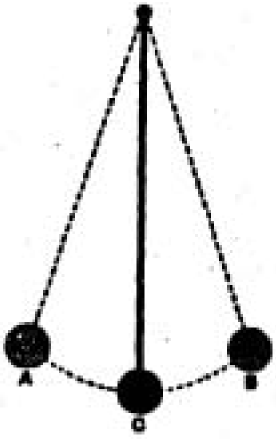
บทที่ ๖
กลสัทศาสตร์ (Acoustic Phonetics)

ความรู้เรื่องเสียงอันเป็นแขนงวิชาหนึ่งของฟิสิกส์ ได้มีการศึกษากันมาแล้วนานปี แต่เมื่อไม่นานมานี้ได้มีการนำเอาความรู้เรื่องเสียงตามแบบวิชาฟิสิกส์มาใช้กับการศึกษาทางภาษาศาสตร์ และในเพียงระยะเวลาอันสั้นนี้ทุกคนต่างก็ลงความเห็นเดียวกันว่ากลสัทศาสตร์มีส่วนช่วยให้เราเข้าใจเรื่องคุณสมบัติของเสียงดีมากขึ้น

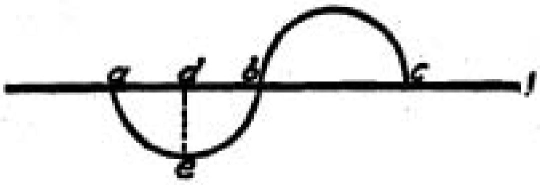
เสียง เสียงเป็นคลื่นซึ่งเดินทางไปในอากาศด้วยความเร็วประมาณ ๑.๑๐๐ ฟุต ต่อวินาที (ถ้าหากเป็นสื่ออย่างอื่น เช่น ของเหลว ก๊าซ ฯลฯ ความเร็วย่อมต่างกันออกไป) คลื่นเกิดขึ้นจากการสั่นสะเทือนอันเกิดขึ้นซ้ำ ๆ ฉะนั้นถ้าเราพิจารณาดูให้ดีจะเห็นว่าจุดกำเนิดของเสียงนั้นไม่ใช่เป็นสิ่งใดสิ่งหนึ่งที่สามารถสร้างเสียงขึ้นมาอย่างต้นกล้วยเป็นแหล่งกำเนิดของลูกกล้วย แต่ทว่าสิ่งที่จะทำให้เกิดเสียงขึ้นได้นั้นจะต้องเป็นสิ่งที่สามารถเปลี่ยนพลังงานรูปอื่นให้มาเป็นพลังงานเสียงได้ อย่างเช่นเราเอามือกระทุ่มน้ำ จะทำให้ผิวน้ำสั่นสะเทือนเกิดเป็นเสียงขึ้น หรืออย่างเวลาคนพูด เมื่อลมผ่านจากปอดเข้าหลอดลม จากหลอดลมเข้ากล่องเสียง ซึ่งมีเส้นเสียงอันเป็นใยกล้ามเนื้อ เมื่อเส้นเสียงสั่นก็จะออกมาเป็นเสียงที่เราพูดกัน เราถือว่าเสียงเป็นพลังงานอย่างหนึ่งที่สามารถแผ่กระจายออกไปได้รอบตัว อย่างพลังงานแสง และความร้อน

ตัวอย่างของการสั่นสะเทือนแบบง่าย ๆ ก็คือการเคลื่อนไหวของลูกตุ้มนาฬิกา



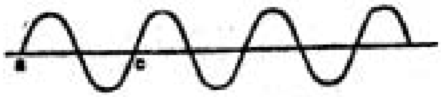
รูปที่ ๑

การเคลื่อนไหวแบบนี้อาจจะเขียนเป็นกราฟได้ดังนี้



รูปที่ ๒

การเคลื่อนที่จาก a ไป b และมาถึง c เรียกว่าหนึ่งรอบ (cycle) ระยะทาง de เรียกว่าความสูงของคลื่น (amplitude) ของการสั่นสะเทือน แกน t แทนเวลา ฉะนั้นการเคลื่อนที่แบบนี้เรียกว่าการเคลื่อนที่ในรูป sinusoidal curve ดูรูปที่ ๓



รูปที่ ๓

วัตถุทุกชิ้นมีความถี่ (frequency) ของการสั่นสะเทือนประจำตัว เป็นต้นว่า วัตถุที่หนัก สั่นสะเทือนช้ากว่าวัตถุที่เบา วัตถุที่ใหญ่ สั่นสะเทือนช้ากว่าวัตถุเล็ก ช่องว่างใดถ้ามีช่องเปิดเล็กก็มีความถี่น้อย ฉะนั้นถ้าต้องการให้มีความถี่สูงขึ้นก็เปิดช่องให้กว้างขึ้น เป็นต้นว่าการออกเสียงพูด เสียง อี ย่อมมีความถี่น้อยกว่า แอ เพราะเราเปิดปากเมื่อออกเสียงอีแคบกว่าเมื่อออกเสียงแอ

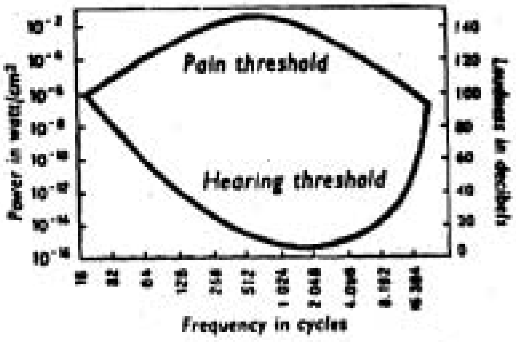
ความเข้มของเสียงและระดับเสียง เสียงที่เราได้ยินนั้นจะดังมากหรือน้อยขึ้นอยู่กับความเข้มของเสียง (intensity) และความไว (sensitivity) ของเครื่องรับเสียง เสียงใดที่มีความเข้มมากเราก็จะรู้สึกว่าดังมาก เสียงใดที่มีความเข้มน้อยก็จะรู้สึกว่าดังน้อย ตัวสำคัญที่ทำให้ความเข้มมากหรือน้อยนี้ก็คือความสูงของคลื่น (amplitude) ถ้าความสูงของคลื่นมากขึ้น ความเข้มก็มากขึ้น ความเข้มของเสียงนี้วัดได้เป็นพลังงานเสียงที่ผ่านออกมาในระยะเวลาหน่วยหนึ่ง ผ่านเนื้อที่ ๑ ตร.ซม. ที่ตั้งฉากกับการเคลื่อนที่ของการสั่นสะเทือน หน่วยที่วัดความเข้มของเสียงนี้ เป็นวัตต์ต่อ ตร.ซม. (watt per cm²)

