

การแยกแยะความผิดปกติของเสียงพูดชนิดสປาสติคในผู้ป่วยทางระบบประสาทออกจากเสียงพูด
ปกติ ด้วยการวิเคราะห์คลื่นเสียงด้วยแบบจำลองทางคอมพิวเตอร์



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาอายุรศาสตร์ ภาควิชาอายุรศาสตร์
คณะแพทยศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2565
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

The Identification of Spastic Dysarthria in Neurological Patients by Speech Analysis
with Computational Modelling



A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Medicine

Department of Medicine

FACULTY OF MEDICINE

Chulalongkorn University

Academic Year 2022

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การแยกแยะความผิดปกติของเสียงพูดชนิดสປาสติกในผู้ป่วยทางระบบประสาทออกจากเสียงพูดปกติ ด้วยการวิเคราะห์คลื่นเสียงด้วยแบบจำลองทางคอมพิวเตอร์
โดย	นายชยุต มฤคทัต
สาขาวิชา	อายุรศาสตร์
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	ดร. นายแพทย์ชัยภัทร ชุณหรัศม์

คณะแพทยศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้หัวข้อวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของ
การศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

..... คณบดีคณะแพทยศาสตร์
(รองศาสตราจารย์ นายแพทย์ฉันทชาย สิทธิพันธุ์)

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ
(ผู้ช่วยศาสตราจารย์ แพทย์หญิงศรีัญญา ภูวนันท์)

..... อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ดร. นายแพทย์ชัยภัทร ชุณหรัศม์)

..... กรรมการ
(รองศาสตราจารย์ นายแพทย์กำพล สุวรรณพิมลกุล)

..... กรรมการภายนอกมหาวิทยาลัย
(ดร.อิทธิ ฉัตรนนทเวช)

ชยุด มฤคทัต : การแยกแยะความผิดปกติของเสียงพูดชนิดสປาสติกในผู้ป่วยทางระบบประสาทออก
จากเสียงพูดปกติ ด้วยการวิเคราะห์คลื่นเสียงด้วยแบบจำลองทางคอมพิวเตอร์. (The
Identification of Spastic Dysarthria in Neurological Patients by Speech Analysis with
Computational Modelling) อ.ที่ปรึกษาหลัก : ดร. นพ.ชัยภัทร ชุณหรัศม์

วัตถุประสงค์ เพื่อประเมินประสิทธิภาพในการวินิจฉัยอาการพูดไม่ชัดชนิดสປาสติกในแง่ความไว
ความจำเพาะ และ AUC จาก กราฟ ROC ของการวินิจฉัยประโยคพูดไม่ชัดที่สร้างขึ้นจากลักษณะเด่นทางคลินิก
และพยาธิสรีรวิทยา และประเมินด้วยเครื่องมือ ASR พร้อมทั้งเปรียบเทียบความแม่นยำในการวินิจฉัยกับ
ประสาทแพทย์

วิธีการวิจัย ผู้ป่วยพูดไม่ชัดชนิดสປาสติกจำนวน 37 คน และกลุ่มควบคุม 37 คน เข้ารับการ
บันทึกเสียงพูด 4 ประโยคที่ถูกสร้างขึ้นจากกลุ่มพยัญชนะต้นที่แตกต่างกันตามการทำงานของกล้ามเนื้อการพูด
และประเมินคะแนนความผิดพลาดของพยางค์แต่ละประโยค (error score of syllable) ด้วย
เครื่องมือ ASR ได้แก่ 'Apple Siri™' และ 'Whisper' แล้ววิเคราะห์ logistic regression analysis และ สร้าง
กราฟ ROC พร้อมทั้ง AUC เพื่อบอกความแม่นยำในการวินิจฉัย พร้อมทั้งให้ประสาทแพทย์วินิจฉัยอาการพูดไม่
ชัดจากไฟล์เสียงเดียวกัน

ผลการศึกษา ค่าเฉลี่ยคะแนนความผิดพลาดของพยางค์แต่ละประโยคระหว่างกลุ่มผู้ป่วยพูดไม่ชัด
ชนิดสປาสติกและกลุ่มควบคุมมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ($p < 0.05$) ทั้งจากทั้งสองเครื่องมือ
โดยที่ AUC สูงที่สุดของเครื่องมือ 'Apple Siri™' และ 'Whisper' เท่ากับ 0.95 และ 0.89 ตามลำดับในการ
วิเคราะห์ประโยคเดียวกันที่เป็นการเคลื่อนไหวของกล้ามเนื้อโคนลิ้น เพดานอ่อนและคอหอย ในขณะที่ประสาท
แพทย์มีความจำเพาะในการวินิจฉัยมากกว่า 0.9 แต่มีความไวที่ไม่แน่นอนตั้งแต่ 0.20 ถึง 0.74

สรุป เครื่องมือ 'Apple Siri™' และ 'Whisper' ASR มีความแม่นยำและน่าเชื่อถือในการวินิจฉัย
แยกแยะอาการพูดไม่ชัดชนิดสປาสติกที่มีความรุนแรงน้อยออกจากเสียงพูดปกติ เมื่อนำมาประยุกต์ใช้กับการ
วิเคราะห์ประโยคพูดที่มีความไม่ชัดชนิดสປาสติกมากที่สุด ซึ่งถูกสร้างขึ้นด้วยความรู้ทางพยาธิสรีรวิทยาและ
สัทศาสตร์

สาขาวิชา อายุรศาสตร์

ลายมือชื่อนิสิต

ปีการศึกษา 2565

ลายมือชื่อ อ.ที่ปรึกษาหลัก

6470015930 : MAJOR MEDICINE

KEYWORD: spastic dysarthria, Automatic Speech Recognition (ASR), incorrect syllable,
lingual-alveolar placement

Chayoot Marukatat : The Identification of Spastic Dysarthria in Neurological Patients
by Speech Analysis with Computational Modelling. Advisor: Chaipat Chunharas, Ph.D.

Objective : To evaluate the diagnostic accuracy of ASR tools in detecting spastic dysarthria using pathophysiologically-informed sentences and comparing the results to those obtained by neurologists through calculating sensitivity, specificity, accuracy, and generating ROC curves with AUC.

Methods : 37 dysarthric patients and 37 healthy controls, were enrolled in four sentence recordings, generated based on group of muscles activated to produce consonant sounds. Recordings were 'scored' by two ASR software, 'Apple Siri™' and 'Whisper'. The scoring was error count of incorrect syllables and was analyzed using logistic regression with ROC/AUC representing diagnostic performance. Neurologists' diagnostic accuracy was compared with ASR result.

Results : Mean errors generated by both ASR tools were shown to be statistically significant difference with $p < 0.05$. Maximum AUC of 'Apple Siri™' and 'Whisper' ASR was 0.95 and 0.89 respectively for the sentence articulated with velopharyngeal and glottis placement. Neurologists showed a specificity exceeding 0.90 and sensitivity varying from 0.20 to 0.74.

Conclusion : 'Apple Siri™' and 'Whisper' ASR were found to be effective and reliable in differentiating mild spastic dysarthria from control. Feature selection, based on pathophysiological and phonological basis, strengthens sentence-specific dysarthria detection.

Field of Study: Medicine

Student's Signature

Academic Year: 2022

Advisor's Signature

กิตติกรรมประกาศ

งานวิจัยครั้งนี้สำเร็จลุล่วงมาได้ด้วยความกรุณาให้คำปรึกษาและชี้แนะแนวทางจากศูนย์วิจัย Cognitive and Clinical & Computational Neuroscience (CCCN) ณ อาคารผู้สูงอายุ สร. โรงพยาบาลจุฬาลงกรณ์ และได้รับการสนับสนุนและความช่วยเหลือเป็นอย่างดีจากอาจารย์ นายแพทย์ ชัยภัทร ชุณหรัศมิ์ ซึ่งเป็นอาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก ซึ่งผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

อีกหนึ่งบุคคลที่เป็นกำลังสำคัญที่ทำให้งานวิจัยชิ้นนี้สำเร็จลุล่วงไปด้วยดีคือคุณชนัญญา ผลลาก นักอรรถบำบัดผู้ป่วยที่มีปัญหาทางด้านกรพูด ผู้คอยให้คำปรึกษาและคำแนะนำในเรื่องสัทศาสตร์ (phonology) และช่วยออกแบบประโยคพูดที่เป็นขั้นตอนสำคัญในงานวิจัยครั้งนี้ อีกทั้งยังเป็นผู้ประเมินความรุนแรงของอาการพูดไม่ชัดผ่าน intelligibility score อีกด้วย

บุคคลสำคัญที่ช่วยในขั้นตอนการเก็บข้อมูลคือนิสิตแพทย์ชั้นปีที่ 5 ประกอบด้วย นางสาวชนนภร ไพจิตร นายชนศิษฐ์ เวชกิจวานิชย์ และ นางสาวกานต์พิชามา หงษ์ทอง งานวิจัยนี้จะไม่สามารถสำเร็จได้เลยหากปราศจากความช่วยเหลือจากบุคคลเหล่านี้

ผู้วิจัยรู้สึกซาบซึ้งในความกรุณาของท่านที่กล่าวมา ตลอดจนผู้ที่ไม่ได้กล่าวนามในที่นี้ซึ่งมีส่วนให้งานวิจัยสำเร็จลุล่วงไปด้วยดี

สุดท้ายนี้ กราบขอบพระคุณบิดา มารดา ที่ให้ความช่วยเหลือและเป็นกำลังใจตลอดมา

สารบัญ

	หน้า
.....	ค
บทคัดย่อภาษาไทย.....	ค
.....	ง
บทคัดย่อภาษาอังกฤษ.....	ง
กิตติกรรมประกาศ.....	จ
สารบัญ.....	ฉ
สารบัญตาราง.....	ฅ
สารบัญรูปภาพ.....	ญ
บทที่ 1 บทนำ.....	1
1.1 ความสำคัญและที่มาของปัญหาการวิจัย.....	1
1.2 คำถามของการวิจัย.....	2
1.3 วัตถุประสงค์งานวิจัย.....	2
1.4 สมมติฐาน.....	2
1.5 ข้อตกลงเบื้องต้น.....	3
1.6 กรอบความคิดแนววิจัย.....	3
1.7 การให้คำนิยามเชิงปฏิบัติที่จะใช้ในการวิจัย.....	3
1.8 ผลหรือประโยชน์ที่คาดว่าจะได้รับจากงานวิจัย.....	5
1.9 อุปสรรคที่อาจเกิดขึ้นระหว่างการวิจัยและมาตรฐานการแก้ไข.....	6
บทที่ 2 ทบทวนวรรณกรรมที่เกี่ยวข้อง.....	7
2.1 สัทศาสตร์ของเสียงพูดไม่ชัด (Phonetics of dysathria).....	7
2.2 การวินิจฉัยอาการพูดไม่ชัดในปัจจุบัน.....	14

2.3 เครื่องมือวิเคราะห์เสียงพูดในปัจจุบันและการนำมาใช้ในคลินิก	16
บทที่ 3 วิธีการดำเนินการวิจัย	20
3.1 รูปแบบการวิจัย	20
3.2 ระเบียบวิธีการวิจัย.....	20
3.3 ขนาดตัวอย่าง	21
3.4 ขั้นตอนการทำวิจัย.....	22
3.5 การรวบรวมข้อมูล.....	27
3.6 ข้อจำกัดในการวิจัย	27
3.7 การเปิดเผยข้อมูลแสดงตัวตนของผู้ป่วย.....	28
3.8 การวิเคราะห์ข้อมูล	28
บทที่ 4 ผลการวิเคราะห์ข้อมูล.....	32
4.1 ข้อมูลพื้นฐานของประชากรที่นำมาศึกษา.....	32
4.2 กราฟคลื่นเสียงที่ได้จากการวิเคราะห์เบื้องต้น	35
4.3 ผลการเปรียบเทียบปัจจัยที่ศึกษา.....	36
บทที่ 5 อภิปรายผล สรุปผลการวิจัย และข้อเสนอแนะ	47
5.1 อภิปรายผล	47
5.2 สรุปผลการวิจัย.....	50
5.3 เปรียบเทียบกับการศึกษาก่อนหน้านี้.....	50
5.4 ข้อดีของการศึกษานี้.....	51
5.5 ข้อด้อยของการศึกษานี้	51
5.6 ข้อเสนอแนะ	52
บรรณานุกรม.....	54
ภาคผนวก.....	58
ประวัติผู้เขียน.....	63



จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

สารบัญตาราง

	หน้า
ตารางที่ 1 แสดงการแบ่งพยัญชนะต้นตามตำแหน่งของกล้ามเนื้อที่ออกเสียง (placement)	12
ตารางที่ 2 แสดงข้อมูลพื้นฐานของผู้ป่วยที่มีอาการพูดไม่ชัดชนิดสປาสติกและกลุ่มควบคุม.....	35
ตารางที่ 3 แสดงจำนวนผู้ป่วยและกลุ่มควบคุมตามคะแนนความผิดพลาดของพยางค์ในแต่ละ ประโยค (error score of syllable) ที่ได้หลังการนับจำนวนพยางค์ที่ผิดในประโยค (incorrect syllable) ที่ถูกวิเคราะห์ด้วย Whisper-Thai-demo	37
ตารางที่ 4 แสดงจำนวนผู้ป่วยและกลุ่มควบคุมตามคะแนนความผิดพลาดของพยางค์ในแต่ละ ประโยค (error score of syllable) ที่ได้รับหลังการนับจำนวนพยางค์ที่ผิดในประโยค (incorrect syllable) ที่ถูกวิเคราะห์ด้วย Apple Siri™	37
ตารางที่ 5 แสดงค่าเฉลี่ยของคะแนนความผิดพลาดของพยางค์ (mean error score of syllable)	38
ตารางที่ 6 แสดงร้อยละของค่าเฉลี่ยคะแนนความผิดพลาดของพยางค์ของทั้ง 4 ประโยคที่ถูก วิเคราะห์และสร้างขึ้นด้วยเครื่องมือ Whisper-Thai-Demo และ Apple Siri™	42
ตารางที่ 7 แสดงความสามารถในการวินิจฉัยอาการพูดไม่ชัดชนิดสປาสติกโดยประสาทแพทย์ทั้งสาม กลุ่ม รวมถึงค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน	43
ตารางที่ 8 แสดง 2x2 contingency table ของความสามารถในการวินิจฉัยอาการพูดไม่ชัดชนิดสປาสติกโดยประสาทแพทย์แต่ละคนจำนวน 10 คน	59

สารบัญรูปภาพ

	หน้า
รูปภาพที่ 1 แสดงกรอบความคิดแนววิจัย.....	3
รูปภาพที่ 2 แสดงตัวอย่างลักษณะเสียงผิดปกติที่พบในผู้ป่วยพูดไม่ชัดทางระบบประสาท.....	9
รูปภาพที่ 3 แสดงตัวอย่างกลุ่มลักษณะเสียงผิดปกติ (speech deviation).....	10
รูปภาพที่ 4 แสดงกล้ำเนื้อการออกเสียงที่สำคัญได้แก่.....	12
รูปภาพที่ 5 แสดงการเปรียบเทียบระหว่างกลุ่มความถี่พื้นฐาน (fundamental formant frequency F0).....	16
รูปภาพที่ 6 ตัวอย่างคลื่นเสียงแบบ Spectrogram	17
รูปภาพที่ 7 แสดงการจัดวางอุปกรณ์บันทึกเสียงในห้องบันทึกเสียง.....	29
รูปภาพที่ 8 แสดงการใช้งานจริงของ Apple Siri™ speech recognition.....	30
รูปภาพที่ 9 แสดงหน้าเว็บไซต์ Hugging face ในส่วน Whisper-Thai-Demo	31
รูปภาพที่ 10 แสดงกราฟคลื่นเสียงประโยคพูด ‘ชาวไร่ตัดต้นสนทำท่อนซุง’	36
รูปภาพที่ 11 แสดงค่าเฉลี่ยของคะแนนความผิดพลาดของพยางค์ (mean error score of syllable) ของทั้ง 4 ประโยคที่ถูกระบุและสร้างขึ้นด้วยเครื่องมือ Apple Siri™	39
รูปภาพที่ 12 แสดงค่าเฉลี่ยของคะแนนความผิดพลาดของพยางค์ (mean error score of syllable) ของทั้ง 4 ประโยคที่ถูกระบุและสร้างขึ้นด้วยเครื่องมือ Whisper-Thai-Demo	40
รูปภาพที่ 13 แสดงความสัมพันธ์ของคะแนนความผิดพลาดของพยางค์ในแต่ละประโยคของผู้ป่วยคนเดียวระหว่างเครื่องมือ Whisper-Thai-Demo และ Apple Siri™.....	41
รูปภาพที่ 14 แสดง Receiver operating characteristic graph (ROC graph) พร้อมทั้ง AUC ของคะแนนในแต่ละประโยคและรวมทุกประโยคจากการวิเคราะห์ด้วย Apple Siri™	44
รูปภาพที่ 15 แสดง Receiver operating characteristic graph (ROC graph) พร้อมทั้ง AUC ของคะแนนในแต่ละประโยคและรวมทุกประโยคจากการวิเคราะห์ด้วย Whisper-Thai-Demo.....	45
รูปภาพที่ 16 แสดงผลการวิเคราะห์ sensitivity analysis หลังคัดตัวอย่างที่มีอาการทางหูคองมูกออก (local voice disturbances).....	58

รูปภาพที่ 17 แสดง threshold (optimal cut-off point) หรือจุดตัดที่เหมาะสมในการวิเคราะห์
 กราฟ ROC จากการวิเคราะห์ด้วย Apple Siri™ 61

รูปภาพที่ 18 แสดง threshold (optimal cut-off point) หรือจุดตัดที่เหมาะสมในการวิเคราะห์
 กราฟ ROC จากการวิเคราะห์ด้วย Whisper-Thai-Demo..... 62



บทที่ 1

บทนำ

1.1 ความสำคัญและที่มาของปัญหาการวิจัย

การตรวจวินิจฉัยร่างกายทางระบบประสาทอย่างแม่นยำสามารถนำไปสู่การวินิจฉัยโรคได้อย่างถูกต้องและได้รับการรักษาอย่างทันท่วงที แต่การตรวจร่างกายทางระบบประสาทนั้นกลับต้องใช้ทักษะของผู้ตรวจค่อนข้างมากทำให้ผลการตรวจมีความหลากหลายโดยเฉพาะการวินิจฉัยอาการพูดไม่ชัดซึ่งต้องอาศัยประสบการณ์การแยกแยะเสียงที่มากเพียงพอทำให้ความถูกต้องและแม่นยำในการวินิจฉัยค่อนข้างต่ำแม้จะเป็นผลการวินิจฉัยจากประสาทแพทย์ (Fonville, van der Worp et al. 2008) โดยที่ประสาทแพทย์มักอาศัยอาการและอาการแสดงจากการตรวจร่างกายในส่วนอื่นๆที่ประเมินได้ง่ายกว่า เช่น การอ่อนแรงของกล้ามเนื้อแขนขา แล้วจึงนำผลการตรวจร่างกายนั้นมาอนุมานความผิดปกติหรือชนิดที่ผิดปกติของเสียงพูด ทำให้ไม่สามารถวินิจฉัยอาการพูดไม่ชัดหรือโรคทางระบบประสาทได้อย่างมั่นใจหากผู้ป่วยมาพบแพทย์ด้วยอาการพูดไม่ชัดเพียงอย่างเดียว นอกจากนี้ในปัจจุบันยังไม่มีเครื่องมือวินิจฉัยใดๆที่นำมาใช้ในทางคลินิกเพื่อเพิ่มประสิทธิภาพในการวินิจฉัยอาการพูดไม่ชัด ทำให้อาการพูดไม่ชัดยังเป็นอุปสรรคในการวินิจฉัยโรคทางระบบประสาทเป็นอย่างมาก

อย่างไรก็ตามในอดีตมีการศึกษามากมายอธิบายรายละเอียดของลักษณะเสียงพูดไม่ชัดชนิดต่างๆไว้ โดยเฉพาะชนิดส pastic (spastic dysarthria) อาทิ ความดิ่งที่ลดลง พยัญชนะที่ผิดเพี้ยน เป็นต้น เนื่องจากอาการพูดไม่ชัดชนิดส pastic เป็นชนิดที่พบได้บ่อยที่สุดและส่งผลกระทบต่อการพูดคุยสื่อสารและคุณภาพชีวิตของผู้ป่วยทางระบบประสาทเป็นจำนวนมากรวมถึงผู้ป่วยในกลุ่มหลอดเลือดสมองตีบที่เป็นกลุ่มที่พบบ่อยมากในผู้ป่วยกลุ่มนี้ นอกจากองค์ความรู้เรื่องเสียงพูดไม่ชัดแล้วยังมีการค้นพบลักษณะทางกายวิภาคและสรีรวิทยาที่ผิดปกติของการออกเสียงพูดไม่ชัดชนิดดังกล่าวมากมาย แต่ยังไม่เคยมีผู้นำมาประยุกต์ใช้จริงในการวินิจฉัยคลินิก การศึกษานี้จึงเป็นการนำองค์ความรู้ดังกล่าวมาประยุกต์ใช้ด้วยวิธีที่เหมาะสมเพื่อการนำไปใช้วินิจฉัยจริงทางคลินิก

นอกจากนี้ในปัจจุบันยังมีการพัฒนาเครื่องมือวิเคราะห์คลื่นเสียงด้วยแบบจำลองทางคอมพิวเตอร์อย่างหลากหลายซึ่งมีความแม่นยำ น่าเชื่อถือและเข้าถึงได้ง่าย จึงสามารถที่จะนำประโยชน์จากเครื่องมือเหล่านี้มาใช้ในทางการแพทย์เพื่อการวินิจฉัยเสียงพูดไม่ชัด การศึกษานี้จึงเป็นการนำความสามารถของเครื่องมือนี้มาประยุกต์ใช้เพื่อพิสูจน์ว่าการนำองค์ความรู้ทางคลินิกมาเพิ่ม

ประสิทธิภาพการวินิจฉัยเสียงพูดไม่ชัดชนิดสປาสติกจะเพิ่มความถูกต้องแม่นยำในการวินิจฉัยได้เป็น
 อย่งดี

1.2 ค้ถามของการวิจัย

ค้ถามหลัก : การสร้างประโยชน์สำหรับผู้ป่วยทางระบบประสาทที่มีอาการพูดไม่ชัดชนิดสປาสติกจากองค์ความรู้ของการออกเสียงพูดไม่ชัด และนำมวิเคราะห์ด้วยเครื่องมือวิเคราะห์คลื่นเสียงด้วยแบบจำลองทางคอมพิวเตอร์ จะสามารถแยกแยะเสียงพูดไม่ชัดชนิดสປาสติกออกจากเสียงพูดปกติได้หรือไม่

Could computational tools differentiate spastic dysarthria from normal speech based on anatomical and physiological background knowledge?

ค้ถามรอง : การวินิจฉัยเสียงพูดไม่ชัดชนิดสປาสติกออกจากเสียงพูดปกติด้วยเครื่องมือวิเคราะห์เสียงทางคอมพิวเตอร์ มีความแม่นยำในการวินิจฉัยแตกต่างจากการวิเคราะห์โดยประสาทแพทย์หรือไม่

Is the diagnostic accuracy of identifying spastic dysarthria using computational tools different from that obtained by neurologists?

1.3 วัตถุประสงค์งานวิจัย

เพื่อหาความแม่นยำของการวินิจฉัยอาการพูดไม่ชัดชนิดสປาสติกจากประโยชน์ที่สร้างขึ้นจากองค์ความรู้ของการออกเสียงพูดไม่ชัดและวิเคราะห์ด้วยเครื่องมือวิเคราะห์คลื่นเสียงด้วยแบบจำลองทางคอมพิวเตอร์ และเปรียบเทียบกับความแม่นยำในการวินิจฉัยอาการพูดไม่ชัดชนิดสປาสติกที่ได้จากการวิเคราะห์โดยประสาทแพทย์

1.4 สมมติฐาน

การวิเคราะห์ประโยชน์ที่สร้างขึ้นจากองค์ความรู้ของการออกเสียงพูดไม่ชัดด้วยเครื่องมือวิเคราะห์คลื่นเสียงด้วยแบบจำลองทางคอมพิวเตอร์สามารถแยกแยะเสียงพูดไม่ชัดชนิดสປาสติกออก

จากเสียงพูดปกติได้อย่างแม่นยำ และมีความแม่นยำในการวินิจฉัยที่ดีกว่าการวิเคราะห์โดยประสาทแพทย์

1.5 ข้อตกลงเบื้องต้น

ผู้ป่วยที่เข้าร่วมการศึกษาต้องมีอาการพูดไม่ชัดชนิดสປาสติกที่เกิดจากความผิดปกติทางระบบประสาทเท่านั้น และต้องไม่มีอาการพูดไม่ชัดที่เกิดจากสาเหตุอื่นๆ ร่วมด้วย รวมถึงต้องเป็นผู้ป่วยที่มีความรู้สึกตัวปกติสามารถให้ความร่วมมือในการอ่านออกเสียงได้โดยสมบูรณ์

ผู้ที่เข้าร่วมการศึกษาทั้งสองกลุ่ม (กลุ่มผู้ป่วยพูดไม่ชัดและกลุ่มควบคุม) ต้องไม่มีความแตกต่างทางด้านการออกเสียงอื่นๆ อาทิ สำเนียงการออกเสียง เพศหรืออายุที่มีผลต่อการออกเสียง เป็นต้น

1.6 กรอบความคิดแนววิจัย

รูปภาพที่ 1 แสดงกรอบความคิดแนววิจัย



1.7 การให้คำนิยามเชิงปฏิบัติที่จะใช้ในการวิจัย

- **Dysarthria** (อาการพูดไม่ชัด) คือกลุ่มอาการออกเสียงพูดที่ผิดปกติที่เกิดจากระบบประสาทที่ทำให้การเคลื่อนไหวของกล้ามเนื้อการออกเสียงผิดปกติไป อาทิ ความแข็งแรง ความเร็ว ความแม่นยำ เสียงสูงต่ำหรือโทนการออกเสียง เป็นต้น ทำให้องค์ประกอบของการออกเสียง

ทั้ง 5 อย่าง ได้แก่ การหายใจออก(breathing) การเปล่งเสียง(phonatory) ความกังวาน (resonance) การออกเสียงสระและพยัญชนะ(articulatory) และโทนน้ำเสียง(prosody) นั้นมีความผิดปกติไป โดยที่อาการไม่ได้เกิดจากศูนย์ควบคุมกล้ามเนื้อการพูดหรือการวางแผนการพูดจากสมองส่วน cerebral cortex และในปัจจุบันยังไม่มีวิธีวินิจฉัยที่ดีที่สุด (gold standard) ที่เป็นที่ยอมรับร่วมกันในทุกวงการ โดยส่วนมากจะเป็นการวินิจฉัยด้วยการฟังและวิเคราะห์ (auditory-perceptual method) โดยนักอรรถบำบัด (speech-language pathologist)

