



การเปรียบเทียบเสียงผู้พูดภาษาไทยด้วยระยะทางแบบยูคลิด: กรณีศึกษาคำทางกลศาสตร์ด้านระยะเวลา

ธณศักดิ์ ศิริคะเนรัตน์ *

คณะมนุษยศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ ประเทศไทย

Speaker voice comparison in Thai using Euclidian distance: a case study of temporal feature

Thanasak Sirikanerat *

Faculty of Humanities, Kasetsart University, Thailand

Article Info

Research Article

Article History:

Received 20 June 2020

Revised 24 August 2020

Accepted 3 September 2020

คำสำคัญ

นิติศาสตร์

กลศาสตร์

การพิสูจน์หลักฐาน

หลักฐานเสียงพูด

การเปรียบเทียบเสียงผู้พูด

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อที่จะศึกษาว่าพารามิเตอร์ด้านระยะสามารถ
ใช้เปรียบเทียบเสียงผู้พูดในภาษาไทยได้หรือไม่ ผู้วิจัยเลือกใช้การวัด
ระยะทางแบบยูคลิดในการเปรียบเทียบเสียงผู้พูด โดยมีสมมติฐานว่าเสียง
ที่พูดด้วยผู้พูดคนเดียวกันจะมีระยะห่างของค่าพารามิเตอร์น้อยกว่าเสียงที่
พูดโดยผู้พูดคนละคนกัน ผู้วิจัยเลือกศึกษาพารามิเตอร์ด้านระยะเวลา 3
พารามิเตอร์ ได้แก่ ΔC , ΔV และ %V ผู้วิจัยเก็บข้อมูลด้วยการสัมภาษณ์
ผู้บอกภาษา 10 คน โดยเก็บข้อมูลเสียงที่มีลักษณะเป็นคำพูดต่อเนื่อง
ผลการวิจัยพบวาระยะทางของค่าพารามิเตอร์ด้านระยะเวลาที่พูดโดยผู้พูด
คนเดียวกันมีค่าอยู่ระหว่าง 0.005 - 0.063 ในขณะที่เสียงที่พูดโดยผู้พูด
คนละคนกันค่าจะอยู่ระหว่าง 0.402 - 8.456 ซึ่งสอดคล้องกับสมมติฐาน
ของผู้วิจัย แสดงให้เห็นว่าพารามิเตอร์ด้านระยะเวลามีประสิทธิภาพในการ
เปรียบเทียบเสียงผู้พูดในภาษาไทยได้

Keywords:

Forensic Phonetics

Acoustic Phonetics

Proof

Voice evidence

Speaker voice

comparison

Abstract

The present study aims to investigate the effectiveness of temporal
features in speaker voice comparison in Thai using Euclidean distance.
The assumption is that the distance of parameter values from
temporal features from the same speakers is smaller than that of
different speakers. Three parameters were used, i.e., ΔC , ΔV and
%V. Acoustic data were obtained by interviewing ten participants.
All data were connected speech. The result revealed that distances
obtained from the same speakers were 0.005 - 0.063 while those
from different speakers were 0.402 - 8.456. The result supported
the assumption and revealed that temporal features were effective
in speaker voice comparison in Thai.

* Corresponding author

E-mail address: thanasak.si@ku.th

1. บทนำ

การเปรียบเทียบเสียงผู้พูด (Forensic voice comparison) คือกระบวนการทางนิติศาสตร์ (Forensic Phonetics) ที่เปรียบเทียบเสียงผู้พูดตั้งแต่ 2 คนขึ้นไปว่าเป็นเสียงของผู้พูดคนเดียวหรือไม่ (Hollien, 2002) ยกตัวอย่างเช่น ในคดีหนึ่ง โจทก์ได้ยื่นหลักฐานเสียงที่บันทึกไว้โดยอ้างว่าเป็นเสียงพูดของจำเลย วิธีการเปรียบเทียบก็คือการนำค่าทางกลศาสตร์ต่างๆ ที่วัดได้จากหลักฐานเสียงมาเทียบกับค่าทางกลศาสตร์ของเสียงพูดจำเลย โดยมีแนวคิดพื้นฐานว่าค่าทางกลศาสตร์จะมีการแปรอยู่แล้วแม้จะพูดโดยคนเดียวกัน แต่การแปรระหว่างบุคคล (between speaker variation) จะสูงกว่าการแปรภายในบุคคล (within speaker variation) (Nolan, 2009; Rose, 2002) การวิเคราะห์การแปรดังกล่าวจะทำให้วิเคราะห์ได้ว่าเสียงพูดทั้ง 2 เสียงเป็นเสียงพูดคนเดียวหรือไม่

กระบวนการดังกล่าวมีบทบาทสำคัญอย่างยิ่งต่อกระบวนการยุติธรรมในประเทศไทย เนื่องจากในปัจจุบันบุคคลทั่วไปสามารถบันทึกเสียงได้อย่างง่ายดาย และเสียงที่บันทึกไว้อาจใช้เป็นหลักฐานในการต่อสู้คดีในชั้นศาลได้ อย่างไรก็ตามการใช้หลักฐานเสียงพูดในชั้นศาลในประเทศไทยยังมีปัญหาในทางปฏิบัติ เนื่องจากปัจจุบันประเทศไทยยังไม่มีเครื่องมือหรือองค์ความรู้ที่จะพิสูจน์หลักฐานเสียงพูด หากคู่ความปฏิเสธว่าเสียงพูดที่คู่ความอีกฝ่ายหนึ่งยื่นต่อศาลไม่ใช่เสียงของตน ศาลก็อาจต้องตีความว่าเสียงพูดนั้นไม่สามารถใช้เป็นหลักฐานในคดีได้ เนื่องจากไม่สามารถพิสูจน์ได้ว่าเสียงพูดที่บันทึกไว้นั้นเป็นเสียงพูดของใครกันแน่ ซึ่งการตีความของศาลลักษณะนี้เป็นไปตามประมวลกฎหมายวิธีพิจารณาความอาญา มาตรา 226/1 และประมวลกฎหมายวิธีพิจารณาความแพ่ง มาตรา 84/1 (เจริญ ภักดีธนากุล, 2561)

นอกจากปัญหาเกี่ยวกับเครื่องมือและองค์ความรู้ที่จะใช้ในการเปรียบเทียบเสียงผู้พูดในกระบวนการยุติธรรมไทยแล้ว การพิสูจน์หลักฐานเสียงพูดยังมีปัญหาในทางปฏิบัติที่สำคัญอีกประการหนึ่งก็คือปัญหาเกี่ยวกับคุณภาพเสียง โดยปกติหลักฐานเสียงพูดมักบันทึกในสิ่งแวดล้อมที่ไม่เป็นอุดมคติ (non-optimal environment) เช่น อาจแอบบันทึกเสียง หรือบันทึกเสียงทางโทรศัพท์โดยใช้แอปพลิเคชันจากโทรศัพท์ ซึ่งเสียงพูดที่บันทึกในลักษณะนี้จะมีเสียงรบกวนหรือเสียงพูดอาจเบาเกินไปจนทำให้การพิสูจน์ยากลำบากมากขึ้น ปัญหาดังกล่าวทำให้ค่าทางกลศาสตร์ที่อาจใช้พิสูจน์หลักฐานเสียงพูดหายไป หรือค่าอาจแปรไปมากจนไม่สามารถนำมาวิเคราะห์ได้

จากปัญหาดังกล่าวผู้วิจัยจึงเห็นว่าการศึกษเกี่ยวกับการเปรียบเทียบเสียงผู้พูดภาษาไทยเชิงนิติศาสตร์เป็นองค์ความรู้ที่สำคัญอย่างยิ่งและมีคุณประโยชน์ต่อกระบวนการยุติธรรมในประเทศไทย และเมื่อพิจารณาถึงข้อจำกัดเกี่ยวกับคุณภาพเสียงที่ได้กล่าวถึงไปแล้ว ผู้วิจัยจึงเห็นว่าค่าทางกลศาสตร์ด้านระยะเวลา (temporal feature) น่าจะเป็นค่าทางกลศาสตร์ที่