ระดับเสียงขึ้นอยู่กับความถี่ ถ้าความถี่สูงเสียงก็สูง ถ้าความถี่น้อยเสียงก็ต่ำ หูคนได้ยินเสียงสั่นสะเทือนตาม logarithmic scale นั่นคือ เสียงที่สั่นสะเทือนเป็นสองเท่าเราจะได้ยินเป็นช่วงที่เท่ากัน ดังที่เรียกว่า octave ในวิชาการดนตรี ความถี่ของเสียงที่เรากล่าวถึงนี้เราวัดเป็นรอบต่อวินาที (cycles per second) หรือเขียนย่อ ๆ ว่า cps เป็นต้นว่าเสียงที่มีความถี่ ๑๐๐ กับ ๒๐๐ cps, ๒๐๐ กับ ๔๐๐ cps, ๔๐๐ กับ ๘๐๐ cps, ๘๐๐ กับ ๑๖๐๐ cps, ๑๖๐๐ กับ ๓๒๐๐ cps ฯลฯ หูเราได้ยินเหมือนกับว่ามีช่วงเท่ากันหมด โดยที่หูเรารับว่าความแตกต่างระหว่าง ๑๐๐ cps กับ ๒๐๐ cps เท่ากับหนึ่ง octave หรือ ๑๓ semitones และระหว่าง ๑๗๐๐ cps กับ ๑๘๐๐ cps ซึ่งมีความแตกต่างกัน ๑๐๐ cps เท่ากับความแตกต่างระหว่าง ๑๐๐ cps กับ ๒๐๐ cps เหมือนกัน แต่หูเรารับว่าความแตกต่างระหว่าง ๑๗๐๐ cps กับ ๑๘๐๐ cps เป็นเพียงหนึ่ง semitone เท่านั้น เพื่อให้เข้าใจเรื่อง octave ดีขึ้น ขอยกตัวอย่าง key เปียโนว่ามีความถี่เป็นอย่างไร

โน้ตแรกบน Keyboard	๒๗.๕	cps
	๕๕	cps
	๑๑๐	cps
	๒๒๐	cps
	๔๔๐	cps
	๘๘๐	cps
	๑๗๖๐	cps
	๓๕๒๐	cps

โน้ตทุกตัวจะต่างกับโน้ตตัวบนและโน้ตตัวล่าง หนึ่ง octave ถ้าดูง่าย ๆ ก็-คือว่าโน้ตตัวบน คูณด้วย ๒ จะเป็นค่าของความถี่ของโน้ตตัวล่าง หรือค่าของความถี่ของโน้ตตัวล่างหารด้วย ๒ จะเป็นความถี่ของโน้ตตัวบน ถ้าหากระดับเสียงต่างกัน ๒ octaves เราอาจจะคูณหรือหารด้วยความถี่ที่กำหนดให้ด้วย $๒^{๒}$ หรือ ๔ ก็จะได้ความถี่ที่ต้องการ หรือถ้าระดับเสียงต่างกัน ๓ octaves เราก็อาจจะคูณหรือหารความถี่ที่กำหนดให้ด้วย $๒^{๓}$ หรือ ๘ ก็จะได้ความถี่ที่ต้องการเหมือนกัน

ความไวของหูต่อความเข้มของเสียงแตกต่างกันไปขึ้นอยู่กับระดับเสียง ปกติหูคนเรารับได้ตั้งแต่ ๑๖ cps จนถึง ๑๖,๐๐๐ cps ดูรูปที่ ๔

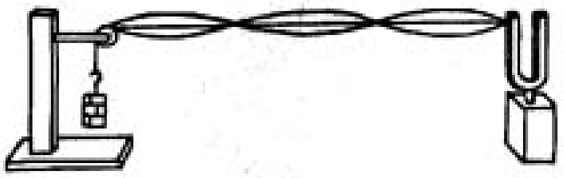


รูปที่ ๔

เสียงที่เหมาะกับหูคนก็คือ เสียงที่มีความถี่ ๖๐๐ cps ถึง ๔,๒๐๐ cps เสียงที่มีความถี่ ๓๐ cps จะให้ได้ยินเหมือนเสียง ๑๐๐๐ cps ก็จะต้องเพิ่มความเข้มขึ้นมากกว่า ๑๐๐๐ เท่า ทั้งนี้ก็เพราะว่าความดังของเสียงซึ่งมีหน่วยเป็น decibel(db) อัน

เป็นหน่วยเทียบความดังกับหน่วยมาตรฐาน ขึ้นอยู่กับค่าของ $\log_{10}$ ของความเข้มของเสียงที่เพิ่มขึ้น เป็นต้นว่าเราเพิ่มความเข้มเป็น ๒ เท่า ของความเข้มเดิม ความดังที่เราจะได้ยินก็จะเป็น logarithm ของ ๒ ซึ่งมีค่าเท่ากับ ๐.๓๐๑ ฉะนั้นแทนที่เราจะได้ยินดังเป็นสองเท่า เราจะได้ยินดังเพียง ๑.๓๐๑ เท่า เท่านั้นเอง

เสียงประสม เสียงที่เราได้ยินส่วนมากไม่ใช่เสียงแบบสามัญอย่างที่ปรากฏในรูปที่ ๒ แต่ว่าเป็นเสียงประสมมากกว่า ซึ่งทำการทดลองตาม Melde ได้โดยเอาเชือกหรือลวดมัดกับขาข้างหนึ่งของส้อมเสียงไฟฟ้า ปลายอีกข้างหนึ่งของเชือกพาดบนร่องของรอก โดยมีตุ้มน้ำหนักผูกปลายเชือกเพื่อถ่วงให้เชือกตึง จากนั้นก็ทำให้ส้อมสั่น เชือกจะสะบัดไปมาเป็นแบบคลื่นตามรูปที่ ๕



