การพัฒนาระบบการถ่ายถอดอักษรสำหรับตัวบทเบรลล์ไทยปนอังกฤษเป็นอักษรไทยและอังกฤษปกติโดยใช้วิธีการแบบผสม

นายวีรชัย อำพรไพบูลย์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาอักษรศาสตรดุษฎีบัณฑิต
สาขาวิชาภาษาศาสตร์ ภาควิชาภาษาศาสตร์
คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2559
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของวิทยานิพนธ์ตั้งแต่ปีการศึกษา 2554 ที่ให้บริการในคลังปัญญาจุฬาฯ (CUIR)
เป็นแฟ้มข้อมูลของนิสิตเจ้าของวิทยานิพนธ์ที่ส่งผ่านทางบัณฑิตวิทยาลัย
The abstract and full text of theses from the academic year 2011 in Chulalongkorn University Intellectual Repository(CUIR)
are the thesis authors' files submitted through the Graduate School.

A DEVELOPMENT OF TRANSLITERATION SYSTEM FOR MIXED THAI AND ENGLISH BRAILLE TEXTS INTO THAI AND ENGLISH PRINTS USING A HYBRID APPROACH

Mr. Weerachai Umpornpaiboon

A Dissertation Submitted in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy Program in Linguistics

Department of Linguistics

Faculty of Arts

Chulalongkorn University

Academic Year 2016

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การพัฒนาระบบการถ่ายถอดอักษรสำหรับตัวบทเบรลล์ไทยปนอังกฤษเป็นอักษรไทยและอังกฤษปกติโดยใช้วิธีการแบบผสม
โดย	นายวีรชัย อำพรไพบูลย์
สาขาวิชา	ภาษาศาสตร์
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	รองศาสตราจารย์ ดร.วิโรจน์ อรุณมานะกุล

คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาดุษฎีบัณฑิต

.....คณบดีคณะอักษรศาสตร์
(รองศาสตราจารย์ ดร.กิ่งกาญจน์ เทพกาญจนา)

คณะกรรมการสอบวิทยานิพนธ์

.....ประธานกรรมการ
(ดร.เทพชัย ทรัพย์นิธิ)

.....อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(รองศาสตราจารย์ ดร.วิโรจน์ อรุณมานะกุล)

.....กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.พิทยาวัฒน์ พิทยาภรณ์)

.....กรรมการ
(อาจารย์ ดร.วรรณชัย คำภีระ)

.....กรรมการภายนอกมหาวิทยาลัย
(อาจารย์ ดร.วีระแมน นิยมพล)

วีรชัย อำพรไพบูลย์ : การพัฒนาระบบการถ่ายถอดอักษรสำหรับตัวบทเบรลล์ไทยปนอังกฤษเป็นอักษรไทยและอังกฤษปกติโดยใช้วิธีการแบบผสม. (A DEVELOPMENT OF TRANSLITERATION SYSTEM FOR MIXED THAI AND ENGLISH BRAILLE TEXTS INTO THAI AND ENGLISH PRINTS USING A HYBRID APPROACH) อ.ที่ปรึกษาวิทยานิพนธ์หลัก : รศ.ดร.วิโรจน์ อรุณมานะกุล, 333 หน้า

วิทยานิพนธ์นี้ศึกษาระบบอักษรเบรลล์ไทย โดยวิเคราะห์อักขระในเบรลล์ไทยและเปรียบเทียบอักขระเบรลล์ไทยกับเบรลล์อังกฤษและเบรลล์ญี่ปุ่น เนื่องจากผู้วิจัยสันนิษฐานว่าอักขระเบรลล์ไทยพื้นฐานนำมาจากอักขระในสองภาษานี้ ผลการวิจัยแสดงให้เห็นว่า ในอักษรเบรลล์ไทยมีสระพื้นฐาน 5 รูป ที่มีรูปและเสียงตรงกับอักขระเบรลล์ญี่ปุ่น และมีพยัญชนะพื้นฐาน 15 รูป ที่มีรูปและเสียงใกล้เคียงกับอักขระเบรลล์อังกฤษ โดยมีการสร้างอักขระขึ้นเพิ่มเติมจากอักขระพื้นฐานเหล่านี้ด้วยการปรับเปลี่ยนจุดภายในเซลล์ และการเพิ่มอักขระเบรลล์รูปแบบเฉพาะหน้าหรือหลังอักขระที่ใช้เป็นฐาน

นอกจากนี้ วิทยานิพนธ์นี้ยังได้พัฒนาโปรแกรมสำหรับถ่ายถอดอักษรเบรลล์ไทยปนอังกฤษเป็นอักษรปกติ และเปรียบเทียบประสิทธิภาพของระบบที่ใช้วิธีการที่แตกต่างกัน 3 วิธี ได้แก่ การใช้กฎ การใช้แบบจำลองเอ็นแกรม และวิธีการแบบผสม โดยทดสอบกับข้อความเบรลล์ 5 ประเภท คือ ข้อความเบรลล์ไทย ข้อความเบรลล์อังกฤษรูปเต็ม ข้อความเบรลล์อังกฤษรูปย่อ ข้อความเบรลล์ไทยปนอังกฤษรูปเต็ม และข้อความเบรลล์ไทยปนอังกฤษรูปย่อ จากการทดลองพบว่า วิธีการแบบผสมถ่ายถอดอักษรได้ถูกต้องมากที่สุดสำหรับข้อความเบรลล์ 4 ประเภท ได้แก่ ข้อความเบรลล์ไทย 99.73 % ข้อความเบรลล์อังกฤษรูปเต็ม 100 % ข้อความเบรลล์อังกฤษรูปย่อ 99.92 % และข้อความเบรลล์ไทยปนอังกฤษรูปย่อ 98.69% นอกจากนี้ผลการทดลองยังแสดงให้เห็นว่า วิธีการถ่ายถอดอักษรทั้ง 3 วิธี ให้ผลความถูกต้องไม่แตกต่างกันมากนัก

ภาควิชา ภาษาศาสตร์ ลายมือชื่อนิสิต .....

สาขาวิชา ภาษาศาสตร์ ลายมือชื่อ อ.ที่ปรึกษาหลัก .....

ปีการศึกษา 2559

##5680506322 : MAJOR LINGUISTICS
KEY WORD : BRAILLE TRANSLITERATION / THAI BRAILLE / ENGLISH BRAILLE / RULED-BASED MODEL / N-GRAM MODEL / HYBRID MODEL

WEERACHAI UMPORNPAIBOON : A DEVELOPMENT OF TRANSLITERATION SYSTEM FOR MIXED THAI AND ENGLISH BRAILLE TEXTS INTO THAI AND ENGLISH PRINTS USING A HYBRID APPROACH. ADVISOR : ASSOC. PROF. WIROTE AROONMANAKUN, Ph.D., 333 pp.

This dissertation aims to investigate the Thai Braille writing system, which is believed to be influenced by English and Japanese braille. The findings show that Japanese Braille and Thai Braille share five vowels similar in form and sound, while English Braille and Thai Braille share fifteen consonants with similar sounds. These shared Braille forms are basic characters, which are then extended into other Braille characters representing vowels, consonants, and tones in Thai. The extensions are made through the alternation of dots within the cell or the addition of a preceding or succeeding Braille cell.

Another objective of this dissertation is to develop a program to transliterate Thai mixed with English Braille into print and compare the performances of the systems based on three methods: the rule-based, the N-gram and the hybrid. The systems were tested with five text categories, namely the Thai Braille, the uncontracted Braille, the contracted Braille, the Thai Braille mixed with uncontracted Braille and the Thai mixed with contracted Braille texts. The results reveal that the hybrid method achieves the highest accuracies for the four text categories. The results are as follows: 99.73 % for the Thai Braille texts, 100% for the uncontracted English Braille texts, 99.92% for the contracted English Braille texts and 98.69% for the Thai mixed with contracted English Braille texts. The study also shows that the three methods do not make significant differences in level of performance.

Department : Linguistics Student's Signature .....

Field of Study : Linguistics Advisor's Signature .....