ประยุกต์ใช้ในสถานการณ์จริงได้เหมาะสมที่สุด เนื่องจากค่าทางกลศาสตร์ด้านระยะเวลาจะได้รับผลกระทบจากปัจจัยอื่นๆ น้อยกว่าค่าทางกลศาสตร์อื่นๆ เช่นค่าความถี่ฟอร์แมนต์หรือค่าความถี่มูลฐาน (Byrne & Foulkes, 2004; Hove & Dellwo, 2012)

การศึกษาวิจัยในครั้งนี้จึงมีวัตถุประสงค์เพื่อที่จะศึกษาว่าค่าทางกลศาสตร์ด้านระยะเวลาสามารถใช้แยกเสียงผู้พูดต่างคนในภาษาไทยได้หรือไม่ โดยใช้วิธีการคำนวณระยะทางแบบยูคลิด ผู้วิจัยหวังว่าองค์ความรู้ที่ได้จากการวิจัยครั้งนี้จะช่วยเสริมสร้างองค์ความรู้ด้านนิติศาสตร์ในประเทศไทยและอาจนำไปสู่การพัฒนาเครื่องมือในการพิสูจน์หลักฐานเสียงในอนาคตเพื่อประโยชน์ต่อกระบวนการยุติธรรมในประเทศไทย

2. งานวิจัยที่เกี่ยวข้อง

2.1 พารามิเตอร์ด้านระยะเวลากับการเปรียบเทียบเสียงผู้พูด

พารามิเตอร์² ด้านระยะเวลาแต่เดิมเป็นพารามิเตอร์ที่ได้จากการคำนวณค่าทางกลศาสตร์ด้านระยะเวลานักภาษาศาสตร์เสนอขึ้นมาเพื่อใช้อธิบายจังหวะ (rhythm) ในภาษา เช่น Dellwo, (2006); Grabe and Low, (2002); White and Mattys, (2007); Ramus et al. (1999) งานวิจัยดังกล่าวได้พิสูจน์ให้เห็นว่าพารามิเตอร์ด้านระยะเวลา เช่น ค่า Δc , Δv , $\%v$ ³ ฯลฯ ซึ่งเป็นค่าที่คำนวณจากการแปรของช่วงเวลาเสียงพยัญชนะหรือสระสามารถแยกกลุ่มภาษาที่มีจังหวะแบบเน้นเสียงหนัก (stress-timed language) และกลุ่มจังหวะแบบพยางค์ (syllable-timed language) ได้เป็นอย่างดี ข้อเสนอการวิเคราะห์จังหวะในภาษาโดยการใช้พารามิเตอร์ด้านระยะเวลาดังกล่าวนับว่าเป็นการเปลี่ยนมุมมองการวิเคราะห์จังหวะในภาษาไปอย่างมากเนื่องจากในอดีตจะใช้วิธีการคำนวณจำนวนพยางค์ต่อวินาทีหรือวัดค่าระยะเวลาของแต่ละพยางค์

ต่อมานักวิจัยกลุ่มที่ศึกษาการเปรียบเทียบเสียงผู้พูดได้นำแนวคิดเรื่องพารามิเตอร์ด้านระยะเวลามาประยุกต์ใช้ในการเปรียบเทียบเสียงผู้พูด โดยมีแนวคิดสำคัญก็คือเสียงผู้พูดจะมีการแปรอยู่แล้วเป็นธรรมชาติ การเปรียบเทียบเสียงผู้พูดคือการเปรียบเทียบค่าทางกลศาสตร์ที่มีการแปร ซึ่งเป็นแนวคิดที่สอดคล้องกับแนวคิดด้านค่าทางกลศาสตร์ด้านระยะเวลา (Rose, 2002)

² ผู้วิจัยใช้คำว่าพารามิเตอร์เมื่อกล่าวถึงตัวแปรที่ได้จากการคำนวณค่าทางกลศาสตร์

³ เครื่องหมาย Δ (เดลต้า) และ $\%$ (เปอร์เซ็นต์) ที่ใช้ในการวิจัยครั้งนี้ใช้ในความหมายที่แตกต่างจากในทางคณิตศาสตร์โดยทั่วไป ผู้วิจัยยึดการใช้สัญลักษณ์ตามงานวิจัยที่ผ่านมาที่ศึกษาเกี่ยวกับพารามิเตอร์ด้านระยะเวลา

ผลจากการนำแนวคิดเรื่องพารามิเตอร์ด้านระยะเวลามาประยุกต์ใช้กับการเปรียบเทียบเสียงผู้พูดพบว่าพารามิเตอร์ด้านระยะเวลาสามารถใช้เปรียบเทียบเสียงผู้พูดได้อย่างมีประสิทธิภาพ เช่น งานวิจัยของ Jessen (2007), Leemann, Kolly and Dellwo (2014) เนื่องจากค่ากลศาสตร์ด้านระยะเวลาเองก็มีคุณสมบัติที่เหมือนกับค่าทางกลศาสตร์ประเภทอื่นๆ คือมีการแปรทั้งภายในบุคคลและระหว่างบุคคล (Cahn, 1990; Dellwo & Koreman, 2008; Dellwo, 2010; Laan, 1997; Wiget, 2010, Yoon, 2010)

อย่างไรก็ตามพารามิเตอร์หรือค่าทางกลศาสตร์ที่มีประสิทธิภาพในการเปรียบเทียบเสียงผู้พูดก็ขึ้นอยู่กับแต่ละภาษาด้วย กล่าวคือค่าที่ใช้เปรียบเทียบเสียงผู้พูดได้ในภาษาหนึ่งอาจไม่สามารถใช้ได้ ในอีกภาษาหนึ่ง (Rose, 2002) หากต้องการเปรียบเทียบเสียงผู้พูดภาษาไทย ก็จำเป็นต้องศึกษากรณีของภาษาไทยเป็นการเฉพาะ ดังนั้นงานวิจัยที่ศึกษาเกี่ยวกับการเปรียบเทียบเสียงผู้พูดด้วยพารามิเตอร์ด้านระยะเวลาในภาษาอื่นๆ จึงยังไม่ใช่เครื่องพิสูจน์ว่าพารามิเตอร์ดังกล่าวก็จะใช้ในภาษาไทยได้ด้วย

แม้งานวิจัยที่ศึกษากรณีภาษาต่างประเทศจะไม่สามารถใช้กับกรณีของภาษาไทยได้โดยตรง แต่ในภาษาไทยก็มีการเปรียบเทียบ (contrast) ที่สัมพันธ์กับค่ากลศาสตร์ด้านระยะเวลาหลักๆ 2 ประการดังนี้ ประการที่ 1 การเปรียบเทียบระหว่างสระสั้นกับสระยาว ประการที่ 2 การเปรียบเทียบด้านค่า VOT (Voice Onset Time) เสียงกักในภาษาไทยมีการเปรียบเทียบค่า VOT 3 ระดับ ได้แก่เสียงกักไม่ก้องไม่พ่นลม เสียงกักไม่ก้องพ่นลม และเสียงกักก้อง ซึ่งลักษณะในประการที่ 1 และ 2 ล้วนเป็นค่าทางกลศาสตร์ด้านระยะเวลาทั้งสิ้น พารามิเตอร์ด้านระยะเวลาจึงน่าจะเป็นค่าที่ใช้เปรียบเทียบเสียงผู้พูดในภาษาไทยได้อย่างมีประสิทธิภาพ

