

บทที่ 3

การวิเคราะห์และถอดสระในภาษาไทย

การแปลภาษาไทยเป็นภาษาเบรลล์ไทยนั้นจำเป็นต้องศึกษาลักษณะของภาษาไทย เนื่องจากต้องมีการวิเคราะห์และถอดสระเพื่อการแปลอักษรไทยเป็นอักษรเบรลล์ไทย ดังนั้นจึงจำเป็นต้องทำการศึกษาลักษณะของคำในภาษาไทย ดังต่อไปนี้

3.1 ลักษณะของภาษาไทย

ภาษาไทย มี ลักษณะ เฉพาะตัว เป็น ภาษา ที่มี ระดับ เสียงโดยประกอบไปด้วย พยัญชนะ สระ และ วรรณยุกต์ ซึ่ง ปัจจุบัน มี อักษร ใช้ 44 ตัว แบ่งออก เป็น

อักษรสูง 11 ตัว ได้แก่ ข ฃ ฉ ฐ ถ ผ ฝ ศ ษ ส ห

อักษรกลาง 9 ตัว ได้แก่ ก จ ด ต ถ ฏ บ ป อ

และ อักษรต่ำ 24 ตัว ได้แก่ ค ฅ ฆ ง ช ซ ฌ ญ ฑ ฒ ท ฐ น พ ฟ ภ ม ย ร ล ว พ ฮ

มีสระ 28 รูป 32 เสียง ประกอบ คำ โดย นำเอา อักษร ผสม กับ สระ โดยสระสามารถแบ่งออกได้ 3 ชนิดด้วยกัน คือ

สระเดี่ยว (สระที่เกิดจากการเคลื่อนไหวของลิ้นและริมฝีปากเพียงส่วนเดียว) ได้แก่ อะ อา อิ อี อื อ้อ อุ อู เอะ เอ แอะ แอ โอะ โอ เอาะ ออ เออะ และ เออ

สระประสม (สระที่เกิดจากการ เคลื่อนไหวของลิ้นทั้งส่วนหน้า - กลางและหลัง ทำให้รูปริมฝีปากเปลี่ยนไป) ได้แก่ อัวะ อัว เอียะ เอีย เอื้อะ เอื้อ

สระเกิน (สระที่มีเสียงพยัญชนะประสมอยู่) ได้แก่ อำ ไอ โอ เอา ฤ ฌ ฎ

สำหรับ วรรณยุกต์ มีเสียง วรรณยุกต์ 5 เสียง และ มี รูป วรรณยุกต์ 4 รูป คือ ไม้เอก โท ตรี และ จัตวา คำกับ บนอักษร

3.2 ระบบการสร้างคำในภาษาไทย

ภาษาไทยเป็นภาษาคำโดด (Isolating) กล่าวคือ คำพื้นฐานของภาษาไทยส่วนใหญ่เป็นคำพยางค์เดียว เช่น วิ่ง นั่ง เดิน ฯลฯ ต่อมาเมื่อพัฒนาการของมนุษย์มีมากขึ้น ภาษาก็มีวิวัฒนาการตามความเจริญของงานของผู้ใช้ภาษา การสร้างคำใหม่ ๆ จึงมีมากขึ้น ๆ เพื่อตอบสนองความ

ต้องการในการสื่อสารของมนุษย์อีกทั้งเมื่อมีการติดต่อไปมาหาสู่กันกับคนกลุ่มอื่น ๆ ภาษาจึงมีการแลกเปลี่ยนหยิบยืมถ่ายทอดไปมาระหว่างกัน

คำมูล

คำมูล คือ คำดั้งเดิมในภาษา ที่ยังมิได้ประสมกับคำอื่น จะเป็นคำที่มาจากภาษาใด ก็ได้ อาจมีพยางค์เดียวหรือหลายพยางค์ก็ได้ หรืออาจกล่าวได้ว่า คำมูลคือคำหนึ่ง ๆ ที่ไม่สามารถแยกออกเป็นคำอื่นได้นั่นเอง ลักษณะของคำมูล มีดังนี้