Academic Year : 2016

กิตติกรรมประกาศ

ในการทำวิจัยครั้งนี้ ข้าพเจ้าได้แรงบันดาลใจจาก มิสเจเนวีฟ คอลฟิลด์ ผู้ก่อกำเนิดการศึกษาสำหรับคนตาบอดไทย ข้าพเจ้าจึงขอกราบขอบพระคุณท่านอย่างสุดซึ้งที่ทำให้งานวิจัยชิ้นนี้เกิดขึ้นได้

ข้าพเจ้าขอกราบขอบพระคุณ รองศาสตราจารย์ ดร.วิโรจน์ อรุณมานะกุล อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่ให้ความเมตตาประสิทธิ์ประสาทความรู้ทางด้านภาษาศาตร์และการเขียนโปรแกรมคอมพิวเตอร์ ซึ่งเป็นเครื่องมือที่สำคัญต่อการวิจัยครั้งนี้ อีกทั้งให้คำแนะนำอันเป็นประโยชน์ต่อการวิจัยอย่างสม่ำเสมอตลอดมา ทำให้งานวิจัยนี้สำเร็จลุล่วงไปได้

ข้าพเจ้าขอกราบขอบพระคุณ ดร.เทพชัย ทรัพย์นิธิ ผู้ช่วยศาสตราจารย์ ดร.พิทยาวัฒน์ พิทยาภรณ์ ดร.วรรณชัย คำภีระ และ ดร.วีระแมน นิยมพล คณะกรรมการสอบวิทยานิพนธ์ ที่ได้เมตตาให้คำแนะนำอันเป็นประโยชน์ ทำให้ข้าพเจ้าเกิดมุมมองที่ลึกซึ้งมากยิ่งขึ้นเกี่ยวกับงานวิจัยนี้

ข้าพเจ้าขอกราบขอบพระคุณ ศาสตราจารย์ ดร.อมรา ประสิทธิ์รัฐสินธุ์ ศาสตราจารย์ ดร.ธีระพันธ์ เหลืองทองคำ และ ผู้ช่วยศาสตราจารย์ยุพา ชวนปรีชา ที่ได้เมตตาให้ความรู้และคำแนะนำอันเป็นประโยชน์ต่อการวิจัยครั้งนี้ และขอกราบขอบพระคุณคณาจารย์ภาควิชาภาษาศาสตร์ทุกท่าน ที่ได้ประสิทธิ์ประสาทความรู้ และทักษะในการวิเคราะห์ทางภาษาศาสตร์อันเป็นรากฐานของการวิจัยแก่ข้าพเจ้า

ขอขอบคุณ คุณสุประภา อำพรไพบูลย์ ผู้ช่วยเก็บข้อมูลและจัดพิมพ์วิทยานิพนธ์ฉบับนี้ คุณศุภเดช อันจรัสวิชัย ที่ได้ให้คำแนะนำเกี่ยวกับการเขียนโปรแกรมการประมวลผลภาษาธรรมชาติ ศูนย์เทคโนโลยีการศึกษาเพื่อคนตาบอดที่ได้อนุญาตให้ใช้ข้อมูลอักษรเบรลล์ในการวิจัย และเจ้าหน้าที่เลขานุการภาควิชาภาษาศาสตร์ทุกท่านที่ให้ความช่วยเหลือเกื้อกูลแก่ข้าพเจ้าด้วยดีตลอดมา

ท้ายที่สุด ข้าพเจ้าขอกราบขอบพระคุณจุฬาลงกรณ์มหาวิทยาลัย ที่เป็นแหล่งประสิทธิ์ประสาทวิชา อันมีค่ายิ่งต่อข้าพเจ้า ทำให้ข้าพเจ้ามีชีวิตที่มีคุณค่าตราบเท่าถึงทุกวันนี้

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ง
บทคัดย่อภาษาอังกฤษ ..... จ
กิตติกรรมประกาศ ..... ฉ
สารบัญ ..... ช
สารบัญตาราง ..... ฎ
สารบัญภาพ ..... ฒ
บทที่ 1 บทนำ 1
1.1 ความเป็นมาและความสำคัญของปัญหา 1
1.2 วัตถุประสงค์ 4
1.3 สมมุติฐาน 4
1.4 เครื่องมือที่ใช้ในการวิจัย 5
1.5 ขอบเขตการวิจัย 5
1.6 ข้อมูลที่ใช้ในการวิจัย 5
1.7 ขั้นตอนการวิจัย 6
1.8 ประโยชน์ที่คาดว่าจะได้รับ 6
1.9 โครงสร้างวิทยานิพนธ์ 6
1.10 รายการนิยามศัพท์ที่เกี่ยวข้อง 7
บทที่ 2 ทบทวนวรรณกรรม 9
2.1 ความรู้ทั่วไปเกี่ยวกับอักษรเบรลล์ 9
2.1.1 อักษรเบรลล์อังกฤษ 10
2.1.2 อักษรเบรลล์ไทย 26
2.1.3 ตัวเลขเบรลล์ 31
2.2 การจำแนกภาษาด้วยคอมพิวเตอร์ 33
2.3 การถ่ายถอดอักษร 39
2.3.1 แนวคิดทั่วไปเกี่ยวกับการถ่ายถอดอักษร 39
2.3.2 แนวทางการถ่ายถอดอักษรด้วยคอมพิวเตอร์ 40
2.4 งานวิจัยที่เกี่ยวข้องกับการถ่ายถอดอักษรเบรลล์ไทยและเบรลล์อังกฤษ 52

2.4.1 การออกแบบและพัฒนาระบบถ่ายถอดอักษรไปกลับ อังกฤษผสมไทยเป็นเบรลล์ ...52
2.4.2 การใช้เอ็นแกรมช่วยในการเพิ่มประสิทธิภาพการตัดสินใจถ่ายถอดอักษรเบรลล์ ในกรณีที่มีคำควบกล้ำ สระประสมและอักษรเบรลล์ 2 เซลล์ ...57
2.5 โปรแกรม Duxbury Barille Translator ...60
บทที่ 3 ระบบหน่วยอักขระเบรลล์อังกฤษและเบรลล์ไทย...62
3.1 ระบบหน่วยอักขระเบรลล์อังกฤษ...62
3.1.1 รูปแบบของหน่วยอักขระเบรลล์อังกฤษ...62
3.1.2 รูปแบบของตัวเลขเบรลล์...64
3.1.3 รูปแบบของเครื่องหมายในเบรลล์อังกฤษ ...65
3.1.4 อักขรวิธีในเบรลล์อังกฤษ ...65
3.2 ระบบหน่วยอักขระเบรลล์ไทย...66
3.2.1 กรอบการวิเคราะห์ระบบหน่วยอักขระเบรลล์ไทย...66
3.2.2 รูปแบบของหน่วยอักขระเบรลล์ไทย...67
บทที่ 4 การออกแบบและพัฒนาระบบ การถ่ายถอดเบรลล์ไทยปนอังกฤษโดยใช้กฎ ...83
4.1 แนวคิดที่ใช้ในการออกแบบและพัฒนาระบบการถ่ายถอดเบรลล์ไทย ...83
4.1.1 ความไม่สอดคล้องระหว่างเบรลล์ไทยและไทยปกติ...83
4.1.2 ความกำกวมของเบรลล์เซลล์เดี่ยว ...87
4.1.3 ความกำกวมของเบรลล์เซลล์คู่...91
4.2 แนวคิดที่ใช้ในการออกแบบและพัฒนาระบบการถ่ายถอดเบรลล์อังกฤษ ...93
4.2.1 เบรลล์อังกฤษรูปเต็ม เบรลล์อังกฤษรูปย่อ และคอมพิวเตอร์เบรลล์...93
4.2.2 แนวคิดที่ใช้ในการถ่ายถอดเบรลล์อังกฤษรูปย่อ...97
4.3 แนวคิดที่ใช้ในการระบุภาษาในอักษรเบรลล์ ...106
4.3.1 การจำแนกภาษาโดยใช้สายอักขระที่เป็นเอกลักษณ์ของแต่ละภาษา...107
4.3.2 การจำแนกภาษาโดยอาศัยบริบทแวดล้อม ...107
4.3.3 การจำแนกภาษาโดยใช้ความยาวของสายอักขระ ...108
บทที่ 5 ระบบการถ่ายถอดเบรลล์ไทยปนอังกฤษโดยใช้กฎ ...109
5.1 ข้อมูลที่ใช้ในงานวิจัย...109
5.1.1 ข้อมูลที่ใช้วิเคราะห์เพื่อสร้างระบบการถ่ายถอดอักษร...109
5.1.2 ข้อมูลที่ใช้ในการทดสอบระบบ...132
5.2 ระบบการถ่ายถอดเบรลล์อักษรโดยใช้กฎ ...134

5.2.1 การระบุภาษา....136
5.2.2 การถ่ายถอดเบรลล์ไทย....137
5.2.3 การถ่ายถอดเบรลล์อังกฤษ....152
5.3 การทดสอบระบบการถ่ายถอดอักษรโดยใช้กฎ....170
5.3.1 การทดสอบระบบกับข้อความเบรลล์ไม่ปนภาษา....173
5.3.2 การทดสอบระบบกับข้อความเบรลล์ปนภาษา....174
5.4 อภิปรายผลการทดสอบระบบ....176
5.4.1 การระบุภาษา....176
5.4.2 การถ่ายถอดอักษร....177
บทที่ 6 ระบบการถ่ายถอดเบรลล์ไทยปนอังกฤษโดยใช้แบบจำลองเอ็นแกรม....183
6.1 การจัดทำข้อมูลเพื่อใช้กับระบบถ่ายถอดอักษร....183
6.1.1 การจัดทำข้อมูลสำหรับการถ่ายถอดเบรลล์ไทย....183
6.1.2 การจัดทำข้อมูลสำหรับการถ่ายถอดเบรลล์อังกฤษ....187
6.2 ระบบการถ่ายถอดอักษรโดยใช้แบบจำลองเอ็นแกรม....196
6.2.1 การสร้างชุดสายอักขระสำหรับภาษาไทยและภาษาอังกฤษ....200
6.2.2 การหาผลลัพธ์สำหรับเบรลล์ไทยปนอังกฤษรูปย่อ....204
6.2.3 การหาค่าสถิติสูงสุดของสายอักขระโดยใช้แบบจำลองไตรแกรมตัวอักษร....205
6.3 การทดสอบระบบการถ่ายถอดอักษรโดยใช้แบบจำลองเอ็นแกรม....206
6.3.1 การทดสอบระบบกับข้อความเบรลล์ไม่ปนภาษา....206
6.3.2 การทดสอบระบบกับข้อความเบรลล์ปนภาษา....207
6.4 อภิปรายผลการทดสอบระบบ....209
6.4.1 การระบุภาษา....209
6.4.2 การถ่ายถอดอักษร....211
บทที่ 7 ระบบการถ่ายถอดเบรลล์ไทยปนอังกฤษโดยใช้วิธีการแบบผสม....216
7.1 ข้อมูลที่ใช้ในงานวิจัย....216
7.1.1 ข้อมูลที่ใช้กับระบบถ่ายถอดอักษร....216
7.1.2 ข้อมูลที่ใช้ในการทดสอบระบบ....218
7.2 ระบบการถ่ายถอดอักษรโดยใช้วิธีการแบบผสม....218
7.2.1 การระบุภาษา....220
7.2.2 การสร้างชุดสายอักขระสำหรับภาษาไทย....221