- Spastic dysarthria (อาการพูดไม่ชัดชนิดส pastic) คือชนิดของอาการพูดไม่ชัดที่เกิดจากรอยโรคทางระบบประสาทบริเวณ corticobulbar tract ซึ่งมักจะเกิดรอยโรคทั้งสองข้าง (bilateral corticobulbar lesions) อย่างไรก็ตามการเกิดรอยโรคข้างเดียว (unilateral) ในระยะแรกก็ทำให้เกิดอาการได้เช่นกัน โดยการเกิด corticobulbar lesion นี้จะทำให้กล้ามเนื้อการออกเสียงนั้นอ่อนแรงและแข็งเกร็ง (weakness and spasticity) ส่งผลให้การออกเสียงมีลักษณะเกร็ง (spastic) ที่เป็นลักษณะเด่นของการออกเสียงพูดไม่ชัดชนิดนี้
- Speech intelligibility คือ ดัชนีวัดความชัดถ้อยชัดคำของข้อความหรือประโยคพูด ซึ่งจะแปรผันตามความรุนแรงของการพูดไม่ชัดและสามารถจำแนกความรุนแรงดังกล่าวได้
- Acute ischemic stroke คือ ภาวะสมองขาดเลือดเฉียบพลันที่เกิดจากการอุดตันของหลอดเลือดที่เลี้ยงสมองทำให้เนื้อสมองบริเวณที่ถูกเลี้ยงด้วยหลอดเลือดนั้นหยุดทำงานหรือเซลล์สมองตายไป ทำให้เกิดอาการต่างๆตามหน้าที่ของสมองส่วนที่เสียไป รวมถึงอาการพูดไม่ชัดชนิดส pasticหากสมองส่วนที่ถูกกระทบประกอบด้วย corticobulbar tract
- TOAST classification คือ การจำแนกภาวะสมองขาดเลือดเฉียบพลันตามสาเหตุการเกิดโรค โดยแบ่งเป็น 5 ชนิดได้แก่ 1. Large vessel atherosclerosis หรือการตีบตันของหลอดเลือดสมองขนาดใหญ่ 2. Small vessel disease หรือการตีบตันของหลอดเลือดสมองขนาดเล็ก 3. Cardioembolic หรือการอุดตันของหลอดเลือดสมองที่เกิดจากลิ่มเลือดที่มาจากหัวใจ 4. Other determined etiologies หรือสาเหตุอื่นๆที่พบได้ไม่บ่อย และ 5. Undetermined etiology ยังไม่พบสาเหตุ
- Lesion sites หรือตำแหน่งของรอยโรค อาทิ เช่น
 - Cortical area หรือผิวสมอง
 - Subcortical area หรือเนื้อสมองใต้ผิวสมอง
 - Internal capsule หรือเนื้อสมองส่วน corticobulbar tract ที่ต่อเนื่องจาก subcortical area

- Cerebral peduncle หรือเนื้อสมองส่วน corticobulbar tract ที่ต่อเนื่องจาก Internal capsule และอยู่ในก้านสมองส่วน Midbrain
- Basis pontis หรือเนื้อสมองส่วน corticobulbar tract ที่ต่อเนื่องจาก cerebral peduncle และอยู่ในก้านสมองส่วน Pons
- Primary lateral sclerosis คือ โรคทางระบบประสาทส่วนกลางที่เกิดจากความเสื่อมถอยแบบค่อยเป็นค่อยไปของเซลล์ประสาทควบคุมการเคลื่อนไหวตั้งแต่ระดับสมองจนถึงไขสันหลัง โดยจะพบการตายของเซลล์ประสาทชนิด upper motor neuron ที่อยู่ใน corticobulbar และ corticospinal tract ทำให้เกิดกล้ามเนื้ออ่อนแรงและแข็งเกร็ง (spasticity) ทั่วร่างกายรวมถึงอาการพูดไม่ชัดชนิดสປาสติกด้วย
- Computer tomography of the brain (CT brain) คือ ภาพถ่ายเอกซเรย์ทางคอมพิวเตอร์บริเวณสมอง
- Magnetic resonance imaging of the brain (MRI brain) คือ ภาพถ่ายทางคลื่นแม่เหล็กไฟฟ้าบริเวณสมอง
- Audiometric test room คือ ห้องตรวจการได้ยินมาตรฐานที่ได้รับการออกแบบให้ปราศจากเสียงรบกวนจากภายนอกโดยสิ้นเชิง ถูกพัฒนาสำหรับการทดสอบการได้ยิน (audiometry) โดยให้ผู้ถูกทดสอบนั่งอยู่ในห้องระหว่างการเปิดเสียงทดสอบ
- Spectrogram คือ กราฟคลื่นเสียงที่มีแกนนอน (X-axis) คือเวลา แกนตั้ง (Y-axis) คือความถี่เสียง และแกนลึก (Z-axis) คือ ความดังของเสียง(amplitude)
- Automatic speech-to-text recognition (ASR) คือ เครื่องมือที่อาศัยเทคโนโลยีทางคอมพิวเตอร์ที่สามารถเปลี่ยนเสียงภาษาพูดของมนุษย์ให้เป็นข้อความและอาจมีความสามารถในการเข้าใจความหมายของประโยค อีกทั้งตอบกลับได้อย่างมีความหมาย โดยเทคโนโลยีดังกล่าวอาศัยกระบวนการเรียนรู้ทางโมเดลคอมพิวเตอร์ (machine learning) ร่วมกับพื้นฐานทางภาษาศาสตร์
- Error of syllable คือ ความไม่ถูกต้องของพยางค์ที่ถูกสร้างขึ้นจากการวิเคราะห์ประโยคพูดด้วยเครื่องมือ Automatic speech-to-text recognition ในองค์ประกอบต่าง ๆ ได้แก่ พยัญชนะ สระและตัวสะกด

1.8 ผลหรือประโยชน์ที่คาดว่าจะได้รับจากงานวิจัย

ได้ค้นหาองค์ประกอบของเสียงพูดไม่ชัดชนิดสປาสติกที่มีความสำคัญต่อการวินิจฉัยอาการพูดไม่ชัดเพื่อนำมาใช้จริงในการวินิจฉัยทางคลินิก และได้ศึกษาความแม่นยำของเครื่องมือการ

วิเคราะห์คลื่นเสียงพูด (Automatic speech-to-text recognition) ในการวินิจฉัยอาการพูดไม่ชัด เพื่อนำเครื่องมือดังกล่าวมาเพิ่มประสิทธิภาพในการวินิจฉัยของแพทย์ได้ในอนาคต

1.9 อุปสรรคที่อาจเกิดขึ้นระหว่างการวิจัยและมาตรฐานการแก้ไข

การดำเนินการวิจัยในช่วงการแพร่ระบาดของโคโรนาไวรัสสายพันธุ์ใหม่อาจทำให้เกิดข้อจำกัดของการเก็บตัวอย่างข้อมูลด้วยความเป็นไปได้หลายประการ อาทิ จำนวนตัวอย่างน้อยกว่าที่คาดการณ์เนื่องจากการไม่ยินยอมเข้าร่วมด้วยความกังวลเรื่องการติดเชื้อไวรัส (การบันทึกเสียงจำเป็นต้องถอดหน้ากากอนามัย) การติดตามนัดผู้ป่วยทางไกลทำให้ไม่สามารถเข้าถึงกลุ่มผู้ป่วยได้ การไม่สามารถควบคุมปัจจัยบางประการ เช่น ระยะเวลาของอาการพูดไม่ชัด ด้วยข้อจำกัดของการคัดเลือกผู้ป่วยเข้าร่วมวิจัย ขนาดของข้อมูลที่จำกัด เช่น จำนวนคำและประโยคที่ต้องพูดต้องกระชับและรวดเร็วเพื่อป้องกันการแพร่ระบาดซึ่งอาจมีผลต่อความละเอียดในการวิเคราะห์ข้อมูล หรือสถานที่บันทึกข้อมูลที่ไม่สามารถควบคุมให้เป็นที่ใดที่หนึ่งได้จากข้อจำกัดในการเคลื่อนย้ายผู้ป่วย เป็นต้น ซึ่งคิดว่าสามารถแก้ไขได้ด้วยการวางระเบียบการวิจัยให้รัดกุม อาทิเช่น การค้นหาวิธีแยกแยะเสียงพูดไม่ชัดออกจากเสียงปกติจากองค์ความรู้ทางประสาทวิทยาที่เชื่อว่าจะทำให้มีความแตกต่างของมาตรวัด (effect size) มากที่สุดเพื่อให้ไม่ต้องใช้จำนวนตัวอย่างที่มากเกินไป หรือการจัดระบบการบันทึกเสียงของตัวอย่างให้เหมือนกันมากที่สุดแม้จะบันทึกในสถานที่ต่างกัน เป็นต้น

ความเสี่ยงต่อผู้ป่วยและกลุ่มควบคุมในการศึกษานี้ ได้แก่ การเสียเวลาและรายได้ อุบัติเหตุระหว่างการเดินทาง เป็นต้น ซึ่งเป็นความเสี่ยงเล็กน้อยที่ผู้เข้าร่วมจะได้รับการชี้แจงโดยละเอียดในเอกสารวิจัยก่อนเข้าร่วมการวิจัย และจะมีการชดเชยตามระเบียบมาตรฐานวิจัย

บทที่ 2

ทบทวนวรรณกรรมที่เกี่ยวข้อง

2.1 สัทศาสตร์ของเสียงพูดไม่ชัด (Phonetics of dysarthria)

อาการพูดไม่ชัดเกิดจากโรคทางระบบประสาท (dysarthria) เป็นอาการที่พบได้บ่อยและ มักถูกวินิจฉัยผิดแม้วินิจฉัยโดยประสาทแพทย์ ตำแหน่งของโรคที่ทำให้เกิดอาการอาจเป็นได้ตั้งแต่ ระบบประสาทส่วนกลาง (central nervous system) ส่วนที่ควบคุมการเคลื่อนไหว ได้แก่ pyramidal system, extrapyramidal system, cerebellum เป็นต้น ไปจนถึงระบบประสาท ส่วนปลาย (peripheral nervous system) ได้แก่ เส้นประสาทสมอง กล้ามเนื้อที่เกี่ยวข้องกับการออกเสียง เช่น กล้ามเนื้อคอหอย (pharyngeal muscles) หรือกล้ามเนื้อลิ้น (tongue muscles) เป็นต้น ทำให้เกิดเสียงพูดไม่ชัดที่มีความหลากหลาย แต่ในขณะเดียวกันก็มีลักษณะเสียงพูดไม่ชัด ที่คล้ายกันหาก รอยโรคในสมองอยู่บริเวณเดียวกัน ดังนั้นการจำแนกเสียงพูดไม่ชัดจากลักษณะเสียง ดังกล่าวอย่าง ถูกต้องจึงสามารถนำไปสู่การวินิจฉัยรอยโรคในสมองได้ (Duffy 2019)

กลไกการออกเสียงพูดโดยทั่วไปประกอบด้วย 5 ส่วนคือ การหายใจออก (breathing) การออกเสียง (phonation) การสร้างคำหรือพยางค์ (articulation) การสะท้อนของเสียงในลำคอ (resonance) และสัทสัมพันธ์ (prosody) ที่ประกอบด้วยองค์ประกอบของการพูดอื่นๆ ที่ทำให้เกิด สำเนียงภาษานั้น เช่น จังหวะ ความเร็ว หรือการเน้นเสียง เป็นต้น โดยกลไกแต่ละส่วนนั้น ประกอบด้วย คุณสมบัติของการทำงานของกล้ามเนื้อการออกเสียงทั้งหมด 6 ปัจจัย ซึ่งมีความ สำคัญ ในการประกอบเป็นประโยคพูด ได้แก่ ความแข็งแรงในการหดตัวของกล้ามเนื้อการพูด (strength) ที่สามารถเกิดความผิดปกติในทุกกลไกที่กล่าวไปข้างต้นตามกล้ามเนื้อที่อ่อนแรงความเร็ว ในการขยับ ของกล้ามเนื้อออกเสียง (speed) ความยาวในการออกเสียง (range) ซึ่งมีผลต่อสัทสัมพันธ์ มากที่สุด ความคงที่ของการออกเสียง (steadiness) ที่มีผลต่อความสั้นของการออกเสียง ความตึงตัวของ กล้ามเนื้อในการออกเสียง (tone) เช่น ความอ่อนตัว (flaccid tone) หรือ ความแข็งเกร็ง (spastic tone) และความแม่นยำในการสร้างเสียงของกล้ามเนื้อ (accuracy or coordination) ทำให้เสียงพูด ออกมาชัดเจนและเป็นจังหวะที่ถูกต้อง (Story 2015) เป็นต้น ความเข้าใจ ในความผิดปกติของ คุณสมบัติเหล่านี้จะสามารถนำไปสู่การวินิจฉัยอาการพูดไม่ชัดได้รวมถึงการ พิจารณาแยกแยะแต่ละ คุณสมบัติที่นั้นอาจนำไปสู่การวินิจฉัยชนิดของอาการพูดไม่ชัดอีกด้วย

ผลลัพธ์ของการเคลื่อนไหวของกล้ามเนื้อเนื้อด้วยคุณสมบัติและกลไกทั้งหมดข้างต้นทำให้เกิดเป็น ประโยคพูดขึ้น และอย่างที่เราทราบกันดีว่าประโยคนั้นประกอบด้วยหน่วยพยางค์ (syllable) ที่ รวมกันเป็นหน่วยคำ (word) ร่วมกับการใช้หลักของไวยากรณ์เพื่อสร้างเป็นประโยค (sentence) อย่างไรก็ตามหน่วยที่เล็กที่สุดในหน่วยพยางค์ซึ่งไม่สามารถแบ่งแยกย่อยได้อีกคือ หน่วยเสียง (phoneme) ซึ่งหน่วยเสียงนั้นประกอบไปด้วยหน่วยเสียงหลักที่เปล่งออกมาคือ ส่วนพยัญชนะซึ่ง รวมถึงตัวสะกด (consonant) ร่วมกับสระ (vowel) โดยพยัญชนะนั้นจะเกิดขึ้น จากตำแหน่งของ กล้ามเนื้อที่ออกเสียง (placement) และกลไกการขยับของกล้ามเนื้อ (manner) (ดังแสดงในตารางที่ 1) และหน่วยเสียงตัวประกอบอื่นคือการเน้นเสียง (stress) และทำนองเสียง (intonation) ดังนั้น อาการพูดไม่ชัดจากระบบประสาทจึงแสดงออกเป็นความผิดปกติของผลลัพธ์เหล่านี้ อาทิเช่น ความไม่ ชัดของพยัญชนะและสระ ความยาวในการออกเสียงสระและจังหวะของประโยคที่ผิดปกติ หรือ ทำนองเสียงที่ผิดปกติ เป็นต้น ซึ่งจะมีความแตกต่างกันในรอยโรคแต่ละตำแหน่งในระบบ ประสาท (Clements and Keyser 1983)

การศึกษาเสียงพูดไม่ชัดทางระบบประสาทโดยนักอรรถบำบัด Darley Aronson Brown และคณะ ในปี ค.ศ.1969 และ ค.ศ.1975 เป็นการศึกษาเสียงพูดไม่ชัดขนาดใหญ่และมีอิทธิพล ถึงปัจจุบัน มี วัตถุประสงค์เพื่อค้นหาลักษณะเสียงพูดผิดปกติที่เด่นชัดจากคุณสมบัติของเสียงพูด 6 ข้อที่กล่าวไป ข้างต้น และจัดกลุ่มลักษณะเสียงผิดปกติตามชนิดของอาการพูดไม่ชัดที่สอดคล้องไปกับ ตำแหน่งของ รอยโรคทางระบบประสาท ภายใต้สมมติฐานที่ว่าเสียงพูดผิดปกติจะมีความ เฉพาะเจาะจงกับ ตำแหน่งและกลไกการเกิดโรค โดยผู้วิจัยพบความผิดปกติของเสียงพูดทั้งหมด 38 ลักษณะและพบว่า โรคที่ตำแหน่งในระบบประสาทและกลไกการเกิดโรคเดียวกันจะมีลักษณะเสียงผิดปกติในกลุ่ม เดียวกันที่แตกต่างจากกลุ่มอื่นๆ (รูปภาพที่ 2 และ 3) เสียงพูดไม่ชัดทั้งหมดได้ถูกจัดกลุ่มได้เป็น ทั้งหมด 5 กลุ่มสอดคล้องกับพยาธิสภาพทางระบบประสาท และใช้ชื่อการจำแนกว่า the Mayo Clinic classification system (Darley, Aronson et al. 1968, Darley, Aronson et al. 1969) โดยมีรายละเอียดดังนี้

1. Spastic dysarthria (หรือ pseudobulbar palsy) เป็นความผิดปกติที่เกิดจาก upper motor neuron lesion
2. Flaccid dysarthria เป็นความผิดปกติที่เกิดจาก lower motor neuron lesion หรือ bulbar muscle pathology

3. Ataxic/cerebellar dysarthria เป็นความผิดปกติที่เกิดจาก cerebellar or cerebellar connection lesion
4. Hypokinetic (parkinsonian) and hyperkinetic dysarthria เป็นความผิดปกติที่เกิดจาก extrapyramidal lesion
5. Mixed types of dysarthria

Dorland's Illustrated Medical Dictionary (24th ed.). Philadelphia: Saunders (1965).
 GENESE, MARGARET C. L., *The Voice and Its Disorders* (2nd ed.). Philadelphia: Lippincott (1964).
 GROWER, F., Classification of dysarthrias. *Acta Psychiat. Scand.*, 32, 325-337 (1957).
 PEACOCK, W. G., The neurological evaluation of delayed speech. *J. Speech Hearing Dis.*, 14, 344-352 (1949).
 PEACOCK, W. G., The etiology and differential diagnosis of dysarthria. *J. Speech Hearing Dis.*, 15, 253-265 (1950).
 WICKHAM, I. S., *Clinical Neurology: With an Introduction to the History of Neurology* (9th ed.). Philadelphia: Saunders (1963).

APPENDIX
 DIMENSIONS USED IN THIS STUDY

No.	Abbreviation	Description
1	Pitch level	Pitch of voice sounds consistently too low or too high for individual's age and sex.
2	Pitch breaks	Pitch of voice shows sudden and uncontrolled variation (falsetto breaks).
3	Monopitch	Voice is characterized by a monopitch or monotone. Voice lacks normal pitch and inflectional changes. It tends to stay at one pitch level.
4	Voice tremor	Voice shows shakiness or tremulousness.
5	Monoloudness	Voice shows monotony of loudness. It lacks normal variations in loudness.
6	Excess loudness variation	Voice shows sudden, uncontrolled alterations in loudness, sometimes becoming too loud, sometimes too weak.
7	Loudness decay	There is progressive diminution or decay of loudness.
8	Alternating loudness	There are alternating changes in loudness.
9	Loudness (overall)	Voice is insufficiently or excessively loud.
10	Harsh voice	Voice is harsh, rough, and raspy.
11	Hoarse (wet) voice	Wet, "liquid sounding" hoarseness.
12	Breathy voice (continuous)	Continuously breathy, weak, and thin.
13	Breathy voice (transient)	Breathiness is transient, periodic, intermittent.
14	Strained-strangled voice	Voice (phonation) sounds strained or strangled (an apparently effortful squeezing of voice through glottis).

DARLEY ET AL.: Diagnostic Patterns of Dysarthria 267

No.	Abbreviation	Description
15	Voice stoppages	There are sudden stoppages of voiced air stream (as if some obstacle along vocal tract momentarily impedes flow of air).
16	Hypernasality	Voice sounds excessively nasal. Excessive amount of air is resonated by nasal cavities.
17	Hyponasality	Voice is denasal.
18	Nasal emission	There is nasal emission of air stream.
19	Forced inspiration-expiration	Speech is interrupted by sudden, forced inspiration and expiration sighs.
20	Audible inspiration	Audible, breathy inspiration.
21	Grunt at end of expiration	Grunt at end of expiration.
22	Rate	Rate of actual speech is abnormally slow or rapid.
23	Phrases short	Phrases are short (possibly due to fact that inspirations occur more often than normal). Speaker may sound as if he has run out of air. He may produce a gasp at the end of a phrase.
24	Increase of rate in segments	Rate increases progressively within given segments of connected speech.
25	Increase of rate overall	Rate increases progressively from beginning to end of sample.
26	Reduced stress	Speech shows reduction of proper stress or emphasis patterns.
27	Variable rate	Rate alternately changes from slow to fast.
28	Intervals prolonged	Prolongation of interword or intersyllable intervals.
29	Inappropriate silences	There are inappropriate silent intervals.
30	Short rushes of speech	There are short rushes of speech separated by pauses.
31	Excess and equal stress	Excess stress on usually unstressed parts of speech, e.g. (1) monosyllabic words and (2) unstressed syllables of polysyllabic words.
32	Imprecise consonants	Consonant sounds lack precision. They show slurring, inadequate sharpness, distortions, and lack of crispness. There is clumsiness in going from one consonant sound to another.
33	Phonemes prolonged	There are prolongations of phonemes.
34	Phonemes repeated	There are repetitions of phonemes.

268 *Journal of Speech and Hearing Research*

13 246-269 1969

รูปภาพที่ 2 แสดงตัวอย่างลักษณะเสียงผิดปกติที่พบในผู้ป่วยพูดไม่ชัดทางระบบประสาท และคำอธิบายของเสียงนั้น อาทิ monotone แสดงถึงความบกพร่องของทำนองเสียง และการเน้นเสียงใน ประโยค Strained-strangled voice แสดงถึงเสียงที่แสดงความแข็งเกร็ง ของกล้ามเนื้อ เป็นต้น (Darley et al., 1969)

Table 1. Relation of Speech Deviations to Certain Neurologic Disorders*

SPEECH DEVIATIONS	MUSCULAR MECHANISM	NEURAL MECHANISM	CERE-BELLAR ATAXIA	CHOREO-ATHETOSIS	DYSTONIA	PSEUDO-BULBAR PALSY	AMYOTROPHIC LATERAL SCLEROSIS	BULBAR PALSY	PARKINSONISM
A. Sudden artic. breakdown	Poor muscular control	Ataxia, involuntary movement, or severe weakness	++	+	+
Prolongation of sounds			++	+	+	...	++
Equalization of stress			++	+	+	+	+
B. Voice tremor	Muscular spasm	Spasticity or involuntary movements	+	+
Strained-strangled voice			...	+	+	++	+
Harsh voice			...	+	+	++	++
Low pitch			++	++
C. Short phrases	Slowness or weakness	Spasticity or flaccidity	...	+	+	+	++	+	...
Hypernasality			+	+	++	...
D. Nasal emission	Weakness	Flaccidity	++	++	...
Audible inhalation			+	++	++	...
E. Breathy voice	Weakness or small movements	Flaccidity or rigidity	+	++	+
Reduced stress			+	+	...	++
Decay of loudness			++
Inadequate loudness			++
F. Fast rate	Fast small movements	Rigidity	+
Short rushes of speech			++
Repetition of initial sounds			+
G. Consistent artic. errors	Multiple	Any, including compensation	+	+	++	++	++	+	++
Vowel distortions			++	++	++	+	++	+	...
Slow rate			+	+	+	++	++	+	...
Monopitch			...	+	+	++	+	+	++
Inappropriate silences			...	+	+	...	+	...	++
Variable rate			...	++	+

*Single plus (+) indicates frequent occurrence of deviation; double plus (++) indicates further prominence

รูปภาพที่ 3 แสดงตัวอย่างกลุ่มลักษณะเสียงผิดปกติ (speech deviation) จะมีความคล้ายกันในกลุ่มโรคทางระบบประสาทเดียวกัน เช่น Pseudobulbar palsy (ความหมายเดียวกับ spastic dysarthria) จะพบลักษณะเสียงที่เป็น Strained-strangled voice Harsh voice และ Low pitch voice (เครื่องหมาย + แทนลักษณะเสียงที่พบได้ในโรคนั้น)

การศึกษาพบว่าอาการพูดไม่ชัดชนิดส pastic นั้นพบได้บ่อยที่สุด (Duffy 2005) และพบได้ถึงร้อยละ 53 ในผู้ป่วยหลอดเลือดสมองตีบ (acute ischemic stroke) ทั้งหมด (Mann, Hankey et al. 1999) โดยความชุกที่พบจะแตกต่างกัน ในแต่ละตำแหน่งและกลไกการขาดเลือด ทั้งนี้ข้อมูลความชุกของโรคหลอดเลือดสมองตีบพบถึงร้อยละ 2 ของประชากรชาวอเมริกันทั้งหมดหรือประมาณ 7 ล้านคน (Virani, Alonso et al. 2020) กล่าวคือจำนวนผู้ป่วยที่มีความลำบากจากการ พูดไม่ชัดชนิดส pastic นั้นมีอยู่จำนวนมาก การวินิจฉัยอาการพูดไม่ชัดชนิดนี้ได้อย่างรวดเร็วเพื่อการแก้ไข ก่อน จะเกิดความพิการถาวรนั้นมีความสำคัญเป็นอย่างมาก

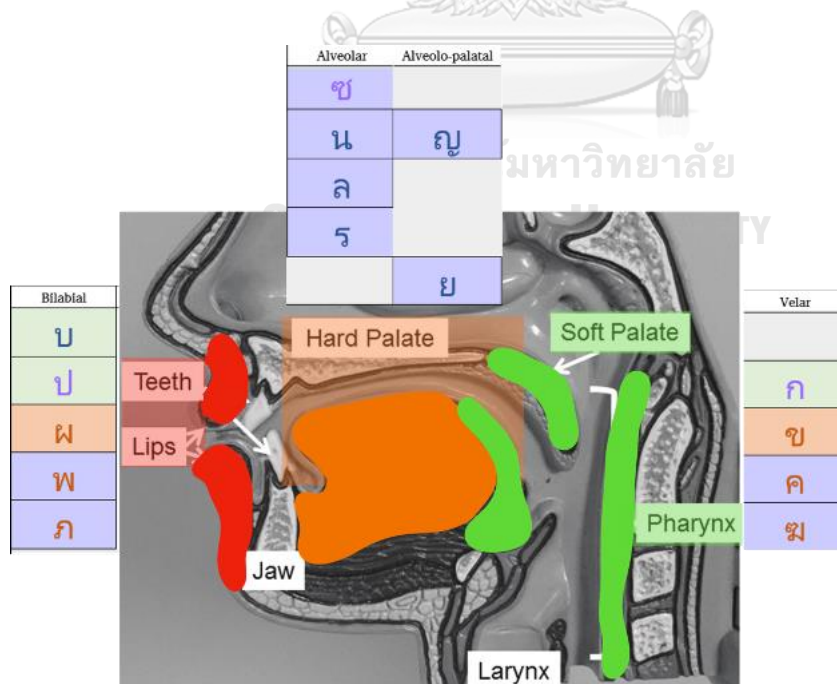
จากการศึกษาชนิดพรรณนาและการศึกษาเปรียบเทียบวิเคราะห์โดยนักอรรถบำบัดในผู้ป่วยหลอดเลือดสมองตีบที่ทำให้เกิดพูดไม่ชัดชนิดส pastic (วินิจฉัยจากรอยโรคใน MRI และการวินิจฉัยโรคหลอดเลือด สมองตีบโดยแพทย์) พบว่าลักษณะเสียงพูดไม่ชัดชนิดส pastic ประกอบด้วย ความผิดปกติที่หลายกลไกการ ออกเสียง ได้แก่ กลไกของการสร้างพยางค์ที่ไม่แม่นยำและไม่ถูกต้อง (imprecise and inaccuracy of articulation) ซึ่งเป็นความผิดปกติที่พบได้มากที่สุด โดยหลาย การศึกษาพบว่าพยาธิขณะเป็นส่วนที่ถูกกระทบมากที่สุดและรองลงมาคือความแม่นยำของสระ

นอกจากนี้กลไกสัทสัมพันธ์ของประโยค เช่น อัตราการพูดที่ช้าลง ความราบเรียบของเสียงพูด (voice monotone) ที่ประกอบด้วยความดังและความสูงต่ำที่คงที่ไม่มี การเน้นเสียง เป็นต้น และกลไกการออกเสียง (phonation) ทำให้เสียงมีความแข็งกระด้าง (harshness) และมีความแข็งเกร็ง (strained-strangled) (Mackenzie 2011) (Dai 1991) ซึ่งกลไกที่ผิดปกติทั้งหมดนี้สัมพันธ์กับความอ่อนแรงของกล้ามเนื้อการออกเสียงที่เกิดจากรอยโรคที่ corticobulbar tract ทำให้ไม่สามารถส่งสัญญาณประสาทจากสมองส่วนกลางมาที่กล้ามเนื้อได้