รูปที่ ๕

เราอาจจะแบ่งการสั่นสะเทือนออกเป็นรูปต่าง ๆ ได้ดังนี้

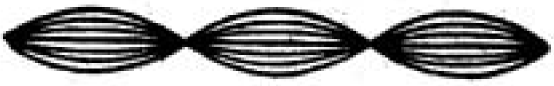
ก. แบบ fundamental คือเชือกจะสะบัดขึ้นลงเพียงวงเดียวดังรูปที่ ๖ ถือว่าให้ความถี่น้อย

รูปที่ ๖

ข. แบบ overtone ที่ ๑ เชือกจะสะบัดขึ้นลง ๒ วง ตามรูปที่ ๗

รูปที่ ๗

ค. แบบ overtone ที่ ๒ เชือกจะสะบัดขึ้นลง ๓ วง ตามรูปที่ ๗



รูปที่ ๗

ดังนั้นเชือกจะไม่สั่นตามแบบ fundamental เท่านั้นจะสั่นตามแบบ over tone ต่าง ๆ ด้วย เป็นลักษณะที่เรียกว่า harmonics ซึ่งเป็นเสียงที่เกิดจากเอาเสียง fundamental คูณด้วย ๒, ๓, ๔, ตามลำดับ ดังตัวอย่างเสียง bass ร้องที่ A (110 cps)

โน้ตร้อง	A	110 cps	fundamental
สูงขึ้นไป 1 octave	A	220 cps	2nd harmonic
	E	330 cps	3rd harmonic
สูงขึ้นไป 2 octaves	A	440 cps	4th harmonic
	C sharp	550 cps	5th harmonic
	E	660 cps	6th harmonic
ราว ๆ	G	770 cps	7th harmonic
สูงขึ้นไป 3 octaves	A	880 cps	8th harmonic
		990-1650 cps	9th, 10th, 11th, 12th, 13th, 14th, 15th, harmonic.

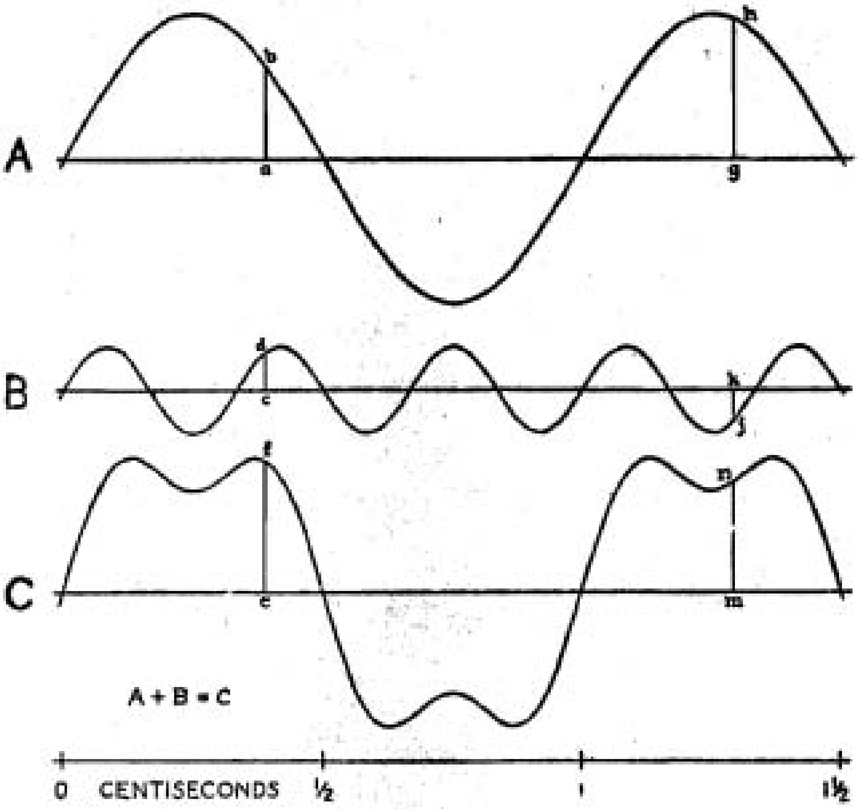
ทั้งหมดที่กล่าวมานี้เราอาจจะสรุปเกี่ยวกับเรื่องสำคัญ ๆ ของเสียงได้ดังนี้

๑. ความถี่มีจำนวนรอบต่อหน่วยเวลา (วินาที) ความถี่ของ fundamental (บอกระดับเสียงดนตรี) และ fundamental อาจจะมีค่าแตกต่างกันไปตามวัตถุที่เป็นต้นกำเนิดเสียง

๒. ความสูงของคลื่น ซึ่งเป็นตัวสำคัญที่ทำให้ความเข้มของเสียงมากหรือน้อย ทั้งนี้โดยถือว่าความถี่คงเดิม

๓. ระดับสูงต่ำของเสียง ขึ้นอยู่กับจำนวนของ harmonic ที่เราได้ยิน

ดังนั้นถ้าเสียงที่มีความถี่คล้ายกันนำมาประสมกัน ผลก็คือว่าความสูงของคลื่น (amplitude) จะเพิ่มขึ้น และความเข้มก็จะเพิ่มขึ้นด้วย ดูรูปที่ ๔ ABC ประกอบ รูปที่ ๔ C เป็นการประสมระหว่างรูปที่ ๔ A และ B



รูปที่ ๔

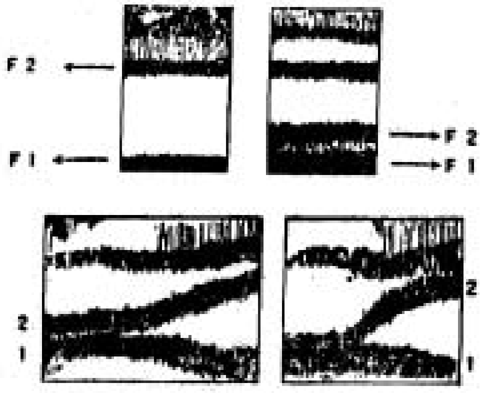
กำทอน* (Resonance) ถ้าเอาซอสองตัวมาวางไว้ใกล้ ๆ กัน สีตัวหนึ่งโดยไม่ต้องเอานิ้วไปกดสายอีกตัวหนึ่งจะดังขึ้นมาเอง อาการเช่นนี้เรียกว่า กำทอน ความจริงมีอยู่ว่าวัตถุทุกชนิดมีความถี่ประจำตัวของมัน ถ้าหากถูกกระตุ้นด้วยแรงที่มีความ

* วิชัย หโยคม "Resonance," *วิทยาศาสตร์* ปีที่ ๒๓ เล่มที่ ๗ กรกฎาคม ๒๕๑๒. หน้า ๕๖๓-๕๖๖.