7.2.3 การสร้างชุดสายอักขระสำหรับภาษาอังกฤษ 224
7.2.4 การหาค่าสถิติสูงสุดด้วยแบบจำลองเอ็นแกรม 225
7.3 การทดสอบระบบการถ่ายถอดอักษรโดยใช้วิธีการแบบผสม 225
7.3.1 การทดสอบระบบกับข้อความเบรลล์ไม่ปนภาษา 225
7.3.2 การทดสอบระบบกับข้อความเบรลล์ปนภาษา 227
7.4 อภิปรายผลการทดสอบระบบ 228
7.4.1 การระบุภาษา 228
7.4.2 การถ่ายถอดอักษร 229
7.5 ผลการเปรียบเทียบวิธีการถ่ายถอดอักษรทั้ง 3 วิธี 230
7.5.1 ผลการเปรียบเทียบการระบุภาษา 230
7.5.2 ผลการเปรียบเทียบการถ่ายถอดอักษร 231
บทที่ 8 สรุปผล อภิปรายผล และข้อเสนอแนะ 234
8.1 สรุปผล 234
8.1.1 หน่วยอักขระเบรลล์ไทยและหน่วยอักขระเบรลล์อังกฤษ 235
8.1.2 ระบบการถ่ายถอดเบรลล์ไทยปนอังกฤษ 237
8.2 อภิปรายผล 239
8.3 ข้อเสนอแนะ 241
รายการอ้างอิง 246
ภาคผนวก 251
ประวัติผู้เขียนวิทยานิพนธ์ 318

สารบัญตาราง

หน้า

ตารางที่ 2.1 ตัวอักษรเบรลล์อังกฤษ 11
ตารางที่ 2.2 การใช้ตัวย่ออักษรโดด 13
ตารางที่ 2.3 รายการตัวย่ออักษรโดด 14
ตารางที่ 2.4 วิธีการใช้ตัวย่อรูปเต็มคงรูป 15
ตารางที่ 2.5 รายการตัวย่อรูปเต็มคงรูป 16
ตารางที่ 2.6 รายการตัวย่อรูปเต็มกลายรูป 16
ตารางที่ 2.7 รายการตัวย่อประกอบคำ 17
ตารางที่ 2.8 วิธีการใช้ตัวย่อประกอบคำ 17
ตารางที่ 2.9 รายการตัวย่อกลางคำหรือท้ายคำ 18
ตารางที่ 2.10 วิธีการใช้ตัวย่อกลางคำหรือท้ายคำ 18
ตารางที่ 2.11 รายการตัวย่อเบรลล์ต่ำย่อคำโดด 19
ตารางที่ 2.12 รายการตัวย่อเบรลล์ต่ำที่ย่อคำโดดพิเศษ 19
ตารางที่ 2.13 รายการตัวย่อเบรลล์ต่ำย่อส่วนของคำ 20
ตารางที่ 2.14 รายการตัวย่อหน่วยหลัก 21
ตารางที่ 2.15 การใช้ตัวย่อหน่วยหลัก 21
ตารางที่ 2.16 รายการตัวย่อหน่วยเสริม 22
ตารางที่ 2.17 รายการคำย่อเบรลล์ 24
ตารางที่ 2.18 การใช้คำย่ออักษรเบรลล์ 25
ตารางที่ 2.19 พยัญชนะในเบรลล์ไทย 27
ตารางที่ 2.20 ตัวอักษรเบรลล์ไทยพื้นฐาน 15 ตัว 28
ตารางที่ 2.21 สระและวรรณยุกต์ในเบรลล์ไทย 29
ตารางที่ 2.22 ตัวเลขในอักษรเบรลล์ 31
ตารางที่ 2.23 เครื่องหมายที่ปรากฏกับตัวเลข 32
ตารางที่ 2.24 เบรลล์เซลล์เดี่ยวที่กำกวม 55
ตารางที่ 2.25 เบรลล์เซลล์คู่ที่กำกวม 57

ตารางที่ 2.26 ตัวอย่างข้อความไทยที่โปรแกรม DUXBURY BRAILLE TRANSLATOR ถ่ายถอดผิด 60
ตารางที่ 2.27 ตัวอย่างข้อความไทยปนอังกฤษที่โปรแกรม DUXBURY BRAILLE TRANSLATOR ถ่ายถอดผิด 61
ตารางที่ 3.1 หน่วยอักขระเบรลล์อังกฤษพื้นฐาน.....63
ตารางที่ 3.2 หน่วยอักขระเบรลล์อังกฤษดัดแปลงที่ได้จากการเติม ‘⠸’ ภายในเซลล์.....63
ตารางที่ 3.3 หน่วยอักขระเบรลล์อังกฤษดัดแปลงที่ได้จากการเติม‘⠿’ ภายในเซลล์.....64
ตารางที่ 3.4 ตัวเลขเบรลล์ 1 - 0 64
ตารางที่ 3.5 เครื่องหมายในเบรลล์อังกฤษ.....65
ตารางที่ 3.6 อักขระเบรลล์ที่มีรูปซ้ำกันระหว่างภาษาไทยและญี่ปุ่น หรือไทยและอังกฤษ 68
ตารางที่ 3.7 อักขระในเบรลล์ญี่ปุ่นและเบรลล์ไทยที่อยู่ในกลุ่มเบรลล์อังกฤษพื้นฐาน 70
ตารางที่ 3.8 รูปของสระในเบรลล์ไทยและเบรลล์ญี่ปุ่นที่มีเสียงใกล้เคียงกัน 70
ตารางที่ 3.9 พยัญชนะเบรลล์ไทยที่มีรูปและเสียงตรงกับเบรลล์อังกฤษ.....71
ตารางที่ 3.10 การสร้างสระขึ้นเพิ่มเติมโดยการเติม ‘⠿’ (จุด 6).....73
ตารางที่ 3.11 การสร้างอักษรโดยการเลื่อนตำแหน่งจุดเบรลล์ให้ต่ำลง 73
ตารางที่ 3.12 การสร้างสระจาก ⠇ (= ะ).....74
ตารางที่ 3.13 อักขระที่เกิดจากการปรับเปลี่ยนจุดเบรลล์แบบอื่น ๆ.....75
ตารางที่ 3.14 การสร้างสระสั้นโดยการเพิ่มอักขระ ‘⠇’ (จุด 1) 76
ตารางที่ 3.15 การสร้างสระเพิ่มเติมโดยการเพิ่มอักขระ ‘⠯’ (จุด 2).....76
ตารางที่ 3.16 พยัญชนะเบรลล์ไทยที่มีรูปและเสียงตรงกับเบรลล์อังกฤษจำแนกตามหมู่อักษร 77
ตารางที่ 3.17 พยัญชนะเบรลล์ไทยที่เกิดจากการดัดแปลงพยัญชนะเบรลล์อังกฤษ.....78
ตารางที่ 3.18 การสร้างพยัญชนะใหม่ที่มีหมู่อักษรต่างจากพยัญชนะฐาน 79
ตารางที่ 3.19 การสร้างพยัญชนะใหม่ที่มีหมู่ตัวอักษรและเสียงตรงกับพยัญชนะฐาน 80
ตารางที่ 5.1 จำนวนสายอักขระที่เป็นเอกลักษณ์ของแต่ละภาษา 112
ตารางที่ 5.2 เครื่องหมายเบรลล์ต้นสายอักขระ 112
ตารางที่ 5.3 คู่เทียบเบรลล์เซลล์เดี่ยวทั่วไป 113
ตารางที่ 5.4 คู่เทียบสระทั่วไป.....114
ตารางที่ 5.5 คู่เทียบสระพิเศษ 114
ตารางที่ 5.6 คู่เทียบสระหน้า.....115