การวิจัยครั้งนี้จะเลือกศึกษาพารามิเตอร์ด้านระยะเวลาทั้งหมด 3 พารามิเตอร์ ได้แก่ ΔC (ค่าความแปรปรวนของค่าระยะเวลาเสียงพยัญชนะ), ΔV (ค่าความแปรปรวนของค่าระยะเวลาเสียงสระ) และ %V (สัดส่วนค่าระยะเวลาเสียงสระต่อค่าระยะเวลาเสียงทั้งหมด) ซึ่งทั้ง 3 พารามิเตอร์เป็นพารามิเตอร์ที่นิยมใช้ในงานวิจัยที่ศึกษาการเปรียบเทียบเสียงผู้พูด ผู้วิจัยเลือกพารามิเตอร์ ΔC และ ΔV เนื่องจากเป็นพารามิเตอร์ที่ได้จากคำนวณความแปรปรวนของค่าระยะเวลาเสียงพยัญชนะและสระ ความแปรปรวนของค่าระยะเวลานี้สัมพันธ์โดยตรงกับลักษณะการเปรียบเทียบด้านค่าระยะเวลาในภาษาไทย ส่วนพารามิเตอร์ %V เป็นพารามิเตอร์ที่พบว่ามีประสิทธิภาพในการสะท้อนลักษณะเฉพาะของเสียงผู้พูดได้ดีเมื่อเทียบกับพารามิเตอร์ด้านระยะเวลาอื่นๆ (Leemann, Kolly, & Dellwo, 2014)

พารามิเตอร์ทั้ง 3 พารามิเตอร์และวิธีการคำนวณแสดงในตารางที่ 1 สูตรการคำนวณในตารางที่ 1 จะมีเพียง %V สูตรเดียวเท่านั้นที่คูณด้วย 10 ในขณะที่อีก 2 สูตรที่เหลือคูณด้วย 100 การคูณ 10 หรือ คูณ 100 เป็นการคูณเพื่อให้สะดวกต่อการนำเสนอข้อมูลเท่านั้น ไม่มี

ผลกระทบต่อผลการวิเคราะห์ต่อไปนี้ การคูณด้วย 10 หรือ 100 จะมีผลทำให้หลักทศนิยมเปลี่ยนไปเท่านั้น การคำนวณค่าของทั้ง 3 พารามิเตอร์ผู้วิจัยใช้สูตรการคำนวณเดียวกับงานวิจัยของ Ramus, Nespor, and Mehler (1999) โดยปรับการคำนวณเพื่อให้สอดคล้องกับการนำเสนอผลการวิจัยในครั้งนี้

ตารางที่ 1

พารามิเตอร์ด้านระยะเวลาและสูตรคำนวณ

ค่าทางกลศาสตร์	สูตรคำนวณ
ΔC (ค่าความแปรปรวนของค่า ระยะเวลาเสียงพยัญชนะ)	$\Delta C = 100 \times \sqrt{\frac{n \sum_{i=1}^n c_i^2 - (\sum_{i=1}^n c_i)^2}{n(n-1)}}$
ΔV (ค่าความแปรปรวนของค่า ระยะเวลาเสียงสระ)	$\Delta V = 100 \times \sqrt{\frac{n \sum_{i=1}^n v_i^2 - (\sum_{i=1}^n v_i)^2}{n(n-1)}}$
$\%V$ (สัดส่วนค่าระยะเวลาเสียง สระต่อค่าระยะเวลาเสียง ทั้งหมด)	$\%V = 10 \times \frac{n \sum_{i=1}^n v_i}{\sum_{i=1}^n c_i + \sum_{i=1}^n v_i}$

c หมายถึงค่าระยะเวลาของเสียงพยัญชนะ v หมายถึงค่าระยะเวลาของเสียงสระ

n หมายถึงจำนวนช่วงเสียงที่วัดค่า (numbers of interval)

เพียงแค่การคำนวณพารามิเตอร์ด้านระยะเวลายังไม่สามารถใช้ในการเปรียบเทียบเสียงผู้พูดได้โดยตรง การเปรียบเทียบเสียงผู้พูดจำเป็นต้องประยุกต์ใช้แนวคิดทางสถิติหรือคณิตศาสตร์ซึ่งในการวิจัยครั้งนี้ผู้วิจัยเลือกใช้การวัดระยะทางแบบยูคลิดในการเปรียบเทียบเสียงผู้พูด

2.2 การคำนวณระยะทางแบบยูคลิด

ระยะทางแบบยูคลิดคือหลักการทางพีชคณิตที่สามารถวัดระยะห่างของค่าต่าง ๆ ได้ โดยกำหนดให้ระยะทางเท่ากับ d และค่าค่าหนึ่งเท่ากับ p และอีกค่าหนึ่งเท่ากับ q การคำนวณระยะห่างของค่า p และ q คำนวณได้จากสูตร $d(p, q) = d(q, p) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2}$ (Anton & Rorres, 2014) ยกตัวอย่างเช่น หากค่า $p = 10$ และค่า $q = 8$ ค่า d ซึ่งแสดงระยะห่างระหว่างค่า p กับ q ก็จะเท่ากับ $\sqrt{(10 - 8)^2} = 4$

การคำนวณระยะทางแบบยูคลิดสามารถคำนวณค่าในหลายมิติ (multi-dimensional) ได้ด้วยการเพิ่มมิติที่ต้องการคำนวณเข้าไป ยกตัวอย่างเช่น หากต้องการคำนวณระยะทาง 2

มิติ จะกำหนดให้ค่าของ p ในมิติที่ 1 เป็น p_1 ส่วนค่าของ q เป็น q_1 และค่าของ p ในมิติที่ 2 เป็น p_2 ส่วนค่าของ q เป็น q_2 โดยสามารถคำนวณหาระยะทางแบบ 2 มิติได้ตามสูตร $\sqrt{(q_1 - p_1)^2 + (p_2 - q_2)^2}$ จะเห็นว่าการเพิ่มมิติเข้าไปในสูตรทำได้ด้วยการ + มิติที่ต้องการคำนวณเพิ่ม โดยการคำนวณระยะทางแบบยูคลิดสามารถเพิ่มมิติได้แบบอนันต์ (infinity) ดังที่ปรากฏในสูตร $d(p, q) = d(q, p) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2}$ (Anton & Rorres, 2014)

งานวิจัยที่ศึกษาการเปรียบเทียบเสียงผู้พูดในทางนิติศาสตร์ยังมีประเด็นที่ถกเถียงกันอยู่มากเกี่ยวกับวิธีการทางคณิตศาสตร์ที่จะนำมาใช้ในการวิเคราะห์ อย่างไรก็ตาม Rose (2002) ได้อธิบายถึงการเปรียบเทียบเสียงผู้พูดโดยใช้ค่าจากหลายมิติร่วมกันในการวิเคราะห์ วิธีการวิเคราะห์ที่ Rose ได้แสดงให้เห็นนั้นมีแนวคิดที่สอดคล้องกับการคำนวณระยะทางแบบยูคลิด นอกจากนี้ยังมีงานวิจัยด้านนิติศาสตร์และการรู้จำเสียง (speech recognition) ที่ประยุกต์ใช้การคำนวณระยะทางแบบยูคลิด เช่นงานวิจัยของ Singh, M. K., Singh, N. and Singh, A. K. (2019); Segundo, Tsanas, and Gómez-Vilda (2017); Thakur and Sahayam (2013) แสดงให้เห็นว่าการคำนวณระยะทางแบบยูคลิดเป็นวิธีการทางคณิตศาสตร์ที่สามารถใช้เปรียบเทียบเสียงผู้พูดได้เช่นเดียวกัน