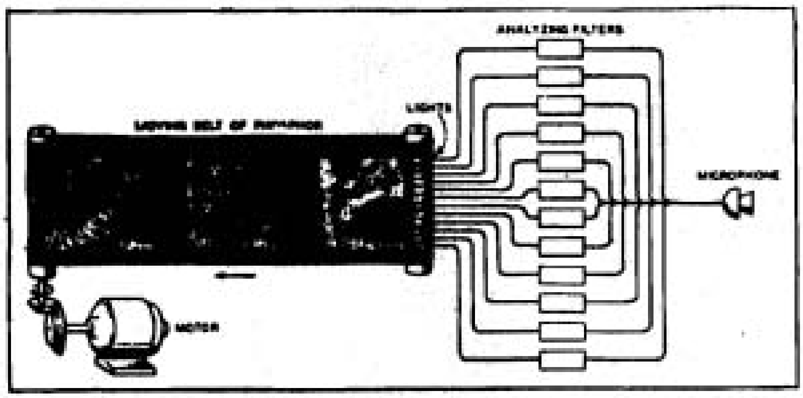
ถี่เท่ากับความถี่ประจำตัวของมันแล้ววัตถุนั้นจะไวต่อแรงกระตุ้นมากที่สุด เมื่อเราสีซอจะเกิดคลื่นเสียงแผ่ไปในอากาศ ไปกระตุ้นซออีกสายหนึ่ง ซึ่งมีความถี่เดียวกัน แต่ถ้าบังเอิญความถี่ประจำตัวต่างกันโดยเป็นซออู้ กับซอด้วงปรากฏการณ์เช่นนี้ก็จะไม่เกิดอะไรก็ตามที่สั่นสะเทือนได้ (เช่น ต้อมเสียง เชือก ช่องว่าง (cavity)) ซึ่งทำให้เสียงที่มีแล้วดังยิ่งขึ้นเรียกว่า resonator ถ้าความแตกต่างของความถี่ของ resonator และของการสั่นสะเทือนมีมาก ผลก้ำทอน (resonance effect) ก็มีน้อย ถ้าความแตกต่างเกินขีดจำกัดขีดหนึ่งจะไม่มีผลใด ๆ เลย นั่นก็คือลักษณะก้ำทอนจะไม่เกิดขึ้น

Filters โดยอาศัยลักษณะก้ำทอนที่เรียนมาแล้ว เราสามารถเสริมความถี่ ซึ่งปรากฏในเสียงประสมแล้วทำให้ความสูงต่ำของเสียงเปลี่ยนไป ถ้าเราเสริมเสียงที่มี harmonic สูง เสียงสูงต่ำก็จะชัดเจนขึ้น ถ้าเราเสริมเสียง fundamental หรือเสียงที่มี harmonic ต่ำ เสียงสูงต่ำก็จะด้อยลง วิธีการเสริมความถี่โดยทำให้ลักษณะอื่นด้อยลงนี้เรียกว่า filter เราสามารถเปลี่ยนรูปและปริมาตรของช่องปากช่องจมูกที่เราใช้เป็นอวัยวะในการพูดด้วยการเคลื่อนไหว กล่องเสียง ลิ้น ริมฝีปาก เพดานอ่อนก็จะทำให้ลักษณะก้ำทอนเปลี่ยนไป เราถือว่าช่องปาก และช่องจมูกเป็น filter

Formants ความถี่หรือความถี่กลุ่มหนึ่งซึ่งเป็นลักษณะของเสียง ๆ หนึ่งซึ่งแตกต่างไปจากเสียงอื่นเรียกว่า formants ปัจจุบันนี้เรามีเครื่องมืออีเลกโทรนิคส์ที่เรียกว่า sound spectrograph และเครื่องมือนี้จะให้ sound spectrogram ซึ่งเป็นภาพการวิเคราะห์เสียงว่ามีความถี่เท่าใด ความสูงของคลื่นเท่าใด ถ้าเราได้พิจารณาภาพใน sound spectrogram จะเห็นว่าจะมีบริเวณที่มี harmonic อยู่หนาแน่นจะเห็นเป็นเส้นดำหนาผิดกว่าบริเวณอื่นได้ถนัด ดูรูปที่ ๑๐ ประกอบ



รูปที่ ๑๐



รูปที่ ๑๑
sound spectrograph

บริเวณที่เห็นเป็นดำ ๆ หนาแน่นนี้แหละเรียกว่า formant บางทีก็เป็นเพียง harmonic เดียว บางทีก็หลาย harmonics

เป็นที่ทราบกันโดยทั่วไปว่า ในภาษาพูดของคน เสียงสระจะมีอย่างน้อยสอง formants และ formants ทั้งสองนี้เกี่ยวพันกับช่องก้าทอน (resonance chambers) สองช่องของอวัยวะออกเสียง คือ ช่องคอ และช่องปาก ส่วน formant ที่สาม F3 และ formant อื่น ๆ บางอันก็ไม่สู้จะมีความสำคัญนัก เนื่องจากประเทศไทยเราในขณะที่แต่งหนังสือเล่มนี้ไม่มี sound spectrograph (ดูรูปที่ ๑๑) เลยสักเครื่องเดียว ปัจจุบันนี้สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัยและสำนักวิทยาการกรมตำรวจมีใช้อยู่แห่งละหนึ่งเครื่อง ท่านที่สนใจอาจจะแวะไปขอชม sound spectrograph ได้ หากท่านต้องการดูรูปเสียงของท่านขอให้ใช้ broad-band ซึ่งมีความกว้างของ band เป็น ๓๐๐ cps ถ้าต้องการจะเห็น formants อย่างชัดเจนใช้ narrow-band ซึ่งมีความกว้างของ band เพียง ๔๕ cps แต่ข้อเสียของ broad-band ก็คือว่ามันยากแก่การวัดหาความถี่ของเสียง แต่ละเสียงขอให้เปรียบเทียบเสียงที่ใช้ broad-band และ narrow-band ดังรูปที่ ๑๒