ตารางที่ 5.7 คู่เทียบวรรณยุกต์....115
ตารางที่ 5.8 คู่เทียบเบรลล์เซลล์คู่ทั่วไป....116
ตารางที่ 5.9 คู่เทียบสระประสม....117
ตารางที่ 5.10 คู่เทียบเบรลล์เซลล์เดี่ยวที่กำกวมในตำแหน่งทั่วไป....118
ตารางที่ 5.11 คู่เทียบเบรลล์ที่กำกวมท้ายสายอักขระ....118
ตารางที่ 5.12 คู่เทียบสำหรับเบรลล์เซลล์คู่ที่กำกวม....119
ตารางที่ 5.13 เซ็ตพยัญชนะตัวสะกดท้ายพยางค์สำหรับการถ่ายถอด ‘ ’ให้เป็น “เ ิ”....120
ตารางที่ 5.14 เซ็ตของพยัญชนะควบกำกวมและพยัญชนะควบไม่กำกวม....121
ตารางที่ 5.15 คู่เทียบเบรลล์ต้นสายอักขระ....124
ตารางที่ 5.16 คู่เทียบเครื่องหมายหลังคำ....125
ตารางที่ 5.17 คู่เทียบตัวย่อต้นคำ....126
ตารางที่ 5.18 คู่เทียบตัวย่อกลางคำ....127
ตารางที่ 5.19 คู่เทียบอักษรเบรลล์หลังตัวย่อระดับคำ....128
ตารางที่ 5.20 คู่เทียบเครื่องหมายที่ใช้คั่นระหว่างคำ....129
ตารางที่ 5.21 คู่เทียบตัวย่อรูปเต็มคงรูป....129
ตารางที่ 5.22 คู่เทียบตัวย่อเบรลล์เซลล์เดี่ยวทั่วไป....130
ตารางที่ 5.23 คู่เทียบตัวเลขเบรลล์....131
ตารางที่ 5.24 ปริมาณของข้อความเบรลล์แต่ละประเภทที่นำมาทดสอบระบบ....134
ตารางที่ 5.25 การทดสอบประสิทธิภาพของกระบวนการต่าง ๆ ในข้อมูลที่ไม่มีการปนภาษา....171
ตารางที่ 5.26 การทดสอบประสิทธิภาพของกระบวนการต่าง ๆ ในข้อมูลปนภาษา....172
ตารางที่ 5.27 ผลการระบุภาษาในข้อความเบรลล์ไม่ปนภาษา....173
ตารางที่ 5.28 ผลการถ่ายถอดอักษรในข้อความเบรลล์ไม่ปนภาษา....174
ตารางที่ 5.29 ผลการระบุภาษาที่ทดสอบกับข้อความเบรลล์ปนภาษา....175
ตารางที่ 5.30 ผลการถ่ายถอดอักษรเฉพาะสายอักขระที่ระบุภาษาได้ถูกต้อง....175
ตารางที่ 5.31 ผลการถ่ายถอดอักษรด้วยการใช้กฎโดยภาพรวม....176
ตารางที่ 6.1 ผลการทดสอบประสิทธิภาพของการระบุภาษาโดยใช้แบบจำลองเอ็นแกรม....207
ตารางที่ 6.2 ผลการทดสอบประสิทธิภาพของการถ่ายถอดอักษรโดยใช้แบบจำลองเอ็นแกรม....207
ตารางที่ 6.3 ผลการระบุภาษาของระบบถ่ายถอดอักษรที่ทดสอบกับข้อมูลปนภาษา....208

ตารางที่ 6.4 ผลการถ่ายถอดอักษรเฉพาะข้อความที่ระบุภาษาถูกต้อง 208
ตารางที่ 6.5 ผลการถ่ายถอดอักษรด้วยการใช้แบบจำลองเอ็นแกรมโดยภาพรวม.....209
ตารางที่ 7.1 ผลการทดสอบประสิทธิภาพของการระบุภาษาโดยใช้วิธีการแบบผสม.....226
ตารางที่ 7.2 ผลการทดสอบประสิทธิภาพของการถ่ายถอดอักษรโดยใช้วิธีการแบบผสม 226
ตารางที่ 7.3 ผลการระบุภาษาของระบบที่ทดสอบกับข้อมูลปนภาษา 227
ตารางที่ 7.4 ผลการถ่ายถอดอักษรเฉพาะข้อความที่ระบุภาษาได้ถูกต้อง.....227
ตารางที่ 7.5 ผลการถ่ายถอดอักษรด้วยการใช้วิธีการแบบผสมโดยภาพรวม.....228
ตารางที่ 7.6 ประสิทธิภาพของการระบุภาษาด้วยวิธีการทั้ง 3 วิธี 230
ตารางที่ 7.7 ประสิทธิภาพของการถ่ายถอดอักษรด้วยวิธีการทั้ง 3 วิธี.....231

สารบัญภาพ

หน้า

ภาพที่ 2.1 การปรากฏของจุดในอักษรเบรลล์.....10
ภาพที่ 2.2 วิธีการ WEIGHTED FINITE-STATE TRANSDUCER 47
ภาพที่ 2.3 โครงสร้างแบ็กพรอพาเกชันนิวรอลเน็ตเวิร์กที่ใช้ในการเรียนรู้คำไทย 49
ภาพที่ 4.1 การจัดเรียงสระพิเศษในอักษรเบรลล์และในอักษรปกติ.....84
ภาพที่ 4.2 การจัดเรียงสระหน้าในอักษรเบรลล์และในอักษรปกติ.....85
ภาพที่ 5.1 ภาพรวมของระบบถ่ายถอดเบรลล์ไทยปนอังกฤษโดยใช้กฎ.....136
ภาพที่ 5.2 กระบวนการถ่ายถอดเบรลล์ไทยโดยใช้กฎ 139
ภาพที่ 5.3 กระบวนการถ่ายถอดเบรลล์อังกฤษโดยใช้กฎ 153
ภาพที่ 5.4 ขั้นตอนในการถ่ายถอดเบรลล์อังกฤษรูปย่อ.....164
ภาพที่ 6.1 ภาพรวมของระบบถ่ายถอดเบรลล์ไทยปนอังกฤษรูปเต็มโดยใช้
แบบจำลองเอ็นแกรม 197
ภาพที่ 6.2 ภาพรววมของระบบถ่ายถอดเบรลล์ไทยปนอังกฤษรูปย่อโดยใช้
แบบจำลองเอ็นแกรม 199
ภาพที่ 7.1 ภาพรวมของระบบถ่ายถอดเบรลล์ไทยปนอังกฤษโดยใช้วิธีการแบบผสม 219

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ผู้พิการทางสายตามีความจำเป็นที่ต้องรับข้อมูลจากสื่อต่าง ๆ เช่นเดียวกับคนสายตาปกติทั่วไป แต่เนื่องจากไม่สามารถจะรับรู้อักษรผ่านการมองเห็นได้ จึงจำเป็นที่จะต้องมีระบบอักษรเบรลล์ ซึ่งเป็นระบบที่แต่ละเซลล์ประกอบด้วยจุดนูนไม่เกินหกจุด (⠿) มาเรียงต่อกัน โดยที่การปรากฏของจุดในตำแหน่งที่ต่างกันใช้แสดงอักขระที่ต่างกัน ทำให้คนที่มองไม่เห็นสามารถรับรู้ข้อมูลได้โดยการสัมผัสด้วยปลายนิ้วมือ

เป็นที่ทราบกันดีว่า อักษรเบรลล์มีต้นกำเนิดมาจากประเทศฝรั่งเศสซึ่งใช้ระบบตัวอักษรร่วมกับภาษาอื่นๆที่ใช้อักษรโรมัน และมีการศึกษาวิจัยเกี่ยวกับตัวอักษรเบรลล์ในภาษาเหล่านี้อย่างกว้างขวางโดยมีคำอธิบายเกี่ยวกับการสร้างตัวอักษรเบรลล์อย่างเป็นระบบ ตัวอักษรเบรลล์ที่ใช้อักษรโรมันเกิดจากการนำเอาอักษรเบรลล์ 10 ตัวแรกคือ ‘⠁’ - ‘⠊’ (=a-j) มาใช้เป็นอักษรหลักเพื่อสร้างอักษรเบรลล์ตัวอื่น ๆ ในระบบ (Englebretson, R. 2009: 3)

ในการเรียนการสอนเบรลล์ไทยนั้น มีตำราสอนอ่านเขียนเบรลล์ไทยเกิดขึ้นบ้างแต่ตำราเหล่านี้เน้นการท่องจำรหัสอักษรเบรลล์แต่ละตัว ในปัจจุบันยังไม่มีเอกสารหรืองานวิจัยชิ้นใดที่อธิบายเบรลล์ไทยอย่างเป็นระบบ แต่ในความเป็นจริงแล้วในระบบเบรลล์ไทยเองนั้นอักษรตัวหนึ่งก็น่าจะมีความสัมพันธ์กับอักษรตัวอื่นด้วยในลักษณะที่คล้ายคลึงกับระบบในภาษาอื่น เช่น '⠎' ใช้แทน “ส” '⠵' ใช้แทน “ซ” พยัญชนะสองตัวนี้มีเสียงคล้ายคลึงกัน ต่างกันตรงที่ “ส” เป็นอักษรสูง ในขณะที่ “ซ” เป็นอักษรต่ำ และผู้วิจัยยังสังเกตเห็นว่าเบรลล์ไทยก็มีความสัมพันธ์กับเบรลล์อังกฤษด้วย

เหตุที่ว่าเบรลล์แต่ละเซลล์เกิดจากการเรียงต่อของจุดไม่เกินหกจุด ทำให้อักษรเบรลล์ในภาษาหนึ่งเกิดการซ้ำรูปกับอีกภาษาหนึ่งอย่างหลีกเลี่ยงไม่ได้ หากพิจารณารูปซ้ำของเบรลล์อังกฤษและเบรลล์ไทย จะพบว่ามีรูปซ้ำจำนวนหนึ่งที่ใช้แทนเสียงเดียวกันหรือเสียงที่ใกล้เคียงกันทั้งสองภาษา เช่น “s” และ “ส” มีคุณสมบัติทางเสียงใกล้เคียงกันมากจึงใช้ '⠎' แทนตัวอักษรทั้งสองตัว