การคำนวณระยะทางแบบยูคลิดจะแสดงให้เห็นถึง “ความคล้าย” (similarity) ในบริบทของการเปรียบเทียบเสียงผู้พูด ความคล้ายจะหมายถึงลักษณะของเสียงที่ใกล้เคียงกัน ความคล้ายนี้จะแสดงให้เห็นจากผลการคำนวณระยะทาง หากค่าระยะห่างน้อยก็หมายความว่าเสียง 2 เสียงที่กำลังเปรียบเทียบกันคล้ายกันมาก หากค่าระยะห่างมากก็จะหมายความว่าเสียง 2 เสียงมีความคล้ายกันน้อย ในแง่ของการเปรียบเทียบเสียงผู้พูด หากเสียง 2 เสียง เป็นเสียงของผู้พูดคนละคนก็จะมีความคล้ายน้อยกว่าเสียงที่พูดโดยคนเดียวกัน (Segundo et al., 2017)

ดังที่ได้อธิบายไป การวิจัยในครั้งนี้ผู้วิจัยเลือกพารามิเตอร์ด้านระยะเวลาที่จะใช้วิเคราะห์ 3 พารามิเตอร์ ได้แก่ ΔC ΔV และ $\%V$ ดังนั้น ในการคำนวณเปรียบเทียบระยะทางแบบยูคลิดของผู้พูด ผู้วิจัยจะกำหนดให้ค่า ΔC ΔV และ $\%V$ ของผู้พูดคนที่ 1 เท่ากับ q_1 q_2 และ q_3 ตามลำดับ ส่วนค่า ΔC ΔV และ $\%V$ ของผู้พูดคนที่ 2 กำหนดเป็น p_1 p_2 และ p_3 ตามลำดับ

2.3 การเปรียบเทียบเสียงผู้พูดในทางนิติศาสตร์ในภาษาไทย

งานวิจัยที่ศึกษาการเปรียบเทียบเสียงผู้พูดในภาษาไทยมีอยู่เพียง 2 งาน ได้แก่งานวิจัยของ Pingjai and Ishihara (2013) และ Thaitechawat and Foulkes (2011) งานวิจัยของ Pingjai and Ishihara (2013) ศึกษาการเปรียบเทียบเสียงผู้พูดโดยใช้ค่า $F0$ และระยะเวลาที่สัมพันธ์กับค่า $F0$ ในขณะที่งานวิจัยของ Thaitechawat and Foulkes (2011) ใช้ค่า ความถี่ฟอร์แมนต์และค่า $F0$ งานวิจัยทั้ง 2 งานได้ผลที่สอดคล้องกันคือค่า $F0$ และค่าความถี่ฟอร์แมนต์มีประสิทธิภาพในการเปรียบเทียบเสียงผู้พูดในภาษาไทย นอกจากนี้ผลการวิจัยของ Thaitechawat and Foulkes

(2011) ยังพบว่าหากใช้ทั้ง 2 พารามิเตอร์ร่วมกันจะสามารถเปรียบเทียบเสียงผู้พูดได้ด้วยความสะดวกที่ง่ายขึ้น แสดงให้เห็นว่าจำนวนของพารามิเตอร์ที่นำมาใช้เปรียบเทียบเสียงผู้พูดมีอิทธิพลต่อผลการวิเคราะห์เป็นอย่างมาก ซึ่งสอดคล้องกับวิธีการเปรียบเทียบเสียงผู้พูดที่ Rose (2002) ได้เสนอแนะไว้ แต่ก็มีข้อจำกัดในการใช้งานจริงที่สำคัญอยู่คือ ค่าความถี่ฟอร์แมนต์เป็นค่าที่แปรไปตามสภาพแวดล้อมได้ง่าย เช่น เสียงแวดล้อม ความดังของเสียงพูดที่บันทึกได้ ฯลฯ ในสถานการณ์จริงมีความเป็นไปได้น้อยมากที่เสียงที่คู่ความในคดีได้บันทึกไว้จะใช้ค่าความถี่ฟอร์แมนต์ในการเปรียบเทียบเสียงผู้พูดได้

นอกจากนี้ งานวิจัยของ Pingjai and Ishihara (2013) และ Thaitechawat and Foulkes (2011) ยังมีข้อจำกัดที่สำคัญประการหนึ่งร่วมกันก็คือทั้ง 2 งานวิจัยเก็บข้อมูลเสียงโดยให้ผู้บอกภาษาออกเสียงเป็นคำพูดเดี่ยว (citation form) ซึ่งในสถานการณ์จริง หลักฐานเสียงพูดแทบจะเป็นไปไม่ได้เลยที่จะปรากฏเป็นคำพูดเดี่ยว ผลการวิเคราะห์ที่ได้จากข้อมูลคำพูดเดี่ยวย่อมแตกต่างจากข้อมูลที่ได้จากคำพูดต่อเนื่อง ผลการวิจัยที่วิเคราะห์จากคำพูดเดี่ยวยังไม่สามารถประยุกต์ใช้ในสถานการณ์จริงได้ เนื่องจากหากต้องพิสูจน์หลักฐานเสียง หลักฐานเสียงนั้นย่อมเป็นลักษณะคำพูดต่อเนื่อง การศึกษาเปรียบเทียบเสียงผู้พูดโดยใช้คำพูดต่อเนื่องจึงเป็นประเด็นสำคัญมากในการประยุกต์ใช้องค์ความรู้ทางนิติวิทยาศาสตร์ในกระบวนการยุติธรรม นอกจากนี้ยังไม่มีงานวิจัยใดที่ศึกษาการเปรียบเทียบเสียงผู้พูดในภาษาไทยด้วยค่ากลศาสตร์ด้านระยะเวลา ณ ปัจจุบันจึงยังไม่มีข้อสรุปว่าค่ากลศาสตร์ด้านระยะเวลาสามารถใช้เปรียบเทียบเสียงผู้พูดในภาษาไทยได้จริงหรือไม่

3. วิธีการวิจัย

การเก็บข้อมูลในครั้งนี้ได้ผ่านการรับรองจากคณะกรรมการจริยธรรมการวิจัยในมนุษย์ เลขที่หนังสือรับรองการยกเว้นพิจารณาจริยธรรมการวิจัยหมายเลข COE 62/043

ผู้วิจัยได้บันทึกเสียงคำพูดต่อเนื่องโดยใช้วิธีการสัมภาษณ์ผู้บอกภาษา จำนวน 10 คน จากนั้นจึงวัดค่ากลศาสตร์ด้านระยะเวลา โดยมีขั้นตอนการวิจัยดังนี้

ขั้นตอนที่ 1 ชี้แจงรายละเอียดการเก็บข้อมูลและให้ผู้บอกภาษาลงชื่อในหนังสือแสดงเจตนาเข้าร่วมการวิจัย

ขั้นตอนที่ 2 ถามคำถามสั้นๆ ให้ผู้เข้าร่วมการทดลองตอบเพื่อให้คุ้นเคยกับการวิจัย

ขั้นตอนที่ 3 บันทึกเสียง

ขั้นตอนที่ 4 วัดค่ากลศาสตร์ด้านระยะเวลา

ขั้นตอนที่ 5 คำนวณพารามิเตอร์ด้านระยะเวลาด้วยสูตรที่แสดงในตารางที่ 1

ขั้นตอนที่ 6 คำนวณระยะทางแบบยูคลิด

ขั้นตอนที่ 7 เปรียบเทียบค่าระยะทางแบบยูคลิดของผู้พูดคนเดียวกันกับผู้พูดต่างคนกัน

3.1 การบันทึกเสียง

ผู้วิจัยใช้โทรศัพท์เคลื่อนที่ยี่ห้อโมโตโรล่า รุ่น G7 โคนใช้แอปพลิเคชัน Voice Recorder ในการอัดเสียง ผู้วิจัยตั้งค่าแอปพลิเคชันตามค่าเริ่มต้นของแอปพลิเคชันเพื่อจำลองสถานการณ์ให้สอดคล้องกับสถานการณ์จริงมากที่สุด โดยผู้วิจัยมีสมมติฐานว่าหลักฐานเสียงที่คู่ความในคดีจะยื่นต่อศาลมีความเป็นไปได้สูงที่บันทึกเสียงด้วยโทรศัพท์เคลื่อนที่โดยไม่ได้ใช้ไมโครโฟนใดๆ และไม่มีการตั้งค่าแอปพลิเคชันก่อนบันทึกเสียง ระหว่างอัดเสียง ผู้วิจัยให้ผู้บอกภาษาเป็นผู้ถือโทรศัพท์เคลื่อนที่เองโดยไม่มีการควบคุมความดังเบาในการออกเสียงและไม่มีการชี้แจงเกี่ยวกับวิธีการถือหรือระยะห่างระหว่างปากกับโทรศัพท์เคลื่อนที่