การออกเสียงพยัญชนะแรกหรือพยัญชนะต้น (initial consonant) โดยเฉพาะเสียงควบกล้ำ (consonant cluster) เป็นความผิดปกติของเสียงที่พบได้บ่อยและมากกว่าส่วนอื่นๆ เช่น สระหรือตัวสะกด ในหน่วยเสียงหนึ่งในผู้ป่วยหลอดเลือดสมองตีบที่มีอาการพูดไม่ชัดชนิดสเปคตริกอ้างอิง การศึกษาวิเคราะห์เชิงพรรณนาโดย Sriwimon et al. และคณะ (Manochiopinig, Thubthong et al. 2008) ซึ่งเป็นกลุ่มนักอรรถบำบัดชำนาญการ พบว่าพยัญชนะที่ ออกเสียงด้วยการหดกล้ามเนื้อปลายลิ้นกระทบกับส่วนเพดานปาก (Lingual-alveolar placement) อาทิ เสียง ‘ร’/r/ ‘ส’/s/ ‘ท’/th/ ‘ช’/ch/ ‘ต’/t/ ‘ด’/d/ นั้นมีความถูกต้องต่ำที่สุด พยัญชนะที่เกิดจาก กลไกการขยับแบบสั่นกล้ามเนื้อลิ้น (trill or flap manner) ที่ต้องใช้ความแม่นยำในการเคลื่อนไหว ของกล้ามเนื้อลิ้นส่วนตรงข้ามกัน (agonist-antagonist coordination) อาทิ เสียง ‘ร’/r/ มีความถูกต้องต่ำ ที่สุด และพยัญชนะที่ออกเสียงแบบเสียดแทรก (fricative) เช่น ‘ส’/s/ ‘ฟ’/f/ หรือกึ่งเสียดแทรก (affricative) เช่น ‘ช’/ch/ ที่เกิดจากออกแรงของกล้ามเนื้อลิ้นเข้าหาเพดานปากและปล่อยลมออก พร้อมกับคลายกล้ามเนื้ออย่างรวดเร็ว นั้นจะมีความถูกต้องน้อยกว่าพยัญชนะที่ออกเสียงเปิด (approximant) เช่น ‘ว’/w/ ที่เป็นการออกแรงของกล้ามเนื้อใบหน้ารอบริมฝีปากเพียงเล็กน้อยเพื่อ การสร้างเสียง (ดังแสดงในตารางที่ 1) สอดคล้องกับการผลการศึกษาศีรษะพูดไม่ชัดชนิดสเปคตริกในผู้ป่วย cerebral palsy โดย Chen and Stevens (2001) (Chen and Stevens 2001) ในขณะที่เสียงสระนั้นได้รับผลกระทบ น้อยที่สุดโดยมีความถูกต้องของการออกเสียงมากที่สุด ซึ่งผู้ป่วยที่ออกเสียงผิดปกติ นั้นจะมีความพยายาม ในการชดเชยพยัญชนะที่ผิดด้วยวิธีต่าง ๆ ได้แก่ การใช้พยัญชนะใกล้เคียงแทนการออกเสียง (substitution) ที่พบบ่อยที่สุด ตามด้วยการละเว้นไม่ออกเสียง (omission) เป็นต้น (รูปภาพที่ 4)

		ริมฝีปากทั้งสอง	ปุ่มเหงือก	เพดานแข็ง	เพดานอ่อน	เส้นเสียง
เสียงนาสิก		[m] ม	[n] ณ, น		[ŋ] ง	
เสียงกัก	ก้อง	[b] บ	[d] ฎ, ต, ท**			
	ไม่ก้อง ไม่มีลม	[p] ป	[t] ฏ, ต	[tʃ] จ	[k] ก	[ʔ] อ
	ไม่ก้อง มีลม	[pʰ] ผ, พ, ภ	[tʰ] ฐ, ท**, ตม, ถ, ท, ถ	[tʃʰ] ฉ, ช, ฉ	[kʰ] ข, ข*, ค, ค*, ฆ	
เสียงเสียดแทรก		[f] ฝ, ฟ	[s] ซ, ศ, ษ, ส			[h] ห, ฮ
เสียงเปิด		[w] ว	[l] ล, ฬ	[j] ญ, ย		
เสียงรัวลิ้น			[r] ร			

ตารางที่ 1 แสดงการแบ่งพยัญชนะต้นตามตำแหน่งของกล้ามเนื้อที่ออกเสียง (placement) ประกอบด้วยริมฝีปาก (bi-labial) ลิ้นกระทบปุ่มเหงือก (Lingual-alveolar) ลิ้นกระทบเพดานแข็ง (Lingual-palatal) ลิ้นกระทบเพดานอ่อน (Velopharyngeal) และเส้นเสียง (Glottis) และกลไกการขยับของกล้ามเนื้อ (manner) ประกอบด้วย เสียงนาสิก เสียงกัก เสียงเสียดแทรก (Fricative) เสียงปิด และเสียงรัวลิ้น (Trill) (ดัดแปลงจาก (รณเกียรติ 2548))



รูปภาพที่ 4 แสดงกล้ามเนื้อการออกเสียงที่สำคัญได้แก่

- กล้ามเนื้อใบหน้าและรอบปาก เช่น *mentalis muscle* (ระบายด้วยสีแดง) ทำหน้าที่ออกเสียงพยัญชนะต้นชนิด *bi-labial placement* เช่น ‘บ’/b/ ‘พ’/p/ เป็นต้น
- กล้ามเนื้อลิ้นที่หัดตัวเคลื่อนไหวกระทบกับส่วนบนของช่องปากคือปุ่มเหงือกหรือเพดานแข็ง (ระบายด้วยสีส้ม) ทำหน้าที่ออกเสียงพยัญชนะต้นชนิด *lingual-alveolar* เช่น ‘ด’/d/ เป็นต้น *lingual-palatal* เช่น ‘จ’/ **tg**/ เป็นต้น
- กล้ามเนื้อส่วนโคนลิ้นด้านหลัง กล้ามเนื้อเพดานอ่อน (*velum muscle*) และกล้ามเนื้อคอหอย (*pharyngeal muscle*) (ระบายด้วยสีเขียว) ออกเสียงชนิด *velopharyngeal* เช่น ‘ก’/k/ เป็นต้น
- กล้ามเนื้อเส้นเสียง (*laryngeal muscle*) ออกเสียง ‘อ’ เช่น เปล่งเสียง อา (เสียงสระเท่านั้น)

การค้นพบจากหลายการศึกษาข้างต้นแสดงให้เห็นว่าผลลัพธ์ของการเกิดอาการพูดไม่ชัดชนิดสเปคตริกนั้นทำให้เกิดเสียงผิดปกติในกลไกต่างๆ ไม่เท่ากัน โดยมักพบความผิดปกติในกลไก การสร้างพยางค์ (*articulation*) และสัทสัมพันธ์ (*prosody*) มากกว่ากลไกอื่นๆ (Ansel and Kent 1992) และในส่วนของ การสร้างหน่วยเสียงนั้นความผิดปกติของการออกเสียงพยัญชนะแรกจะพบได้ชัดเจนที่สุด สอดคล้องไปกับพยาธิสภาพและตำแหน่งของการเกิดโรคที่ *corticobulbar tract* ซึ่งการอ่อนแรง และแข็งเกร็งของกล้ามเนื้อที่มีความรุนแรงของกล้ามเนื้อแต่ละส่วนไม่เท่ากันรวมถึงกล้ามเนื้อการกลืนและการออกเสียง (*Bulbar muscles*) ดังแสดงออกในรูปแบบของกล้ามเนื้อออกเสียงพยัญชนะ ต้นดัดการศึกษาข้างต้นที่พบว่ากล้ามเนื้อลิ้นที่อ่อนแรงมากที่สุดทำให้ไม่สามารถออกเสียงด้วยลิ้นกระทบกับส่วนเพดานปาก (*lingual-alveolar placement*) ได้อย่างถูกต้องเมื่อเทียบกับเสียงพยัญชนะ จากกล้ามเนื้อส่วนอื่น และวิธีการเคลื่อนไหวของกล้ามเนื้อแบบเสียดแทรก (*fricative manner*) นั้นใช้การเคลื่อนไหวอย่างละเอียดที่ต้องใช้แรงกล้ามเนื้อและจังหวะการหดตัวของกล้ามเนื้อที่พอดี (*coordination of movement*) ร่วมกับการปล่อยลมหายใจออกในเวลาที่เหมาะสม ซึ่งความผิดปกติชัดเจนที่พบในผู้ป่วยกลุ่มนี้เกิดจากรอยโรคใน *corticobulbar tract* ทำให้ขาดสัญญาณประสาทชุดการเคลื่อนไหวอย่างมีแบบแผนละเอียดจากศูนย์สั่งการเคลื่อนไหวจากสมองส่วนกลางส่งผลให้ไม่สามารถเคลื่อนไหวได้อย่างละเอียดแม่นยำ (*fine movement*) (Honda 2015, Hoole and Pouplier 2015) นอกจากนี้ก็อีกหลายการศึกษาเกี่ยวกับกลไกการขยับกล้ามเนื้อการพูดได้สรุปไปในทางเดียวกันคือ กล้ามเนื้ออ่อนแรงในส่วนต่างๆของผู้ป่วยกลุ่มนี้นั้นไม่เท่ากัน โดยที่การอ่อนแรงของกล้ามเนื้อ ลิ้นและใบหน้าพบมากที่สุด ตามด้วยอาการอ่อนแรงของกล้ามเนื้อส่วนคอหอย (Duffy 2019) กล่าวโดยสรุปหาก พิจารณาข้อมูลทั้งหมดจะสามารถอนุมานได้ว่าอาการพูดไม่ชัดชนิดสเปคตริกมีกลไกที่ทำให้การออกเสียงพยัญชนะต้นบกพร่องชัดเจนเฉพาะในบางตัวมากกว่าตัว

อื่นๆ หรือในทางสรีรวิทยาคือกล้ามเนื้อ (ลิ้น) อ่อนแรงและเกร็งมากกว่ากล้ามเนื้อส่วนอื่น ต่างกับชนิดอาการพูดไม่ชัดอื่นๆ ที่มีลักษณะ เสียงไม่ชัดเฉพาะตัว (Kim, Martin et al. 2010)

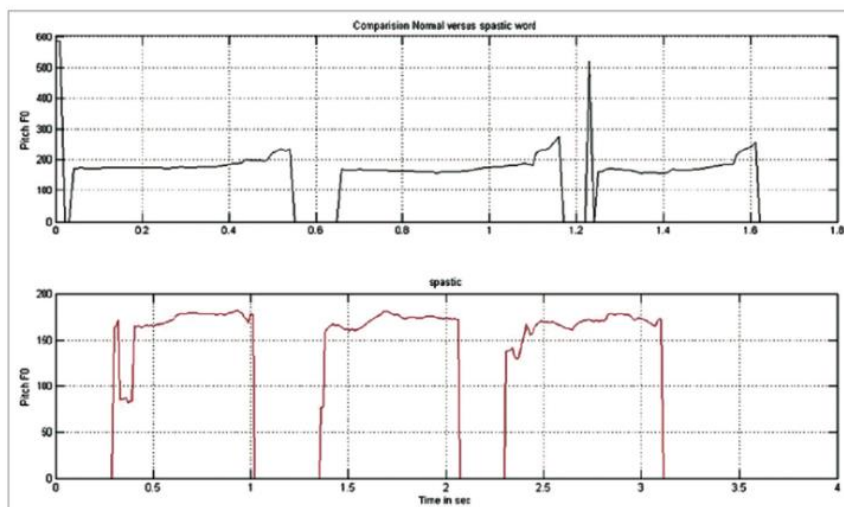
2.2 การวินิจฉัยอาการพูดไม่ชัดในปัจจุบัน

ปัจจุบันการวินิจฉัยอาการพูดไม่ชัดยังอาศัยการวิเคราะห์แยกแยะกลไกเสียงพูดไม่ชัดจากการฟัง (auditory-perceptual assessment) เป็นการตรวจอ้างอิงที่เป็นมาตรฐาน (Gold standard) ทำให้เกิดความหลากหลายและไม่แม่นยำของการวินิจฉัยโดยแพทย์เนื่องจากแพทย์ไม่ได้รับการฝึกฝนด้านการแยกแยะเสียงเป็นอย่างดีและความสามารถนี้มักอยู่ในบทบาทของนักอรรถบำบัดที่ไม่ได้พบผู้ป่วยในช่วงการวินิจฉัยอาการทำให้เกิดช่องว่างระหว่างความรู้ทางสรีรวิทยาจากแพทย์และความรู้ทางศาสตร์และสัทวิทยาจากนักอรรถบำบัด กล่าวคือจนถึงปัจจุบันยังไม่มีการบูรณาการความรู้ทั้งสองสาขามาเพื่อช่วยแพทย์วินิจฉัย แพทย์ยังคงใช้วิธีแยกแยะเสียงพูดได้ต่อระหว่างการตรวจด้วยความคุ้นเคยและไม่มีหลักการชัดเจน ดังนั้นการนำองค์ประกอบเสียงพูดไม่ชัดที่แตกต่างจากปกติอย่างชัดเจนมาใช้ในการวินิจฉัยอาการพูดไม่ชัด อาทิ การตรวจการพูดไม่ชัดโดยให้ผู้ป่วยพูดตามคำที่ขึ้นต้นด้วยพยัญชนะต้นชนิด lingual-alveolar placement เป็นต้น อาจเพิ่มความถูกต้องและแม่นยำในการแยกแยะเสียงพูดไม่ชัดชนิดสปาสติก ไม่ว่าจะเป็นการวินิจฉัยโดยแพทย์หรือโดยเครื่องมือวิเคราะห์เสียงต่างๆ ได้ (Duffy 2019)

ถึงแม้เกณฑ์การวินิจฉัยชนิดอาการพูดไม่ชัด ‘Mayo Clinic classification system’ ที่กล่าวข้างต้นจะเป็นที่ยอมรับในวงกว้างว่าเป็นระบบที่รัดกุมและครอบคลุมทุกแง่มุมเสียงมากที่สุด ถูกนำไปใช้ในทางการแพทย์และศึกษาต่อยอดความเข้าใจทางสรีรวิทยาจากกลุ่มเสียงที่ผิดปกตินี้ (Kent, Kent et al. 1998) อย่างไรก็ตามเสียงผิดปกติ 38 รูปแบบที่ถูกแจกแจงในระบบนี้สามารถนำไปใช้ได้จริงในผู้ที่เชี่ยวชาญ ด้านการแยกแยะเสียงเท่านั้นและไม่สามารถนำมาใช้ได้จริงในทางคลินิกโดยแพทย์ ความไม่ถูกต้อง ไม่แม่นยำและไม่น่าเชื่อถือในการวินิจฉัยของแพทย์นั้นเป็นที่ประจักษ์ในหลายการศึกษาในอดีต ยกตัวอย่างการศึกษาโดย Van der Graaff และคณะพบว่าเมื่อให้ประสาทแพทย์และแพทย์ประจำบ้านวินิจฉัยชนิดของอาการพูดไม่ชัดจากการวิเคราะห์ด้วยการฟังไม่ให้ผู้ฟังทราบถึงประวัติและการวินิจฉัยของผู้ป่วยมาก่อนนั้นมีความถูกต้องไม่เกินร้อยละ 50 ในทุกชนิดของอาการพูดไม่ชัดและมีความถูกต้องเพียงร้อยละ 20-30 ในเฉพาะอาการพูดไม่ชัดชนิดสปาสติก (Van der Graaff, Kuiper et al. 2009) และการศึกษาโดย Fonville และคณะพบว่าความเชื่อมั่นระหว่างผู้ประเมินโดยใช้สถิติแคปปา (inter-rater agreement) อยู่ในระดับต่ำ ($K < 0.2$) (Fonville, van der

Worp et al. 2008) ข้อมูลแสดงให้เห็นว่าประสิทธิภาพการวินิจฉัยจากการวิเคราะห์ด้วยการฟังโดยอาศัย Mayo classification and rating system นั้นต่ำเกินไปกว่าจะนำมาใช้ในทางปฏิบัติได้

อย่างไรก็ตามในช่วงทศวรรษที่ 2000 มีการศึกษาสรีรวิทยาในแง่ของวิทยาศาสตร์ฟิสิกส์คลื่นเสียงพูดไม่ชัดทำให้เกิดการวิเคราะห์ในรูปแบบของคลื่นเสียงมากขึ้น (acoustic signal analysis) โดยการสร้างกราฟคลื่นเสียงและคำนวณได้เป็นตัวแปรทางเสียงที่มีความตรงไปตรงมามากขึ้นและเปรียบเทียบได้ในทางความถี่ ขนาดคลื่นเสียง ช่วงเวลาของคลื่นเสียง (objective assessment) เช่น ตัวแปรค่าความถี่พื้นฐานในเสียงพูด (fundamental frequency : F0) ความไม่คงที่ของความถี่เสียง (F0 jitter) เป็นต้น ทำให้เข้าใจธรรมชาติของคลื่นเสียงในกลไกการพูดไม่ชัดมากขึ้น บอกความแตกต่างระหว่างเสียงพูดไม่ชัดและเสียงปกติได้ชัดเจน สามารถระบุความแตกต่างของแต่ละชนิดของอาการพูดไม่ชัดได้ (Lansford and Liss 2014) และยังมีแนวโน้มมากพอที่จะบอกความแตกต่างระหว่างโรค บอกความรุนแรงของอาการพูดไม่ชัด และแยกแยะประเภทของการพูดไม่ชัดได้ถูกต้องแม่นยำมากกว่าการวิเคราะห์แยกแยะกลไกเสียงพูดไม่ชัดจากการฟังโดยอาศัย Mayo classification and rating system (auditory-perceptual assessment) (Kim, Kent et al. 2011) ยกตัวอย่างจากการศึกษาหนึ่งว่าผู้ป่วยพูดไม่ชัดชนิด spastic dysarthria จะพบลักษณะของความไม่คงที่ของความถี่เสียง (F0 jitter) และ fundamental frequency (F0) แตกต่างกับลักษณะความถี่พื้นฐานในเสียงพูดปกติอย่างชัดเจน (ดังแสดงในรูปภาพที่ 5) (Thoppil, Kumar et al. 2017) ดังนั้นเครื่องมือการศึกษาทางคลื่นเสียงนี้จึงสามารถเพิ่มประสิทธิภาพในการช่วยวินิจฉัยอาการพูดไม่ชัดของแพทย์ได้แม่นยำมากขึ้น แต่ในทางปฏิบัตินั้นยังไม่ได้ถูกนำไปใช้จริงทางคลินิกเนื่องจากความซับซ้อนของตัวแปรและไม่สามารถนำมาประยุกต์เชื่อมโยงเข้ากับลักษณะเสียงทางคลินิกได้โดยตรงและใช้การวิเคราะห์ด้วยการฟังเป็นหลัก ทั้งนี้แม้จะไม่ได้ถูกใช้ในการวินิจฉัยแต่ก็ถูกนำไปใช้ในการวางแผนการรักษา ใช้ระหว่างการรักษา เช่น visual feedback therapy หรือใช้ติดตามการกายภาพบำบัดแล้วในเบื้องต้น (Ansel and Kent 1992)



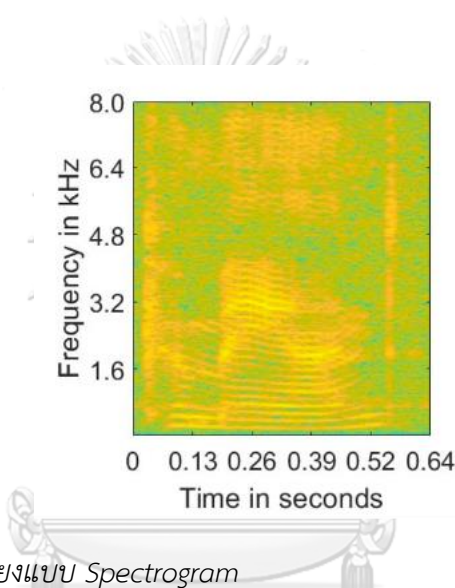
รูปภาพที่ 5 แสดงการเปรียบเทียบระหว่างกลุ่มความถี่พื้นฐาน (fundamental formant frequency F0)

ในเสียงพูดปกติ (บน) และเสียงพูดไม่ชัดชนิด spastic dysarthria (ล่าง) โดยพบลักษณะของความถี่ไม่คงที่ (F0 jitter) ในเสียงพูดไม่ชัด

2.3 เครื่องมือวิเคราะห์เสียงพูดในปัจจุบันและการนำมาใช้ในคลินิก

มีความพัฒนาอย่างก้าวกระโดดในช่วง 10 ปีที่ผ่านมาในการวิเคราะห์คลื่นเสียงในมิติที่ซับซ้อนขึ้น เช่น การวิเคราะห์จากกราฟ Spectrogram (ดังแสดงในรูปภาพที่ 6) ที่พิจารณามิติของความถี่ เวลา และน้ำหนักหรือ amplitude ของเสียงแต่ละย่านความถี่ที่แทนด้วยแถบสีต่างกันผ่านการคำนวณโดยใช้หลักการทางคณิตศาสตร์ ทำให้บอกรายละเอียดได้แม่นยำมากยิ่งขึ้นและสามารถนำไปวิเคราะห์ในรูปแบบอื่นๆ และใช้กระบวนการวิเคราะห์ทางคอมพิวเตอร์ที่มีความสามารถเรียนรู้ด้วยตัวเองและศึกษาข้อมูลขนาดใหญ่ได้ เช่น การวิเคราะห์ข้อมูลเสียงเป็นภาพแทนคลื่นเสียง (Convolutional neural network (CNN)) ด้วยเทคโนโลยีปัญญาประดิษฐ์ (artificial intelligence) หรือเทคนิคการเรียนรู้ข้อมูลเสียงโดยคอมพิวเตอร์อื่นๆ (machine learning) อย่างไรก็ตามการค้นพบส่วนมากยังจำกัดอยู่ในวงการวิทยาศาสตร์และวิศวกรรมศาสตร์มากกว่าการแพทย์ที่มีวัตถุประสงค์เพื่อการพัฒนาทางอุตสาหกรรมการสื่อสารโดยเฉพาะการตรวจจับเสียงพูด (automatic speech recognition) ในเครื่องมือสื่อสารต่างๆ ตัวอย่างที่เห็นได้ชัดคือความสามารถในการตรวจจับเสียงพูดและสร้างเป็นประโยคข้อความในโทรศัพท์มือถือ (automatic speech-to-text recognition) ได้ อย่างแม่นยำ ในขณะที่การนำมาวิจัยประยุกต์ทางทางการแพทย์นั้นพบได้น้อยและไม่ถูกนำมาใช้จริงจัง อาทิ การศึกษาโดย Chandrashekar et al. จากการวิเคราะห์ภาพ spectrogram ผ่าน CNN

เพื่อจำแนกความรุนแรงของอาการพูดไม่ชัดในพบว่ามีความแม่นยำมากกว่าร้อยละ 90 ในการบอก ระดับความรุนแรงและความเข้าใจของประโยค (intelligibility assessment of dysarthria) เปรียบเทียบกับการประเมินโดยผู้เชี่ยวชาญ (H, Karjigi et al. 2020) เป็นต้น ทั้งนี้งานวิจัยลักษณะ ข้างต้นมีวัตถุประสงค์เพื่อแก้ไขปัญหาในการใช้งานการตรวจจับเสียงพูดมากกว่าการนำมาใช้ทาง คลินิกและยังไม่มี ความพยายามนำมาประยุกต์ใช้ด้วยเหตุผลที่ว่าแพทย์ผู้วินิจฉัยไม่สามารถเข้าใจ ความหมายทางคลินิกของตัวแปรต่างๆและไม่สามารถเชื่อมโยงการค้นพบเหล่านี้กับการใช้ตรวจ ร่างกายได้



รูปภาพที่ 6 ตัวอย่างคลื่นเสียงแบบ Spectrogram

โดยแกนนอนแทนเวลา แกนตั้งแทนความถี่เสียง และแถบสีแทนแอมพลิจูดของคลื่นเสียง

CHULALONGKORN UNIVERSITY

จะเห็นได้ว่าถึงแม้ความเข้าใจเชิงลึกในเสียงพูดจะพัฒนาไปเป็นอย่างมากและสามารถนำมา ประยุกต์ใช้ประโยชน์ได้มหาศาลแต่ยังไม่สามารถนำมาใช้ในทางคลินิกได้ เนื่องจากแก่นในการ นำมาใช้ทางการแพทย์นั้นคือความเข้าใจเชิงสรีรวิทยาการทำงานของร่างกายสอดคล้องกับอาการ แสดงที่เกิดขึ้นมากกว่าตัวแปรทางคอมพิวเตอร์ที่อาจไม่มีนัยสำคัญทางคลินิก จึงมีความจำเป็นอย่าง ยิ่งที่จะต้องเกิดการวิจัยเพื่อผสมผสานความรู้ทางพยาธิสรีรวิทยาเข้ากับความรู้ด้านการแสดงทางเสียงพูดไม่ ชัดในเบื้องต้นเพื่อก่อให้เกิดความเข้าใจในอาการพูดไม่ชัดและใช้ประโยชน์จากการศึกษาเสียงพูดไม่ ชัดในอดีตเพื่อการวินิจฉัยโรคได้ดียิ่งขึ้น อย่างไรก็ตามการค้นพบเทคโนโลยีที่มีความสามารถในการ วิเคราะห์ข้อมูลได้อย่างถูกต้องแม่นยำจากโมเดลทางคอมพิวเตอร์ที่ประสิทธิภาพสูง เช่น เครื่องมือ Automatic speech-to-text recognition (ASR) เพื่อตรวจจับเสียงพูด เปลี่ยนเป็นคลื่นเสียงทาง

ดิจิทัล และสร้างเป็นประโยคพูดในภาษาเขียนได้อย่างความแม่นยำและได้ถูกนำไปใช้ทางการสื่อสารอย่างแพร่หลายนั้น ยังมีความเป็นไปได้ที่จะถูกนำมาใช้ทางการแพทย์ โดยแพทย์อาจใช้ประโยชน์จากความเที่ยงตรงและน่าเชื่อถือของเครื่องมือเพื่อที่จะลดความไม่เที่ยงตรง (subjective assessment) และลดปัญหาความเชื่อมั่นระหว่างแพทย์ผู้ประเมิน (low inter-rater reliability) ของการวินิจฉัยโดยแพทย์ที่มีทักษะในการแยกแยะคลื่นเสียงที่ไม่เท่ากัน หรืออีกนัยคือสร้างเครื่องมือประเมินเสียงพูดไม่ชัดจากการวิเคราะห์เสียงพูดด้วยอัลกอริทึมเดิมที่ใช้สร้างประโยคขึ้นมาจากเสียงพูดชัดปกติเพื่อตรวจหาหน่วยเสียงที่เครื่องมือคำนวณและสร้างขึ้นมาอย่างผิดพลาด โดยที่แพทย์ไม่จำเป็นต้องรู้กระบวนการวิเคราะห์ของเครื่องมือใดๆ อาทิ การสร้าง spectrogram หรือการใช้โมเดลปัญญาประดิษฐ์ (artificial intelligence) แต่นำองค์ความรู้ทางการพูด ไม่ชัดทั้งหมด ที่ได้กล่าวไปมาทดสอบกับผู้ป่วยแล้วประเมินผลด้วยเครื่องมือวิเคราะห์เสียงด้วยโมเดล Automatic speech recognition ที่แม่นยำก็สามารถทำให้เราเข้าใจและนำองค์ความรู้เหล่านั้นมากขึ้น เช่น พยาธิขณะต้นที่ออกเสียงไม่ชัดมากที่สุด และนำมาเพิ่มความสามารถในการวินิจฉัยได้เป็นอย่างดี