หลักฐานข้างต้นแสดงให้เห็นว่าเบรลล์ไทยมีรูปที่ปรากฏอย่างเป็นระบบ จึงน่าสนใจว่าเบรลล์ไทยมีความสัมพันธ์กับเบรลล์อังกฤษหรือไม่อย่างไรและเบรลล์ไทยแต่ละตัวมีความสัมพันธ์

กันหรือไม่ ในลักษณะใด วิทยานิพนธ์ฉบับนี้จึงมุ่งที่จะหาคำตอบในประเด็นเหล่านี้ ซึ่งเป็นประเด็นพื้นฐานทางภาษาศาสตร์สำหรับการเรียนการสอนอักษรเบรลล์ไทย

นอกจากนี้ พัฒนาการของเทคโนโลยีทำให้มีการใช้อักษรเบรลล์สำหรับภาษาต่างๆในคอมพิวเตอร์อย่างแพร่หลาย แต่อักษรเบรลล์มีการซ้ำรูประหว่างภาษาจึงยากที่เครื่องคอมพิวเตอร์จะบอกได้ว่ากำลังทำงานกับอักษรเบรลล์ของภาษาใดอยู่ แม้ว่าจะมีโปรแกรมอยู่จำนวนหนึ่งที่สามารถถ่ายถอดอักษรเบรลล์เป็นอักษรปกติได้ แต่น่าเสียดายว่าในการถ่ายถอดอักษรนั้น ต้องให้ผู้ใช้กำกับภาษาเพื่อบอกให้คอมพิวเตอร์รู้ว่ากำลังทำงานกับภาษาใด จึงทำให้ในปัจจุบันยังไม่มีโปรแกรมใดที่รองรับการถ่ายถอดข้อความเบรลล์ไทยปนอังกฤษเป็นอักษรปกติโดยอัตโนมัติได้

เพื่อแก้ปัญหาดังกล่าว ได้มีนักวิจัยไทยพยายามเสนอวิธีการจำแนกภาษาของอักษรเบรลล์ดังที่ปรากฏในงานวิจัยของรวิศม์ วงษ์สมาน (2551) ในส่วนหนึ่งของการวิจัย รวิศม์ได้ศึกษาการถ่ายถอดอักษรเบรลล์เป็นอักษรปกติ โดยใช้แบบจำลองเอ็นแกรมเข้ามาช่วยในการจำแนกอักษรเบรลล์ของภาษาไทยและภาษาอังกฤษออกจากกันให้ได้ ก่อนที่จะทำการถ่ายถอดแต่ละภาษาเป็นอักษรปกติ รวิศม์พบว่ามีข้อผิดพลาดที่เกิดจากการระบุภาษาผิด เช่นการถ่ายถอด ‘⠇⠁⠉⠗’ ซึ่งควรเป็นคำว่า “ละคร” ในภาษาไทย แต่กลับได้เป็น "laur" ในภาษาอังกฤษ (รวิศม์ วงษ์สมาน, 2551: 60)

เป็นที่น่าสังเกตว่าในการวิจัย รวิศม์ได้ใช้ข้อมูลอักษรเบรลล์ในการฝึกระบบ ซึ่งก่อให้เกิดความยุ่งยากหลายประการ ประการแรก การใช้ข้อมูลอักษรเบรลล์ฝึกระบบทำให้สิ้นเปลืองเวลาและทรัพยากร เพราะในการใช้คลังข้อมูลอักษรเบรลล์ต้องอาศัยผู้เชี่ยวชาญตรวจสอบข้อมูลอักษรเบรลล์ให้ถูกต้องเสียก่อน มิฉะนั้นแล้วจะทำให้คอมพิวเตอร์ประมวลผลได้ไม่ถูกต้อง

นอกจากนี้การใช้ข้อมูลอักษรเบรลล์มาฝึกระบบ ยังส่งผลให้ไม่สามารถแยกภาษาไทยและอังกฤษออกจากกันได้อย่างเป็นเอกเทศ เนื่องจากอักษรเบรลล์ทั้ง 2 ภาษานี้ต่างก็ใช้ระบบหกจุด ทำให้สายอักขระบางส่วนในคลังข้อมูลภาษาหนึ่งตรงกับสายอักขระในคลังข้อมูลของอีกภาษาหนึ่งได้ หากข้อมูลที่ต้องการถ่ายถอดตรงกับสายอักขระที่ตรงกันดังกล่าว ทำให้ไม่สามารถบ่งชี้ว่าเป็นภาษาใดได้ทันที เมื่อพิจารณาสิ่งที่กล่าวมาข้างต้นจะเห็นได้ว่าคอมพิวเตอร์ไม่สามารถบ่งชี้ว่า ‘⠇⠁⠉⠗’ เป็น "ละคร" หรือ"laur" เนื่องจาก ‘⠇⠁⠉⠗’ ปรากฏอยู่ในคลังข้อมูลทั้งภาษาไทยและภาษาอังกฤษ

ดังนั้น วิทยานิพนธ์ฉบับนี้จึงมุ่งที่จะนำเสนอวิธีการจำแนกเบรลล์อังกฤษและเบรลล์ไทยออกจากกัน โดยใช้กฎทางภาษาร่วมกับแบบจำลองเอ็นแกรม ที่ผ่านการฝึกระบบด้วยคลังข้อมูลอักษรปกติของทั้ง 2 ภาษา แทนที่จะใช้คลังข้อมูลอักษรเบรลล์โดยตรง

ในการถ่ายถอดเบรลล์ที่มีการปนกันระหว่างภาษาไทยและอังกฤษ นอกจากปัญหาในการจำแนกทั้งสองภาษาออกจากกันแล้ว อักษรเบรลล์ไทยเองก็ยังพบปัญหาความกำกวมในการถ่ายถอดให้เป็นไทยปกติ เหตุที่เป็นเช่นนี้เพราะในเบรลล์ไทยมีการนำเบรลล์หนึ่งเซลล์มาแทนอักขระ

หลายตัวในการเขียนสระประสม เช่น ‘⠽’ ใช้แทน “ัว” หรือ ‘⠿’ ใช้แทน “เื อ” เป็นต้น ในการเขียนเบรลล์ไทยจะนำสระประสมไปวางไว้หลังพยัญชนะต้น เช่นคำว่า “เรียน” ในเบรลล์ไทยจะเขียนว่า ‘⠗⠑⠽⠝’ (=ร-เ◌ีย-น) คำว่า “เรื่อง” เขียนว่า ‘⠗⠑⠿⠝⠷⠛’ (=ร-เ◌ื อ-◌่-ง) เป็นต้น ปัญหาความกำกวมจะเกิดขึ้นในกรณีที่พยัญชนะควบกล้ำเกิดหน้าสระประสมเหล่านี้ ดังจะเห็นได้จากงานของ รวิศม์ วงษ์สมาน (2551) ที่ศึกษาการถ่ายถอดเบรลล์ไทยแบบไฟล์ต่อไฟล์โดยใช้กฎทางภาษา และวรพล ทินกรสูติบุตร (2551) ที่ศึกษาการถ่ายถอดเบรลล์ไทยแบบทันทีทันใดโดยใช้กฎ ทั้งรวิศม์และวรพล พบปัญหาในการถ่ายถอดสระที่ตัวอักษรนำหน้าสองตัวมีรูปคล้ายพยัญชนะควบกล้ำ แต่ไม่ใช่พยัญชนะควบกล้ำ เช่น ‘⠃⠷⠓⠔⠔⠕⠗⠞⠁⠇⠁⠵’ (บ-◌่-อ-ห-◌ิ-น-น-อ-ก-ล-แ-ะ) คือ “บ่อหินนอกและ” ถูกถ่ายถอดเป็น “บ่อหินนอแกละ” เนื่องจากคอมพิวเตอร์ตีความให้ “กล” เป็นพยัญชนะควบกล้ำ แต่ที่จริง “ก” เป็นตัวสะกดของคำแรก และ “ล” เป็นพยัญชนะต้นของคำถัดมา

ทศวัฒน์ ชุณหวิทยะธีระ (2555) ได้พยายามแก้ปัญหาดังกล่าว โดยเสนอให้มีการตัดคำในอักษรเบรลล์ด้วยเทคนิคเอ็นแกรมที่ผ่านการฝึกระบบด้วยคลังข้อมูลอักษรเบรลล์ แล้วนำคำเบรลล์แต่ละคำมาเทียบกับพจนานุกรมเบรลล์เพื่อถ่ายถอดเป็นอักษรปกติ ทศวัฒน์พบว่าวิธีการนี้สามารถแก้ปัญหาพยัญชนะที่กำกวมได้ในระดับหนึ่ง แต่วิธีการนี้ไม่เหมาะสมในการนำมาใช้กับแฟ้มข้อมูลที่มีขนาดใหญ่ เพราะต้องใช้เวลาในการประมวลผลมากอีกทั้งวิธีการนี้ต้องอาศัยทั้งพจนานุกรมและคลังข้อมูลเบรลล์ที่ถูกต้องสมบูรณ์จึงจะประมวลผลได้ถูกต้อง ซึ่งเป็นไปได้ยากเนื่องจากต้องใช้เวลาและมีค่าใช้จ่ายสูง