เสียงผู้ชาย ใช้ narrow-band

เสียงผู้ชาย ใช้ broad-band หรือ wide band

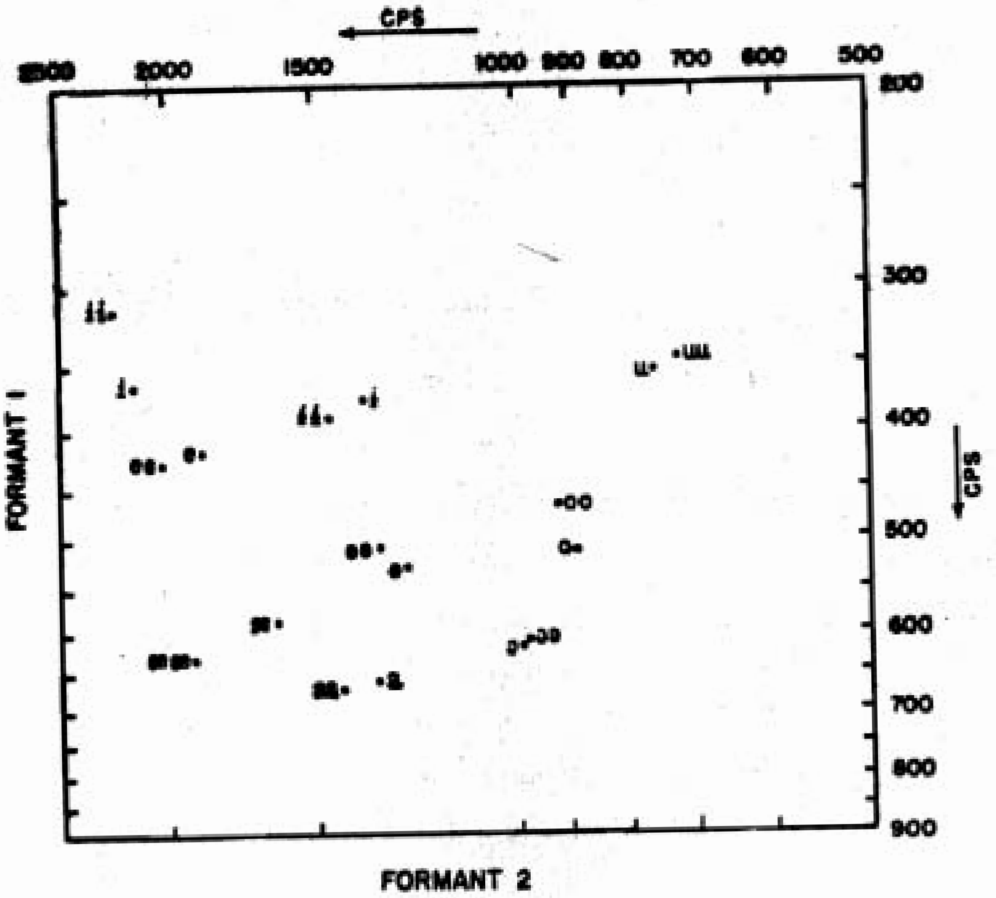
รูปที่ ๑๒

เสียงสระ ตามที่กล่าวมาข้างต้นแล้วว่าเสียงสระมี formants ที่สำคัญ ๒ formants เราเรียกว่า formant ที่หนึ่ง (F1) คือ บริเวณดำ ๆ อันล่างใน sound spectrogram และ formant ที่สอง (F2) คือบริเวณดำ ๆ อันถัดขึ้นไป ดูได้จากรูปที่ ๑๐

เสียงสระในภาษาไทย ศาสตราจารย์ ดร. อาเธอร์ แอบแรมสัน ผู้ซึ่งเคยเป็นอาจารย์ฟุลไบร์ทที่วิทยาลัยวิชาการศึกษาประสานมิตรได้ทำการวิเคราะห์ภาษาไทยโดยใช้ sound spectrograph จากเสียงของ ดร. เอกวิทย์ ณ ถลาง ขณะที่ศึกษาอยู่ที่มหาวิทยาลัยโคลัมเบียสหรัฐอเมริกาว่ามี formants ดังนี้

[i]	๓๕๐, ๒๑๒๐	[ɨ]	๓๕๐,๑๓๓๐	[u]	๓๗๐,๖๔๐
[ii]	๓๓๐, ๒๒๖๐	[ɨɨ]	๔๐๐, ๑๔๗๐	[uu]	๓๗๐,๘๗๐
[e]	๔๖๐, ๑๗๖๐	[ə]	๕๖๐, ๑๒๔๐	[o]	๕๐๐,๘๓๐
[ee]	๔๖๐, ๒๐๕๐	[əə]	๔๗๐,๑๓๓๐	[oo]	๔๕๐,๗๖๐
[ɛ]	๖๑๐,๑๖๒๐	[a]	๖๖๐,๑๒๗๐	[ɔ]	๖๑๐,๗๕๐
[ɛɛ]	๖๔๐,๑๔๗๐	[aa]	๖๕๐,๑๔๐๐	[ɔɔ]	๖๔๐,๑๐๗๐

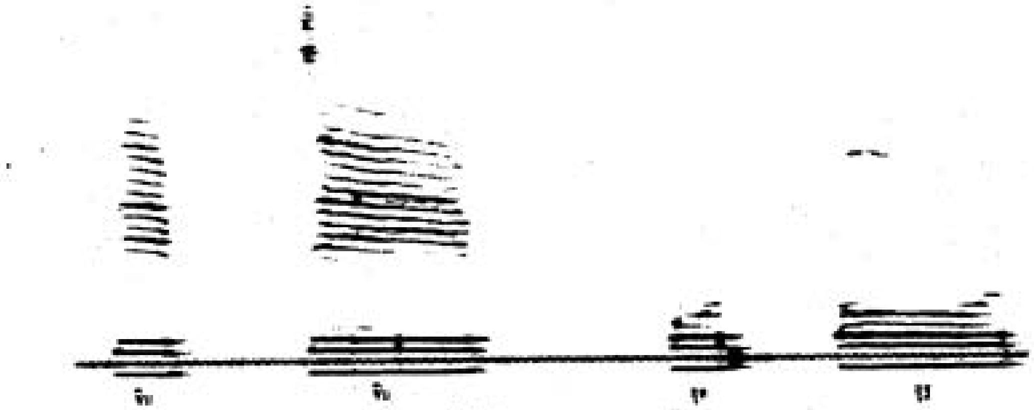
ตัวแรกตัวหน้าเป็น F1 และตัวหลังเป็น F2 ซึ่งเขียนเป็นรูปได้ดังนี้