คำมูลพยางค์เดียวที่เป็นคำไทยแท้ เช่น พ่อ แม่ พี่ น้อง กิน ดื่ม เล่น คิด

คำมูลพยางค์เดียวที่เป็นคำยืมจากภาษาอื่น เช่น เดิน บวช นาค กอล์ฟ ฟาล์ว

คำมูลหลายพยางค์ที่เป็นคำไทยแท้ เช่น กระจอก คะนอง จำเพาะ สะเก็ด

คำมูลหลายพยางค์ที่เป็นคำยืมภาษาอื่น เช่น สติ นาฬิกา สับประรด กาแฟ การะเกด

พยายาม

คำประสม

คำประสม คือ การนำเอาคำมูลที่มีความหมายต่างกัน ๒ คำ หรือมากกว่า ๒ คำขึ้นไปมา ประสมกัน ทำให้เกิดคำใหม่ขึ้นอีกคำหนึ่งในภาษา คำประสมในภาษาไทยมีประเด็นที่น่าพิจารณาดังนี้

(1) ลักษณะทางความหมายของคำที่นำมาประสมกันมี 2 ลักษณะ ได้แก่

คำประสมที่มีความหมายตรง คำประสมประเภทนี้มีความหมายสำคัญอยู่ที่คำตั้งหรือคำหลัก ส่วนขยายมีความสำคัญรองลงไป เช่น ผลผลิต ธารน้ำ แม่น้ำ ทางเดิน เรือหางป๋อตา ลายมือ เมืองนอก พัดลม

คำประสมที่มีความหมายโดยนัย หรือความหมายเปรียบเทียบ เช่น มือเท้า ใก้อ่อน หัวสูง เส้นสาย ตาขาว วิจารณ์ ออกแตก หัวหมุน หัวแข็ง หน้าม้า ใจแคบ คอสูง

(2) วิธีการสร้างคำประสม คำประสมมีวิธีการสร้างคำ โดยนำคำมูลที่อาจเป็นภาษาเดียวกัน หรือต่างภาษากันก็ได้ ได้แก่

การนำคำที่ไม่สามารถปรากฏตามอิสระได้ เช่น คำว่า “ช่าง ชาว นัก หมอ การ ความ เครื่อง ของ ที่” เป็นคำตั้งประสมกับคำที่สามารถปรากฏตามอิสระได้ เช่น ช่างทอง ชาวานา นักเขียน หมอความ การกิน ความดี เครื่องเขียน ของคาว ที่พักผ่อน

คำไทยแท้ประสมคำไทยแท้ เช่น พ่อตา โรงเรียน คนงาน คอตก แงงไก่อ เส้นสาย

คำไทยแท้ประสมคำบาลีสันสกฤต เช่น หลักฐาน แจกขาม นางเยาว์ ราชวัง ภูมิลำเนา

อนึ่ง การนำคำ บาลี – สันสกฤต มาประสมกัน เช่น คำว่า อักษรศาสตร์ พุทธสาวก โจรภัย กิตติศัพท์ วัฒนธรรม ราชโอรส เป็นต้น การสร้างคำเหล่านี้ เป็นการสร้างคำแบบคำประสม แต่เนื่องจากการยืมวิธีการสร้างคำบาลี-สันสกฤตมาใช้ จึงจะนำไปกล่าวต่างหากในเรื่องการสร้าง คำสมาส

(3) ลักษณะคำประสม

คำประสมที่ใช้เป็นคำนาม เช่น คนกรุง หน้ายักษ์ น้ำพริก รถด่วน ปลาเค็ม ปากกล้า ฯลฯ

คำประสมที่ใช้เป็นคำกริยา เช่น ยิงธนู ตัดเสื้อ ตัดถนน กินที่ ถือหาง ฯลฯ

คำประสมที่ใช้เป็นคำวิเศษณ์ ได้แก่ (คน) ชั้นต่ำ (นก) สองหัว (เครื่อง) คิดเลข (ผ้า) กั้นเปื้อน (กีฬา) ในร่ม (ถนน) นอกเมือง ฯลฯ

3.3 การวิเคราะห์คำ และถอด สระ เพื่อการแปลเป็นอักษรเบรลล์

สำหรับการแปลภาษาไทยเป็นภาษาเบรลล์ไทยนั้น การวิเคราะห์คำ และ การถอดสระถือเป็นกระบวนการส่วนที่สำคัญอย่างยิ่ง เพราะจากลักษณะของอักษรไทยปกติกับอักษรเบรลล์ไทย นั้นมีความแตกต่างกัน อักษรไทยปกตินี้จะมีระดับของตัวอักษรที่แตกต่างกัน (Level) แบ่งออกได้ เป็นกลุ่มดังนี้ คือ 1. พยัญชนะ (Consonant) เป็นอักษร 44 ตัว ดังที่ได้กล่าวมาแล้วข้างต้น