ในงานวิจัยนี้ ผู้วิจัยจึงขอเสนอวิธีการถ่ายถอดอักษรเบรลล์ไทยปนอังกฤษโดยใช้คลังข้อความภาษาไทยและภาษาอังกฤษเป็นหลัก ร่วมกับการใช้กฎทางภาษาและความน่าจะเป็นทางสถิติ ซึ่งจะนำเสนอต่อไป

นอกจากนี้ในการใช้เบรลล์ไทยปนอังกฤษนั้น ผู้วิจัยยังพบว่าเบรลล์อังกฤษมีการใช้ใน 2 รูปแบบ คือ เบรลล์อังกฤษรูปเต็ม (English braille grade 1) และ เบรลล์อังกฤษรูปย่อ (English braille grade 2) ซึ่งงานวิจัยการถ่ายถอดเบรลล์ไทยปนอังกฤษที่ผ่านมาศึกษาเฉพาะเบรลล์อังกฤษรูปเต็มเท่านั้น แต่ในความเป็นจริงมีการใช้ภาษาอังกฤษทั้งสองรูปแบบ โดยเฉพาะหากเป็นการจดบันทึกจะนิยมใช้รูปย่อมากกว่า งานวิจัยนี้จึงสนใจศึกษาการถ่ายถอดเบรลล์ไทยปนอังกฤษทั้งรูปเต็มและรูปย่อเพราะจะเป็นประโยชน์ต่อคนตาบอดในการนำไปใช้งานมากยิ่งขึ้น

สำหรับเบรลล์ไทยรูปย่อ (Thai Braille grade 2) ได้มีการคิดค้นโดยกลุ่มคนในสถานศึกษาเพื่อคนตาบอด ด้วยหวังว่าจะช่วยให้คนตาบอดไทยจดบันทึกด้วยอักษรเบรลล์ได้สะดวกและรวดเร็ว แต่ยังไม่มีผู้ใดศึกษาความเหมาะสมในการใช้อักษรเบรลล์ระบบนี้เลย จากการศึกษาเบื้องต้นผู้วิจัยได้ทำการสุ่มตัวอย่างโดยสัมภาษณ์คนตาบอดกลุ่มหนึ่งในสถานศึกษา พบว่า

เบรลล์ไทยรูปย่อที่ประดิษฐ์ขึ้นใหม่นี้มีความซับซ้อนในการใช้เป็นอย่างมาก ตัวอย่างเช่นการนำเบรลล์ 1 เซลล์มาใช้ย่อส่วนของคำ เช่น '⠒' แทนส่วนของคำ "-าน" โดยวางสัญลักษณ์นี้ไว้หน้าพยัญชนะต้นของพยางค์ อีกทั้งคำที่ปรากฏในเบรลล์ไทยรูปย่อก็ไม่ได้เกิดจากการศึกษาวิจัยอย่างเป็นระบบ คำย่อหลายคำมีความถี่ในการใช้น้อยมาก เช่น '⠩⠱⠩⠱' (ศ-ฐ) ใช้แทน "เศรษฐี" '⠏⠏⠒' (พพ-) ใช้แทน "พหูพจน์" และ '⠏⠊⠊⠊' (ป-ชช) ใช้แทน "ประชาชาติ" ฯลฯ เบรลล์ไทยรูปย่อจึงไม่เป็นที่นิยมของกลุ่มคนตาบอดไทยที่ผู้วิจัยทำการสัมภาษณ์ นอกจากนี้ผู้วิจัยได้ทำการสุ่มตัวอย่างเพื่อหาว่ามีการใช้เบรลล์ไทยรูปย่อในสื่อสิ่งพิมพ์อักษรเบรลล์บ้างหรือไม่ โดยสำรวจจากห้องสมุดอักษรเบรลล์หลายแห่ง พบว่ายังไม่มีการใช้อักษรเบรลล์รูปแบบนี้ในเอกสารที่ทำการสุ่มแต่ประการใด ด้วยเหตุที่กล่าวมาข้างต้น ผู้วิจัยจึงไม่นำเบรลล์ไทยรูปย่อมาใส่ไว้ในขอบเขตของงานวิจัยครั้งนี้

1.2 วัตถุประสงค์

1. วิเคราะห์ระบบอักษรเบรลล์ไทยและระบบอักษรเบรลล์อังกฤษ

2. พัฒนาระบบการถ่ายถอดอักษรสำหรับตัวบทเบรลล์ไทยปนอังกฤษโดยการใช้กฎทางภาษา ร่วมกับแบบจำลองเอ็นแกรม (n-gram)

1.3 สมมุติฐาน

1. หน่วยอักขระเบรลล์ไทยสามารถจำแนกเป็น 3 ประเภท ได้แก่ หน่วยอักขระเบรลล์ไทยพื้นฐาน หน่วยอักขระเบรลล์ไทยดัดแปลง และหน่วยอักขระเบรลล์ไทยจำเพาะ หน่วยอักขระเบรลล์ไทยพื้นฐานแต่ละตัวมีรูปและเสียงสอดคล้องกับหน่วยอักขระเบรลล์ในภาษาอังกฤษและภาษาญี่ปุ่น หน่วยอักขระเบรลล์ไทยดัดแปลงแต่ละตัวได้มาจากการปรับหน่วยอักขระเบรลล์ไทยพื้นฐานที่มีเสียงตรงกันโดยการกำกับด้วยเครื่องหมาย (glyph) และหน่วยอักขระเบรลล์ไทยจำเพาะเป็นหน่วยอักขระเบรลล์ที่ปรากฏเฉพาะภาษาไทยเท่านั้น

2. หน่วยอักขระเบรลล์อังกฤษสามารถจำแนกเป็น 2 ประเภท ได้แก่ หน่วยอักขระเบรลล์อังกฤษพื้นฐาน และหน่วยอักขระเบรลล์อังกฤษดัดแปลง หน่วยอักขระเบรลล์อังกฤษพื้นฐานประกอบด้วยหน่วยอักขระเบรลล์ 10 ตัวแรกในภาษาอังกฤษ (a - j) และหน่วยอักขระเบรลล์อังกฤษดัดแปลงได้มาจากการกำกับหน่วยอักขระเบรลล์อังกฤษพื้นฐานแต่ละตัวด้วยเครื่องหมาย (glyph)

3. ระบบการถ่ายถอดอักษรสำหรับตัวบทเบรลล์ไทยปนอังกฤษเป็นอักษรปกติที่อาศัยกฎทางภาษาร่วมกับแบบจำลองเอ็นแกรม จะให้ผลลัพธ์ที่ถูกต้องมากกว่าการใช้วิธีการใดวิธีการหนึ่ง

1.4 เครื่องมือที่ใช้ในการวิจัย

1. เครื่องไมโครคอมพิวเตอร์ส่วนบุคคล
2. โปรแกรมภาษา Perl ของบริษัท Active Perl
3. เครื่อง Braille Display ยี่ห้อ Focus ของบริษัท Freedom Scientific

1.5 ขอบเขตการวิจัย

1. ศึกษาการถ่ายถอดเบรลล์ไทยรูปเต็ม (Thai Braille Grade1) เบรลล์อังกฤษรูปเต็ม(English Braille Grade1) และเบรลล์อังกฤษรูปย่อ (English Braille Grade2) เป็นอักษรปกติ

2. ศึกษาเฉพาะสัญลักษณ์เบรลล์ที่ปรากฏในข้อความทั่วไป ไม่รวมถึงสัญลักษณ์ที่ปรากฏเฉพาะด้าน เช่น ดนตรี และคณิตศาสตร์

1.6 ข้อมูลที่ใช้ในการวิจัย

ข้อมูลที่ใช้ในงานวิจัยนี้ แบ่งออกเป็น 2 กลุ่ม ได้แก่

1.6.1 ข้อมูลในการวิเคราะห์ระบบตัวอักษรเบรลล์ไทยและระบบตัวอักษรเบรลล์อังกฤษ

1.6.1.1 รายการตัวอักษรเบรลล์ไทย ได้มาจาก ตำราสอนอ่านเขียนอักษรเบรลล์ (สมทรง พันธ์สุวรรณ, 2538)

1.6.1.2 รายการตัวอักษรเบรลล์อังกฤษ ได้มาจากตำรา Instruction Manual for Braille Transcribing (Risjord, Constance, 2009)

1.6.2.3 รายการตัวอักษรเบรลล์ญี่ปุ่น ได้จากเว็บไซต์การเรียนการสอนอักษรเบรลล์ญี่ปุ่น (www.jfbkk.or.th/pdf/JL/2014)

1.6.2 ข้อมูลที่ใช้ในระบบการถ่ายถอดอักษรเบรลล์

1.6.2.1 ข้อมูลที่ใช้วิเคราะห์เพื่อสร้างระบบการถ่ายถอดอักษร

ข้อมูลชุดนี้ประกอบด้วย ข้อมูลที่ใช้ในการระบุภาษา ข้อมูลที่ใช้ในการถ่ายถอดเบรลล์ไทย และข้อมูลที่ใช้ในการถ่ายถอดเบรลล์อังกฤษ

