะได้ว่าความเร็วเชิงปริมาตรบริเวณรอยต่อต้องมีค่าเท่ากันดังนั้น

$$(f-b) = (v-w) \tag{2.11}$$

และทำนองเดียวกันความดันที่เกิดขึ้นบริเวณรอยต่อของท่อ A และท่อ B จะต้องเท่ากัน ดังนั้นแทนค่าในสมการที่ (2.6) จะได้ว่า

$$\rho c/A(f+b) = \rho c/B(v+w) \tag{2.12}$$

เขียนในรูปเมทตริกซ์ได้เป็น

$$\begin{pmatrix} 1 & -1 \\ B & B \end{pmatrix}\begin{pmatrix} f \\ b \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ A & A \end{pmatrix}\begin{pmatrix} v \\ w \end{pmatrix} \tag{2.13}$$

ดังนั้น

$$\begin{pmatrix} f \\ b \end{pmatrix} = \frac{1}{2B}\begin{pmatrix} B & 1 \\ -B & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ A & A \end{pmatrix}\begin{pmatrix} v \\ w \end{pmatrix}$$

$$= \frac{1}{2B}\begin{pmatrix} A+B & A-B \\ A-B & A+B \end{pmatrix}\begin{pmatrix} v \\ w \end{pmatrix}$$

กำหนดให้ $r = (B-A) / (B+A)$

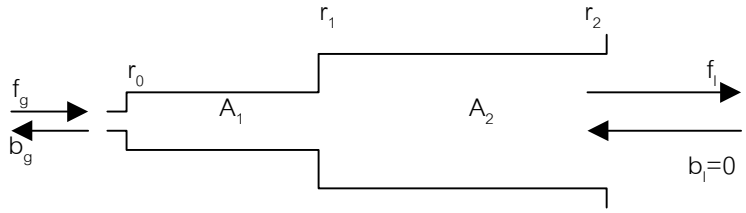
ดังนั้น

$$\begin{pmatrix} f \\ b \end{pmatrix} = \frac{1}{2B}\begin{pmatrix} A+B & A-B \\ A-B & A+B \end{pmatrix}\begin{pmatrix} v \\ w \end{pmatrix} = \frac{1}{1+r}\begin{pmatrix} 1 & -r \\ -r & 1 \end{pmatrix}\begin{pmatrix} v \\ w \end{pmatrix} \qquad (2.14)$$

2.5.2.3 ช่องทางเดินเสียง 2 เซกเมนต์

ดังภาพประกอบ 2-26 แบบจำลองจะประกอบด้วยรอยต่อ 3 รอยต่อ กำหนดให้สัมประสิทธิ์การสะท้อนกลับ (reflection coefficient)

$$r_k = \frac{A_{k+1} - A_k}{A_{k+1} + A_k} \qquad (2.15)$$



ภาพประกอบ 2-26 ช่องทางเดินเสียง 2 เซกเมนต์

พิจารณาทีละรอยต่อและโดยการแทนค่าในสมการที่ (2.10) และ (2.14) จะได้ว่า

$$\frac{1}{1+r_2}\begin{pmatrix} 1 & -r_2 \\ -r_2 & 1 \end{pmatrix}\begin{pmatrix} F_l \\ 0 \end{pmatrix}$$

$$\frac{1}{1+r_1}\begin{pmatrix} 1 & -r_1 \\ -r_1 & 1 \end{pmatrix} \times z^{1/2}\begin{pmatrix} 1 & 0 \\ 0 & z^{-1} \end{pmatrix} \times$$

$$\begin{pmatrix} F_g \\ B_g \end{pmatrix} = \frac{1}{1+r_0}\begin{pmatrix} 1 & -r_0 \\ -r_0 & 1 \end{pmatrix} \times z^{1/2}\begin{pmatrix} 1 & 0 \\ 0 & z^{-1} \end{pmatrix} \times \qquad (2.16)$$

คูณ (2.16) จะได้

$$\begin{pmatrix} F_g \\ B_g \end{pmatrix} = \frac{z}{\prod_{k=0}^{2}\left(1+r_k\right)} \begin{pmatrix} 1+(r_0r_1+r_1r_2)\,z^{-1}+r_0r_2\,z^{-2} \\ -r_0-(r_1+r_0r_1r_2)\,z^{-1}-r_2\,z^{-2} \end{pmatrix} F_l \qquad (2.17)$$