ลักษณะการบันทึกเสียงในงานวิจัยครั้งนี้ทำให้เสียงที่บันทึกได้มีคุณภาพต่ำเมื่อเทียบกับการอัดเสียงในหลายงานวิจัย เช่น ความดังเบา ความยาวในการพูด เสียงรบกวนระหว่างการบันทึกเสียง ความชัดเจนในการออกเสียงของผู้บอกภาษา ความเร็วในการพูดของผู้บอกภาษาที่แตกต่างกัน อย่างไรก็ตามคุณภาพเสียงที่ไม่เป็นอุดมคติในงานวิจัยครั้งนี้เกิดจากความตั้งใจของผู้วิจัยที่ต้องการลักษณะเสียงที่คล้ายคลึงกับหลักฐานเสียงที่อาจเกิดขึ้นจริงมากที่สุด

3.2 ผู้บอกภาษา

ผู้บอกภาษามีจำนวนทั้งสิ้น 10 คน (เพศหญิงทั้งหมด) มีอายุอยู่ในช่วง 18-24 ปี ผู้บอกภาษาทุกคนเกิดและอาศัยอยู่ในกรุงเทพฯ และไม่เคยอาศัยอยู่ที่ต่างจังหวัดหรือต่างประเทศนานเกินกว่า 3 เดือน ไม่มีผู้บอกภาษาคนใดที่เคยมีรายงานความผิดปกติด้านการพูด

3.3 การสัมภาษณ์

ผู้วิจัยบันทึกเสียงคำพูดต่อเนื่องโดยใช้วิธีการสัมภาษณ์ ก่อนสัมภาษณ์ผู้วิจัยได้ชี้แจงกับผู้บอกภาษาว่าเนื้อหาในการตอบไม่ใช่ประเด็นสำคัญ เพื่อให้ผู้วิจัยไม่กังวลในการตอบคำถาม การสัมภาษณ์ผู้วิจัยใช้คำถามที่เกี่ยวข้องกับเรื่องทั่วไปในชีวิตประจำวัน (เช่น ปกติวันหยุดมักทำอะไร และเหตุใดจึงเลือกทำกิจกรรมนั้น) กับคำถามที่ให้ผู้บอกภาษาแสดงความคิดเห็น (เช่น หากต่อไปนี้มีอินเทอร์เน็ต ชีวิตผู้คนจะเปลี่ยนไปหรือไม่ อย่างไร) รวมทั้งสิ้น 8 คำถาม หากคำถามใดที่ผู้บอกภาษาตอบคำถามสั้นเกินไป เช่น ตอบเป็นคำคำเดียว ผู้วิจัยจะถามรายละเอียดเพิ่มเติมเพื่อให้ผู้บอกภาษาพูดยาวขึ้น ระยะเวลาในการสัมภาษณ์ผู้บอกภาษาอยู่ที่ประมาณ 4-5 นาทีต่อคน อย่างไรก็ตามผู้วิจัยไม่ได้ควบคุมรูปแบบประโยคหรือความยาวในการพูดเพื่อให้ได้ข้อมูลคำพูดต่อเนื่องที่ใกล้เคียงกับสถานการณ์จริงมากที่สุด

3.4 การตัดค่าทางกลศาสตร์ก่อนเริ่มการวิเคราะห์

ผู้วิจัยจะวัดค่ากลศาสตร์ด้านระยะเวลาของเสียงพยัญชนะและสระ โดยจะวัดเฉพาะเสียงพูดของผู้บอกภาษาเท่านั้น ก่อนเริ่มการวัดค่าทางกลศาสตร์ ผู้วิจัยจึงตัดบริเวณที่เป็น

เสียงพูดของผู้วิจัยเอง จากนั้นผู้วิจัยได้ทดสอบไฟล์เสียงด้วยการฟังและตัดช่วงที่ไม่นำมาวิเคราะห์ออก ลักษณะของเสียงที่ตัดออกมีลักษณะดังต่อไปนี้ 1) ช่วงที่ผู้วิจัยไม่สามารถจับใจความได้ว่าผู้บอกภาษาออกเสียงว่าอะไร เนื่องจากลักษณะเสียงเช่นนี้จะไม่สามารถตัดสินได้ว่าเป็นเสียงพยัญชนะหรือสระ 2) ช่วงที่ผู้บอกภาษาลงเลและลากเสียงยาว เช่น การพูดว่า “คือ...” หรือ “อืม...” เนื่องจากลักษณะการพูดเช่นนี้จะส่งผลกระทบต่อค่าระยะเวลาเป็นอย่างมาก การวัดค่าทางกลศาสตร์ทำโดยโปรแกรม PRAAT

3.5 การแบ่งข้อมูล

ผู้วิจัยได้แบ่งข้อมูลค่าทางกลศาสตร์ด้านระยะเวลาของผู้บอกภาษาแต่ละคนออกเป็น 2 ชุด แต่ละชุดมีจำนวนช่วงเสียง (interval) เท่ากับ 50% ของจำนวนช่วงเสียง การแบ่งข้อมูลค่าทางกลศาสตร์ของผู้บอกภาษาแต่ละคนทำให้สามารถคำนวณหาระยะห่างของพารามิเตอร์ของเสียงผู้พูดคนเดียวกันได้ จากนั้นจึงเปรียบเทียบกับค่าระยะห่างของผู้บอกภาษาที่ต่างกัน โดยข้อมูลชุดแรกและชุดที่ 2

3.6 การคำนวณระยะทางแบบยุคลิด

หลังจากวัดค่าทางกลศาสตร์ด้านระยะเวลาแล้วผู้วิจัยได้คำนวณค่า ΔC ΔV และ $\%V$ จากนั้นจึงนำค่าทั้ง 3 ของผู้บอกภาษาแต่ละคนมาคำนวณระยะทางแบบยุคลิด การคำนวณระยะทางจะคำนวณโดยแบ่งออกเป็นการคำนวณระยะทางภายในบุคคลเดียวกันกับระยะทางระหว่างบุคคล

การคำนวณระยะทางภายในบุคคลเดียวกันจะคำนวณระยะห่างของข้อมูลชุดที่ 1 และ 2 โดยทำเช่นนี้กับข้อมูลของผู้บอกภาษาทั้ง 10 คน ผลจากการคำนวณจะได้ค่าระยะทางภายในบุคคลทั้งหมด 10 ค่า การคำนวณระยะทางระหว่างบุคคลจะคำนวณโดยแยกข้อมูล 2 ชุด ออกจากกัน และจับคู่คำนวณระยะทางของผู้บอกภาษาต่างคนกันจนครบทุกคู่ เช่น ข้อมูลชุดที่ 1 จะเริ่มจากการคำนวณระยะทางระหว่างผู้บอกภาษาคนที่ 1 กับคนที่ 2 จากนั้นจึงคำนวณระยะทางระหว่างผู้บอกภาษาคนที่ 1 กับคนที่ 3 ทำเช่นนี้ไปเรื่อย ๆ จนครบทุกคู่ รวมจะได้ข้อมูลระยะทางระหว่างบุคคลทั้งสิ้น 90 ค่า ($45 \text{ คู่} \times \text{ข้อมูล 2 ชุด}$)