รูปที่ ๑๓

เสียงสระของ ดร. เอกวิทย์ ณ ถลาง

ถ้าสังเกตจากรูปที่ ๑๔ จะเห็นว่าสระเสียงยาวจะมีความยาว หรืออีกนัยหนึ่งจะต้องใช้เวลานานเกือบ ๓ เท่าของสระเสียงสั้น ฉะนั้นการที่นักภาษาศาสตร์บางคนเขียนสระเสียงสั้นด้วยสัญญลักษณ์ตัวเดียว เช่น [i] และสระเสียงยาวด้วยสัญญลักษณ์ของสระเสียงนั้นสองตัว [ii] อาจจะทำให้บางคนคิดว่าสระเสียงยาวมีความยาวเป็นสองเท่าของสระเสียงสั้นซึ่งผิดความจริง



เสียงผู้หญิง ใช้ narrow-band

รูปที่ ๑๔

เสียงสระภาษาอังกฤษ จากการทดลองเสียงสระในภาษาอังกฤษปรากฏผลออกมาโดยเฉลี่ยดังนี้

[i]	๓๙๐,๑๙๙๐	[ɨ]	๓๐๐,๑๕๐๐	[u]	๔๔๐,๑๐๒๐
[e]	๕๓๐,๑๘๔๐	[ə]	๖๐๐,๑๓๐๐	[o]	๕๕๐,๙๐๐
[æ]	๖๖๐,๑๗๒๐	[a]	๗๓๐,๑๐๙๐	[ɔ]	๕๗๐,๘๔๐

เวลาเราได้ยินเสียงพูดโดยไม่ได้เห็นตัวผู้พูดเราก็ทราบได้ว่าผู้พูดเป็นหญิงหรือผู้ชาย[๒] ทั้งนี้ก็เพราะว่าความถี่ของเสียงผู้หญิงสูงกว่าความถี่ของเสียงผู้ชายซึ่งอาจจะเทียบได้ดังนี้

ความถี่ของ formant ที่หนึ่ง

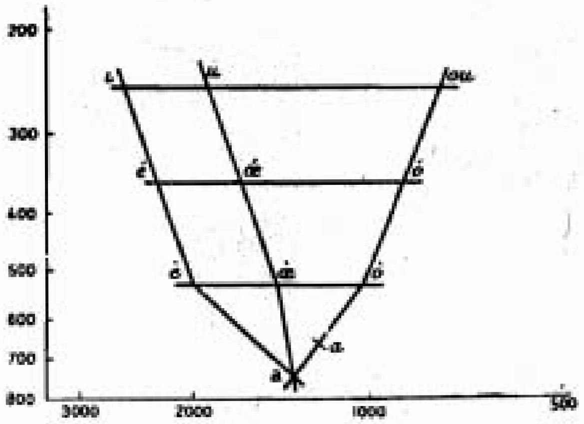
	i	e	æ	ɨ	ə	a	u	o	ɔ
ชาย	๓๙๐	๕๓๐	๖๖๐	๓๐๐	๕๑๐	๗๓๐	๔๔๐	๕๕๐	๕๗๐
หญิง	๔๓๐	๖๑๐	๘๖๐	๓๗๐	๕๕๐	๘๕๐	๔๗๐	๖๐๐	๕๙๐

ความถี่ของ formant ที่สอง

ชาย	๑๙๙๐	๑๘๔๐	๑๗๒๐	๑๕๐๐	๑๓๕๐	๑๐๙๐	๑๐๒๐	๙๐๐	๘๔๐
หญิง	๒๔๘๐	๒๓๓๐	๒๐๕๐	๑๗๒๐	๑๖๔๐	๑๒๒๐	๑๑๖๐	๑๐๑๐	๙๒๐

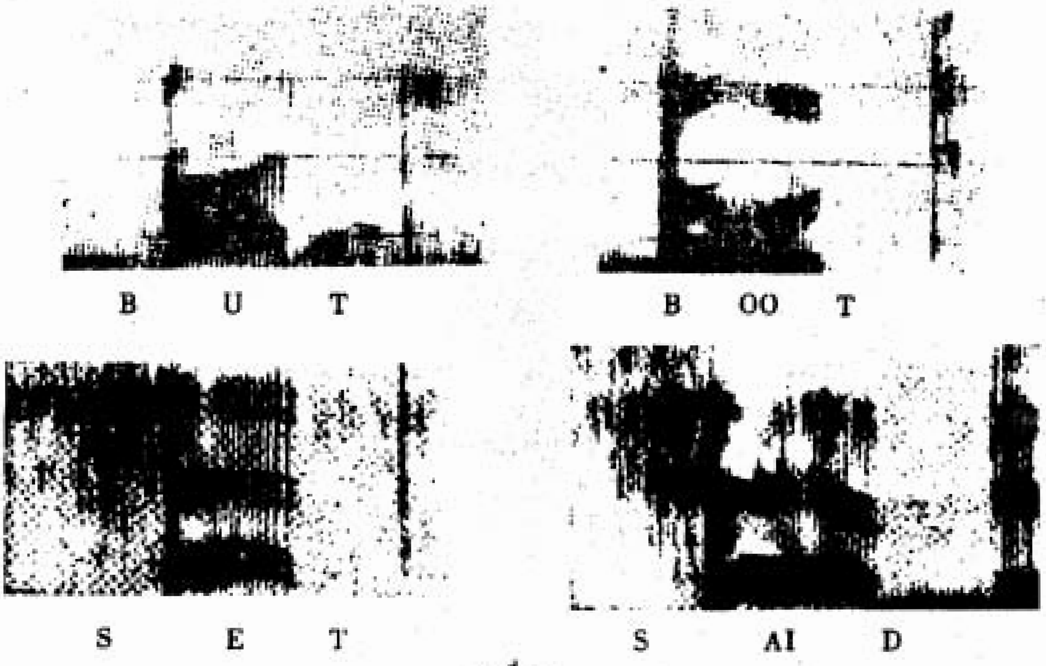
ขอให้สังเกตกราฟข้างล่างในรูปที่ ๑๕ ที่แสดง F1 และ F2 ของสระในภาษาฝรั่งเศสแกน y แสดง F1 และ x แสดง F2

[๒] ความถี่เฉลี่ยของเสียงผู้ชายอยู่ระหว่าง ๑๐๐ ถึง ๑๕๐ cps ของผู้หญิงระหว่าง ๒๐๐ ๓๐๐ cps เสียงผู้ชายที่ต่ำที่สุด ๖๐ cps เสียงโซปราโนอยู่ระหว่าง ๑.๒๐๐ ถึง ๑๓๐๐ cps.