ในช่วง 10 ปีที่ผ่านมาโมเดลการตรวจจับเสียงพูดเปลี่ยนเป็นข้อความ (Automatic speech-to-text recognition (ASR)) มีความแม่นยำเพิ่มขึ้นอย่างมากจึงถูกนำมาใช้ในเครื่องมือสื่อสารในชีวิตประจำวันไม่ว่าจะเป็นโทรศัพท์มือถือ แท็บเล็ต และอุปกรณ์อิเล็กทรอนิกส์มากมายช่วยให้ผู้ใช้สามารถใช้ งานเครื่องมือต่างๆ ได้สะดวกขึ้นและสื่อสารผ่านแอปพลิเคชันข้อความโดยไม่ต้องใช้มือพิมพ์แป้น นอกจากนี้เครื่องมือยังสามารถเข้าใจภาษาเบื้องต้นและโต้ตอบกลับได้ผ่านเทคโนโลยี natural language processing ได้อีกด้วย การพัฒนาอย่างรวดเร็วนี้เกิดจากความสามารถในการคำนวณและวิเคราะห์ข้อมูลปริมาณมหาศาลได้อย่างรวดเร็วแม่นยำจากการเรียนรู้โดยคอมพิวเตอร์ผ่านเทคนิคการจำลองรูปแบบการประมวลผลของสมองมนุษย์โดยใช้โครงข่ายประสาทเทียมหลายชั้นหรือที่รู้จักกันในนาม Deep learning ที่เป็นเทคโนโลยีในกลุ่มปัญญาประดิษฐ์เช่นกัน โดยสองเครื่องมือที่เป็นที่นิยมอย่างมากได้แก่ Google assistant™ ที่มีผู้ใช้มากถึงร้อยละ 29 ของจำนวนการสื่อสารทางออนไลน์ (statistics) และ Apple's Siri™ ที่มีการทดสอบเพื่อยืนยันความแม่นยำในการเปลี่ยนเสียงพูดเป็นประโยคพิมพ์มากกว่า 90% (Team 2017) นอกจากนี้ที่มิวิจัยยังมีการพัฒนากระบวนการวิเคราะห์ของเครื่องมืออย่างต่อเนื่องเพื่อเพิ่มความแม่นยำและรวดเร็วในการโต้ตอบผู้ใช้งาน อาทิ พัฒนาการวิเคราะห์ข้อมูลเสียงให้จำเพาะกับรายบุคคล (Yang, Hu et al. 2023) เป็นต้น และมีการให้ข้อมูลกระบวนการวิเคราะห์ที่จะถูกนำมาใช้กับระบบ ASR ไร้ชัดเจน (Kalinli, Bhattacharya et al. 2019) แต่เนื่องจากมีผู้ใช้จำนวนมากที่เป็นผู้ป่วยมีอาการพูดไม่ชัดต้องพบกับ

ปัญหาในการแปลข้อความเสียงเป็นประโยคที่ไม่ถูกต้องเนื่องจากลักษณะเสียงที่เปลี่ยนไปทำให้ข้อมูลที่ป้อนเข้ามีความผิดพลาดส่งผลต่อการวิเคราะห์ของเครื่องมือ การศึกษาหนึ่งพบว่าประสิทธิภาพในการเปลี่ยนคลื่นเสียงเป็นข้อความโดย Apple's Siri™ ในโทรศัพท์มือถือ iPhone รุ่นต่างๆ ได้อย่างถูกต้องลดลงถึงร้อยละ 30-40 และมีความสม่ำเสมอของข้อความที่สร้างขึ้นค่อนข้างต่ำ (Ballati, Corno et al. 2018, Rohlfing, Buckley et al. 2021) อย่างไรก็ตามหากมองในมุมกลับ ผู้ใช้อาจนำข้อบกพร่องของการวินิจฉัยนี้ของเครื่องมือ ASR ทางการค้าเหล่านี้มาใช้ประโยชน์จากการตรวจจับและแยกแยะเสียงพูดไม่ชัดทางคลินิกออกจากเสียงพูดปกติเพื่อการวินิจฉัยโรคได้เป็นอย่างดีหากผู้ใช้หรือแพทย์ออกแบบกระบวนการวัดที่เหมาะสม เช่น การสร้างประโยคพูดที่ประกอบด้วยลักษณะเด่นของเสียงพูดไม่ชัดชนิดนั้นอย่างทีกล่าวไปข้างต้น หรือการสร้างมาตรวัดคะแนนที่เป็นตัวบ่งชี้การพูดไม่ชัดนั้นๆ ได้เป็นอย่างดี เป็นต้น

กล่าวโดยสรุป ประสิทธิภาพในการวินิจฉัยความผิดปกติของการออกเสียงพูดไม่ชัดทั้งความแม่นยำและความถูกต้องยังอยู่ในเกณฑ์ที่ไม่น่าพึงพอใจและไม่สามารถนำมาใช้ได้จริงในทางคลินิก เนื่องจากปัจจุบันยังอาศัยการวิเคราะห์แยกแยะเสียงพูดไม่ชัดจากการฟังและประเมินโดยแพทย์ (auditory-perceptual assessment) เป็นหลักซึ่งไม่ใช่ผู้เชี่ยวชาญในการแยกแยะเสียงที่ควรเป็นนักอรรถบำบัดที่ไม่ได้มีบทบาทในการวินิจฉัย ทั้งๆที่ผ่านมามีการค้นพบองค์ความรู้ในด้านสรีรวิทยา ระบบประสาทการออกเสียง ด้านสัทศาสตร์การออกเสียง ด้านฟิสิกส์ของคลื่นเสียงพูดไม่ชัดอย่างมากาย ดังนั้นการพัฒนาความสามารถในการวินิจฉัยของแพทย์ด้วยการบูรณาการความรู้นี้เข้าเป็นส่วนหนึ่งของกระบวนการวินิจฉัยจึงเป็นสิ่งที่จะต้องมีการริเริ่มปฏิบัติและไม่ใช้กระบวนการที่ซับซ้อนแต่อย่างใด ประกอบกับปัจจุบันมีการพัฒนาเครื่องมือทางด้านการฟัง แปลความ วิเคราะห์ แสดงผลหรือแม้กระทั่งเข้าใจเสียงพูดของมนุษย์ที่มีความแม่นยำอย่างยิ่งยวดนั้น จึงมีความเป็นไปได้แน่นอนที่จะนำเครื่องมือนี้มาเป็นมาตรวัดสำหรับการประยุกต์วิธีการวินิจฉัยการพูดไม่ชัดดังกล่าวให้มีประสิทธิภาพให้ดีมากพอที่จะนำมาใช้ได้จริงในคลินิก

บทที่ 3

วิธีการดำเนินการวิจัย

3.1 รูปแบบการวิจัย

การศึกษานี้เป็นชนิด Analytical, cross-sectional, diagnostic study

3.2 ระเบียบวิธีการวิจัย

ประชากรที่ศึกษา

กลุ่มผู้ป่วยที่มีอาการปวดไม่ชัดชนิดสไปนัลที่ได้รับการยืนยันการวินิจฉัยโรคด้วยประสาทแพทย์ ที่เข้ามารับการรักษาที่โรงพยาบาลจุฬาลงกรณ์จำนวน 37 ราย และกลุ่มควบคุมที่เสี่ยงปวดปกติ และไม่มีโรคทางระบบประสาท

เกณฑ์การคัดเลือกผู้ป่วยที่มีอาการปวดไม่ชัดชนิดสไปนัลที่เข้าร่วมการศึกษาวิจัย (inclusion criteria)

1. อายุเท่ากับหรือมากกว่า 18 ปี และน้อยกว่าหรือเท่ากับ 90 ปี
2. ผู้ป่วยได้รับการวินิจฉัยอาการปวดไม่ชัดชนิดสไปนัลซึ่งยืนยันการวินิจฉัยโดยต้องเข้าเกณฑ์ทุกข้อดังนี้ (Gold standard)
 - a. อาการปวดไม่ชัดจากผู้ป่วยและได้รับการยืนยันจากญาติใกล้ชิด
 - b. การวินิจฉัยรอยโรคทางระบบประสาทจากภาพถ่ายทางรังสีวิทยา ได้แก่ Computer tomography of brain (CT brain) และ Magnetic resonance imaging of brain (MRI brain) ที่ยืนยันรอยโรคตำแหน่ง Corticobulbar tract (Upper motor neuron)
 - c. การวินิจฉัยโรคทางระบบประสาทที่ทำให้เกิดอาการในข้อ a. และรอยโรคในข้อ b. ยืนยันโดยประสาทแพทย์ เช่น ภาวะสมองขาดเลือดอย่างเฉียบพลันที่ทำให้เกิดรอยโรคที่ internal capsule เป็นต้น
3. ผู้ป่วยเชื้อชาติ สัญชาติไทย และพูดสำเนียงภาษากลาง

เกณฑ์การคัดผู้ป่วยที่มีอาการพูดไม่ชัดชนิดสປาสติกออกจากการศึกษาวิจัย (exclusion criteria)

1. ผู้ป่วยมีอาการพูดไม่ชัดทางระบบประสาทชนิดอื่นร่วมด้วย โดยวินิจฉัยจากภาพถ่ายทางรังสีวิทยาและโรคทางระบบประสาทที่ทำให้เกิดอาการพูดไม่ชัดชนิดอื่น เช่น ภาวะสมองขาดเลือดอย่างเฉียบพลันที่ทำให้เกิดรอยโรคที่ cerebellum ที่ทำให้เกิดอาการพูดไม่ชัดชนิด Cerebellar dysarthria
2. ผู้ป่วยมีอาการพูดไม่ชัดโดยกำเนิดร่วมด้วย
3. ผู้ป่วยที่มีความบกพร่องทางภาษา (aphasia)
4. ผู้ป่วยที่ระดับการรู้ตัวที่ลดลงหรือไม่สามารถให้ความร่วมมือในการตรวจ เช่น ไม่เข้าใจคำสั่งของผู้ตรวจ เป็นต้น

เกณฑ์การคัดกลุ่มควบคุมเข้าร่วมการศึกษาวิจัย (inclusion criteria)

1. อายุเท่ากับหรือมากกว่า 18 ปี และน้อยกว่าหรือเท่ากับ 90 ปี
2. ผู้ป่วยเชื้อชาติ สัญชาติไทย และพูดสำเนียงภาษากลาง
3. ไม่มีอาการพูดไม่ชัดจากสาเหตุใดๆ ทั้งในอดีตและปัจจุบัน

3.3 ขนาดตัวอย่าง

เนื่องจากไม่เคยมีงานวิจัยที่นำเครื่องมือตรวจจับเสียงพูดและเปลี่ยนเป็นข้อความ (ASR) มาใช้ในการวินิจฉัยแยกอาการพูดไม่ชัดชนิดสປาสติกออกจากเสียงพูดปกติมาก่อน การศึกษานี้จึงอ้างอิงขนาดตัวอย่างที่เหมาะสมจากการศึกษาของการศึกษาในอดีตที่มีความใกล้เคียงกันมากที่สุด โดย Irene และคณะซึ่งเป็นการศึกษาประสิทธิภาพของเครื่องมือ ASR ในการตรวจจับเสียงพูดไม่ชัดล่าสุดพบว่าความแม่นยำโดยเฉลี่ยเครื่องมือ ASR ในการเปลี่ยนเสียงพูดไม่ชัดเป็นข้อความอยู่ที่ร้อยละ 60 ในขณะที่ข้อความที่แปลงจากเสียงพูดปกติมีความแม่นยำโดยเฉลี่ยอยู่ที่ร้อยละ 90 (Calvo, Tropea et al. 2021) (อย่างไรก็ตามโมเดลการสร้างคลื่นเสียงและโมเดลทางภาษาที่ใช้ในการสร้างข้อความมีความแตกต่างกับโมเดลในการศึกษานี้ชัดเจน แต่ก็เป็นการล่าสุดและใกล้เคียงกับการศึกษานี้มากที่สุดที่ได้รับการตีพิมพ์)

จึงคำนวณขนาดตัวอย่างจากพื้นฐานอัตราความผิดพลาดในการสร้างข้อความจากเสียงพูดที่ต่างกันของสองกลุ่ม (เพื่อให้สอดคล้องกับวิธีการประเมินผลในการศึกษานี้) ด้วยสูตรคำนวณ

$$n = (Z_{\alpha/2} + Z_{\beta})^2 * (p_1(1-p_1) + p_2(1-p_2)) / (p_1 - p_2)^2$$

โดยที่ $Z_{\alpha/2}$ คือ critical value ของ Normal distribution ที่ $\alpha/2$ (ยกตัวอย่างหาก confidence level ที่ 95%, α คือ 0.05 และ critical value คือ 1.96)

Z_{β} คือ critical value ของ Normal distribution ที่ β (ยกตัวอย่างหาก power ที่ 80%, β คือ 0.2 และ critical value คือ 0.84)

p_1 and p_2 คือ สัดส่วนที่คาดการณ์ของตัวอย่างกลุ่มที่ 1 และกลุ่มที่ 2 ตามลำดับ

คำนวณผ่านสูตรคำนวณออนไลน์ทาง website

<https://epitools.ausvet.com.au/twoproportions>

กำหนดค่า $\alpha = 0.05$ และ $\beta = 0.20$

คำนวณจำนวนกลุ่มตัวอย่างได้ กลุ่มละ 38 คน

3.4 ขั้นตอนการทำวิจัย

ผู้ป่วยที่มีอาการปวดไมซัดชนิดสไปนัลติกที่ได้รับการยืนยันการวินิจฉัยโรคทางระบบประสาท โดยประสาทแพทย์และได้รับการตรวจวินิจฉัยยืนยันด้วยภาพถ่ายทางรังสี (CT brain และ MRI brain) ที่เข้าเกณฑ์ inclusion และ exclusion criteria ในการศึกษาได้รับชักชวนให้เข้าร่วมการศึกษา โดยเป็นผู้ป่วยจากหอผู้ป่วยในแผนกโรคหลอดเลือดสมองขาดเลือดชั้น 18 โชน A อาคาร ภูมิสิริมังคลานุสรณ์ (เก็บตัวอย่างชนิด prospective data collection) และจากแผนกผู้ป่วยนอก แผนกประสาทวิทยาอาคารปร. (เก็บตัวอย่างชนิด retrospective data collection) ประกอบด้วยผู้ป่วยทั้งหมด 37 คน (เพศชาย 29 คนและเพศหญิง 8 คน)

ผู้เข้าร่วมการวิจัยในกลุ่มควบคุมจากคลินิกผู้สูงอายุสุขภาพดี (อาคารสธ. ชั้น 4) และจากกลุ่มเจ้าหน้าที่โรงพยาบาลจุฬาลงกรณ์ที่เข้าเกณฑ์ inclusion criteria ในการศึกษา มีอายุและสัดส่วนของเพศเทียบเท่ากับกลุ่มผู้ป่วย (sex-age matched samples) ได้รับชักชวนให้เข้าร่วมการศึกษา ทั้งหมด 37 คน (เพศชาย 27 คนและเพศหญิง 10 คน)

โดยที่

1. ผู้วิจัยให้ข้อมูลการศึกษาทั้งหมดอย่างละเอียดรวมถึงข้อดี ข้อเสียและค่าตอบแทนของการเข้าร่วมการศึกษา

2. หลังได้รับความยินยอมจากผู้ป่วยในใบยินยอมเข้าร่วมการวิจัย ผู้ป่วยจะได้รับการเชิญไปบันทึกเสียงที่ห้องบันทึกเสียงที่จัดเตรียมไว้ในโรงพยาบาลจุฬาลงกรณ์ตามวันและเวลาที่นัดหมาย (ผู้ป่วยจากหอผู้ป่วยในจะได้รับการบันทึกเสียงเมื่ออาการทางระบบประสาทคงที่ในวันที่แพทย์ผู้รักษาจำหน่ายผู้ป่วยจากโรงพยาบาล)

หลังจากบันทึกข้อมูลเสียงแล้ว ข้อมูลจะถูกเก็บในรูปแบบของไฟล์เสียงและถูกนำไปวิเคราะห์ด้วยเครื่องมือ ASR ต่อไปตามรายละเอียดด้านล่าง

ก. การออกแบบประโยคพูดเพื่อเพิ่มความสามารถในการวินิจฉัย

- การทบทวนวรรณกรรมข้างต้นพบว่า การออกเสียงพยัญชนะต้นได้รับผลกระทบมากที่สุดในการพูดไม่ชัดชนิดสປาสติคสอดคล้องกับความรู้กล้ามเนื้อการออกเสียงที่ถูกกระทบจากรอยโรคที่ corticobulbar tract ผู้วิจัยจึงออกแบบประโยคพูดจากหลักการดังกล่าว เพื่อเพิ่มความสามารถในการแยกแยะเสียงพูดไม่ชัดออกจากเสียงปกติ
- ปรึกษานักอรรถบำบัด ภาควิชาเวชศาสตร์ฟื้นฟูโรงพยาบาลจุฬาลงกรณ์ ในการสร้างประโยคพูดเพื่อให้เป็นไปตามหลักการข้างต้น กล่าวคือสร้างประโยคตามพยัญชนะต้น รวมถึงพยัญชนะสะกดที่จะสอดคล้องไปกับตำแหน่งของกล้ามเนื้อที่ออกเสียง (placement) ในกลุ่มเดียวกันที่ถูกควบคุมด้วยเส้นประสาทสมองและโครงสร้างระบบประสาทเดียวกัน และครอบคลุมกลไกการขยับของกล้ามเนื้อ (manner) ให้ได้มากที่สุด (ดังตารางที่ 1) ในขณะที่เดียวกันต้องประกอบด้วยคำพื้นฐานที่เข้าใจง่าย ออกเสียงง่าย เช่น ไม่มีเสียงควบกล้ำ จดจำและพูดตามได้โดยง่าย จึงเกิดเป็นประโยคดังนี้
 1. ประโยคที่เกิดจากการเคลื่อนไหวกล้ามเนื้อริมฝีปาก (bilabial placement) คือ ‘ปุม้าวิ่งไปมาบนไปไม้’
 2. ประโยคที่เกิดจากการเคลื่อนไหวกล้ามเนื้อลิ้นกระทบเพดานปาก (lingual-alveolar placement) คือ ‘ชาวไร่ตัดต้นสนทำท่อนซุง’
 3. ประโยคที่เกิดจากการเคลื่อนไหวกล้ามเนื้อคอหอยและกล่องเสียง (velopharyngeal and glottis placement) คือ ‘อีกาคอยคาบงูคาบไก่’
 4. ประโยคที่เกิดจากการเคลื่อนไหวกล้ามเนื้อหลายส่วนรวมกัน (mixed placements) คือ ‘เพียงแค่ฝนตกลงที่หน้าต่างในบางครา’

ข. วิธีการบันทึกเสียง

1. ผู้ป่วยจะได้รับการบันทึกเสียงในห้องที่มีเสียงรบกวนภายนอกน้อยกว่า 40 dB สอดคล้องกับงานวิจัยศึกษาคลื่นเสียงในผู้ป่วยพูดไม่ชัดในอดีต (Kim, Kent et al. 2011, Lansford and Liss 2014) โดยแบ่งเป็นทั้งหมดสองห้องได้แก่ ห้องทดสอบการได้ยิน (audiometric test room) สำหรับผู้ป่วยนอกที่ได้รับการนัดหมาย และห้องที่มียุติเสียงภายในหอผู้ป่วยในแผนกโรคหลอดเลือดสมองขาดเลือดสำหรับผู้ป่วยใน
2. จัดตั้งอุปกรณ์บันทึกเสียงโดยอ้างอิงจากศึกษาคลื่นเสียงในผู้ป่วยพูดไม่ชัดโดย Lansford และ Kim (Kim, Kent et al. 2011, Lansford and Liss 2014) โดยที่จัดตั้งเหมือนกันในการบันทึกแต่ละครั้ง (รูปภาพที่ 7)
 - ก. บันทึกเสียงด้วยไมโครโฟนในโทรศัพท์มือถือรุ่น iPhoneXR ผ่านแอปพลิเคชันบันทึกเสียง Voice recorder กำหนดอัตราการสุ่มข้อมูลอยู่ที่ 44.1 kHz และความละเอียดของบิตอยู่ที่ 32-bit quantization
 - ข. โทรศัพท์มือถือจะถูกตั้งบนขาตั้งที่แข็งแรงและวางบนโต๊ะปรับความสูงได้ โดยไมโครโฟนโทรศัพท์มือถือจะมีระยะห่างจากผู้ป่วย 20 เซนติเมตรและทำมุม 45 องศากับตำแหน่งปากของผู้ป่วยเพื่อลดเสียงรบกวนจากการหายใจ
 - ค. แผ่นโฟมดูดซับเสียงและป้องกันเสียงสะท้อนถูกติดตั้งด้านหลังโทรศัพท์มือถือด้วยระยะห่าง 20 เซนติเมตรเท่ากัน
3. ทำการบันทึกเสียงทั้งหมดภายในนัดหมายครั้งเดียวกัน
 - ก. ให้ผู้เข้าร่วมวิจัยพูดตามประโยคทั้ง 4 ประโยคที่เตรียมไว้ข้างต้น โดยผู้วิจัยจะทวนแต่ละประโยคให้ฟังสองครั้งก่อนให้ผู้เข้าร่วมวิจัยหายใจเข้าลึกและพูดประโยคนั้นๆ ได้มากที่สุด 3 ครั้ง และผู้วิจัยเลือกประโยคที่จำนวนคำในแต่ละประโยคครบถ้วนที่สุดโดยไม่ได้คำนึงถึงความถูกต้องของการออกเสียง (เป็นตัวแปรของการวิจัย) โดยจะบันทึกเสียงทีละประโยค (จะได้ไฟล์เสียงจำนวน 4 ไฟล์ต่อผู้เข้าร่วมวิจัย 1 คน)
 - ข. หลังจากบันทึกเสียงที่ใช้ในการวิเคราะห์ทั้งหมดแล้วจึงให้ผู้เข้าร่วมวิจัยทวนคำทั้งหมด 36 คำใน ‘แบบประเมินความชัดเจนของการพูดในผู้ป่วยดิสอาเทรียไทย (Intelligibility scores in Thai dysarthria patient’s speech)’ ออกแบบโดย นันทนา(2000) ที่ถูกใช้อย่างแพร่หลายในคลินิกอรรถบำบัด เพื่อประเมินความถูกต้องของการพูดซึ่งเป็นตัวแทนของความรุนแรงในการพูดไม่ชัด (dysarthria severity) โดยนักอรรถบำบัดโรงพยาบาลจุฬาลงกรณ์จะเป็นผู้ประเมินความถูกต้องของแต่ละคำและคิดคะแนนเป็นร้อยละของคำที่ถูกต้อง

และผู้ป่วยจะถูกแบ่งกลุ่มตามความรุนแรงออกเป็น ความรุนแรงน้อย (mild dysarthria) ที่ได้คะแนนระหว่างร้อยละ 75 ถึง 100 ความรุนแรงปานกลาง (moderate dysarthria) ได้คะแนนอยู่ที่ร้อยละ 45 ถึงร้อยละ 75 และความรุนแรงมาก (severe dysarthria) ได้คะแนนอยู่ที่ 0 ถึงร้อยละ 44

ค. วิธีการประเมินคะแนนประโยคพูด

1. เลือกวิธีการวินิจฉัยประโยคพูดไม่ชัดชนิดสเปคตริกด้วยเครื่องมือการเปลี่ยนเสียงพูดเป็นข้อความด้วยโมเดลทางคอมพิวเตอร์และปัญญาประดิษฐ์ หรือ Automatic speech-to-text recognition (ASR) ที่มีรูปแบบกระบวนการวิเคราะห์คลื่นเสียงที่ทันสมัย เป็นรูปธรรม และมีความน่าเชื่อถือประกอบกับมีโมเดลทางภาษาในการสร้างข้อความภาษาไทยขึ้นมาได้เป็นอย่างดี รวมถึงมีงานวิจัยรองรับและได้รับการยอมรับจากผู้ใช้งานจริงดังกล่าวในบทความวรรณกรรม เป็นที่แพร่หลายในสังคมแพทย์และสามารถเข้าถึงได้ง่าย เพื่อให้การประเมินมีความเป็นรูปธรรม (objective) ที่เที่ยงตรง (precise) และน่าเชื่อถือ (reliability) เครื่องมือที่ถูกเลือกใช้ได้แก่
 - ก. Apple Siri™ speech recognition ในระบบปฏิบัติการ Apple iOS 16.3 (ปรับเป็นเวอร์ชันล่าสุดเดือนกุมภาพันธ์ 2023) (รูปภาพที่ 8)
 - ข. Whisper-Thai-Demo เป็นซอฟต์แวร์ ASR ที่สามารถใช้งานบนหน้าเว็บไซต์ '<https://huggingface.co/spaces/biodatlab/whisper-thai-demo>' ซึ่งถูกพัฒนาเพื่อรองรับภาษาไทยโดยเฉพาะโดยทีมนักวิจัยทางโมเดลคอมพิวเตอร์และปัญญาประดิษฐ์จากหลายประเทศที่รู้จักในนาม Hugging face (รูปภาพที่ 9) โดยผู้วิจัยเลือกโมเดลนี้เพื่อเป็นการเปรียบเทียบความสามารถในการวินิจฉัยของเครื่องมือที่ใช้หลักการวิเคราะห์เดียวกันแต่มีความหลากหลายของกระบวนการวิเคราะห์ว่ามีประสิทธิภาพที่ใกล้เคียงกันหรือไม่ และยังสามารถยืนยันสมมติฐานการนำองค์ความรู้การพูดไม่ชัดมาเพิ่มประสิทธิภาพในการวินิจฉัยของประโยคหากผลการทดสอบของทั้งสองเครื่องมือสอดคล้องกันได้อีกด้วย
2. ป้อนข้อมูลเสียงที่บันทึกไว้ในรูปแบบไฟล์เสียงดิจิทัลให้กับเครื่องมือ ASR ทั้งสองเครื่องมือ ซึ่งในแต่ละเครื่องมือจะให้ผลลัพธ์เป็นข้อความทั้งหมด 4 ประโยคต่อผู้เข้าร่วมวิจัย 1 คน

- ก. เปิดใช้งานระบบ Apple Siri™ speech recognition ในโทรศัพท์ iPhone และเล่นไฟล์เสียงกลุ่มตัวอย่างที่บันทึกไว้ด้วยความดังที่มากพอ จากนั้นรอการประมวลผลออกมาเป็นข้อความ (ใช้เวลาไม่เกิน 1 วินาที)
- ข. นำเข้าไฟล์เสียงไปยังหน้าเว็บไซต์ Whisper-Thai-Demo และรอการประมวลผลไฟล์เสียงออกมาเป็นข้อความโดยที่ไม่ต้องเปิดเล่นไฟล์เสียงผ่านลำโพง (ใช้เวลา 3-5 วินาที)

(ข้อความที่สร้างขึ้นโดยเครื่องมือมีความเที่ยงตรงเป็นข้อความเดิมในทุกครั้งของการวิเคราะห์)

3. ประเมินความผิดพลาดในการสร้างข้อความโดยการนับจำนวนพยางค์ที่ผิดในประโยค (incorrect syllable) และรวมเป็นคะแนนความผิดพลาดของพยางค์แต่ละประโยค (error score of syllable) โดยหลักเกณฑ์ของพยางค์ที่ผิดคือ พยางค์ที่อ่านออกเสียงต่างออกไปจากที่ควรจะเป็นด้วยองค์ประกอบที่ไม่ถูกต้องในพยัญชนะต้น สระวรรณยุกต์ หรือพยัญชนะตัวสะกด และไม่นับพยางค์ที่อ่านออกเสียงเหมือนกันแต่สะกดต่างกัน อาทิ ‘ทาวไล้ตัดต้นสนธรรมท่อนสูง’ ประกอบด้วยพยางค์ที่ผิดคือ ‘ทาว’ ‘ไล้’ และ ‘สูง’ แต่ไม่นับพยางค์ ‘ธรรม’ เป็นพยางค์ที่ผิดเนื่องจากออกเสียงเหมือน ‘ท่า’ แต่สะกดต่างกัน จึงสรุปได้ว่าประโยคข้างต้นมีคะแนนความผิดพลาด 3 คะแนน จาก 8 คะแนน ดังนั้นผู้เข้าร่วมวิจัย 1 คนจะมีคะแนนความผิดพลาดของพยางค์ทั้งหมด 4 ค่า

ง. การวินิจฉัยอาการหูไม่ชัดโดยประสาทแพทย์

ทำการประเมินความสามารถในการวินิจฉัยอาการหูไม่ชัดชนิดประสาทหู โดยประสาทแพทย์ที่จะเป็นผู้ฟังและประเมินว่าเสียงพูดที่บันทึกไว้เป็นเสียงจากผู้ป่วยที่มีอาการหูไม่ชัดหรือไม่ โดยผู้ฟังประกอบด้วยประสาทแพทย์ทั้งหมด 10 คน เป็นอาจารย์แพทย์จำนวน 5 คน แพทย์ประจำบ้านต่อยอดสาขาหลอดเลือดสมองจำนวน 2 คน และแพทย์ประจำบ้านจำนวน 3 คน