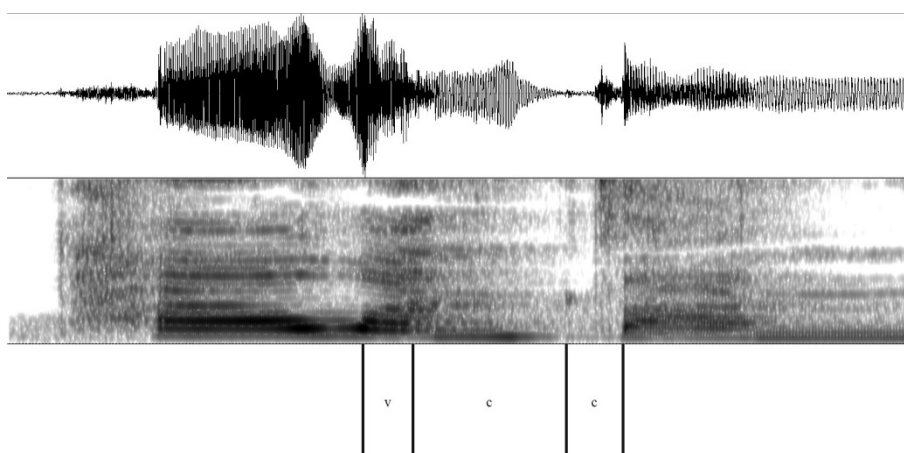
3.7 การวัดค่าทางกลศาสตร์

ผู้วิจัยเลือกใช้วิธีการวัดค่าทางกลศาสตร์ด้านระยะเวลาดำเนินการของงานวิจัยที่ศึกษาการเปรียบเทียบเสียงผู้พูดด้วยค่าทางกลศาสตร์ด้านระยะเวลา เช่น งานวิจัยของ Dellwo (2006); Dellwo and Koreman (2008); Grabe and Low (2002); Ramus et al. (1999)

การวัดค่าระยะเวลาเสียงพยัญชนะจะวัดค่าระยะเวลาตามลักษณะการออกเสียง (manner of articulation) เช่น หากเป็นเสียงกักจะวัดค่า VOT หากเป็นเสียงนาสิกจะวัดค่าตลอดช่วงเสียง

นาสิก ส่วนการวัดค่าเสียงสระจะวัดค่าระยะเวลาตลอดช่วงเสียงสระโดยไม่แยกลักษณะของเสียงสระ กรณีที่เป็นเสียงพยัญชนะต่อเนื่อง เช่นเสียง /n/ แล้วตามด้วยเสียง /c/ จะวัดค่าระยะเวลาเสียงพยัญชนะทั้ง 2 เสียงต่อเนื่องกันโดยถือว่าเป็น 1 ช่วงเสียงพยัญชนะดังที่แสดงในภาพที่ 1

ภาพที่ 1 การวัดค่าระยะเวลาเสียงพยัญชนะในกรณีที่เป็นเสียงพยัญชนะต่อเนื่อง



ภาพที่ 1 คือแผ่นภาพคลื่นเสียง (spectrograph) ของวลี “วันจันทร์” โดยที่ตัว c ตัวแรกคือช่วงเสียง /n/ ในคำว่า “วัน” ส่วน c ตัวที่ 2 คือช่วงเสียง /c/ ในคำว่า “จันทร์” ในกรณีดังกล่าวค่าระยะเวลาของ c ทั้ง 2 ช่วงจะนับรวมกันเป็น 1 ช่วง

4. ผลการวิจัย

ข้อมูลจากการวิเคราะห์ทางกลศาสตร์เบื้องต้นพบว่าจำนวนช่วงเสียงสระและพยัญชนะของผู้บอกภาษาแต่ละคนแตกต่างกันตามที่แสดงในตารางที่ 2 ช่วงเสียงสระมีจำนวนอยู่ระหว่าง 492-1,016 ช่วงเสียง ส่วนช่วงเสียงพยัญชนะมีจำนวนอยู่ระหว่าง 171-1,406 ช่วงเสียง แสดงให้เห็นว่าความยาวในการพูดของผู้บอกภาษาก่อนข้างหลากหลาย ซึ่งสอดคล้องกับวัตถุประสงค์ของการวิจัยที่ต้องการข้อมูลที่ใกล้เคียงกับสถานการณ์จริงมากที่สุด

ตารางที่ 2

จำนวนช่วงเสียงสระและพยัญชนะของผู้บอกภาษาแต่ละคน

หมายเลขผู้บอกภาษา	ช่วงเสียงสระ	ช่วงเสียงพยัญชนะ
1	492	725
2	983	1,380
3	798	171
4	1,016	1,406
5	517	725
6	684	978
7	620	878
8	1,006	1,397
9	646	959
10	766	1,083

4.1 ผลการวิเคราะห์พารามิเตอร์ด้านระยะเวลา

ตารางที่ 3 แสดงผลการวิเคราะห์พารามิเตอร์ด้านระยะเวลาของผู้บอกภาษาทั้ง 10 คน การนำเสนอข้อมูลแบ่งข้อมูลออกเป็น 2 ชุด ข้อมูลชุดที่ 1 แสดงผลการคำนวณค่า ΔC ΔV และ %V สำหรับข้อมูล 50% แรก ส่วนข้อมูลชุดที่ 2 แสดงผลการคำนวณของข้อมูล 50% หลัง ตามที่ได้กล่าวไว้ในหัวข้อที่ 3.5 การแบ่งข้อมูลออกเป็น 2 ส่วนนี้จะแสดงให้เห็นการแปรภายในบุคคลได้ เมื่อพิจารณาข้อมูลเบื้องต้นจะเห็นว่าแม้เป็นผู้บอกภาษาคนเดียวกันแต่ค่า ΔC ΔV และ %V ก็อาจแตกต่างกันได้ นอกจากนี้ความแตกต่างภายในบุคคลยังอาจสูงกว่าความแตกต่างระหว่างบุคคลด้วย เช่น กรณีของค่า ΔV ของผู้บอกภาษาคนที่ 5 ค่า ΔV ของข้อมูลชุดที่ 1 และ 2 ต่างกัน 0.595 ในขณะที่ค่า ΔV ของผู้บอกภาษาคนที่ 4 และ 5 ในข้อมูลชุดที่ 1 กลับต่างกันเพียง 0.242 แสดงให้เห็นการใช้เพียงค่า ΔC ΔV และ %V ไม่สามารถใช้แยกเสียงผู้พูดแต่ละคนออกจากกันได้อย่างมีประสิทธิภาพ จึงยังไม่อาจสรุปได้ว่าพารามิเตอร์ด้านระยะเวลาใช้เปรียบเทียบเสียงผู้พูดในภาษาไทยได้อย่างมีประสิทธิภาพหรือไม่

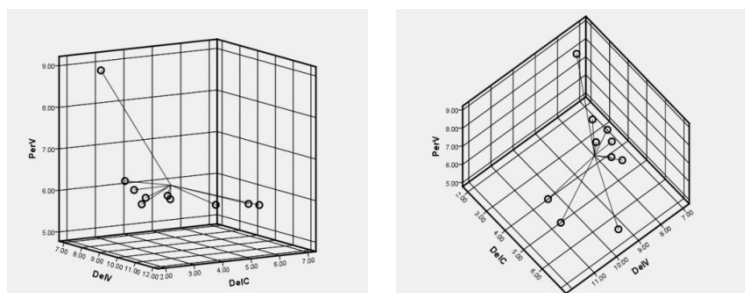
ตารางที่ 3

ผลการวิเคราะห์พารามิเตอร์ด้านระยะเวลา

หมายเลขผู้ บอกภาษา	ข้อมูลชุดที่ 1			ข้อมูลชุดที่ 2		
	ΔC	ΔV	%V	ΔC	ΔV	%V
1	6.942	9.631	5.605	8.070	8.787	5.404
2	3.907	8.096	5.822	3.858	7.740	5.326
3	2.515	7.804	8.957	2.429	6.938	8.965
4	4.109	7.551	5.587	4.933	7.176	5.225
5	4.960	7.793	5.673	5.820	8.388	5.706
6	4.636	10.873	5.931	5.637	8.382	5.362
7	5.628	11.128	5.910	5.942	10.257	5.892
8	3.851	7.536	5.948	3.616	7.140	5.929
9	4.698	8.065	5.809	4.830	8.162	5.708
10	3.364	7.799	6.231	3.817	8.053	6.040