2. สระหรืออักษรที่อยู่ด้านบน (Upper Vowel) เป็นสระหรืออักษรที่อยู่ด้านบนของพยัญชนะ ได้แก่ ะ ุ ู ฌ ๓ + ๗ ๘ ๙ ๑๐ ๑๑

3. สระหรือตัวอักษรที่อยู่ด้านล่าง (Lower Vowel) เป็นสระหรืออักษรที่อยู่ด้านล่างของพยัญชนะ ได้แก่ ๓ ๔

4. สระหรือตัวอักษรที่อยู่ด้านหน้า (Front Vowel) เป็นสระหรืออักษรที่อยู่ด้านหน้าของพยัญชนะ ได้แก่ ๓ ๔ ๕ ๖ ๗ ๘ ๙ ๑๐ ๑๑

5. สระหรือตัวอักษรที่อยู่ด้านหลัง (Rear Vowel) เป็นสระหรืออักษรที่อยู่ด้านหลังของพยัญชนะ ได้แก่ ๓ ๔ ๕ ๖ ๗ ๘ ๙ ๑๐ ๑๑

เมื่อมีการประสมคำระหว่างอักษรปกติจะทำให้คำที่ได้มีระดับที่แตกต่างกันไป ส่วน

สำหรับอักษรเบรลล์ไทยนั้น จะมีรูปแบบลักษณะโดยทั่วไปเหมือนกับอักษรเบรลล์ คือ มีจุด 6 จุด

เรียงกันเป็นอักษรที่มีระดับเดียวกันตลอด โดยที่อักษรก็แยกกลุ่มได้เป็น พยัญชนะ สระ และ

วรรณยุกต์ ซึ่งในอักษรเบรลล์ไทยนั้นจะมีการแทนสระ หรือตัวอักษรในระดับบนและล่าง เพื่อให้

สามารถที่จะทำให้ตัวอักษรอยู่ในระดับเดียวกันทั้งหมด ดังนั้นการวิเคราะห์คำจึงจำเป็นอย่างมาก เพื่อช่วยให้เกิดความสะดวกในการแปลอักษรไทยเป็นอักษรเบรลล์ไทย

การวิเคราะห์คำนั้น สังเกตได้ว่าหากคำที่เป็นคำเดียวนั้นจะทำให้สามารถถอดสระได้ง่ายกว่าคำประสม ด้วยเหตุนี้เองในการวิเคราะห์คำจึงจำเป็นที่จะต้องอาศัยแนวทางหรือวิธีการในการแบ่งคำ หรือที่เรียกกันว่า การตัดคำ (Word Segmentation) เข้ามา ช่วย ในการแบ่งคำ ออกจากข้อความ หรือสายอักขระ(ซึ่งในบทต่อไปจะได้กล่าวถึง วิธีการและลักษณะของการตัดคำข้อดี ข้อเสียของแต่ละวิธี) โดยที่จะทำการเริ่มพิจารณาจากข้อความที่สั้นที่สุดก่อน จากนั้นก็ทำข้อความถัดไปจนหมด เมื่อตัดคำเสร็จแล้ว ก็จะได้เป็นคำเดี่ยว ซึ่งจะไม่พิจารณาถึงความหมายเพราะในการแปลไทยเป็นเบรลล์ไทยนั้น ไม่คำนึงถึงเพราะต้องการแค่แปลอักษรเท่านั้น โดยเมื่อทำการแปลเสร็จแล้ว ผู้ใช้หรือผู้พิการทางสายตาจะสามารถรู้จากบริบทได้เองว่าคำๆนั้นหมายความว่าสิ่งใด