ประสาทแพทย์ทุกคนต้องฟังไฟล์เสียงครบทั้ง 4 ไฟล์ของผู้เข้าร่วมวิจัยแต่ละคน ทั้งจากกลุ่มผู้ป่วยและกลุ่มควบคุมโดยที่ไม่ทราบข้อมูลใดๆของผู้เข้าร่วมวิจัย และต้องตัดสินใจว่าผู้เข้าร่วมวิจัยแต่ละคนมีอาการหูไม่ชัดชนิดประสาทหูหรือหูตึงปกติ จากนั้นจึงประเมินประสิทธิภาพในการวินิจฉัยโดยคำนวณเป็นความไว (sensitivity) ความจำเพาะ (specificity) และความถูกต้อง (accuracy) ของประสาทแพทย์แต่ละคน

3.5 การรวบรวมข้อมูล

วิธีการเก็บข้อมูล

- ผู้เข้าร่วมวิจัยได้รับการบันทึกข้อมูลพื้นฐาน ได้แก่ อายุ เพศ ดรรชนีมวลกาย (BMI) ที่มีผลต่อลักษณะของเสียงพูด (Dehqan, Scherer et al. 2012, Souza and Santos 2018) ภาวะที่อาจทำให้มีการเปลี่ยนแปลงของลักษณะเสียงพูดในวันที่บันทึกเสียง ได้แก่ มีน้ำมูกหรือเสมหะ มีภาวะเหนื่อย เป็นต้น และประวัติโรคหรือประวัติการผ่าตัดหู คอ จมูกที่อาจส่งผลต่อลักษณะเสียงโดยที่ผู้ป่วยและญาติไม่ได้ตระหนักถึง
- ผู้เข้าร่วมวิจัยกลุ่มผู้ป่วยพูดไม่ชัดจะได้รับการบันทึกข้อมูลที่เกี่ยวข้องกับโรคทางระบบประสาท ได้แก่ ประวัติการพูดไม่ชัดในอดีต การวินิจฉัยทางระบบประสาทในปัจจุบัน รอยโรคทางระบบประสาทจากภาพรังสีวินิจฉัย (CT brain และ MRI brain) สาเหตุของการเกิดโรคสมองขาดเลือดเฉียบพลันหรือ TOAST classification (เฉพาะผู้ป่วยที่ได้รับการวินิจฉัย acute ischemic stroke) ความรุนแรงของอาการพูดไม่ชัด ระยะเวลาตั้งแต่เริ่มมีอาการพูดไม่ชัด
- หลังจากทำการบันทึกเสียงข้อมูลจะถูกบันทึกไว้ในรูปไฟล์เสียงดิจิทัลและถูกโอนย้ายเพื่อเก็บข้อมูลใน External hard drive สำหรับบันทึกข้อมูลวิจัยโดยเฉพาะ
- เมื่อนำข้อมูลเสียงไปวิเคราะห์ด้วยเครื่องมือ ASR ได้ผลลัพธ์เป็นประโยค พร้อมทั้งคำนวณคะแนนความผิดพลาดของพยางค์แต่ละประโยคแล้ว ผลลัพธ์จะถูกเก็บในรูปแบบของไฟล์ excel ใน External hard drive ข้างต้นเช่นกัน

จุฬาลงกรณ์มหาวิทยาลัย

CHULALONGKORN UNIVERSITY

3.6 ข้อจำกัดในการวิจัย

- ผู้ป่วยที่เข้าร่วมการวิจัย เป็นผู้ป่วยโรคสมองขาดเลือดเฉียบพลันเป็นส่วนมาก (acute ischemic stroke) อาจไม่ใช่ตัวแทนที่ดีของประชากรผู้ป่วยพูดไม่ชัดชนิดสเปคตริกทั้งหมด
- แบบประเมินความชัดเจนของการพูดในผู้ป่วยดิสอาเทรียไทยที่นำมาใช้นั้น ไม่มีการระบุความแม่นยำของการระบุความรุนแรงของอาการพูดไม่ชัด และความน่าเชื่อถือ (interrater and intrarater reliability) ของเครื่องมือไว้ชัดเจน เนื่องจากถูกนำไปใช้ในการติดตามการบำบัดการพูดมากกว่าการวินิจฉัย และมีผู้ประเมินเป็นนักรรณบำบัดเพียงผู้เดียว อาจทำให้เกิดความคลาดเคลื่อนและอคติในการบอกความรุนแรงของอาการพูดไม่ชัดได้

- เครื่องมือ ASR มีความหลากหลายของกระบวนการวิเคราะห์และเครื่องมือที่นำมาใช้มีการพัฒนาอย่างต่อเนื่อง ดังนั้นควรมีความระมัดระวังในการนำผลการศึกษาวิจัยในครั้งนี้ไปใช้กับเครื่องมือ ASR ในอนาคตหรือเครื่องมืออื่นๆที่ไม่อยู่ในการวิจัย

3.7 การเปิดเผยข้อมูลแสดงตัวตนของผู้ป่วย

ข้อมูลที่แสดงตัวตนของผู้ป่วยจะถูกเก็บไว้เป็นความลับ และไม่มีการนำข้อมูลดังกล่าวไปเปิดเผยโดยเด็ดขาด สำหรับการวิเคราะห์ข้อมูลจะใช้รหัสแทนตัวผู้ป่วยแต่ละราย นอกจากนี้การตีพิมพ์หรือนำเสนอผลงานการวิจัยในงานวิชาการจะนำเสนอในภาพรวมของผลการวิจัยโดยที่ไม่แสดงตัวตนของผู้ป่วย หากมีความจำเป็นต้องแสดงข้อมูลตัวตนของผู้ป่วยจะต้องได้รับความยินยอมจากผู้ป่วยอย่างเป็นทางการลายลักษณ์อักษรเท่านั้น

3.8 การวิเคราะห์ข้อมูล

ข้อมูลปัจจัยพื้นฐานและปัจจัยอื่นๆที่เกี่ยวข้องกับเสียงพูดของผู้ป่วยจะถูกแสดงเป็นค่าเฉลี่ยและค่าความคลาดเคลื่อนมาตรฐานสำหรับข้อมูลแบบต่อเนื่อง ได้แก่ อายุและ BMI และแสดงเป็นจำนวนและร้อยละสำหรับข้อมูลแบบไม่ต่อเนื่อง จากนั้นใช้สถิติเชิงวิเคราะห์ด้วย ‘Independent samples T-test’ สำหรับข้อมูลแบบต่อเนื่อง และ ‘Fisher’s exact test’ สำหรับข้อมูลไม่ต่อเนื่อง เพื่อพิจารณาความแตกต่างทางสถิติของปัจจัยต่างๆระหว่างทั้งสองกลุ่ม

ข้อมูลคะแนนความผิดพลาดของประโยคทั้งหมดจะถูกวิเคราะห์เชิงพรรณนาโดยแสดงเป็นค่าเฉลี่ยและค่าความคลาดเคลื่อนมาตรฐานของแต่ละประโยคในผู้ป่วยแต่ละกลุ่ม และถูกแสดงผลในลักษณะเดียวกันสำหรับทั้งสองเครื่องมือ (ตารางที่)

จากนั้นใช้สถิติเชิงวิเคราะห์ด้วย ‘Independent samples T-test’ เพื่อพิจารณาคะแนนความผิดพลาดระหว่างกลุ่มผู้ป่วยและกลุ่มตัวอย่างในแต่ละเครื่องมือว่ามีความแตกต่างอย่างมีนัยสำคัญหรือไม่ พร้อมทั้งแสดงให้เห็นว่าประโยคใดที่บอกความแตกต่างได้มากที่สุด

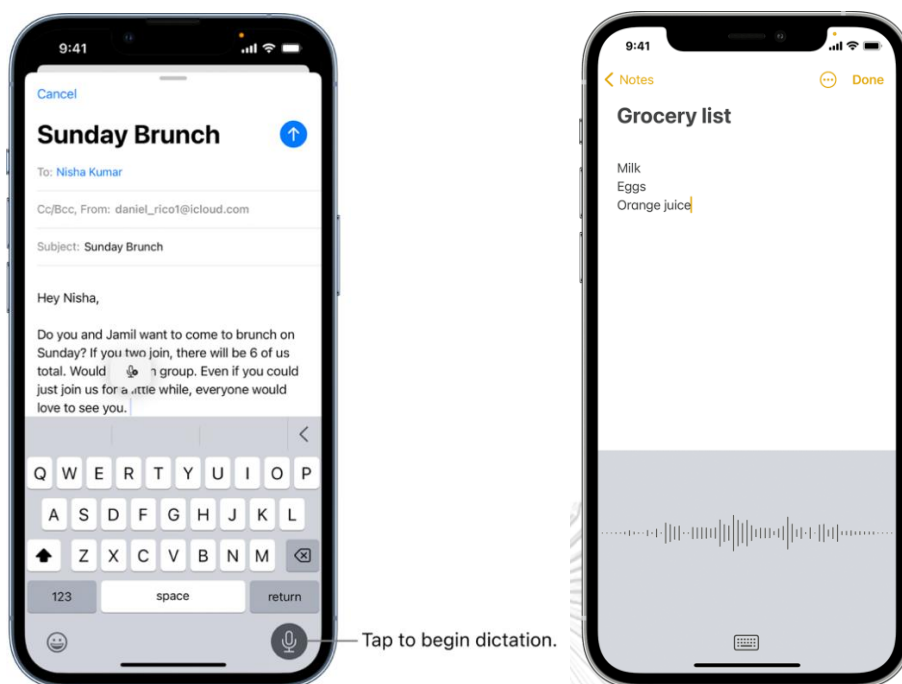
นำข้อมูลคะแนนของแต่ละประโยคในทุกกลุ่มมาวิเคราะห์ทางสถิติด้วยวิธีการวิเคราะห์ ‘การถดถอยโลจิสติก (Logistic regression analysis)’ เพื่อสร้างสมการความสัมพันธ์เชิงเส้นระหว่างคะแนนความผิดพลาดและโอกาสของอาการพูดไม่ชัดของแต่ละประโยคในแต่ละเครื่องมือ จากนั้นจึงทำการสร้าง ‘เส้นโค้ง Receiver Operating Characteristic (ROC)’ จากความสัมพันธ์เชิงเส้น และ

คำนวณพื้นที่ใต้เส้นโค้ง (Area under the curve) พร้อมทั้งค่าความเชื่อมั่น (95% confidence interval) เพื่อเป็นตัวบ่งชี้ประสิทธิภาพและความแม่นยำในการวินิจฉัยอาการพูดไม่ชัดด้วยเครื่องมือ ASR ร่วมกับประโยคที่สร้างขึ้น และสามารถเปรียบเทียบกันระหว่างเครื่องมือได้อีกด้วย และพิจารณาจุดตัดที่เหมาะสมในเส้นโค้ง ROC เพื่อกำหนดจุดตัดของคะแนนความผิดพลาดที่จะทำให้ได้ค่าความไวและความจำเพาะสูงที่สุดสำหรับการวินิจฉัยอาการพูดไม่ชัด อนึ่งผลการวิเคราะห์ทางสถิติเบื้องต้นทั้งหมดจะมีความแตกต่างอย่าง มีนัยสำคัญที่ $p\text{-value} < 0.05$



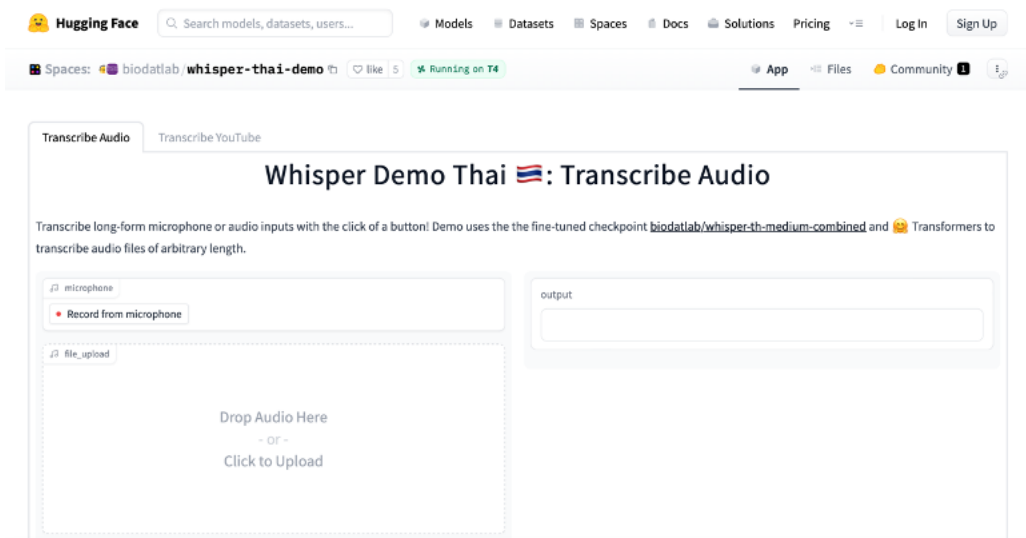
CHULALONGKORN UNIVERSITY

รูปภาพที่ 7 แสดงการจัดวางอุปกรณ์บันทึกเสียงในห้องบันทึกเสียง



รูปภาพที่ 8 แสดงการใช้งานจริงของ Apple Siri™ speech recognition

ภาพซ้ายแสดงแป้นสัมผัสด้านล่างของหน้าจอโทรศัพท์เพื่อเริ่มใช้งานโดยผู้ใช้งานพูดและโมเดลจะทำการเปลี่ยนเป็นข้อความในอีเมลได้อย่างถูกต้อง ภาพขวาแสดงหน้าจอแสดงผลขณะพูดในส่วนสี่เหลี่ยมด้านล่าง



รูปภาพที่ 9 แสดงหน้าเว็บไซต์ Hugging face ในส่วน Whisper-Thai-Demo โดยมี address คือ <https://huggingface.co/spaces/biodatlab/whisper-thai-demo> ซึ่งผู้ใช้นำเข้าไฟล์เสียงที่กล่องด้านซ้ายล่าง (file_upload) จากนั้นรอการประมวลผลแสดงออกในกล่องข้อความด้านขวาบน (output)

บทที่ 4

ผลการวิเคราะห์ข้อมูล

4.1 ข้อมูลพื้นฐานของประชากรที่นำมาศึกษา

อยู่ในช่วงเดือนกันยายน 2565 ถึงเดือนกุมภาพันธ์ 2566 มีประชากรที่นำมาศึกษาทั้งหมด 74 คนประกอบด้วยผู้ป่วยหูตึงไม่ชัดชนิดประสาทหูจำนวน 37 คน (เพศชาย 29 คนและเพศหญิง 8 คน) และผู้เข้าร่วมกลุ่มควบคุมจำนวน 37 คน (เพศชาย 27 คนและเพศหญิง 10 คน) ในอายุที่ใกล้เคียงกัน คือผู้ป่วยหูตึงไม่ชัดชนิดประสาทหูมีอายุอยู่ระหว่าง 37 ถึง 91 ปี (อายุเฉลี่ย 65.4 ± 13.5 ปี) และผู้เข้าร่วมกลุ่มควบคุมมีอายุระหว่าง 40 ถึง 83 ปี (อายุเฉลี่ย 62.7 ± 10.8 ปี) โดยพบว่าไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติของอายุในทั้งสองกลุ่ม ($p=0.33 > 0.05$) นอกจากนี้สัดส่วนของเพศในทั้งสองกลุ่มก็ไม่มี ความแตกต่างกันอย่างมีนัยสำคัญทางสถิติเช่นกัน ($p=0.787$) โดยเพศชายมีจำนวนมากกว่าเพศหญิงในทั้งสองกลุ่มด้วยข้อจำกัดของช่วงเวลาการเก็บข้อมูล กลุ่มตัวอย่างที่เข้าร่วมการศึกษาทั้งหมดสามารถพูดและอ่านตามประโยคที่เตรียมไว้ได้โดยไม่มีข้อจำกัดทางการตื่นรู้ตัวทางการมองเห็น หรือทางการอ่านใดๆ และกลุ่มตัวอย่างทุกคนสามารถพูดภาษาไทยสำเนียงกลางได้ แม้มารจากต่างภูมิภาค

ปัจจัยอื่นที่อาจส่งผลกระทบต่ออาการหูตึงของผู้ป่วย ปัจจัยแรกคือปัจจัยภาวะทางหู คอ จมูกที่มีผลกระทบต่ออาการหูตึงได้ ซึ่งผู้เข้าร่วมวิจัยทั้งหมดไม่มีโรคทางหู คอ จมูกหรือประวัติการผ่าตัดบริเวณดังกล่าวที่ส่งผลกระทบต่อเสียงพูดชัดเจน แต่มีภาวะทางหูคอจมูกพื้นฐานบางอย่างที่อาจส่งผลกระทบต่อเสียงพูด อาทิ การได้ยินที่ลดลงแต่ยังสามารถสื่อสารได้ปกติและภาวะฟันไม่ครบซี่ที่มักพบได้ในผู้สูงอายุ เป็นต้น (Relevant ENT conditions) โดยในกลุ่มผู้ป่วยหูตึงไม่ชัดพบภาวะทางหูคอจมูกดังกล่าวทั้งหมด 17 คน (ร้อยละ 37.84) และในกลุ่มควบคุมพบผู้มีภาวะดังกล่าวทั้งหมด 9 คน (ร้อยละ 24.32) การคำนวณทางสถิติพบว่าไม่มีความแตกต่างของสัดส่วนภาวะทางหูคอจมูกที่อาจเกี่ยวข้องกับเสียงพูดอย่างมีนัยสำคัญทางสถิติ ($\chi^2=0.315$, df 1, p-value at 0.05)

ปัจจัยที่สองคืออาการทางหูคอจมูกในปัจจุบันที่ส่งผลกระทบต่อเสียงพูด ได้แก่ มีน้ำมูก มีเสมหะ เสียงแหบ ภาวะเหนียวที่ไม่สามารถออกเสียงได้ยาว มีสายยางให้อาหาร เป็นต้น พบว่าในกลุ่มผู้ป่วยหูตึงไม่ชัดมีอาการทางหูคอจมูกในปัจจุบันทั้งหมด 8 คน (ร้อยละ 21.62) และกลุ่มควบคุมพบอาการดังกล่าวทั้งหมด 2 คน (ร้อยละ 5.41) โดยพบว่าไม่มีความแตกต่างอย่างมีนัยสำคัญทางสถิติของอาการดังกล่าวระหว่างผู้ป่วยทั้งสองกลุ่ม ($\chi^2=0.085$, df = 1, p-value at 0.05) อย่างไรก็ตามแม้ไม่

เห็นความแตกต่างทางสถิติอย่างมีนัยสำคัญแต่อาจมีความแตกต่างทางคลินิกเนื่องจากสัดส่วนผู้ที่
อาการทางหูคอจมูกในกลุ่มผู้ป่วยมากกว่ากลุ่มควบคุมถึง 4 เท่า ผู้วิจัยจึงได้วิเคราะห์ความอ่อนไหว
โดยการคัดผู้ป่วยที่มีปัจจัยดังกล่าวออกเพิ่มเติม (sensitivity analysis) ดังแสดงในภาคผนวก
(รูปภาพที่ 16)

ในกลุ่มผู้ป่วยพูดไม่ชัดชนิดสเปคตริกมีผู้ป่วยที่เคยพูดไม่ชัดชนิดสเปคตริกในอดีตมาก่อน 6 คน
(ร้อยละ 16.22) ซึ่งผู้ป่วยและญาติยืนยันว่าเสียงพูดกลับมาปกติหรือใกล้เคียงปกติมากแล้วทั้งหมด
ผู้ป่วยเกือบทุกรายได้รับการวินิจฉัยเป็นภาวะสมองขาดเลือดเฉียบพลัน (acute ischemic stroke)
จำนวน 35 คน (ร้อยละ 94.59) ผู้ป่วย 1 คน ได้รับการวินิจฉัยเป็นโรคความเสื่อมถอยของเซลล์
ประสาทควบคุมการเคลื่อนไหวแบบค่อยเป็นค่อยไป (primary lateral sclerosis) และผู้ป่วย 1 คน
ได้รับการวินิจฉัยเป็นก้อนยังไม่ระบุชนิดในเนื้อสมอง (undetermined brain mass) ช่วงเวลาตั้งแต่
ผู้ป่วยมีอาการไปจนถึงเวลาที่บันทึกเสียงมีความแตกต่างกันตั้งแต่ 3 วันถึง 526 วัน (ค่ามัธยฐานของ
ช่วงเวลาที่บ้านทักเสียงหรือ Median time-to-recording คือ 6 วัน)

ความรุนแรงของอาการพูดไม่ชัด (dysarthria severity) ได้ถูกประเมินโดยคะแนนความ
ถูกต้องและเข้าใจได้ของประโยคที่พูด (intelligibility) โดยใช้แบบประเมินและนักออรรถบำบัด
ดังกล่าวในขั้นตอนการทำวิจัย พบว่ากลุ่มผู้ป่วยประกอบด้วยอาการพูดไม่ชัดที่รุนแรงน้อย (mild
dysarthria) จำนวน 34 คน (ร้อยละ 91.89) พูดไม่ชัดรุนแรงปานกลาง (moderate dysarthria)
จำนวน 2 คน (ร้อยละ 5.41) และพูดไม่ชัดรุนแรงมาก (severe dysarthria) จำนวน 1 คน (ร้อยละ
2.70) หรืออาจกล่าวได้ว่าการศึกษานี้เป็นการศึกษาเครื่องมือการวินิจฉัยอาการพูดไม่ชัดชนิดสเปคตริ
กออกจากเสียงพูดปกติ โดยเฉพาะกลุ่มที่มีความรุนแรงเพียงเล็กน้อย

ในกลุ่มผู้ป่วยพูดไม่ชัดชนิดสเปคตริกที่ได้รับการวินิจฉัยเป็นภาวะสมองขาดเลือดเฉียบพลัน
(acute ischemic stroke) 35 คน มีสาเหตุการเกิดโรคตาม TOAST classification ได้แก่ ภาวะขาด
เลือดจากการตีบตันของหลอดเลือดขนาดเล็ก (small vessel disease) มากที่สุดจำนวน 23 คน
(ร้อยละ 65.71) ภาวะขาดเลือดจากหลอดเลือดใหญ่ตีบ (large vessel atherosclerosis) จำนวน 4
คน (ร้อยละ 11.43) ภาวะหลอดเลือดอุดตันจากลิ้มเลือดหัวใจ 4 คน (ร้อยละ 11.43) ภาวะขาดเลือด
จากสาเหตุอื่นๆ 1 คน (ร้อยละ 2.86) และไม่ทราบสาเหตุการขาดเลือดจำนวน 3 คน (ร้อยละ 8.57)

ผู้ป่วยโรคทางระบบประสาทส่วนกลาง (ผู้ป่วยภาวะสมองขาดเลือดเฉียบพลันและผู้ป่วยที่มี
ก้อนในสมอง) ที่มีอาการพูดไม่ชัดชนิดสเปคตริกมีตำแหน่งรอยโรคต่างกัน ได้แก่ ใต้ผิวสมอง

(subcortical) มากที่สุดจำนวน 20 คน (ร้อยละ 54.05) ที่ผิวสมอง (cortical) จำนวน 8 คน (ร้อยละ 21.62) ตำแหน่ง Internal capsule จำนวน 9 คน (ร้อยละ 24.32) ตำแหน่งก้านสมองส่วน cerebral peduncle จำนวน 2 คน (ร้อยละ 5.41) ตำแหน่งก้านสมองส่วน basis pontis จำนวน 13 คน (ร้อยละ 35.14)

Characteristics	Results		P-value	χ^2
	Spastic dysarthria	Control		
Demographic characteristics				
Male gender, n (%)	29 (78.38)	27 (72.97)	NS	0.787
Mean age, years (SD)	65.41 (13.45)	62.65 (10.80)	0.332	
Mean BMI, kg/m ² (SD)	23.75 (3.09)	24.49 (3.73)	0.361	
Voice disturbances				
Relevant ENT conditions, n (%)	14 (37.84)	9 (24.32)	NS	0.315
Local voice disturbances, n (%)	8 (21.62)	2 (5.41)	NS	0.085
Clinical characteristics				
History of spastic dysarthria, n (%)	6 (16.22)			
Diagnosis, n (%)				
- Stroke	35 (94.59)			
- Primary lateral sclerosis	1 (2.70)			
- Undetermined brain mass	1 (2.70)			
TOAST classification, n (%)				

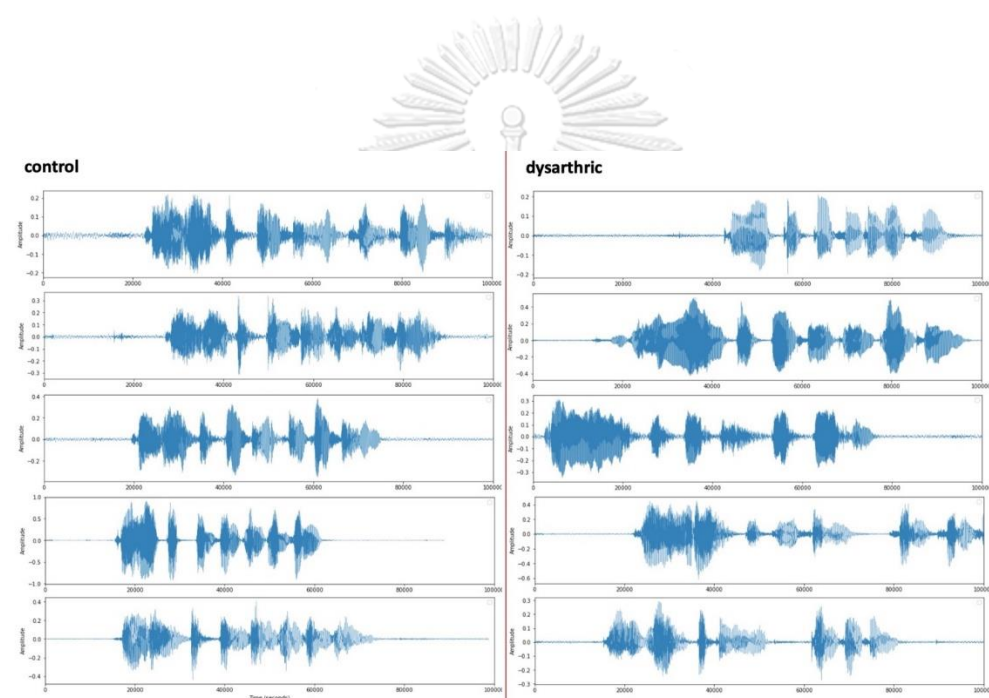
- Large-vessel atherosclerosis	4 (11.43)
- Small vessel disease	23 (65.71)
- Cardioembolic	4 (11.43)
- Other determined	1 (2.86)
- Undetermined	3 (8.57)
Median time-to-recording, days (IQR)	6.00 (4-90)
Intelligibility, n (%)	
- Mild	34 (91.89)
- Moderate	2 (5.41)
- Severe	1 (2.70)
Lesion sites, n (%)	
- Cortical	8 (21.62)
- Subcortical	20 (54.05)
- Internal capsule	9 (24.32)
- Cerebral peduncle	2 (5.41)
- Basis pontis	13 (35.14)

ตารางที่ 2 แสดงข้อมูลพื้นฐานของผู้ป่วยที่มีอาการพูดไม่ชัดชนิดส pastic และกลุ่มควบคุม และลักษณะทางคลินิกของผู้ป่วยที่มีอาการพูดไม่ชัดชนิดส pastic

4.2 กราฟคลื่นเสียงที่ได้จากการวิเคราะห์เบื้องต้น

จากการวิเคราะห์ไฟล์เสียงพูดไม่ชัดและแสดงผลเป็นกราฟคลื่นเสียงเบื้องต้นด้วยโมเดลทางคอมพิวเตอร์ผ่านโปรแกรม 'Python programming' เบื้องต้น และสุ่มเลือกไฟล์เสียงของผู้ป่วยและ