รูปที่ ๑๕

เสียงพยัญชนะ ความรู้เรื่องเสียงพยัญชนะตามหลักกลสัทศาสตร์ นี้เรายังรู้น้อยมากเราทราบว่าเสียงพยัญชนะหยุด พยัญชนะเสียดสีไม่มี formants (F1, F2, F3, ฯลฯ) อย่างสระดังที่กล่าวมาแล้ว เสียงเสียดสีมีความถี่สูง เช่น [s] มีความถี่ถึง ๔,๐๐๐-๘,๐๐๐ cps เมื่อเทียบกับ [ʃ] มีความถี่ราว ๖,๐๐๐-๗,๐๐๐ cps เราอาจจะสังเกตความแตกต่างระหว่างเสียงหยุดกับเสียงเสียดสีจาก sound spectrogram ได้ ดูรูปที่ ๑๖ เราจะสังเกตเห็นว่าเสียงหยุดจะเป็นเส้นตรง



รูปที่ ๑๖

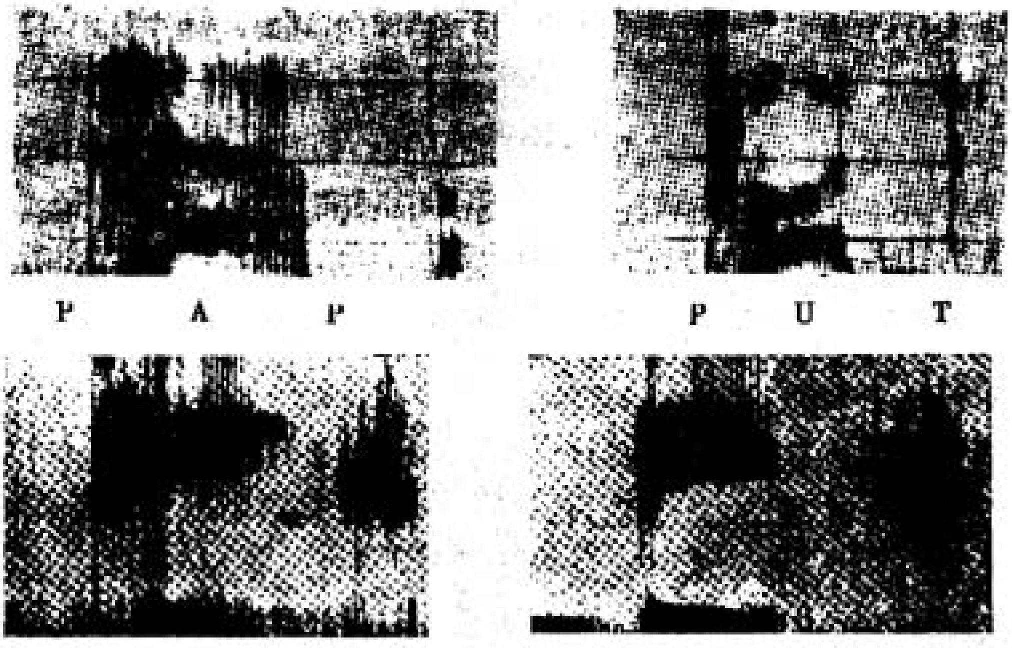
M I L N EA T

T U N

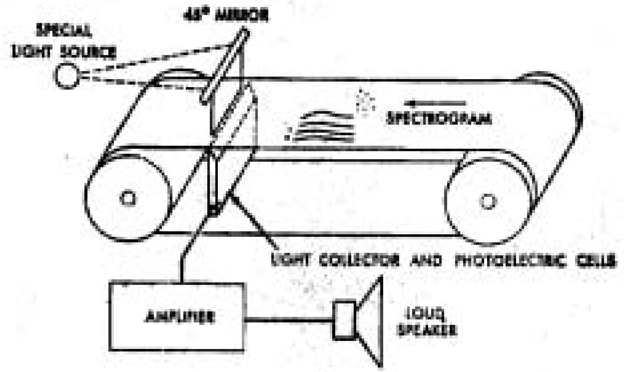
รูปที่ ๑๗

นำหน้าสระ ส่วนเสียงเสียดสีจะมีเหมือนกลุ่มควันลอยอยู่หน้าสระ ถ้าสังเกตดูเสียงนาสิกจะเห็น F2 ได้ถนัด ส่วนบริเวณดำๆข้างล่างนั้นแสดงเสียงโฆษะดังจะเห็นได้จากความแตกต่างระหว่าง [t] และ [d] ในรูปที่ ๑๖ข้างบน ที่จริงแล้วเสียงนาสิก [m, n, ŋ] และเสียงพยัญชนะเหลว [l, r] ที่เราจัดไว้เป็นเสียงพยัญชนะ ที่จริงแล้วมีลักษณะทางสัทศาสตร์คล้ายสระมากดูรูปที่ ๑๗ นอกจากนั้นแล้ว เรายังอาศัยการบิดเบนของ F2 ของสระที่ตามมาเป็นเครื่องบอกว่าพยัญชนะที่นำหน้าสระนั้นมีฐานกรณ์อยู่ที่ใด ขอให้สังเกตดูรูปที่ ๑๘

จะสังเกตเห็นได้ว่า เสียงที่มีฐานกรณ์ที่ริมฝีปากทั้งสอง เช่น [p] ก็ดี หรือ [m] ก็ดี F2 หักลง ส่วนเสียงที่มีฐานกรณ์ที่ปุ่มเหงือกเช่น [t] ก็ดีหรือ [n] ก็ดี F2 จะหักขึ้นเล็กน้อย แต่เสียงที่มีฐานกรณ์อยู่ที่เพดานอ่อนอย่าง [k] หรือ [ŋ] F2 หักขึ้นมาก