1.6.2.2 ข้อมูลที่นำมาใช้ทดสอบระบบ

ข้อมูลชุดนี้ประกอบด้วยข้อความ 5 ประเภทได้แก่ ข้อความเบรลล์ไทยล้วน ข้อความเบรลล์อังกฤษรูปเต็ม ข้อความเบรลล์อังกฤษรูปย่อ ข้อความเบรลล์ไทยปนอังกฤษรูปเต็ม และ ข้อความเบรลล์ไทยปนอังกฤษรูปย่อ ซึ่งได้มาจากศูนย์เทคโนโลยีการศึกษาเพื่อคนตาบอด

1.7 ขั้นตอนการวิจัย

1.7.1 เก็บรวบรวมข้อมูลที่ใช้ในการวิจัย

1.7.2 วิเคราะห์ระบบอักษรเบรลล์ไทย และระบบอักษรเบรลล์อังกฤษ

1.7.3 ดำเนินการวิจัยตามหัวข้อการออกแบบและพัฒนาระบบการถ่ายถอดอักษรเบรลล์ไทยปนอังกฤษ ซึ่งมีขั้นตอนดังนี้

1.7.3.1 รวบรวมข้อมูลที่ใช้ในการพัฒนาระบบ ได้แก่ รายการตัวย่อและคำย่อเบรลล์อังกฤษ ตัวบทเบรลล์ไทย ตัวบทเบรลล์อังกฤษ และตัวบทเบรลล์ไทยปนอังกฤษ

1.7.3.2 จัดทำคลังข้อความชุดฝึกสอนระบบ และดำเนินการฝึกสอนระบบโดยใช้ข้อมูลจากคลังข้อความภาษาไทยและคลังข้อความภาษาอังกฤษ

1.7.3.3 จัดทำชุดข้อความชุดทดสอบระบบ ได้แก่ กลุ่มที่ 1 ข้อความเบรลล์ไทย กลุ่มที่ 2 ข้อความเบรลล์อังกฤษ และกลุ่มที่ 3 ข้อความเบรลล์ไทยปนอังกฤษ

1.7.4 ออกแบบและพัฒนาระบบการถ่ายถอดอักษร เบรลล์ไทย เบรลล์อังกฤษ และเบรลล์ไทยปนอังกฤษ โดยแต่ละระบบจะใช้วิธีการ 3 รูปแบบคือ

1.7.4.1 ระบบการถ่ายถอดอักษรโดยใช้กฎทางภาษา

1.7.4.2 ระบบการถ่ายถอดอักษรโดยใช้แบบจำลองเอ็นแกรม

1.7.4.3 ระบบการถ่ายถอดอักษรโดยใช้กฎทางภาษาร่วมกับแบบจำลองเอ็นแกรม

1.7.5 ทดสอบการทำงานของแต่ละระบบ (ข้อ 1.7.4.1 – 1.7.4.3)

1.7.6 เปรียบเทียบประสิทธิภาพด้านความถูกต้องของแต่ละวิธีการ

1.7.7 วิเคราะห์ผลการวิจัย และสรุปผลการวิจัย

1.8 ประโยชน์ที่คาดว่าจะได้รับ

1. ทำให้ทราบระบบของอักษรเบรลล์ไทยและระบบอักษรเบรลล์อังกฤษ
2. เป็นแนวทางในการศึกษาเปรียบเทียบระบบอักษรเบรลล์ในภาษาอื่น ๆ
3. เป็นแนวทางในการพัฒนาระบบการถ่ายถอดอักษรเบรลล์ในภาษาอื่นๆ

1.9 โครงสร้างวิทยานิพนธ์

ในวิทยานิพนธ์ฉบับนี้จะแบ่งเนื้อหาในการนำเสนอออกเป็น 5 ส่วน ได้แก่

ในบทที่ 2 จะกล่าวถึงทฤษฎีที่เกี่ยวข้อง ได้แก่ ความรู้ทั่วไปเกี่ยวกับอักษรเบรลล์ การจำแนกภาษา การถ่ายถอดอักษร และงานวิจัยที่เกี่ยวข้องกับการถ่ายถอดเบรลล์ไทยปนอังกฤษ

ในบทที่ 3 จะกล่าวถึงการวิเคราะห์ระบบอักษรเบรลล์ไทยโดยการเปรียบเทียบตัวอักษรเบรลล์ไทยกับเบรลล์อังกฤษและเบรลล์ญี่ปุ่น และการวิเคราะห์อักขรวิธีของเบรลล์ไทยตามหลักภาษาศาสตร์

ในบทที่ 4 จะกล่าวถึง แนวคิดในการออกแบบระบบการถ่ายถอดอักษร ได้แก่ แนวคิดในการถ่ายถอดเบรลล์ไทย แนวคิดในการถ่ายถอดเบรลล์อังกฤษ และแนวคิดในการระบุภาษา

ในบทที่ 5 – บทที่ 7 จะได้นำเสนอ ระบบการถ่ายถอดเบรลล์ไทย ระบบการถ่ายถอดเบรลล์อังกฤษ และระบบการถ่ายถอดเบรลล์ไทยปนอังกฤษ โดยบทที่ 5 เสนอระบบการถ่ายถอดอักษรโดยใช้กฎ บทที่ 6 เสนอระบบการถ่ายถอดอักษรโดยใช้แบบจำลองเอ็นแกรม และบทที่ 7เสนอระบบการถ่ายถอดอักษรโดยวิธีการแบบผสม

ในบทที่ 8 จะสรุปผลที่ได้จากวิทยานิพนธ์นี้ รวมทั้งให้ข้อเสนอแนะในการพัฒนาระบบการถ่ายถอดอักษรเบรลล์ให้เป็นอักษรปกติ ต่อไป

1.10 รายการนิยามศัพท์ที่เกี่ยวข้อง

1. ระบบการถ่ายถอดอักษร (Transliteration system) หมายถึง กระบวนการในการเปลี่ยนข้อความอิเล็กทรอนิกส์จากอักษรเบรลล์เป็นอักษรปกติ โดยอาศัยคอมพิวเตอร์

2. อักษรเบรลล์ (Braille systems) หมายถึง ระบบตัวเขียนที่คนตาบอดใช้สื่อสารโดยรับรู้ผ่านการสัมผัสด้วยปลายนิ้วมือ ประกอบด้วยเซลล์ต่างๆนำมาเรียงต่อกันจากซ้ายไปขวาโดยที่แต่ละเซลล์ประกอบด้วยจุดนูนไม่เกิน 6 จุด การปรากฏของจุดในตำแหน่งที่ต่างกันใช้แทนอักขระต่างกัน

3. อักษรเบรลล์ไทยปนอังกฤษ (English-Thai braille) หมายถึง ข้อความอักษรเบรลล์ที่มีการสลับกันของภาษาไทยและภาษาอังกฤษ ในข้อความเดียวกัน

4. อักษรปกติ (Print systems) หมายถึง ระบบตัวเขียนสำหรับคนสายตาปกติที่ประกอบด้วย ตัวอักษร (letters) ตัวเลข (numbers) หรือสัญลักษณ์ (signs) ที่ปรากฏบนวัสดุที่มีพื้นผิวเรียบ เช่น กระดาษ หน้าจอแสดงผลของอุปกรณ์อิเล็กทรอนิกส์ต่าง ๆ เป็นต้น

5. อักษรเบรลล์รูปเต็ม หรืออักษรเบรลล์เกรด 1 (Braille grade 1) หมายถึง อักขระเบรลล์หนึ่งหรือสองตัว ที่ใช้แทนตัวอักษรปกติหนึ่งตัว

6. อักษรเบรลล์รูปย่อ หรืออักษรเบรลล์เกรด 2 (Braille grade 2) หมายถึง อักขระเบรลล์หนึ่งตัวหรือมากกว่าหนึ่งตัวที่ใช้แทนตัวอักษรปกติหนึ่งกลุ่ม

7. คอมพิวเตอร์เบรลล์ (Computer braille) หมายถึง อักขระเบรลล์แต่ละตัวที่ใช้แทนอักขระแอสกีโดยตรงแบบหนึ่งต่อหนึ่ง

8. สัญลักษณ์แอสกี (ASCII symbols) หมายถึง สัญลักษณ์ที่ปรากฏบนหน้าจอคอมพิวเตอร์ ซึ่งใช้แทนอักษรเบรลล์เช่น “g” ในคอมพิวเตอร์ใช้แทน ‘⠛’ และ “(” ในคอมพิวเตอร์ใช้แทน ‘⠣’

9. เครื่องหมายนำเลข (Number sign; #) หมายถึง เครื่องหมายที่ใช้นำหน้าตัวอักษร 10 ตัวแรกในภาษาอังกฤษ (a-j) เพื่อบ่งชี้ว่าตัวอักษรที่ปรากฏหลังเครื่องหมายนี้เป็นตัวเลข ไม่ใช่ตัวอักษร