กลุ่มควบคุมมาเปรียบเทียบเชิงพรรณนา พบว่ากราฟคลื่นเสียงที่ถูกสร้างขึ้นจากประโยคพูดไม่ชัดชนิดสປาสติคมีความแตกต่างของลักษณะคลื่นเสียงอย่างชัดเจนที่สามารถเห็นได้ด้วยตาเปล่า อาทิ ตัวอย่างจากประโยคที่เกิดจากการเคลื่อนไหวกลิ้มเนื้อลิ้นกระทบเพดานปาก (lingual-alveolar placement) หรือ ‘ชาวไร่ตัดต้นสนทำท่อนซุง’ จากผู้ป่วยและกลุ่มควบคุมอย่างละ 5 คน พบความแตกต่างอย่างเห็นได้ชัดในหลายแ่งมุมของคลื่นเสียง ได้แก่ จังหวะของพยางค์ที่พูด การรวมกันของสระและพยัญชนะของพยางค์ที่แยกกัน (co-articulation) ความสม่ำเสมอของความดังในการพูด เป็นต้น



รูปภาพที่ 10 แสดงกราฟคลื่นเสียงประโยคพูด ‘ชาวไร่ตัดต้นสนทำท่อนซุง’ จากผู้ป่วยพูดไม่ชัดชนิดสປาสติคและกลุ่มควบคุมอย่างละ 5 คน

4.3 ผลการเปรียบเทียบปัจจัยที่ศึกษา

จากการวิเคราะห์เสียงประโยคพูดของผู้เข้าร่วมวิจัยทั้งหมดด้วยเครื่องมือ Whisper-Thai-demo ออกมาเป็นประโยคและนับจำนวนพยางค์ที่ผิดในประโยค (incorrect syllable) พบว่าจำนวนผู้ป่วยและกลุ่มควบคุมที่มีคะแนนความผิดพลาดของพยางค์ในแต่ละประโยค (error score of syllable) ตั้งแต่ไม่ผิดพลาดเลย (ได้คะแนนเท่ากับ 0) จนถึงผิดพลาดทั้งหมด (ได้คะแนนเท่ากับ 7, 8

หรือ 12 ตามจำนวนพยางค์ในประโยคนั้น) ถูกแสดงในตารางที่ 3 และการวิเคราะห์เสียงประโยคพูดของผู้เข้าร่วมวิจัยทั้งหมดด้วยเครื่องมือ Apple Siri™ ถูกแสดงในลักษณะเดียวกันในตารางที่ 4

Sentence by Whisper	LAP								BP								VAGP							MP															
Number of errors	0	1	2	3	4	5	6	7	8	0	1	2	3	4	5	6	7	8	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	8	9	10	11	
Number of subjects, n																																							
Spastic dysarthria	7	4	4	9	5	3	0	4	1	11	13	5	1	5	0	1	0	1	13	2	9	9	3	0	0	1	1	20	11	2	1	0	0	0	0	0	1	1	
Control	19	10	6	1	1	0	0	0	0	15	22	0	0	0	0	0	0	0	22	4	11	0	0	0	0	0	6	29	2	0	0	0	0	0	0	0	0	0	

ตารางที่ 3 แสดงจำนวนผู้ป่วยและกลุ่มควบคุมตามคะแนนความผิดพลาดของพยางค์ในแต่ละประโยค (error score of syllable) ที่ได้หลังการนับจำนวนพยางค์ที่ผิดในประโยค (incorrect syllable) ที่ถูกวิเคราะห์ด้วย Whisper-Thai-demo โดย

LAP คือ Lingual-Alveolar Placement หรือประโยค ‘ชาวไร่ตัดต้นสนทำท่อนซุง’

BP คือ Bilabial Placement หรือประโยค ‘ปทุมวิจิ้งไปมาบนใบไม้’

VAGP คือ Velopharyngeal and Glottis Placement หรือประโยค ‘อีภาคอยคาบุงคาบไก่’

MP คือ Mixed Placement หรือประโยค ‘เพียงแค่มนต์กลางที่หน้าต่างในบางครา’

Sentence by Siri	LAP								BP								VAGP							MP															
Number of errors	0	1	2	3	4	5	6	7	8	0	1	2	3	4	5	6	7	8	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	8	9	10	11	
Number of subjects, n																																							
Spastic dysarthria	4	3	10	5	2	2	3	2	6	14	9	2	3	4	0	2	0	3	4	7	2	9	4	4	4	3	13	10	3	4	1	3	0	0	0	0	0	3	
Control	27	5	3	1	1	0	0	0	0	29	5	3	0	0	0	0	0	0	5	16	6	8	2	0	0	0	29	7	0	1	0	0	0	0	0	0	0	0	

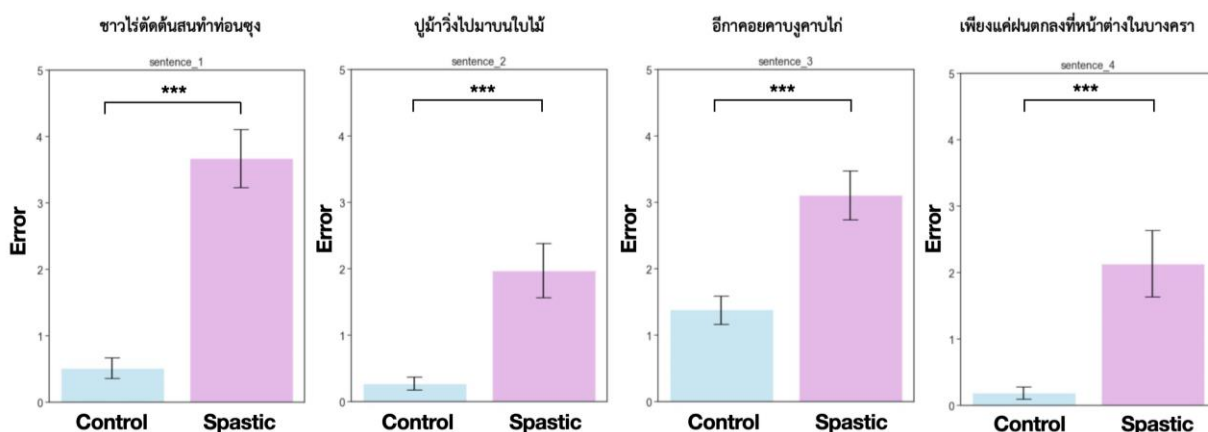
ตารางที่ 4 แสดงจำนวนผู้ป่วยและกลุ่มควบคุมตามคะแนนความผิดพลาดของพยางค์ในแต่ละประโยค (error score of syllable) ที่ได้รับหลังการนับจำนวนพยางค์ที่ผิดในประโยค (incorrect syllable) ที่ถูกวิเคราะห์ด้วย Apple Siri™

เมื่อวิเคราะห์ด้วยเครื่องมือ Whisper-Thai-demo พบว่าค่าเฉลี่ยของคะแนนความผิดพลาดของพยางค์ (error score of syllable) ในประโยค ‘ชาวไร่ตัดต้นสนทำท่อนซุง’ (LAP) ในกลุ่มผู้ป่วยพูดไม่ชัดชนิดส pastic เท่ากับ 2.97 (SEM = 0.39) เทียบกับค่าเฉลี่ยของคะแนนความผิดพลาดในประโยคเดียวกันในกลุ่มควบคุมเท่ากับ 0.78 (SEM = 0.17) นั้นพบว่ามีความแตกต่างกันอย่างมี

นัยสำคัญทางสถิติ ($t = -5.22$, $p\text{-value} < 0.001$) ระหว่างผู้เข้าร่วมวิจัยทั้งสองกลุ่ม ค่าเฉลี่ยของคะแนนความผิดพลาดของพยางค์ (error score of syllable) ในประโยค ‘ปุม้าวิ่งไปมาบนใบไม้’ (BP) ในกลุ่มผู้ป่วยพูดไม่ชัดชนิดส pastic เท่ากับ 1.62 (SEM = 0.31) เทียบกับค่าเฉลี่ยของคะแนนความผิดพลาดในประโยคเดียวกันในกลุ่มควบคุมเท่ากับ 0.59 (SEM = 0.09) นั้นพบว่ามี ความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ($t = -3.24$, $p\text{-value} < 0.01$) ค่าเฉลี่ยของคะแนนความผิดพลาดของพยางค์ (error score of syllable) ในประโยค ‘อีภาคอยคาบงูคาบไก่’ (VAGP) ในกลุ่มผู้ป่วยพูดไม่ชัดชนิดส pastic เท่ากับ 1.78 (SEM = 0.28) เทียบกับค่าเฉลี่ยของคะแนนความผิดพลาดในประโยคเดียวกันในกลุ่มควบคุมเท่ากับ 0.70 (SEM = 0.16) นั้นพบว่ามี ความแตกต่างอย่างมีนัยสำคัญทางสถิติ ($t = -3.49$, $p\text{-value} < 0.001$) ค่าเฉลี่ยของคะแนนความผิดพลาดของพยางค์ (error score of syllable) ในประโยค ‘เพียงแค่ฝนตกลงที่หน้าต่างในบางครา’ (MP) ในกลุ่มผู้ป่วยพูดไม่ชัดชนิดส pastic เท่ากับ 1.97 (SEM = 0.37) เทียบกับค่าเฉลี่ยของคะแนนความผิดพลาดในประโยคเดียวกันในกลุ่มควบคุมเท่ากับ 0.91 (SEM = 0.08) นั้นพบว่ามี ความแตกต่างอย่างมีนัยสำคัญทางสถิติ ($t = -2.10$, $p\text{-value} < 0.05$) (ตารางที่ 5, รูปภาพที่ 11)

Mean error \pm SEM	Whisper				Apple Siri™			
	LAP	BP	VAGP	MP	LAP	BP	VAGP	MP
Spastic dysarthria	2.97 \pm 0.39	1.62 \pm 0.31	1.78 \pm 0.28	1.97 \pm 0.37	3.68 \pm 0.45	2.00 \pm 0.42	3.22 \pm 0.37	2.16 \pm 0.52
Control	0.78 \pm 0.17	0.59 \pm 0.09	0.70 \pm 0.16	0.91 \pm 0.08	0.49 \pm 0.17	0.30 \pm 0.11	1.62 \pm 0.19	0.27 \pm 0.10

ตารางที่ 5 แสดงค่าเฉลี่ยของคะแนนความผิดพลาดของพยางค์ (mean error score of syllable) ของทั้ง 4 ประโยคที่ถูกวิเคราะห์และสร้างขึ้นด้วยเครื่องมือ Whisper-Thai-Demo และ Apple Siri™



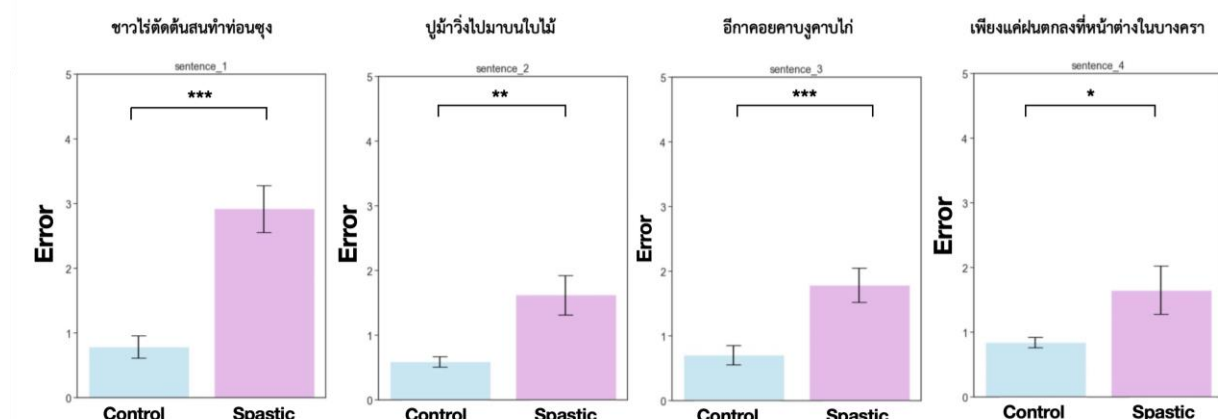
รูปภาพที่ 11 แสดงค่าเฉลี่ยของคะแนนความผิดพลาดของพยางค์ (mean error score of syllable) ของทั้ง 4 ประโยคที่ถูกวิเคราะห์และสร้างขึ้นด้วยเครื่องมือ Apple Siri™ และความแตกต่างอย่างมีนัยสำคัญทางสถิติระหว่างกลุ่มผู้ป่วยและกลุ่มควบคุม

โดยแกนตั้งคือค่าเฉลี่ยของคะแนนความผิดพลาดของพยางค์ และแกนนอนคือกลุ่มของผู้เข้าร่วมวิจัย

* แทน $p\text{-value} < 0.05$, ** แทน $p\text{-value} < 0.01$, *** แทน $p\text{-value} < 0.001$

เมื่อวิเคราะห์ด้วยเครื่องมือ Apple Siri™ พบว่าค่าเฉลี่ยของคะแนนความผิดพลาดของพยางค์ (error score of syllable) ในประโยค ‘ชาวไร่ตัดต้นสนทำท่อนซุง’ (LAP) ในกลุ่มผู้ป่วยพูดไม่ชัดชนิดสປาสຕิกเท่ากับ 3.68 (SEM = 0.45) เทียบกับค่าเฉลี่ยของคะแนนความผิดพลาดในประโยคเดียวกันในกลุ่มควบคุมเท่ากับ 0.49 (SEM = 0.17) นั้นพบว่ามี ความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ($t = -6.74$, $p\text{-value} < 0.001$) ระหว่างผู้เข้าร่วมวิจัยทั้งสองกลุ่ม ค่าเฉลี่ยของคะแนนความผิดพลาดของพยางค์ (error score of syllable) ในประโยค ‘ปทุมมาวิ่งไปมาบนใบไม้’ (BP) ในกลุ่มผู้ป่วยพูดไม่ชัดชนิดสປาสຕิกเท่ากับ 2.00 (SEM = 0.42) เทียบกับค่าเฉลี่ยของคะแนนความผิดพลาดในประโยคเดียวกันในกลุ่มควบคุมเท่ากับ 0.30 (SEM = 0.11) นั้นพบว่ามี ความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ($t = -4.01$, $p\text{-value} < 0.001$) ค่าเฉลี่ยของคะแนนความผิดพลาดของพยางค์ (error score of syllable) ในประโยค ‘อีกาคอยคาบงูคาบไก่’ (VAGP) ในกลุ่มผู้ป่วยพูดไม่ชัดชนิดสປาสຕิกเท่ากับ 3.22 (SEM = 0.37) เทียบกับค่าเฉลี่ยของคะแนนความผิดพลาดในประโยคเดียวกันในกลุ่มควบคุมเท่ากับ 1.62 (SEM = 0.19) นั้นพบว่ามี ความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ($t = -4.01$, $p\text{-value} < 0.001$) ค่าเฉลี่ยของคะแนนความผิดพลาดของพยางค์ (error score of syllable) ใน ‘เพียงแค่ฝนตกลงที่หน้าต่างในบางครา’ (MP) ในกลุ่มผู้ป่วยพูดไม่ชัด

ชนิดสเปคตริกเท่ากับ 2.16 (SEM = 0.52) เทียบกับค่าเฉลี่ยของคะแนนความผิดพลาดในประโยคเดียวกันในกลุ่มควบคุมเท่ากับ 0.27 (SEM = 0.10) นั้นพบที่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ($t = -3.77$, $p\text{-value} < 0.001$) (ตารางที่ 5, รูปภาพที่ 12)

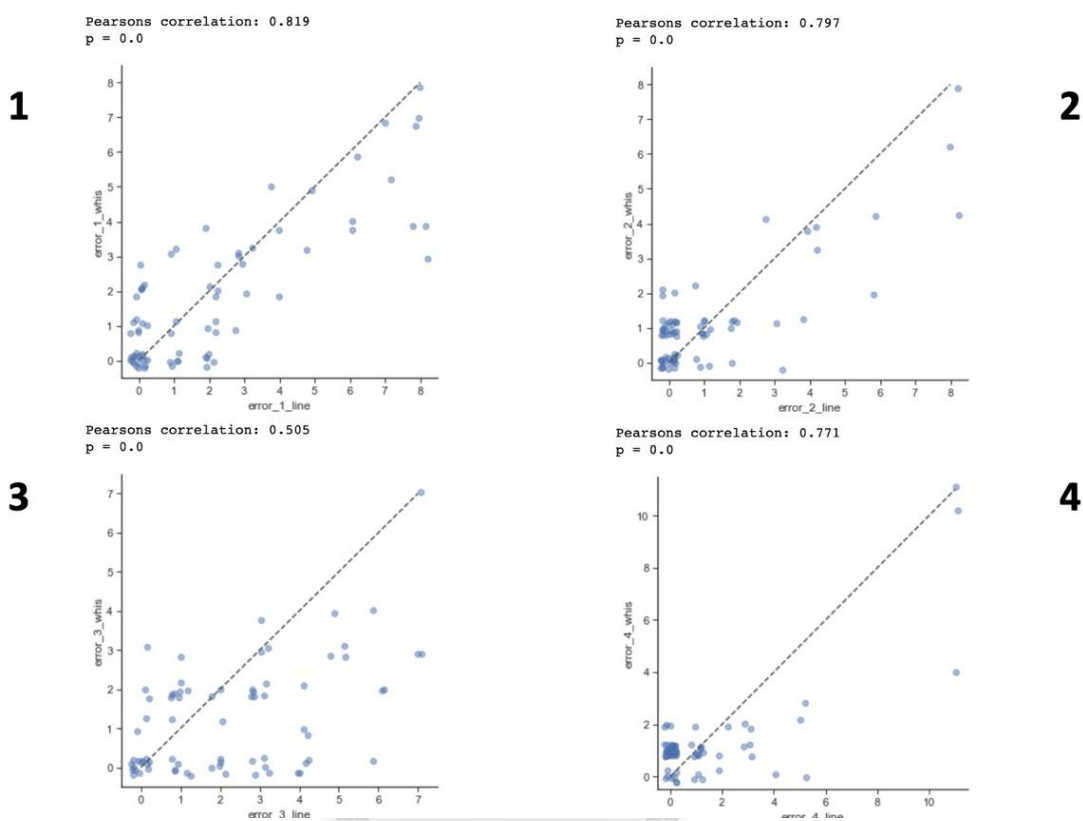


รูปภาพที่ 12 แสดงค่าเฉลี่ยของคะแนนความผิดพลาดของพยางค์ (mean error score of syllable) ของทั้ง 4 ประโยคที่ถูกวิเคราะห์และสร้างขึ้นด้วยเครื่องมือ Whisper-Thai-Demo และความแตกต่างอย่างมีนัยสำคัญทางสถิติระหว่างกลุ่มผู้ป่วยและกลุ่มควบคุม

การศึกษาความสัมพันธ์ของคะแนนความผิดพลาดของพยางค์ในแต่ละประโยค (error score of syllable) ของผู้ป่วยคนเดียวกันระหว่างเครื่องมือการวินิจฉัยทั้งสอง (Whisper-Thai-Demo และ Apple Siri™) โดยการวิเคราะห์ทางสถิติสหสัมพันธ์แบบเพียร์สัน (Pearson correlation coefficient) พบว่าเครื่องมือทั้งสองมีคะแนนที่สัมพันธ์กันอย่างมีนัยสำคัญทางสถิติ ($p\text{-value} < 0.001$) โดยประโยค ‘ชาวไร่ตัดต้นสนทำท่อนซุง’ (LAP) มีค่าสัมประสิทธิ์สหสัมพันธ์ (r) เท่ากับ 0.819 ประโยค ‘ปุม้าวิ่งไปมาบนใบไม้’ (BP) มีค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ 0.797 ประโยค ‘อีกาคอยคาบงูคาบไก่’ (VAGP) มีค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ 0.505 ประโยค ‘เพียงแค่มองตลงที่หน้าต่างในบางครา’ (MP) มีค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ 0.771 (รูปภาพที่ 13)

ตารางที่ 6 แสดงร้อยละของค่าเฉลี่ยคะแนนความผิดพลาดของพยางค์ในแต่ละประโยค (error score of syllable) หรือร้อยละของสัดส่วนของจำนวนพยางค์ที่ผิดในประโยค (incorrect syllable) ต่อจำนวนพยางค์ทั้งหมดของประโยคนั้นเปรียบเทียบกันระหว่างผู้ป่วยพูดไม่ชัดชนิดสเป

สติและกลุ่มควบคุม ซึ่งจะสามารถเปรียบเทียบความแตกต่างของจำนวนพยางค์ที่ผิดในแต่ละประโยคที่มีจำนวนพยางค์ไม่เท่ากันได้



รูปภาพที่ 13 แสดงความสัมพันธ์ของคะแนนความผิดพลาดของพยางค์ในแต่ละประโยคของผู้ป่วยคนเดียวกันระหว่างเครื่องมือ Whisper-Thai-Demo และ Apple Siri™ โดยการวิเคราะห์ทางสถิติสัมพันธ์แบบเพียร์สัน (Pearson correlation coefficient)

โดยหมายเลข 1 คือประโยค ‘ชาวไร่ตัดต้นสนทำท่อนซุง’ หมายเลข 2 คือประโยค ‘ปทุมำวิ่งไปมาบนไผ่ไม้’ หมายเลข 3 คือประโยค ‘อีกาคอยคาบงูคาบไก่’ และหมายเลข 4 คือประโยค ‘เพียงแค่ฝนตกลงที่หน้าต่างในบางครา’

Mean percentage of error \pm SEM	Whisper				Apple Siri™			
	LAP	BP	VAGP	MP	LAP	BP	VAGP	MP
Spastic dysarthria	37.13 \pm 4.71	20.25 \pm 3.83	25.43 \pm 3.88	17.91 \pm 3.29	46.00 \pm 5.52	25.00 \pm 5.13	46.00 \pm 5.08	19.64 \pm 4.59
Control	9.75 \pm 2.06	7.38 \pm 1.023	10.00 \pm 2.13	8.27 \pm 0.69	6.13 \pm 1.97	3.75 \pm 1.27	23.14 \pm 2.67	2.46 \pm 0.91

ตารางที่ 6 แสดงร้อยละของค่าเฉลี่ยคะแนนความผิดพลาดของพยางค์ของทั้ง 4 ประโยคที่ถูกวิเคราะห์และสร้างขึ้นด้วยเครื่องมือ Whisper-Thai-Demo และ Apple Siri™

ผลการวินิจฉัยอาการพูดไม่ชัดชนิดสປาสติกโดยผู้ฟังและวินิจฉัยอาการประกอบด้วยประสาทแพทย์ทั้งหมด 10 คน เป็นอาจารย์แพทย์จำนวน 5 คน แพทย์ประจำบ้านต่อยอดสาขาหลอดเลือดสมองจำนวน 2 คน และแพทย์ประจำบ้านชั้นปีที่สามจำนวน 3 คน พบว่าความสามารถในการวินิจฉัยที่แตกต่างกันมากมีค่าเฉลี่ยความแม่นยำในการวินิจฉัยอาการพูดไม่ชัดชนิดสປาสติก (accuracy) อยู่ที่ร้อยละ 72.86 (SD = 7.38) ค่าเฉลี่ยความไวของการวินิจฉัย (sensitivity) อยู่ที่ร้อยละ 48.57 (SD = 15.94) ค่าเฉลี่ยความจำเพาะของการวินิจฉัย (specificity) อยู่ที่ 97.14 (SD = 3.30) ค่าเฉลี่ยค่าทำนายผลบวก (positive predictive value/PPV) อยู่ที่ 95.20 (SD = 5.75) ค่าเฉลี่ยค่าทำนายผลลบ (negative predictive value/NPV) อยู่ที่ 65.46 (SD = 7.62) (ตารางที่ 7) โดยรายละเอียดการวินิจฉัยอาการพูดไม่ชัดชนิดสປาสติกของประสาทแพทย์แต่ละคนถูกแสดงไว้ในรูปแบบของตารางการณ์จร (2x2 contingency table) ในภาคผนวก (ตารางที่ 8)

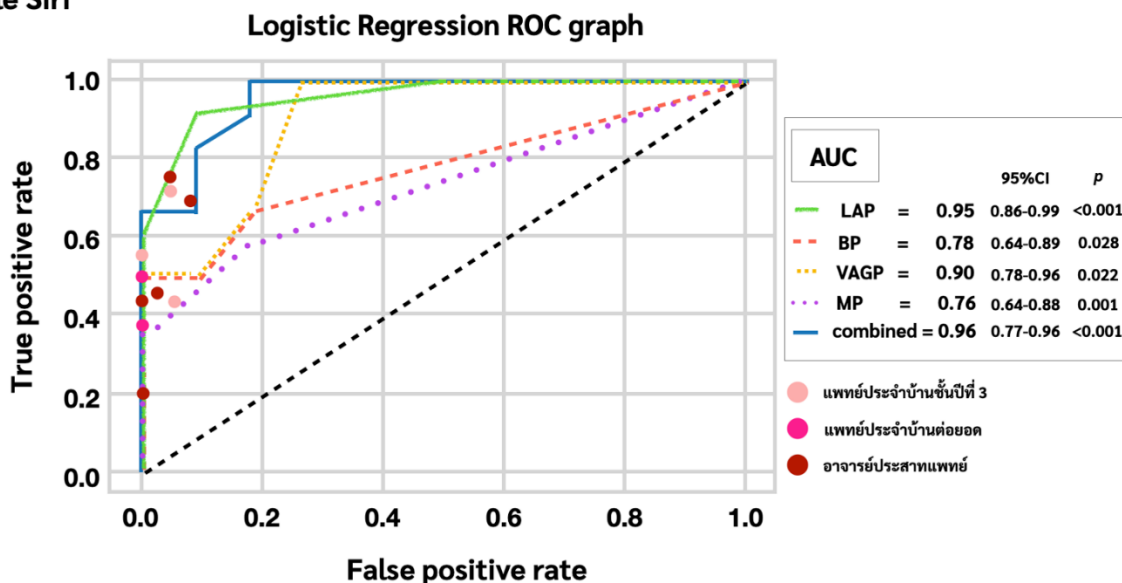
ข้อสังเกตที่น่าสนใจคือความไวในการวินิจฉัยโรคโดยเฉลี่ยค่อนข้างต่ำ (น้อยกว่าร้อยละ 50) และมีความแตกต่างกันชัดเจนระหว่างประสาทแพทย์แต่ละคนตั้งแต่ร้อยละ 20 ถึงร้อยละ 72.29 และสังเกตได้จากค่าเบี่ยงเบนมาตรฐานที่สูง (SD = 15.94) ในขณะที่ค่าเฉลี่ยความจำเพาะและค่าทำนายผลบวกค่อนข้างสูงต่างจากค่าความไวอย่างชัดเจนและมีความแตกต่างกันไม่มากในประสาทแพทย์แต่ละคน อย่างไรก็ตามเนื่องจากจำนวนประสาทแพทย์ในแต่ละกลุ่มมีน้อยและไม่เท่ากันจึงไม่สามารถเปรียบเทียบความสามารถในการวินิจฉัยของประสาทแพทย์ในแต่ละกลุ่มเชิงสถิติได้ แต่จะสังเกตได้ว่าไม่ได้มีความแตกต่างกันชัดเจนในระหว่างกลุ่ม ทั้งนี้ความแตกต่างในความสามารถในการวินิจฉัยระหว่างกลุ่มไม่ใช่วัตถุประสงค์ของการวิจัยครั้งนี้

ลำดับที่	อาจารย์ประสาทแพทย์					แพทย์ประจำบ้านตอยอด		แพทย์ประจำบ้านชั้นปีที่ 3			Mean	SD
	1	2	3	4	5	1	2	1	2	3		
Accuracy	84.29	71.43	71.43	68.57	60.00	74.29	68.57	82.86	78.57	68.57	72.86	7.38
Sensitivity	74.29	45.71	42.86	45.71	20.00	48.57	37.14	71.43	57.14	42.86	48.57	15.94
Specificity	94.29	97.14	100.00	91.43	100.00	100.00	100.00	94.29	100.00	94.29	97.14	3.30
PPV	92.86	94.12	100.00	84.21	100.00	100.00	100.00	92.59	100.00	88.24	95.20	5.75
NPV	78.57	64.15	57.14	62.75	55.56	66.04	61.40	76.74	70.00	62.26	65.46	7.62