รูปที่ ๑๘



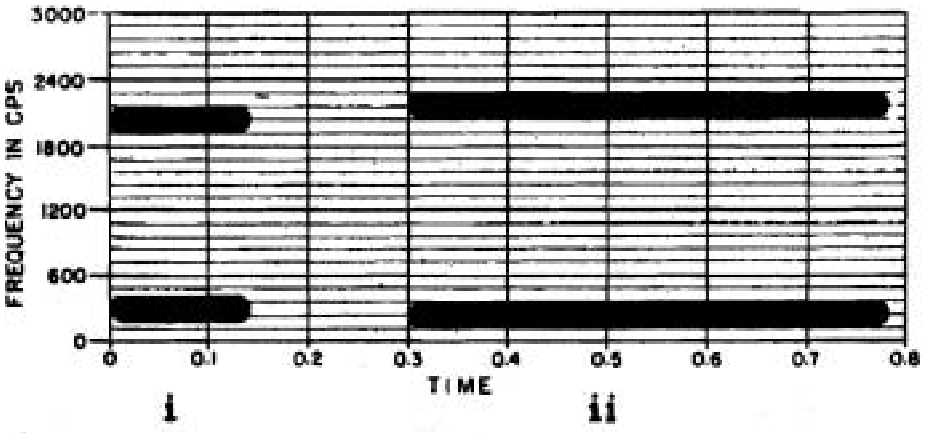
รูปที่ ๑๙
speech synthesizer

เสียงสังเคราะห์ (synthetic sounds) เป็นการเปลี่ยนจากภาพที่ปรากฏใน sound spectrogram ให้เป็นเสียงโดยมีเครื่องมือที่จะอ่านภาพออกมาเป็นเสียง เครื่องมือนั้นเรียกว่า speech synthesizer ดูรูปที่ ๑๙ นักภาษาศาสตร์แทนที่จะใช้ sound spectrogram จริง ๆ ใส่เข้าไปใน speech synthesizer กลับเอากระดาษมาป้ายด้วยสีดำ ๆ ให้เป็นรูป formants ต่าง ๆ แล้วใส่ให้เครื่องอ่านออกมาเป็นเสียงเพื่อเป็นการศึกษาว่าเสียงที่ออกมานั้น เจ้าของภาษายอมรับว่าเป็นเสียงที่

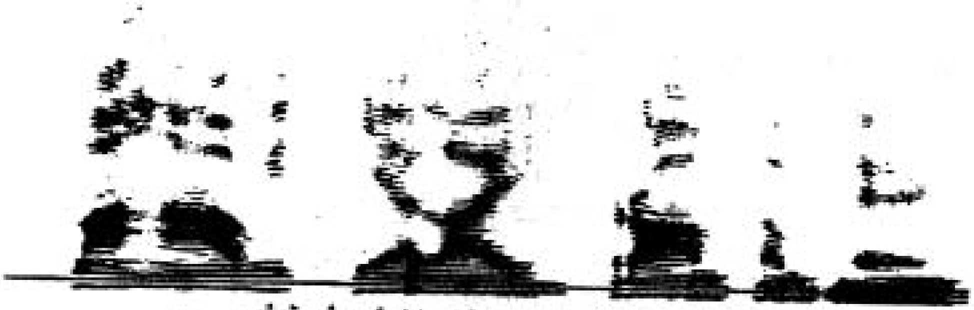
ต้องการหรือเปล่า วิธีทดสอบก็ทำได้ดังนี้ คือเราทราบว่าเสียง [อี] ในภาษาไทย F1 = 330, F2 = 2260 ก็ป้ายกระดาษด้วยสีดำในตำแหน่งดังกล่าว ส่งเข้า speech synthesizer แล้วก็จะออกมาเป็นเสียง คนไทยก็คอยฟังว่าเสียงที่ออกมาคล้ายกับเสียง [อี] ที่คนไทยออกไหม ถ้าไม่คล้ายก็ย่อมแสดงว่านอกเหนือจาก formants ทั้งสองแล้วอาจจะมีลักษณะอื่นที่จะต้องเติมเข้าไปอีก เป็นต้นว่า F3 เพื่อให้เสียงใกล้เคียงกับเสียงเจ้าของภาษายิ่งขึ้น ขอให้สังเกตรูปที่ ๒๐ และรูปที่ ๒๑ ข้างล่างนี้เป็นเป็นการเปรียบระหว่างภาพที่ปรากฏใน sound spectrogram และภาพที่สังเคราะห์ขึ้น

ดร. อาเธอร์ เอส. แอบแรมสัน ได้ทำการวิจัย และค้นพบว่า สระสังเคราะห์สำหรับภาษาไทยที่คนไทยฟังแล้วรู้สึกว่าใกล้เคียงกับภาษาไทยที่สุดควรจะมี formants ดังนี้

[i]	๓๖๐,๒๑๐๐	[ɨ]	๓๐๐,๑๓๔๐	[u]	๓๖๐,๗๒๐
[ii]	๓๐๐,๒๒๒๐	[ɨɨ]	๓๐๐,๑๓๔๐	[uu]	๓๐๐,๖๖๐
[e]	๔๕๐,๑๙๔๐	[ə]	๔๕๐,๑๒๐๐	[o]	๔๔๐,๘๕๐
[ee]	๔๔๐,๑๘๔๐	[əə]	๔๕๐,๑๒๖๐	[oo]	๔๔๐,๘๔๐
[ɛ]	๗๔๐,๑๗๐๐	[a]	๗๒๐,๑๒๐๐	[ɔ]	๖๖๐,๙๖๐
[ɛɛ]	๗๒๐,๑๗๐๐	[aa]	๗๔๐,๑๒๖๐	[ɔɔ]	๖๖๐,๙๖๐



รูปที่ ๒๐
ภาพเสียงที่สังเคราะห์ขึ้น



เสียงผู้ชายเมื่อออกเสียงคำว่า มหาวิทยาลัยรามคำแหง

เสียงผู้หญิงเมื่อออกเสียงคำว่า มหาวิทยาลัยรามคำแหง