ยกตัวอย่างคำที่มีความกำกวม “ตากลม” อาจเกิดความสับสนได้ว่า เป็น ตาก-ลม หรือ ตา-ลม ซึ่งเป็นปัญหาในการตัดคำโดยทั่วไป ในที่นี้สำหรับการแปลจากอักษรไทยปกติ เป็นอักษรเบรลล์ไทยนั้น ไม่ว่าจะตัดคำออกมาเป็นตา-ลม หรือ ตาก-ลม ผลลัพธ์ที่ได้จากการแปลก็จะได้ ตากลม ซึ่งผู้อ่านจะรู้ได้จากบริบทว่าข้อความหรือประโยคนั้นกล่าวถึง ตาก-ลม หรือ ตา-ลม หรือแม้แต่คำว่า ไปห้ามเหสี ผลลัพธ์จากการแปลก็คือ ไปห้ามเหสี ซึ่งผู้อ่านจะสามารถรู้ได้ว่า ข้อความนี้หมายถึง ไป-ห้ามเหสี ไม่ได้หมายความว่า ไป-ห้าม-เห-สี

อาจสรุปได้ว่า ปัญหาคำกำกวมที่ เกิดจากคำที่ประกอบไปด้วยสระเดี่ยว (อะ อา อิ อี อือ อุ อู เอ แอ โอ ออ) สระเกิน (อำ ไอ โอ ฤ ฦ ฎ ฏ) นั้นจะไม่มีผลต่อการแปลอักษรไทยเป็นอักษรเบรลล์ไทยกล่าวคือสามารถแปลได้อักขระต่ออักขระ แต่สำหรับคำที่ประกอบไปด้วยสระเดี่ยว (อะ แอะ โอะ เออะ เออะ) สระเกิน (เอา) และ สระประสม (อัวะ อัว เอียะ เอีย เอือะ เอือ) จะเป็นปัญหาในเรื่องของการถอดสระ เช่นยกตัวอย่างการถอดสระ จาก คำว่า เพลา ซึ่งเป็นคำพ้องรูป ซึ่งมีได้สองรูปนั้นคือ

1. เปลา ในความหมายที่หมายถึง เพลารถ
2. เปลา ในความหมายที่หมายถึง เปลาๆ หน่อย
3. เปลา ในความหมายที่หมายถึง เปลาที่บ่งบอกเวลา เช่น เพลายืน

ซึ่งเมื่อทำการตัดคำแล้วได้เป็น เพ-ลา ผลลัพธ์ที่ได้ออกมา จะเป็น เพ-ลา จะแต่ถ้าหากตัดคำได้ว่า เพลา ก็จะเกิดปัญหาในการถอดสระ เพราะคำว่าเพลาประกอบไปด้วยสระเกิน ซึ่งอาจจะทำให้ได้คำว่า พล-เ-า ซึ่งได้มาจาก เพลา หรือ อาจจะได้คำว่า เพ-ลา จากคำว่า เพลา ส่วนสำหรับคำว่า เวลา นั้น ก็ไม่เป็นปัญหาในการถอดสระ เพราะว่ามี การอ่านในลักษณะของสระเกิน คือไม่มีการอ่านว่า ว ล-เ-า แต่จะมีการอ่านเพียงแบบเดียวคือคำว่า เว-ลา เท่านั้น

ดังนั้นจะเห็นได้ว่า การถอดสระนั้นจะมีความถูกต้องมากน้อยเพียงไร สิ่งที่สำคัญที่จะช่วยให้มีความถูกต้องก็ต้องมาจากผลของการตัดคำ นั่นคือ การตัดคำให้ได้คำเดียว ก็จะทำให้ง่ายต่อการถอดสระ และง่ายต่อการแปลอักษรไทยเป็นอักษรไทยเบรลล์ด้วยเช่นกัน

สำหรับความถูกต้องของการตัดคำนั้น ก็ขึ้นอยู่กับการเลือกวิธีการเทคนิคที่มีความเหมาะสมกับงาน ซึ่งในปัจจุบันได้มีการวิจัย ถึงเรื่องการตัดคำในภาษาไทย อยู่มากมายหลายวิธี ซึ่งก็มีขั้นตอนและกระบวนการที่แตกต่างกันออกไปตามแต่วัตถุประสงค์ของวิธีการนั้นๆ โดยบทต่อไปจะกล่าวถึงเทคนิควิธีการตัดคำที่ใช้กันอยู่ในปัจจุบัน ตลอดจนถึงข้อดี ข้อเสียของแต่ละวิธี เพื่อหาแนวทางในการนำมาประยุกต์ใช้ในการแปลภาษาไทยเป็นภาษาไทยเบรลล์ต่อไป