ตารางที่ 7 แสดงความสามารถในการวินิจฉัยอาการพูดไม่ชัดชนิดสປาสตึกโดยประสาทแพทย์ทั้งสามกลุ่ม รวมถึงค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน

จากการวิเคราะห์สมการ Logistic regression analysis เพื่อทำนายโอกาสของการมีอาการพูดไม่ชัดชนิดสປาสตึกจากคะแนนความผิดพลาดของพยางค์ (error score of syllable) ในแต่ละประโยคและนำเสนอการดังกล่าวมาคำนวณและสร้างเป็น Receiver operating characteristic graph (ROC graph) เพื่อคำนวณหา Area Under the Curve (AUC) ที่จะเป็นตัวแทนของประสิทธิภาพในการวินิจฉัยอาการพูดไม่ชัดชนิดสປาสตึกของเครื่องมือการวิเคราะห์และสร้างประโยคทั้งสองเครื่องมือพบว่าเมื่อพิจารณาทีละประโยคที่ถูกสร้างขึ้นจากเครื่องมือ Apple Siri™ AUC ของการวินิจฉัยอาการพูดไม่ชัดจากประโยค ‘ชาวไร่ตัดต้นสนทำท่อนซุง’ หรือประโยคที่เกิดจากการเคลื่อนไหวกลิ้มเนื้อลิ้นกระแทกเพดานปาก (lingual-alveolar placement/LAP) นั้นสูงที่สุดเท่ากับ 0.95 ตามด้วย AUC ของประโยค ‘อีก้าคอยคาบุงคาบโก’ ที่เกิดจากการเคลื่อนไหวกลิ้มเนื้อคอหอยและกล่องเสียง (velopharyngeal and glottis placement/VAGP) เท่ากับ 0.90 ประโยค ‘ปุม้าวิ่งไปมาบนใบไม้’ (bilabial placement/BP) มี AUC เท่ากับ 0.78 และประโยค ‘เพียงแค่ฝนตกลงที่หน้าต่างในบางครา’ ที่ไม่ได้เน้นการเคลื่อนไหวกลิ้มเนื้อที่ส่วนใด (mixed placements/MP) มี AUC ต่ำที่สุดเท่ากับ 0.73 นอกจากนี้ยังได้ทำการวิเคราะห์สมการ Multivariate logistic regression analysis เพื่อพิจารณาตัวแปรคะแนนความผิดพลาดของพยางค์ของทั้ง 4 ประโยค (4 ตัวแปร) ข้างต้นในการทำนายโอกาสการมีอาการพูดไม่ชัดชนิดสປาสตึกและสร้างเป็น ROC graph พบว่ามี AUC เท่ากับ 0.96 ใกล้เคียงกับ AUC จากประโยค ‘ชาวไร่ตัดต้นสนทำท่อนซุง’ (รูปภาพที่ 14)

Apple Siri™



รูปภาพที่ 14 แสดง Receiver operating characteristic graph (ROC graph) พร้อมทั้ง AUC ของคะแนนในแต่ละประโยคและรวมทุกประโยคจากการวิเคราะห์ด้วย Apple Siri™ พร้อมทั้งจุดแสดงความสามารถในการวินิจฉัยอาการพูดไม่ชัดโดยประสาทแพทย์ 10 คน โดยที่

LAP คือ Lingual-Alveolar Placement หรือประโยค ‘ชาวไร่ตัดต้นสนทำท่อนซุง’

BP คือ Bilabial Placement หรือประโยค ‘ปุม้าวิ่งไปมาบนใบไม้’

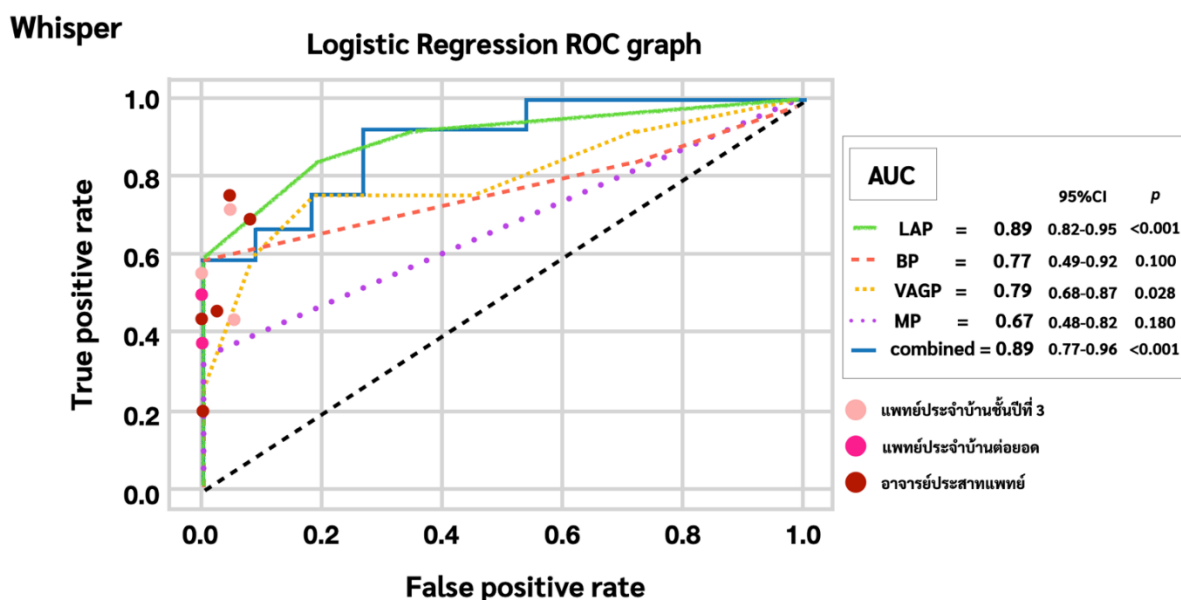
VAGP คือ Velopharyngeal and Glottis Placement หรือประโยค ‘อีกาคอยคาบงูคาบไก่’

MP คือ Mixed Placement หรือประโยค ‘เพียงแค่นกตกลงที่หน้าต่างในบางครา’

COMBINED คือ นำทุกประโยคมาวิเคราะห์ร่วมกันโดย Multivariate logistic regression analysis จากนั้นสร้างเป็น ROC graph และคำนวณ AUC

เมื่อพิจารณาทีละประโยคที่ถูกสร้างขึ้นจากเครื่องมือ Whisper-Thai-Demo พบว่า AUC ของการวินิจฉัยอาการพูดไม่ชัดจากประโยค ‘ชาวไร่ตัดต้นสนทำท่อนซุง’ (lingual-alveolar placement/LAP) นั้นสูงที่สุดเช่นกันเท่ากับ 0.89 ตามด้วย AUC ของประโยค ‘อีกาคอยคาบงูคาบไก่’ (velopharyngeal and glottis placement/VAGP) เท่ากับ 0.79 ตามด้วยประโยค ‘ปุม้าวิ่งไปมาบนใบไม้’ (bilabial placement/BP) ที่มี AUC เท่ากับ 0.77 และประโยค ‘เพียงแค่นกตกลงที่หน้าต่างในบางครา’ ที่ไม่ได้เน้นการเคลื่อนไหวของกล้ามเนื้อที่ส่วนใด (mixed placements/MP) มี

AUC เท่ากับ 0.67 เมื่อพิจารณาตัวแปรคะแนนความผิดพลาดของพยางค์ของทั้ง 4 ประโยค และสร้างเป็น ROC graph พบว่ามี AUC เท่ากับ 0.89 เท่ากับ AUC จากประโยค ‘ชาวไร่ตัดต้นสนทำท่อนซุง’ (รูปภาพที่ 15)



รูปภาพที่ 15 แสดง Receiver operating characteristic graph (ROC graph) พร้อมทั้ง AUC ของคะแนนในแต่ละประโยคและรวมทุกประโยคจากการวิเคราะห์ด้วย Whisper-Thai-Demo พร้อมทั้งจุดแสดงความสามารถในการวินิจฉัยอาการพูดไม่ชัดโดยประสาทแพทย์ 10 คน

นอกจากนี้ผู้วิจัยยังได้ทำการพิจารณาตัวบ่งชี้ในการประเมินความถูกต้องของวิธีการวินิจฉัยอาการพูดไม่ชัดหรือการหาจุดตัดที่เหมาะสม (optimal cut-off point/threshold) ในการวินิจฉัยว่ามีอาการพูดไม่ชัดชนิดส pastic หรือไม่ สำหรับการวิเคราะห์ด้วยเครื่องมือ Apple Siri™ พบว่าเมื่อพิจารณารวมทั้ง 4 ประโยค (combined) นั้นความไวอยู่ที่ร้อยละ 80 ความจำเพาะอยู่ที่ร้อยละ 86 และจุดตัดที่เหมาะสม (คะแนนความผิดพลาดของพยางค์ในประโยค) อยู่ที่ 0.5 และเมื่อพิจารณาแยกทีละประโยคพบว่าประโยคที่มี AUC สูงที่สุดสองประโยคได้แก่ ประโยค ‘ชาวไร่ตัดต้นสนทำท่อนซุง’ (LAP) มีความไวอยู่ที่ร้อยละ 89 ความจำเพาะอยู่ที่ร้อยละ 86 และจุดตัดที่เหมาะสมอยู่ที่ 1 ในขณะที่ประโยค ‘อีกาคอยคาบงูคาบไก่’ (VAGP) มีความไวอยู่ที่ร้อยละ 64 ความจำเพาะอยู่ที่ร้อยละ 79 และจุดตัดที่เหมาะสมอยู่ที่ 2 (ภาคผนวก, รูปภาพที่ 17)

จุดตัดที่เหมาะสม (optimal cut-off point/threshold) สำหรับการวิเคราะห์ด้วยเครื่องมือ Whisper-Thai-Demo พบว่าเมื่อพิจารณารวมทั้ง 4 ประโยค (combined) นั้นความไวอยู่ที่ร้อยละ 65 ความจำเพาะอยู่ที่ร้อยละ 82 และจุดตัดที่เหมาะสมอยู่ที่ 0.5 และเมื่อพิจารณาแยกทีละประโยค พบว่าประโยคที่มี AUC สูงที่สุดสองประโยคได้แก่ ประโยค ‘ชาวไร่ตัดต้นสนทำท่อนซุง’ (LAP) มีความไวอยู่ที่ร้อยละ 81 ความจำเพาะอยู่ที่ร้อยละ 82 และจุดตัดที่เหมาะสมอยู่ที่ 2 ในขณะที่ ประโยค ‘อึกาคอยคาบงูคาบไก่’ (VAGP) มีความไวอยู่ที่ร้อยละ 63 ความจำเพาะอยู่ที่ร้อยละ 81 และจุดตัดที่เหมาะสมอยู่ที่ 1 (ภาคผนวก, รูปภาพที่ 18)

เมื่อนำความสามารถในการวินิจฉัยของประสาทแพทย์ทั้งสามกลุ่ม (10 คน) มาเปรียบเทียบกับเชิงพรรณนาด้วยการสังเกตในกราฟ ROC พบว่ามีเพียง 3 คนที่มีความไวและความจำเพาะในการวินิจฉัยอาการพูดไม่ชัดชนิดส pastic ที่ดีใกล้เคียงกับจุดตัดที่เหมาะสม (optimal point for ROC graph) ของเครื่องมือทั้ง Apple Siri™ และ Whisper-Thai-Demo อย่างไรก็ตามความไวจากการวินิจฉัยอาการพูดไม่ชัดโดยประสาทแพทย์ทุกคนจำกัดอยู่ที่ไม่เกินร้อยละ 80 ในขณะที่การวิเคราะห์ด้วยเครื่องมือมีความไวเกินร้อยละ 80

บทที่ 5

อภิปรายผล สรุปผลการวิจัย และข้อเสนอแนะ

5.1 อภิปรายผล

ผลการศึกษาวินิจฉัยครั้งนี้แสดงให้เห็นอย่างชัดเจนว่าเครื่องมือวิเคราะห์เสียงพูดสร้างเป็นประโยคหรือ ASR สามารถนำมาใช้แยกแยะเสียงพูดไม่ชัดชนิดสປາສຕິคได้อย่างแม่นยำและถูกต้อง โดยไม่ขึ้นกับโมเดลหรืออัลกอริทึมของเครื่องมือ ซึ่งได้รับการพิสูจน์ด้วยค่า AUC ที่มากกว่า 0.75 แสดงถึงประสิทธิภาพในการแยกแยะประโยคพูดไม่ชัดเมื่อพิจารณาทุกประโยคพร้อมกัน (combined) ได้เป็นอย่างดีและคล้ายคลึงกันในทั้งสองเครื่องมือ (maximum $AUC_{combined} = 0.96$ โดย Apple Siri™ และ maximum $AUC_{combined} = 0.89$ โดย Whisper) โดยเฉพาะการแยกแยะเสียงพูดไม่ชัดที่มีความรุนแรงน้อยที่ประสาทแพทย์มักมีความลำบากในการวินิจฉัยและมีความแม่นยำในการวินิจฉัยต่ำ กล่าวคือการนำเทคโนโลยีการวิเคราะห์คลื่นเสียงพูดผ่านโมเดลที่ก้าวหน้าและซับซ้อน อาทิ deep learning และ natural language processing ที่มีความแม่นยำในการตรวจจับและสร้างประโยคเสียงพูดปกตินั้น อาจนำมาใช้ตรวจจับเสียงพูดไม่ชัดชนิดสປາສຕິคได้ด้วย

ข้อสรุปที่สำคัญอีกประการที่ได้จากการศึกษานี้คือความแตกต่างของ AUC หรือประสิทธิภาพการวินิจฉัยอาการพูดไม่ชัดชนิดสປາສຕິคระหว่างประโยคที่มีพยัญชนะต้นเกิดจากการเคลื่อนไหวของกล้ามเนื้อต่างกลุ่มกันและรูปแบบการเคลื่อนไหวที่ต่างกัน เครื่องมือช่วยวินิจฉัยเดียวกัน โดยกล้ามเนื้อที่สัมผัสกับเพดานปากที่ทำให้เกิดเสียงพยัญชนะต้น ‘ร’/r/ ‘ส’/s/ ‘ท’/th/ ‘ช’/ch/ ‘ต’/t/ ในประโยค ‘ชาวไร่ตัดต้นสนทำท่อนซุง’ นั้นมีการเคลื่อนไหวที่ผิดปกติมากที่สุดทำให้เกิดความผิดพลาดในการสร้างประโยคจากเสียงพูดมากที่สุด และมีความแตกต่างจากประโยคพูดเสียงปกติมากที่สุด ประกอบกับการที่ประโยคดังกล่าวมีพยัญชนะที่เกิดขยับแบบสั้นกล้ามเนื้อ (trill or flap manner) เช่น ‘ร’/r/ และพยัญชนะที่เป็นเสียงเสียดแทรก (fricative) เช่น ‘ส’/s/ อยู่มากกว่าประโยคอื่นๆ ทำให้ความสามารถในการแยกแยะอาการพูดไม่ชัดชนิดสປາສຕິคด้วยประโยคพูดที่มีความแม่นยำที่สุดสอดคล้องกับผลการศึกษาในอดีตที่พบว่ากล้ามเนื้อกลุ่มนี้ได้รับผลกระทบมากที่สุดในผู้ป่วยพูดไม่ชัดประเภทนี้

อย่างไรก็ตามกล้ามเนื้อคอคอหอย (Velopharyngeal muscle) ก็เป็นกล้ามเนื้อที่กลุ่มได้รับผลกระทบเช่นกัน สอดคล้องกับการพบปัญหาการกลืนในผู้ป่วยภาวะสมองขาดเลือดเฉียบพลันที่มีอาการพูดไม่ชัดชนิดสປາສຕິคร่วมด้วย ซึ่งผลการศึกษานี้ก็บ่งชี้ว่ากล้ามเนื้อส่วนดังกล่าวที่ทำให้

เกิดเสียงพยัญชนะต้นในประโยค ‘อีกาคอยคาบงคาบไก่’ ได้รับผลกระทบรองลงมาจากกล้ำมเนื้อลิ้น กระทบเพดานปากจากค่า AUC ที่ต่ำกว่า AUC จากประโยค ‘ชาวไร่ตัดต้นสนทำท่อนซุง’ เพียงเล็กน้อย (AUC ของ VAGP = 0.90 เทียบกับ AUC ของ LAP = 0.95 โดย Apple Siri™, AUC ของ VAGP = 0.79 เทียบกับ AUC ของ LAP = 0.89 โดย Whisper) ในขณะที่กล้ำมเนื้อสร้างพยัญชนะ ส่วนไบหน้าที่ทำหน้าที่ขยับริมฝีปากเกิดเป็นประโยค ‘ปุม้าวิ่งไปมาบนใบไม้’ ได้รับผลกระทบน้อยที่สุด

ความสามารถในการแยกแยะเสียงพูดไม่ชัดจากประโยคที่ไม่ได้เน้นการออกเสียงพยัญชนะต้น ส่วนใดๆที่เกิดจากการผสมผสานของกล้ำมเนื้อทุกส่วนและรูปแบบการออกเสียงที่หลากหลาย (mixed placements and manners) หรือประโยค ‘เพียงแค่ฝนตกลงที่หน้าต่างในบางครา’ นั้นต่ำที่สุด (AUC ของ MP = 0.73 โดย Apple Siri™ และ AUC ของ MP = 0.67 โดย Whisper) เน้นย้ำให้เห็นถึงความสำคัญของการเลือกกล้ำมเนื้อที่เคลื่อนไหวผิดปกติมากที่สุดในการออกเสียงเพื่อเพิ่มประสิทธิภาพการแยกแยะเสียงพูดไม่ชัด

นอกจากนี้การนำคะแนนความผิดพลาดของพยางค์จากทุกประโยคมาใช้แยกแยะเสียงพูดไม่ชัดชนิดสเปคตริกนั้นพบว่าความถูกต้องในการวินิจฉัยไม่ได้แตกต่างอย่างชัดเจนจากการแยกแยะเสียงพูดไม่ชัดจากประโยค ‘ชาวไร่ตัดต้นสนทำท่อนซุง’ เพียงประโยคเดียว (AUC ของ COMBINED = 0.96 เทียบกับ AUC ของ LAP = 0.95 โดย Apple Siri™, AUC ของ COMBINED = 0.89 เทียบเท่ากับ AUC ของ LAP = 0.89 โดย Whisper) บอกเป็นนัยว่ากล้ำมเนื้อลิ้นที่มีปัญหาการเคลื่อนไหวมากที่สุดมีอิทธิพลต่อความสามารถในการวินิจฉัยอาการพูดไม่ชัดชนิดสเปคตริกมากที่สุดในขณะที่ปัญหาการเคลื่อนไหวของกล้ำมเนื้อกลุ่มอื่นๆอาจไม่ได้มีผลต่อการแยกแยะเสียงพูดไม่ชัดออกจากเสียงพูดปกติมากนัก

ข้อสังเกตที่สำคัญอีกประการจากผลการศึกษานี้คือความสามารถในการวินิจฉัยอาการพูดไม่ชัดชนิดสเปคตริกของทั้งสองเครื่องมือนี้เป็นไปในทางเดียวกัน กล่าวคือมีความสอดคล้องของลำดับของความสามารถในการวินิจฉัยอาการพูดไม่ชัดเหมือนกัน (AUC สูงจากมากไปน้อยในประโยค LAP, VAGP, BP และ MP ตามลำดับ) หมายความว่า การเคลื่อนไหวของกล้ำมเนื้อที่ได้รับผลกระทบมากที่สุดมีผลต่อความถูกต้องแม่นยำในการวินิจฉัยอาการพูดไม่ชัดโดยที่ไม่ได้ขึ้นอยู่กับเครื่องมือวิเคราะห์คลื่นเสียงที่มีความถูกต้องแม่นยำสูงที่ใช้ในการวินิจฉัย

ความสามารถในการวินิจฉัยอาการหูตึงไม่ชัดชนิดประสาทที่ความรุนแรงน้อยโดยประสาทแพทย์ทั้ง 10 คนแสดงให้เห็นว่าความถูกต้องแม่นยำอยู่ที่ร้อยละ 72.86 โดยเฉลี่ยเท่านั้น ซึ่งไม่เพียงพอต่อการวินิจฉัยทางคลินิก และพบว่าไม่มีความแตกต่างกันชัดเจนระหว่างกลุ่มประสาทแพทย์ที่มีประสบการณ์สูงและขาดประสบการณ์ โดยที่การวินิจฉัยโดยประสาทแพทย์จะมีลักษณะเด่นอยู่ที่ความจำเพาะและค่าทำนายผลบวกที่สูง หรือผลบวกลวงที่ต่ำมาก (ค่าเฉลี่ยความจำเพาะในการวินิจฉัยเท่ากับ 97.14 และค่าการทำนายผลบวกอยู่ที่ 95.20) ในขณะที่มีความไวในการวินิจฉัยที่ค่อนข้างต่ำและไม่แน่นอน ตั้งแต่ร้อยละ 20 ไปจนถึงร้อยละ 74.29 (ค่าเฉลี่ยความไวในการวินิจฉัยอยู่ที่ 48.57) แสดงให้เห็นว่าประสาทแพทย์มักวินิจฉัยอาการหูตึงไม่ชัดเฉพาะเสียงหูตึงไม่ชัดชนิดประสาทที่ผิดปกติแน่นอนและมีความรุนแรงมากพอเท่านั้น และอาจบอกเป็นนัยได้ว่าการจะวินิจฉัยอาการหูตึงไม่ชัดชนิดประสาทนั้นไม่ได้ขึ้นกับประสบการณ์ทางคลินิกในการแยกแยะเสียงหูตึงไม่ชัด นอกจากนี้ผลการศึกษาความแม่นยำในการวินิจฉัยเสียงหูตึงไม่ชัดโดยประสาทแพทย์นี้ยังเน้นย้ำถึงความไม่เหมาะสมในการอาศัยการวิเคราะห์แยกแยะกลไกเสียงหูตึงไม่ชัดจากการฟัง (auditory-perceptual assessment) เป็นการตรวจวินิจฉัยที่จะเป็นมาตรฐานทางคลินิก (Gold standard) อีกด้วย

ด้วยผลการศึกษาทั้งปวงที่ได้กล่าวไปทำให้สามารถสรุปได้ว่า ถึงแม้การใช้เครื่องมือวิเคราะห์เสียงที่ถูกพัฒนาขึ้นจากเทคโนโลยีใหม่ที่มีความแม่นยำสูงในการแปลงเสียงพูดเป็นประโยคเพียงอย่างเดียวนั้น จะสามารถแยกแยะเสียงหูตึงไม่ชัดชนิดประสาทได้ แต่อาจไม่มากพอต่อการวินิจฉัยอาการหูตึงไม่ชัดชนิดประสาททางคลินิก และการเพิ่มความสามารถในการวินิจฉัยแยกแยะเสียงหูตึงไม่ชัดด้วยปัจจัยความผิดปกติของการเคลื่อนไหวกล้ามเนื้อการออกเสียงที่ได้รับผลกระทบมากที่สุด ตามกลไกการควบคุมกล้ามเนื้อได้แก่ corticobulbar tract ที่เสียหายจากโรคในสมองส่วนกลาง ทำให้กล้ามเนื้อได้รับผลกระทบและมีผลต่อการออกเสียงพยัญชนะต้นอย่างชัดเจนนั้น จะยิ่งทำให้มีความแม่นยำในการวินิจฉัยเพิ่มมากขึ้นกว่าขีดจำกัดการวินิจฉัยโดยประสาทแพทย์ที่เป็นมาตรฐานการวินิจฉัยในปัจจุบันได้

ดังนั้นจึงเป็นที่ประจักษ์ว่าการผสมความรู้ทางพยาธิสรีรวิทยาคู่ขนานกับความรู้ทางศาสตร์เสียงหูตึงไม่ชัด ร่วมกับการใช้เทคโนโลยีเครื่องมือวิเคราะห์เสียงสมัยใหม่เป็นมาตรฐานวัด ‘จำนวน’ ของพยางค์ที่หูตึงไม่ชัดนี้ ทำให้สามารถวินิจฉัยอาการหูตึงไม่ชัดได้อย่างแม่นยำถูกต้องและนำไปใช้ได้จริงในทางคลินิก และที่สำคัญยังก่อให้เกิดความเข้าใจในอาการหูตึงไม่ชัดชนิดประสาทในแง่กลไกการเคลื่อนไหวของกล้ามเนื้อที่ได้รับผลกระทบมากขึ้น

5.2 สรุปผลการวิจัย

การศึกษานี้แสดงให้เห็นว่าเครื่องมือวิเคราะห์เสียงพูดแปลงเป็นประโยคข้อความ (Automatic speech recognition) จะมีความถูกต้องแม่นยำ มีความไวและความจำเพาะที่สูงมากพอที่จะนำมาใช้ในการวินิจฉัยอาการพูดไม่ชัดชนิดสປาสติกโดยเฉพาะกลุ่มที่มีความรุนแรงน้อยออกจากเสียงพูดปกติได้จริงในทางคลินิก เมื่อใช้แยกแยะประโยคพูดที่เกิดจากการเคลื่อนไหวของกล้ามเนื้อที่ผิดปกติมากที่สุดในกลุ่มโรคที่มีอาการพูดไม่ชัดชนิดสປาสติกตามหลักการพยาธิสรีรวิทยาได้แก่ กล้ามเนื้อลิ้นกระหนบเพดานปากที่มีผลต่อความผิดพลาดของเสียงพยัญชนะต้นและข้อความที่เครื่องมือสร้างขึ้น อีกทั้งเครื่องมือมีความแม่นยำในการวินิจฉัยอาการพูดไม่ชัดมากกว่าความสามารถของประสาทแพทย์อย่างชัดเจน นอกจากนี้เครื่องมือดังกล่าวยังมีความเที่ยงตรงในการวินิจฉัยโดยไม่ขึ้นกับโมเดลหรือกระบวนการวิเคราะห์ของเครื่องมือทำให้กระบวนการวินิจฉัยอาการพูดไม่ชัดชนิดสປาสติกในการศึกษานี้มีความน่าเชื่อถือและสามารถนำไปใช้ได้จริง

5.3 เปรียบเทียบกับการศึกษาก่อนหน้านี้

ยังไม่เคยมีการศึกษาประสิทธิภาพในการวินิจฉัยอาการพูดไม่ชัดชนิดใดๆรวมถึงชนิดสປาสติกด้วยเครื่องมือ ASR มาก่อนเนื่องจากในอดีตอาศัยการวิเคราะห์แยกแยะกลไกเสียงพูดไม่ชัดจากการฟัง (auditory-perceptual assessment) มาโดยตลอด และยังไม่เคยมีการศึกษาที่นำลักษณะเด่นทางคลินิกของการพูดไม่ชัดชนิดสປาสติกมาเพิ่มความสามารถในการวินิจฉัยอาการพูดไม่ชัดมาก่อน ทำให้ไม่สามารถเปรียบเทียบได้โดยตรง อย่างไรก็ตามการศึกษานี้มีรายละเอียดผลการวิเคราะห์บางส่วนที่สอดคล้องกับการศึกษาในอดีต ได้แก่ ผลการศึกษาขั้นต้นที่พบว่าร้อยละของความถูกต้องของข้อความที่เกิดจากเครื่องมือ ASR อยู่ที่ 60-80 โดยประมาณในกลุ่มผู้ป่วยมีอาการพูดไม่ชัด เทียบกับร้อยละ 90-100 ในกลุ่มผู้ที่มีเสียงพูดปกติคล้ายกับการศึกษาของ Calvo และคณะ (Calvo, Tropea et al. 2021) การศึกษาลักษณะเด่นทางคลินิกของอาการพูดไม่ชัดพบว่ากล้ามเนื้อปลายลิ้นกระหนบกับส่วนเพดานปาก (Ligual-alveolar placement) ถูกกระหนบมากที่สุดสอดคล้องกับการศึกษาของ Sriwimon et al. และคณะ (Manochiopinig, Thubthong et al. 2008) และความแม่นยำในการวินิจฉัยอาการพูดไม่ชัดโดยประสาทแพทย์ที่ต่ำเกินกว่าที่จะสามารถนำไปใช้จริงในคลินิก คล้ายกับการศึกษาโดย Van Der Graaff และ คณะ (Van der Graaff, Kuiper et al. 2009) เป็นต้น