กำหนด Bg มีค่าเป็นศูนย์เนื่องจากถูกดูดซับไว้โดยปอดทั้งหมด จะได้สมการการถ่ายโอน (transfer function) เป็น

$$\frac{F_l}{F_g} = \frac{\prod_{k=0}^{2}\left(1+r_k\right)z^{-1}}{1+(r_0r_1+r_1r_2)\,z^{-1}+r_0r_2\,z^{-2}}$$

$$= \frac{Gz^{-1}}{1+(r_0r_1+r_1r_2)\,z^{-1}+r_0r_2\,z^{-2}}$$

$$= \frac{Gz^{-1}}{1-a_1\,z^{-1}-a_2\,z^{-2}} \qquad (2.18)$$

2.5.2.4 ช่องทางเดินเสียง p เซกเมนต์

กำหนดให้

$$\frac{1}{1+r}\begin{pmatrix} 1 & -r \\ -r & 1 \end{pmatrix} z^{1/2} \begin{pmatrix} 1 & 0 \\ 0 & z^{-1} \end{pmatrix} = \frac{z^{1/2}}{1+r}\begin{pmatrix} 1 & -r\,z^{-1} \\ -r & z^{-1} \end{pmatrix}$$

จากหัวข้อ 2.5.2.3 พิจารณาทีละรอยต่อโดยกำหนด Bg มีค่าเป็นศูนย์จะได้

$$\begin{pmatrix} F_g \\ 0 \end{pmatrix} = \frac{z^{\frac{1}{2}p}}{\prod_{k=0}^{p}\left(1+r_k\right)} \prod_{k=0}^{p-1}\begin{pmatrix} 1 & -r_k\,z^{-1} \\ -r_k & z^{-1} \end{pmatrix}\begin{pmatrix} 1 \\ -r_p \end{pmatrix} F_l$$

$$\begin{pmatrix} F_l \\ F_g \end{pmatrix} = \frac{\prod_{k=0}^{p}\left(1+r_k\right)z^{-\frac{1}{2}p}}{\prod_{k=0}^{p-1}\begin{pmatrix} 1 & -r_k\,z^{-1} \\ -r_k & z^{-1} \end{pmatrix}\begin{pmatrix} 1 \\ -r_p \end{pmatrix}}$$

$$\left(\frac{F_l}{F_g}\right) = \frac{\prod_{k=0}^{p}\left(1 + r_k\right) z^{-\frac{1}{2}p}}{\left(1 - r_0\right)\prod_{k=1}^{p}\begin{pmatrix} 1 & -r_k \\ -r_k z^{-1} & z^{-1} \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix}} \qquad (2.19)$$

(Deller, 2000)

2.5.3 แบบจำลองของการแพร่เสียง (radiation model)

เนื่องจากเสียงพูดที่ได้ยินกันนั้นมีความชันของสเปคตรัมประมาณ -6 dB/octave และแบบจำลองที่ได้นำเสนอไปนั้นมีความชันของแบบจำลองของกลอททัลพัลส์ (glottal pulse model) ประมาณ -12 dB/octave ดังนั้นเพื่อให้ได้เสียงออกมาใกล้เคียงกับเสียงธรรมชาติ ส่วนของการแพร่เสียง ควรจะมีความชันทางสเปคตรัม +6 dB/octave จึงแทนส่วนของการแพร่เสียงด้วยตัวกรองชนิดความถี่สูงผ่าน (high pass filter) โดยจะมีฟังก์ชันถ่ายโอนในสมการดังนี้ (ไชยยันต์ สุวรรณชีวะศิริ, 2541)

$$R(z) = 1\text{-}0.95z^{-1} \qquad (20.20)$$

อธิบายเป็นลักษณะทางกายภาพได้ว่าแบบจำลองส่วนนี้แทนการสูญเสียของพลังงานเนื่องจากแรงเสียดทานและความร้อนในระหว่างการออกเสียงซึ่งในแบบจำลองที่ได้นำเสนอไม่นำส่วนนี้มาคิด