ภาพที่ 2 แสดงผลการคำนวณค่าระยะทางแบบยูคลิดของผู้บอกภาษาทั้ง 10 คน โดยที่แกน X แสดงค่า ΔV แกน Y แสดงค่า %V และแกน Z แสดงค่า ΔC ภาพด้านซ้ายและขวาเป็นภาพเดียวกันเพียงแต่ผู้วิจัยปรับมุมมองโดยการหมุนแกน Y เพื่อแสดงภาพจากมุมมองด้านบน วงกลมในภาพแทนผู้บอกภาษาแต่ละคน จากภาพจะเห็นว่าเมื่อแสดงค่า ΔC ΔV และ %V แบบ 3 มิติจะสามารถแสดงให้เห็นได้ว่าตำแหน่งของผู้บอกภาษาทั้ง 10 คนกระจายอยู่กันคนละที่ภายในพื้นที่ 3 มิติ ไม่มีผู้บอกภาษาคนใดที่ทับซ้อนกัน แต่อย่างไรก็ตามภาพที่ 2 ก็ยังแสดงให้เห็นว่าผู้บอกภาษาบางคนอาจมีค่า ΔC ΔV และ %V ใกล้เคียงกัน กรณีนี้ตีความได้ว่าเสียงผู้บอกภาษาบางคนมีความคล้ายกัน

ภาพที่ 2 ค่า ΔC ΔV และ %V ของผู้บอกภาษาทั้ง 10 คน



4.2 ผลการคำนวณระยะทางแบบยุคลิด

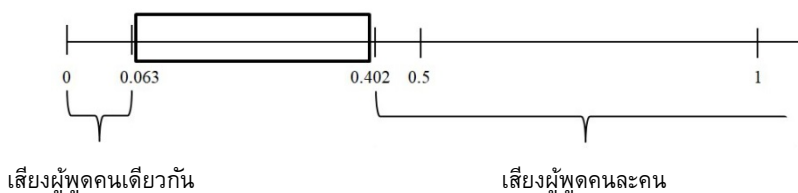
ผลการคำนวณระยะทางแบบยุคลิดพบว่าระยะทางภายในบุคคลเดียวกันมีค่าอยู่ระหว่าง 0.005-0.063 ในขณะที่ระยะทางระหว่างบุคคลมีค่าอยู่ระหว่าง 0.402-8.456 ค่าดังกล่าวแสดงให้เห็นว่าหากเป็นเสียงผู้พูดคนเดียวกัน ค่าระยะทางที่คำนวณได้จะน้อยกว่ากรณีของเสียงผู้พูดคนละคน แม้ว่าค่าระยะทางของทั้งเสียงผู้พูดคนเดียวกันและคนละคนจะมีการแปร แต่ค่าระยะทางสูงสุดของเสียงผู้พูดคนเดียวกัน (0.063) ก็ยังคงน้อยกว่าเสียงที่พูดโดยคนละคนอย่างชัดเจน โดยที่ค่าระยะทางสูงสุดของเสียงผู้พูดคนเดียวกันน้อยกว่าค่าระยะทางต่ำสุดของเสียงผู้พูดคนละคน (0.402) อยู่ประมาณ 6.38 เท่า

5. สรุปผลและอภิปราย

จากผลการวิจัยเบื้องต้นแสดงให้เห็นว่าค่า ΔC ΔV และ $\%V$ มีการแปรทั้งภายในบุคคลและระหว่างบุคคล และบางกรณีอาจพบว่าการแปรภายในบุคคลของค่าบางค่าอาจสูงกว่าการแปรระหว่างบุคคล การเปรียบเทียบเสียงผู้พูดโดยใช้ค่าใดค่าหนึ่งเพียงค่าเดียวไม่สามารถใช้เปรียบเทียบเสียงผู้พูดได้อย่างมีประสิทธิภาพ แต่เมื่อนำค่า ΔC ΔV และ $\%V$ ทั้ง 3 ค่ามาแสดงร่วมกันบนแกน 3 มิติดังที่ปรากฏในภาพที่ 2 จะเห็นว่าสามารถแสดงความแตกต่างของเสียงผู้พูดทั้ง 10 คนได้ ผลดังกล่าวสอดคล้องผลการวิจัยของ Thaitachawat and Foulkes (2011) ที่ศึกษาการเปรียบเทียบเสียงผู้พูดในภาษาไทยและพบว่าการใช้ตัวแปรหลายตัววิเคราะห์ร่วมกันมีประสิทธิภาพกว่าการใช้ตัวแปรเดียว

แม้ภาพที่ 2 จะแสดงให้เห็นว่าเมื่อนำค่าทั้ง 3 ค่ามาพิจารณาร่วมกันจะแสดงให้เห็นความแตกต่างของผู้พูดแต่ละคนได้ แต่ก็ยังมีปัญหาว่าเสียงของผู้พูดบางคนอยู่ใกล้กัน ผู้วิจัยจึงประยุกต์ใช้แนวทางการคำนวณระยะทางแบบยุคลิดเพื่อเปรียบเทียบเสียงผู้พูด ดังที่ได้กล่าวไปในหัวข้อที่ 2.2 ผลการคำนวณระยะทางแบบยุคลิดพบว่าหากเป็นเสียงผู้พูดคนเดียวกัน ค่าระยะทางจะมีค่าน้อยกว่า (อยู่ระหว่าง 0.005-0.063) เมื่อเทียบกับเสียงผู้พูดคนละคน (อยู่ระหว่าง 0.402-8.456) ค่าระยะทางของเสียงผู้พูดคนเดียวกันและคนละคนสามารถแสดงได้ตามภาพที่ 3

ภาพที่ 3 เส้นแสดงระยะทางของเสียงผู้พูดคนเดียวกันกับผู้พูดคนละคน



จากภาพที่ 3 อาจสรุปผลการวิจัยได้ว่าหากเป็นเสียงผู้พูดคนเดียวกัน ค่าระยะทางจะน้อยกว่าหรือเท่ากับ 0.063 และหากเป็นเสียงผู้พูดคนละคนกัน ค่าระยะทางจะมากกว่าหรือเท่ากับ 0.402 ผลดังกล่าวแสดงให้เห็นว่าค่าระยะทางของเสียงผู้พูดคนเดียวกันกับเสียงผู้พูดคนละคนจะมีค่าระยะทางอยู่คนละช่วง ผลการวิเคราะห์นี้ชี้ให้เห็นว่าค่า ΔC ΔV และ %V ซึ่งเป็นพารามิเตอร์ด้านระยะเวลาที่มีประสิทธิภาพในการเปรียบเทียบเสียงผู้พูดในภาษาไทยดังที่ผู้วิจัยได้ตั้งสมมติฐานไว้ และสอดคล้องกับงานวิจัยของ Leemann et al. (2014) ที่ศึกษาประสิทธิภาพของพารามิเตอร์ด้านระยะเวลาในการเปรียบเทียบเสียงผู้พูดภาษาเยอรมัน