5.4 ข้อดีของการศึกษานี้

การศึกษาเครื่องมือการวินิจฉัยนี้จัดว่าเป็นการศึกษาเชิงนวัตกรรมที่ได้ผลลัพธ์เป็นเครื่องมือและกระบวนการวินิจฉัยอาการพูดไม่ชัดชนิดสປาสติกที่สามารถนำไปใช้ได้จริงและทันทีเนื่องจากเป็นเครื่องมือที่แพร่หลายที่แพทย์ในเวชปฏิบัติส่วนมากมักมีครอบครองอยู่แล้ว นอกจากนี้ยังเข้าถึงได้โดยง่ายและพร้อมใช้งานได้อย่างไม่มีขั้นตอนที่ยุ่งยาก ทำให้เกิดประโยชน์กับแพทย์ทุกสาขาโดยไม่จำเป็นต้องเป็นประสาทแพทย์

การศึกษานี้เป็นการแยกแยะผู้ป่วยที่มีอาการพูดไม่ชัดที่มีความรุนแรงน้อยเป็นส่วนมากออกจากเสียงพูดปกติซึ่งเป็นความลำบากในการวินิจฉัยอาการพูดไม่ชัดของประสาทแพทย์ โดยที่มีผู้ป่วยความรุนแรงปานกลางถึงมากเป็นส่วนน้อยซึ่งประสาทแพทย์มักวินิจฉัยได้อย่างถูกต้องแม่นยำมากกว่า

การศึกษานี้พบลักษณะเด่นทางคลินิกของการพูดไม่ชัดชนิดสປาสติก คือการใช้พยัญชนะต้นที่เกิดจากการเคลื่อนไหวของลิ้นกระทบเพดานปากจะทำให้แยกแยะเสียงพูดไม่ชัดได้ดีขึ้น ซึ่งประสาทแพทย์สามารถนำไปใช้ในการตรวจร่างกายเพื่อการวินิจฉัยได้แม้จะไม่มีเครื่องมือ ASR ก็ตาม

การศึกษานี้เป็นการพิสูจน์แนวคิดว่าการนำเทคโนโลยีใกล้ตัวที่มีความก้าวหน้าแต่แพทย์อาจมองข้ามไปมาเป็นเครื่องมือเพื่อชดเชยข้อจำกัดการวินิจฉัยของแพทย์ได้ หากมีกระบวนการนำไปใช้ที่ถูกต้องดังเช่นการนำหลักการทางพยาธิสรีรวิทยาและสรีศาสตร์มาประกอบใช้ในขั้นตอนการวินิจฉัยในการศึกษา

สิ่งที่สำคัญมากที่ได้จากการศึกษานี้คือการทำให้การวินิจฉัยอาการพูดไม่ชัดมีความเป็นรูปธรรมและสามารถวัดเชิงปริมาณได้ (objective assessment) ซึ่งไม่เคยมีมาก่อน เนื่องจากในอดีตเป็นการวินิจฉัยจากการฟัง (auditory-perceptual assessment) ซึ่งเป็นนามธรรม (subjective) และไม่มีหลักเกณฑ์ที่แน่ชัด

5.5 ข้อดีของการศึกษานี้

ข้อจำกัดของการศึกษานี้คือเครื่องมือและกระบวนการวินิจฉัยอาการพูดไม่ชัดนี้ใช้แยกแยะอาการพูดไม่ชัดชนิดสປาสติกออกจากเสียงพูดปกติเท่านั้น ไม่สามารถใช้จำแนกประเภทของการพูดไม่ชัดซึ่งอาจมีลักษณะเด่นทางคลินิกอื่นๆ และอาจต้องมีวิธีการประเมินประโยคพูดไม่ชัดที่แตกต่าง

ออกไป อย่างไรก็ตามการศึกษานี้ก็เป็นการพิสูจน์แนวคิดที่สามารถนำเครื่องมือ ASR มาใช้ในการช่วยวินิจฉัยได้เป็นอย่างดีเพื่อการต่อยอดการวิจัยในอนาคต

การศึกษานี้ประกอบด้วยผู้ป่วยเพศชายมากกว่าร้อยละ 70 เนื่องด้วยข้อจำกัดทางการเก็บข้อมูล ทำให้ผลการศึกษาความแม่นยำของเครื่องมือและกระบวนการวินิจฉัยนี้อาจไม่สามารถนำไปอ้างอิงกับความแม่นยำในการวินิจฉัยเสียงพูดไม่ชัดของเพศหญิงได้ทั้งหมดและอาจเป็นข้อจำกัดในการนำไปใช้จริงได้

เนื่องจากการศึกษานี้เป็นการศึกษาเบื้องต้นเพื่อพิสูจน์แนวคิดและทดสอบสมมติฐาน จึงเริ่มการศึกษาในกลุ่มผู้ป่วยกลุ่มเฉพาะที่มีเงื่อนไขข้อบังคับทางเสียงค่อนข้างมาก อาทิ ศึกษาในผู้ป่วยพูดภาษาไทยสำเนียงภาคกลางเท่านั้น จึงเกิดข้อกีดในการนำเครื่องมือไปใช้กับผู้ป่วยชาวไทยที่พูดสำเนียงอื่นๆ

อัลกอริทึมการวิเคราะห์เสียงเปลี่ยนเป็นข้อความโดย ASR มีการเปลี่ยนแปลงพัฒนาอย่างต่อเนื่องเพื่อเพิ่มความแม่นยำในการสร้างประโยคให้มากยิ่งขึ้น เป็นผลให้ความถูกต้องแม่นยำในการวินิจฉัยอาการพูดไม่ชัดที่ได้จากการศึกษาครั้งนี้อาจเปลี่ยนแปลงไปได้ ทำให้ต้องมีการศึกษาความแม่นยำของเครื่องมืออย่างต่อเนื่องควบคู่ไปกับการพัฒนาของเครื่องมือ

5.6 ข้อเสนอแนะ

การศึกษาขั้นต่อไปเพื่อต่อยอดจากการศึกษานี้คือการศึกษากายวิภาคและการพูดไม่ชัดชนิดอื่นๆนอกจากเสียงพูดปกติ หรือการจำแนกประเภทของการพูดไม่ชัดทางระบบประสาทออกจากกัน โดยสามารถใช้หลักการและขั้นตอนกระบวนการวิจัยที่คล้ายเดิมคือผสมผสานองค์ความรู้ทางการแพทย์และสัตววิทยา แต่เปลี่ยนแปลงในรายละเอียดโดยเฉพาะลักษณะเด่นทางคลินิกและทางการออกเสียงที่ต่างออกไปตามพยาธิสรีรวิทยาที่เกิดขึ้นในระบบประสาทนั้น อาทิเช่น การแยกแยะอาการพูดไม่ชัดชนิดอ่อนปวกเปียกหรือ flaccid dysarthria ที่มีลักษณะเด่นคือกล้ามเนื้อใบหน้าและกล้ามเนื้อคอหอยที่อ่อนแรงแต่ไม่แข็งเกร็ง ทำให้เกิดลักษณะเด่นทางสัทศาสตร์การออกเสียงคือเป็นเสียงขึ้นจมูกชัดเจน (Hypernasality) จึงควรออกแบบประโยคพูดที่ประกอบด้วยพยัญชนะหรือสระที่มีเสียงนาสิก (เสียงขึ้นจมูก) มากที่สุดเพื่อเพิ่มความสามารถในการวินิจฉัยอาการพูดไม่ชัด เป็นต้น

คะแนนความผิดพลาดในการออกเสียงในผู้ป่วยที่มีอาการพูดไม่ชัดชนิดสປາສຕิกนี้ มักแปรผันตรงกับความอ่อนแรงและแข็งเกร็งของกล้ามเนื้อลิ้น กล้ามเนื้อเพดานอ่อน และกล้ามเนื้อคอหอย ซึ่ง

เป็นกล้ามเนื้อที่ออกแรงทำงานเพื่อการกลืนด้วย ดังนั้นคะแนนความผิดพลาดที่มากขึ้นอาจแปรผันตรงกับความเสี่ยงในการเกิดอาการสำลักจากการกลืนที่อ่อนแอกว่ามากขึ้นด้วย การศึกษาในอนาคตอาจวิเคราะห์หาความสัมพันธ์ของทั้งสองปัจจัยนี้เพิ่มเติม เพื่อให้ได้ประโยชน์สูงสุดทางคลินิกนอกจากการใช้ในการวินิจฉัยอาการพูดไม่ชัด

การศึกษาในอนาคตอาจเพิ่มจำนวนและความหลากหลายของตัวอย่างที่เข้าร่วมการศึกษา เช่น เพิ่มจำนวนตัวอย่างเพศหญิงมากขึ้น คัดตัวอย่างที่พูดภาษาไทยสำเนียงอื่นทั้งหมดเข้ามาในการศึกษา เป็นต้น เพื่อให้เครื่องมือที่ศึกษาสามารถนำไปใช้กับผู้ป่วยที่มีอาการพูดไม่ชัดได้อย่างกว้างขวางมากขึ้น



บรรณานุกรม

Ansel, B. M. and R. D. Kent (1992). "Acoustic-phonetic contrasts and intelligibility in the dysarthria associated with mixed cerebral palsy." J Speech Hear Res 35(2): 296-308.

Ballati, F., et al. (2018). "Hey Siri, do you understand me?": Virtual Assistants and Dysarthria. Intelligent Environments 2018, IOS Press: 557-566.

Calvo, I., et al. (2021). "Evaluation of an Automatic Speech Recognition Platform for Dysarthric Speech." Folia Phoniatr Logop 73(5): 432-441.

Chen, H. and K. N. Stevens (2001). "An acoustical study of the fricative /s/ in the speech of individuals with dysarthria." J Speech Lang Hear Res 44(6): 1300-1314.

Clements, G. N. and S. J. Keyser (1983). CV Phonology: A Generative Theory of the Syllable, MIT Press.

Dai, Z. H. (1991). "[Olivopontocerebellar atrophy: clinical analysis of 100 cases]." Zhonghua Shen Jing Jing Shen Ke Za Zhi 24(2): 111-113, 126.

Darley, F. L., et al. (1968). "Motor Speech Signs in Neurologic Disease." Medical Clinics of North America 52(4): 835-844.

Darley, F. L., et al. (1969). "Differential diagnostic patterns of dysarthria." J Speech Hear Res 12(2): 246-269.

Dehqan, A., et al. (2012). "The effects of aging on acoustic parameters of voice." Folia Phoniatr Logop 64(6): 265-270.

Duffy, J. R. (2005). Motor Speech Disorders: Substrates, Differential Diagnosis, and Management, Elsevier Mosby.

Duffy, J. R. (2019). Motor Speech Disorders E-Book: Substrates, Differential Diagnosis, and Management, Elsevier Health Sciences.

Fonville, S., et al. (2008). "Accuracy and inter-observer variation in the classification of dysarthria from speech recordings." J Neurol 255(10): 1545-1548.

H, M. C., et al. (2020). "Investigation of Different Time-Frequency Representations for Intelligibility Assessment of Dysarthric Speech." IEEE Trans Neural Syst Rehabil Eng 28(12): 2880-2889.

Honda, K. (2015). Supralaryngeal Articulators in the Oropharyngeal Region. The Handbook of Speech Production: 59-78.

Hoole, P. and M. Pouplier (2015). Interarticulatory Coordination. The Handbook of Speech Production: 131-157.

Kalinli, O., et al. (2019). Parametric Cepstral Mean Normalization for Robust Speech Recognition.

Kent, R. D., et al. (1998). "The dysarthrias: speech-voice profiles, related dysfunctions, and neuropathology." Journal of Medical Speech-language pathology.

Kim, H., et al. (2010). "Frequency of consonant articulation errors in dysarthric speech." Clin Linguist Phon 24(10): 759-770.

Kim, Y., et al. (2011). "An acoustic study of the relationships among neurologic disease, dysarthria type, and severity of dysarthria." J Speech Lang Hear Res 54(2): 417-429.

Lansford, K. L. and J. M. Liss (2014). "Vowel acoustics in dysarthria: speech disorder diagnosis and classification." J Speech Lang Hear Res 57(1): 57-67.

Mackenzie, C. (2011). "Dysarthria in stroke: a narrative review of its description and the outcome of intervention." Int J Speech Lang Pathol 13(2): 125-136.

Mann, G., et al. (1999). "Swallowing function after stroke: prognosis and prognostic factors at 6 months." Stroke 30(4): 744-748.

Manochiopinig, S., et al. (2008). "Dysarthric speech characteristics of Thai stroke patients." Disabil Rehabil Assist Technol 3(6): 332-338.

Rohlfing, M. L., et al. (2021). "Hey Siri: How effective are common voice recognition systems at recognizing dysphonic voices?" The Laryngoscope 131(7): 1599-1607.

Souza, L. B. R. d. and M. M. d. Santos (2018). "Body mass index and acoustic voice parameters: is there a relationship?" Brazilian Journal of Otorhinolaryngology 84(4): 410-415.

statistics, V. s. m. u. "Think with Google." Retrieved Accessed March 1, 2020.

Story, B. H. (2015). Mechanisms of Voice Production. The Handbook of Speech Production: 34-58.

Team, S. (2017). "Hey Siri: An On-device DNN-powered Voice Trigger for Apple's Personal Assistant." Retrieved May 30, 2023.

Thoppil, M. G., et al. (2017). "Speech Signal Analysis and Pattern Recognition in Diagnosis of Dysarthria." Ann Indian Acad Neurol 20(4): 352-357.

Van der Graaff, M., et al. (2009). "Clinical identification of dysarthria types among neurologists, residents in neurology and speech therapists." Eur Neurol 61(5): 295-300.

Virani, S. S., et al. (2020). "Heart Disease and Stroke Statistics-2020 Update: A Report From the American Heart Association." Circulation 141(9): e139-e596.

Yang, K., et al. (2023). Text is all You Need: Personalizing ASR Models Using Controllable Speech Synthesis. ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE.

รณเกียรติ, น. (2548). สัตศาสตร์ภาคทฤษฎีและภาคปฏิบัติ. กรุงเทพฯ มหาวิทยาลัยธรรมศาสตร์.

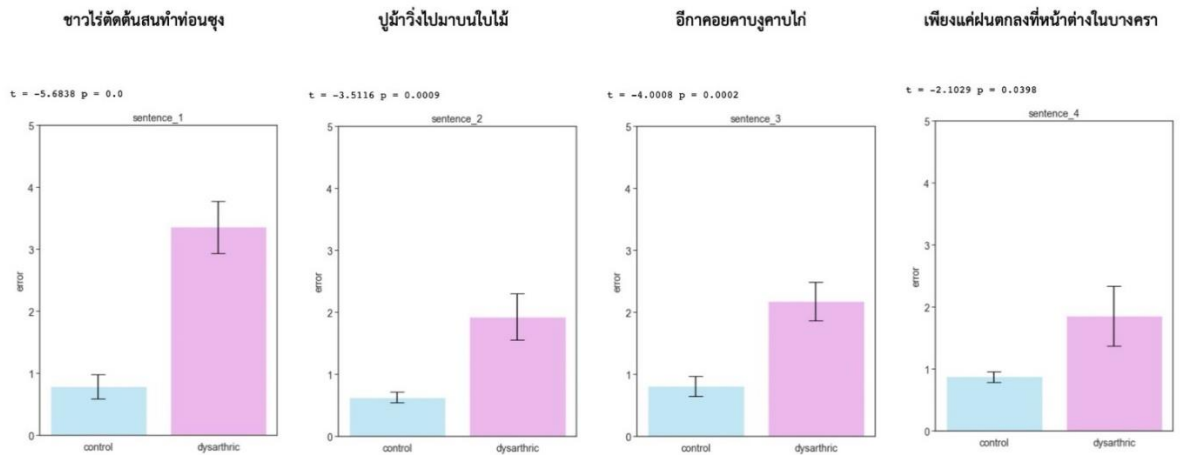


ภาคผนวก

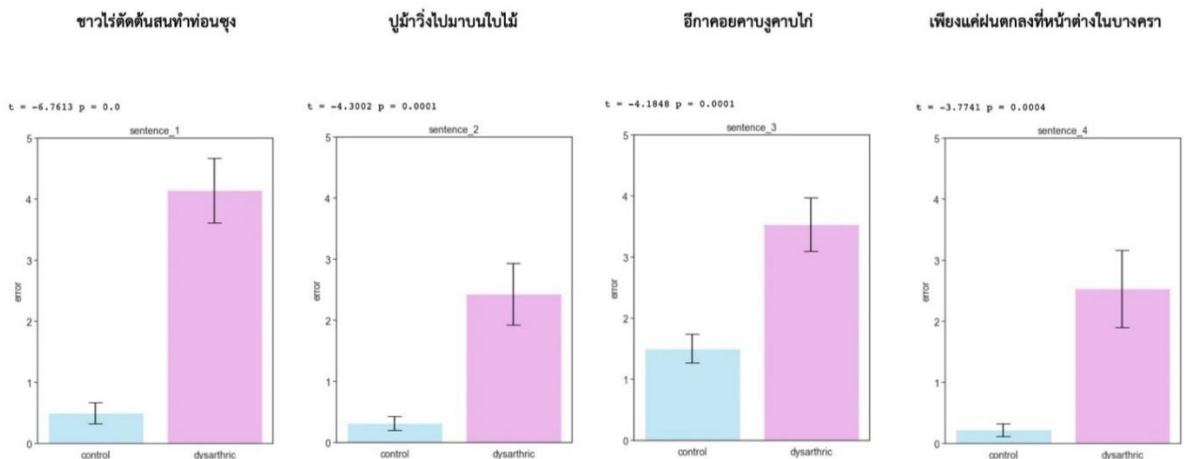
รูปภาพที่ 16 แสดงผลการวิเคราะห์ sensitivity analysis หลังคัดตัวอย่างที่มีอาการทางหูคอจมูก ออก (local voice disturbances)

โดยแสดงเป็นค่าเฉลี่ยของคะแนนความผิดพลาดของทั้ง 4 ประโยคที่ถูกวิเคราะห์ด้วยเครื่องมือ Whisper-Thai-Demo และ Apple Siri™ และแสดงความแตกต่างอย่างมีนัยสำคัญทางสถิติระหว่างกลุ่มผู้ป่วยและกลุ่มควบคุมในทุกประโยคของทั้งสองเครื่องมือ

Whisper



Apple Siri™



ตารางที่ 8 แสดง 2x2 contingency table ของความสามารถในการวินิจฉัยอาการพูดไม่ชัดชนิดสไปสติคโดยประสาทแพทย์แต่ละคนจำนวน 10 คน

โดย Staff คืออาจารย์ประสาทแพทย์

Fellow คือแพทย์ประจำบ้านต่อยอดสาขาสมองขาดเลือด

Resident คือแพทย์ประจำบ้านสาขาประสาทวิทยา

Staff 1		Gold standard		
		Spastic	Control	Total
Neurologist's diagnosis	Positive	8	0	8
	Negative	29	37	66
Total		37	37	

Staff 2		Gold standard		
		Spastic	Control	Total
Neurologist's diagnosis	Positive	16	0	16
	Negative	21	37	58
Total		37	37	

Staff 3		Gold standard		
		Spastic	Control	Total
Neurologist's diagnosis	Positive	17	3	20
	Negative	20	34	54
Total		37	37	

Staff 4		Gold standard		
		Spastic	Control	Total
Neurologist's diagnosis	Positive	28	2	30
	Negative	9	35	44
Total		37	37	

Staff 5		Gold standard		
		Spastic	Control	Total
Neurologist's diagnosis	Positive	17	1	18
	Negative	20	36	56
Total		37	37	

Fellow 1		Gold standard		
		Spastic	Control	Total
Neurologist's diagnosis	Positive	18	0	18
	Negative	19	37	56
Total		37	37	

Fellow 2		Gold standard		
		Spastic	Control	Total
Neurologist's diagnosis	Positive	14	0	14
	Negative	23	37	60
Total		37	37	

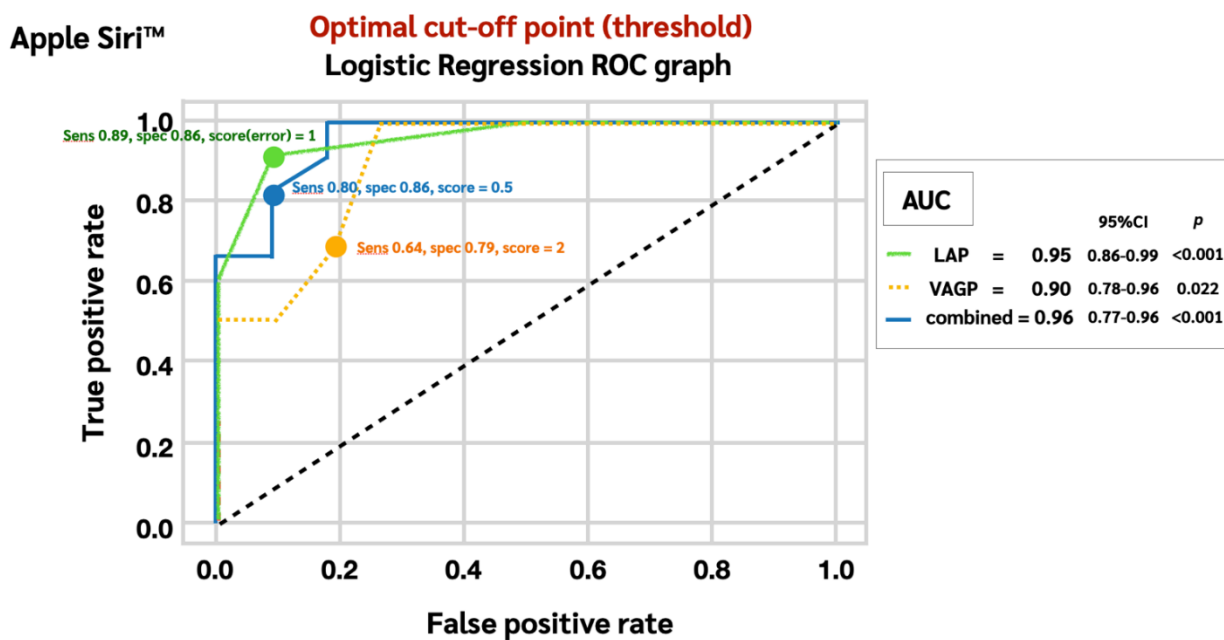
Resident 1		Gold standard		
		Spastic	Control	Total
Neurologist's diagnosis	Positive	26	2	28
	Negative	11	35	46
Total		37	37	

Resident 2		Gold standard		
		Spastic	Control	Total
Neurologist's diagnosis	Positive	17	2	19
	Negative	20	35	55
Total		37	37	

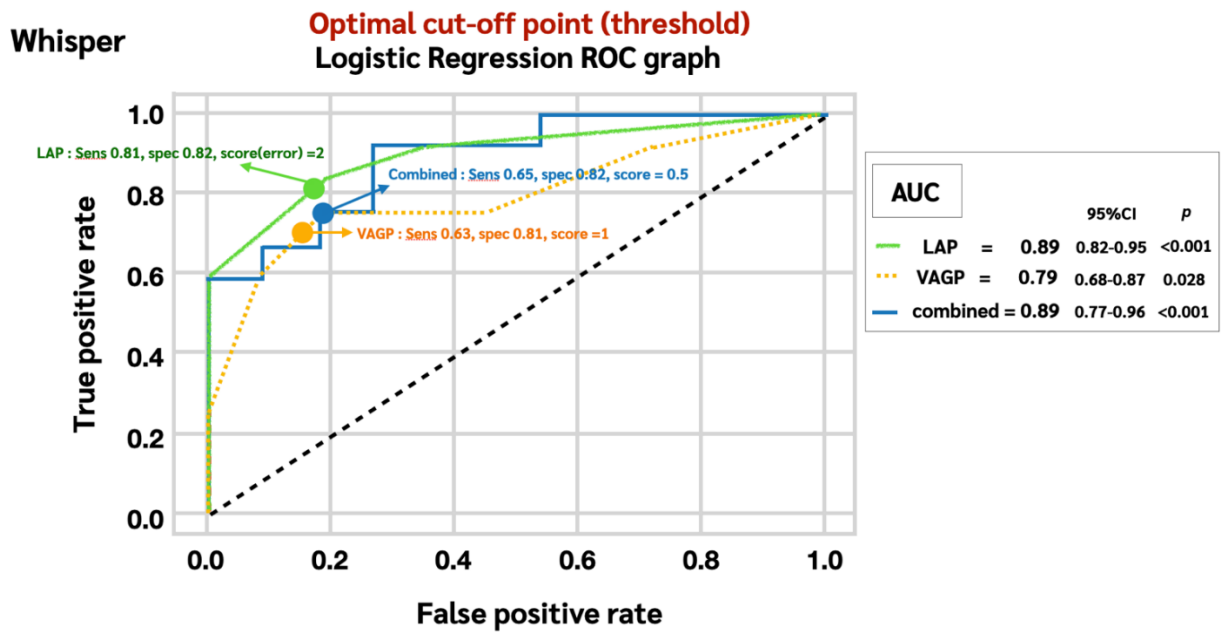
Resident 3		Gold standard		
		Spastic	Control	Total
Neurologist's diagnosis	Positive	21	0	21
	Negative	16	37	53
Total		37	37	

รูปภาพที่ 17 แสดง threshold (optimal cut-off point) หรือจุดตัดที่เหมาะสมในการวิเคราะห์กราฟ ROC จากการวิเคราะห์ด้วย Apple Siri™

พร้อมทั้ง sensitivity (sens), specificity (spec) และ คะแนนความผิดพลาดของพยากรณ์ของจุดตัดนั้น (score) ของประโยคที่มี AUC สูงที่สุดสองประโยค (LAP และ VAGP) และรวมพิจารณาทั้ง 4 ประโยค (combined)



รูปภาพที่ 18 แสดง threshold (optimal cut-off point) หรือจุดตัดที่เหมาะสมในการวิเคราะห์กราฟ ROC จากการวิเคราะห์ด้วย Whisper-Thai-Demo พร้อมทั้ง sensitivity (sens), specificity (spec) และ คะแนนความผิดพลาดของพยากรณ์ของจุดตัดนั้น (score) ของประโยคที่มี AUC สูงที่สุดสองประโยค (LAP และ VAGP) และรวมพิจารณาทั้ง 4 ประโยค (combined)



ประวัติผู้เขียน

ชื่อ-สกุล	ชยุต มฤคทัต
วัน เดือน ปี เกิด	25 กรกฎาคม 2536
สถานที่เกิด	กรุงเทพมหานคร
วุฒิการศึกษา	พ.ศ. 2554 - 2560 นิสิตคณะแพทยศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย พ.ศ. 2560 - 2561 แพทย์เพิ่มพูนทักษะโรงพยาบาลดำเนินสะดวก พ.ศ. 2562 - 2563 อาจารย์แพทย์ คณะแพทยศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง พ.ศ. 2563 - 2566 แพทย์ประจำบ้าน สาขาประสาทวิทยา ภาควิชาอายุศาสตร์ คณะแพทยศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ที่อยู่ปัจจุบัน	ปริญญาและประกาศนียบัตร พ.ศ.2560 แพทยศาสตรบัณฑิต สมาชิกแพทยสภา สมาชิกราชวิทยาลัยอายุรแพทย์แห่งประเทศไทย 472 ซอย ชานเมือง 2 ถนน รัชดาภิเษก ดินแดง ดินแดง กรุงเทพฯ 10400