อย่างไรก็ตามข้อสรุปนี้เป็นเพียงข้อสรุปที่เกิดจากข้อมูลในการวิจัยในครั้งนี้เท่านั้น ยังคงมีประเด็นที่ต้องศึกษาต่อไปว่าหากค่าระยะทางที่คำนวณได้เป็นค่าที่อยู่ในบริเวณพื้นที่สีเหลี่ยมในภาพที่ 3 ซึ่งมีค่าตั้งแต่ 0.064 จนถึง 0.401 จะตีความว่าเป็นเสียงผู้พูดคนเดียวกันหรือคนละคนกัน จากข้อมูลที่ได้จากการวิจัยในครั้งนี้ยังไม่สามารถตอบประเด็นปัญหาดังกล่าวได้ เพื่อที่จะตอบประเด็นปัญหาดังกล่าวจำเป็นต้องศึกษาจากกลุ่มผู้บอกภาษาจำนวนมากขึ้นเพื่อวิเคราะห์ว่าเสียงผู้พูดคนเดียวกันควรจะมีความอยู่ระหว่างเท่าใด และเสียงผู้พูดคนละคนควรจะมีความอยู่ระหว่างเท่าใด

นอกจากนี้ยังคงมีประเด็นที่ต้องศึกษาเพิ่มเติมว่าเป็นไปได้หรือไม่ที่เสียงผู้พูดคนละคนจะมีค่าระยะทางอยู่ระหว่าง 0–0.063 ซึ่งเป็นค่าระยะทางของเสียงผู้พูดคนเดียวกันที่พบในงานวิจัยนี้ และเป็นไปได้หรือไม่ที่เสียงผู้พูดคนเดียวกันจะมีค่าระยะทางตั้งแต่ 0.402 ขึ้นไป ซึ่งเป็นค่าระยะทางของเสียงผู้พูดคนละคนที่พบในการวิจัยครั้งนี้ การศึกษากลุ่มตัวอย่างจำนวนมากขึ้นจะช่วยตอบประเด็นปัญหาทั้ง 2 ได้เป็นอย่างดี การตอบประเด็นปัญหาทั้ง 2 นี้จะนำไปสู่การประยุกต์ใช้องค์ความรู้ทางนิติวิทยาศาสตร์ในการพิสูจน์หลักฐานเสียงพูดในกระบวนการยุติธรรมได้อย่างมีประสิทธิภาพมากยิ่งขึ้น

การวิจัยในครั้งนี้ผู้วิจัยได้ใช้วิธีการคำนวณระยะทางแบบยุคลิดซึ่งเป็นการเปรียบเทียบเสียงผู้พูดด้วยวิธีการคณิตศาสตร์รูปแบบหนึ่งเท่านั้น แม้ผลจากการวิจัยจะชี้ให้เห็นว่าวิธีการดังกล่าวสามารถใช้เปรียบเทียบเสียงผู้พูดได้ แต่ก็ไม่ได้หมายความว่าวิธีการนี้เป็นวิธีการที่ดีที่สุด ข้อจำกัดที่เห็นได้ชัดในวิธีการที่ผู้วิจัยใช้ในครั้งนี้ก็คือไม่สามารถตีความผลในรูปแบบความน่าจะเป็นได้ แตกต่างจากการวิเคราะห์ผลด้วยภาวะความน่าจะเป็น (Likelihood ratio) ซึ่งเป็นวิธีการวิเคราะห์ที่เปรียบเทียบผลต่างทางสถิติที่ใช้ในงานวิจัยด้านนิติศาสตร์หลาย ๆ งานวิจัย การศึกษาเพิ่มเติมเกี่ยวกับวิธีการทางคณิตศาสตร์และสถิติที่ใช้ในการเปรียบเทียบเสียงผู้พูดในภาษาไทยก็เป็นประเด็นที่สำคัญที่จะต้องศึกษาวิจัยต่อไป

เอกสารอ้างอิง/References

- จรรย์ ภักดีธนากุล. (2561). *กฎหมายลักษณะพยานหลักฐาน* (พิมพ์ครั้งที่ 13). กรุงเทพฯ: สำนัก
อบรมศึกษากฎหมายแห่งเนติบัณฑิตยสภา.
- Byrne, C., & Foulkes, P. (2004). The 'mobile phone effect' on vowel formants. *Journal
of Speech Language and the Law*, 11(1), 83-102.
- Cahn, J. E. (1990). The generation of affect in synthesized speech. *Journal of the American
Voice I/O Society*, (electronic Publication: <https://eprints.kfupm.edu.sa/70011/1/70011.pdf>)
- Dellwo, V. (2006). Rhythm and speech rate: a variation coefficient for _C. *Language and
Language-processing*. In P. Karnowski, & I. Szigeti (Eds). Frankfurt am Main:
Peter Lang, pp. 231-241.
- Dellwo, V. (2010). *Influences of speech rate on the acoustic correlates of speech rhythm:
An experimental phonetic study based on acoustic and perceptual evidence*
(Doctoral Dissertation). University of Bonn.
- Dellwo, V., & Koreman, J. (2008). How speaker idiosyncratic is acoustically measurable
speech rhythm?, *Electronic Proceedings of the annual meeting of the International
Association of Forensic Phonetics and Acoustics (IAFPA)*, Lausanne/Switzerland.
- Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class
hypothesis, in *Papers in Laboratory Phonology 7*. In C. Gussenhoven, & N. Warner
(Eds.). Berlin, New York: Mouton de Gruyter.
- Hollien, H. (2002). *Forensic voice identification*. London: Taylor and Francis.
- Hove, I., & Dellwo, V., (2012). The effect of articulatory obstruction on temporal characteristics
of speech. *Abstract Presented at the IAFPA Conference 2012*, Santander.
- Jessen, M. (2007). Forensic reference data on articulation rate in German. *Science and
Justice*, (47), 50–67.
- Laan, G. (1997). The contribution of intonation, segmental durations, and spectral features
to the perception of a spontaneous and a read speaking style. *Speech
Communication*, 22, 43-65.
- Leemann, A., Kolly, M.-J., & Dellwo, V. (2014). Speaker-individuality in suprasegmental
temporal features: Implications for forensic voice comparison. *Forensic Science
International*, 238, 59–67.
- Nolan, F. (2009). *The Phonetic Bases of Forensic Speaker Identification* (2nd ed.) CUP:
Cambridge.

- Pingjai, S., & Ishihara, S. (2013). Forensic voice comparison in Thai: a likelihood ratio-based approach using tonal acoustics. *Journal of Humanities Naresuan University*, 9(13), 51-66.
- Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, 73, 265-292.
- Rose, P. (2002). *Forensic speaker identification*. London: Taylor and Francis.
- Segundo, E., San, Tsanas, A., & Gómez-Vilda, P. (2017). Euclidean Distances as measures of speaker similarity including identical twin pairs: A forensic investigation using source and filter voice characteristics. *Forensic Science International*, (270), 25-38.
- Singh, M. K., Singh, N., & Singh, A. K. (2019) Speaker's voice characteristics and similarity measurement using Euclidean Distances, International Conference on Signal Processing and Communication (ICSC), NOIDA, India, 2019, (pp. 317-322), doi: 10.1109/ICSC45622.2019.8938366.
- White, L., & Mattys, S. (2007). Calibrating rhythm: first language and second language studies. *Journal of Phonetics*, 35, 501-522.
- Wiget, L., White, L., Schuppler, B., Grenon, I., Rauch, O., & Mattys, S. L. (2010). How stable are acoustic metrics of contrastive speech rhythm? *Journal of the Acoustical Society of America*, 127, 1559-1569.
- Yoon, T. J. (2010) Capturing inter-speaker invariance using statistical measures of speech rhythm, *Proceedings of Speech Prosody*, 5, Chicago.
- Taitechawat, S., & Foulkes, P. (2011). Discrimination of speakers using tone and formant dynamics in Thai. *Proceeding of 17th International Congress of Phonetic Sciences (ICPhS)*, pp.1975–1981.
- Thakur, A. S., & Sahayam, N. (2013). Speech recognition using Euclidean Distance. *International Journal of Emerging Technology and Advanced Engineering*, 3(3), 587-590.