10. เครื่องหมายประกอบคำ (Composition sign) หมายถึงสัญลักษณ์ที่ปรากฏเฉพาะในอักษรเบรลล์อังกฤษเพื่อกำกับประเภทของอักษร ประกอบด้วย ‘⠨’ (เครื่องหมายแสดงตัวพิมพ์ใหญ่เดี่ยว) เป็นเครื่องหมายที่ใช้ระบุว่าอักษรตัวแรกที่ตามมาเป็นอักษรตัวพิมพ์ใหญ่ ‘⠨⠨’ (เครื่องหมายแสดงตัวพิมพ์ใหญ่คู่) ใช้ระบุว่าอักษรทั้งหมดที่ตามมาเป็นอักษรตัวพิมพ์ใหญ่ และ ‘⠸’ (เครื่องหมายกำกับตัวอักษร) เป็นเครื่องหมายที่บ่งชี้ว่าอักษรที่ตามมาเป็นตัวอักษรไม่ใช่ตัวย่อ

11. การจำแนกภาษา หรือการระบุภาษา (Language identification) หมายถึง การบอกว่าสายอักขระใดสายอักขระหนึ่งในอักษรเบรลล์นั้นต้องตีความให้เป็นภาษาใด โดยใช้คอมพิวเตอร์

12. วิธีการใช้กฎ (Rule-based method) หมายถึง การนำกฎทางภาษามาใช้ในการถ่ายถอดอักษรเบรลล์ให้เป็นอักษรปกติ ซึ่งกฎที่ใช้ประกอบด้วย กฎในการจับคู่หน่วยอักขระระหว่างภาษาและกฎในการจัดลำดับของหน่วยอักขระในอักษรปกติ

13. วิธีการใช้แบบจำลองเอ็นแกรม (N-gram method) หมายถึง การถ่ายถอดอักษรเบรลล์ให้เป็นอักษรปกติโดยอาศัยข้อมูลทางสถิติ โดยที่แบบจำลองเอ็นแกรมจะเป็นตัวเลือกรูปแบบที่จะนำมาใช้เป็นผลลัพธ์ของการถ่ายถอดอักษร

14. วิธีการแบบผสม (Hybrid method) หมายถึง การถ่ายถอดอักษรเบรลล์ให้เป็นอักษรปกติโดยการใช้กฎร่วมกับแบบจำลองเอ็นแกรม

15. คู่เทียบ หมายถึง อักษรเบรลล์และอักษรปกติที่ตรงกัน ซึ่งระบบจะใช้ในการถ่ายถอดจากอักษรเบรลล์ให้เป็นอักษรปกติ

16. สายอักขระ หมายถึง ข้อความเบรลล์แต่ละวรรคที่นำมาถ่ายถอดเป็นอักษรปกติ

17. เซ็ตของสายอักขระที่เป็นเอกลักษณ์ของแต่ละภาษา หมายถึง เซ็ตของอักขระเบรลล์ 1ตัว 2 ตัว หรือ 3 ตัว ในภาษาหนึ่งที่แตกต่างจากอีกภาษาหนึ่ง ซึ่งระบบนำมาใช้ในการระบุภาษา

18. เซ็ตไตรแกรมท้ายพยางค์ หมายถึง เซ็ตของอักขระเบรลล์ 3 ตัว ที่ระบบนำมาใช้ในการตัดสินใจ กรณีที่พบอักขระเบรลล์ที่ตีความได้มากกว่าหนึ่งกรณี

19. สระพิเศษ หมายถึงสระที่จับคู่กับอักษรปกติได้แบบตัวต่อตัว แต่เมื่อสระเหล่านี้ปรากฏกับวรรณยุกต์จะมีการจัดเรียงที่แตกต่างจากอักษรปกติ ได้แก่สระ 2 ตัวคือ ‘⠡’(ะ) และ ‘⠳’(ำ)

20. สระหน้า หมายถึง สระที่ปรากฏหน้าพยัญชนะต้นหรือปรากฏร่วมกับ ‘⠡’(ะ) ได้ ในเบรลล์ไทย ได้แก่สระ 3 ตัว คือ ‘⠝’ (เ-) ‘⠕’(แ-) และ ‘⠥’ (โ-)

บทที่ 2

ทบทวนวรรณกรรม

ในการวิจัยครั้งนี้ ผู้วิจัยจะนำเสนอเนื้อหาที่ได้ทำการศึกษามาโดยแบ่งออกเป็นหัวข้อต่าง ๆ ได้แก่ หัวข้อที่ 2.1 ความรู้ทั่วไปเกี่ยวกับอักษรเบรลล์ เพื่อให้ผู้อ่านมีความเข้าใจพื้นฐานของการใช้อักษรเบรลล์ในการแทนตัวอักษรภาษา หัวข้อที่ 2.2 จะทบทวนงานด้านการจำแนกภาษาด้วยคอมพิวเตอร์ เนื่องจากข้อความเบรลล์ที่ใช้จริงมีการปนระหว่างภาษาซึ่งการจำแนกให้รู้ว่าเป็นเบรลล์ไทยหรืออังกฤษถือเป็นขั้นตอนแรกของการถ่ายถอดอักษร จากนั้นหัวข้อที่ 2.3 จะทบทวนงานด้านการถ่ายถอดอักษรระหว่างภาษาด้วยคอมพิวเตอร์ ในหัวข้อที่ 2.4 จะได้ทบทวนงานวิจัยที่เกี่ยวข้องกับการถ่ายถอดเบรลล์ไทยปนอังกฤษเป็นอักษรปกติ ที่ได้มีการศึกษาวิจัยมาก่อนหน้านี้ และในหัวข้อที่ 2.5 จะได้ทบทวนข้อมูลพื้นฐานเกี่ยวกับโปรแกรม Duxbury Braille Translator เนื่องจากผู้วิจัยจะนำโปรแกรมนี้มาใช้เป็นเครื่องมือในการแปลงข้อความจากเบรลล์อังกฤษรูปย่อเป็นเบรลล์อังกฤษรูปเต็มเพื่อใช้ทดสอบระบบ

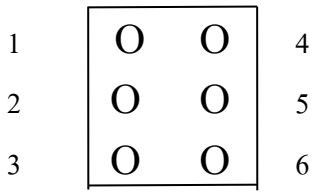
2.1 ความรู้ทั่วไปเกี่ยวกับอักษรเบรลล์

ในการรับรู้ข้อมูลจากสื่อสิ่งพิมพ์ต่าง ๆ นั้น ผู้ที่มีสายตาปกติมักจะรับรู้ผ่านตัวอักษร เนื่องจากตัวอักษรเหล่านี้มักจะเป็นหมึกพิมพ์อยู่บนหน้ากระดาษ ทำให้ไม่สามารถรับรู้ด้วยประสาทสัมผัสอื่น ๆ ได้นอกจากสายตาเท่านั้น จึงจำเป็นที่จะต้องมีระบบอักษรเบรลล์สำหรับคนตาบอด ซึ่งเป็นระบบที่ประกอบด้วยเซลล์ต่างๆนำมาเรียงต่อกัน แต่ละเซลล์ประกอบด้วยจุดนูนไม่เกิน 6 จุด โดยที่การปรากฏของจุดในตำแหน่งที่ต่างกันใช้แสดงอักขระที่ต่างกัน ทำให้คนตาบอดสามารถรับรู้ข้อมูลภาษาได้โดยการสัมผัสด้วยปลายนิ้วมือ

อักษรเบรลล์แต่ละตัวหรือเรียกว่าแต่ละเซลล์ เป็นอักษรที่ใช้เนื้อที่ในการปรากฏเท่า ๆ กัน คือมีขนาดเนื้อที่พอที่จุดนูนทั้ง 6 จุด จะสามารถเรียงต่อกันแบบ 3 แถว 2 คอลัมน์ได้[1] เพื่อให้เกิดความสะดวกในการอ้างถึงจุดต่าง ๆ จึงมีการเรียกชื่อจุดแต่ละจุด โดยเริ่มจากคอลัมน์ทางซ้ายจากบนลงล่างเป็นจุด 1, 2, 3 และ คอลัมน์ทางขวาจากบนลงล่างเป็นจุด 4, 5, 6 ตามลำดับ (Petzold, 2000: 17) ดังปรากฏในภาพที่ 2.1

---

[1] อักษรเบรลล์อังกฤษประกอบด้วยระบบ 6 จุด และ 8 จุด ที่นิยมใช้คือระบบ 6 จุด ส่วนระบบ 8 จุด จะนำมาใช้ในอุปกรณ์คอมพิวเตอร์แบบเบรลล์ (Sullivan, 2014:7) งานวิจัยชิ้นนี้จะศึกษาเฉพาะอักษรเบรลล์ระบบ 6 จุด เท่านั้น



ภาพที่ 2.1 การปรากฏของจุดในอักษรเบรลล์

เนื่องจากอักษรเบรลล์เกิดจากการเรียงต่อกันของจุดไม่เกิน 6 จุด จึงใช้แทนตัวอักขระได้เพียง 64 ตัวเท่านั้น แต่ในความเป็นจริงแล้วภาษาในโลกเป็นจำนวนมากมีหน่วยอักขระ (grapheme) มากกว่า 64 ตัว ทำให้อักษรเบรลล์ของแต่ละภาษามีโครงสร้างซับซ้อน เนื่องจากวิทยานิพนธ์ฉบับนี้มุ่งศึกษาเฉพาะภาษาอังกฤษและภาษาไทยเท่านั้น จึงจะกล่าวถึงการใช้อักษรเบรลล์สำหรับภาษาอังกฤษและภาษาไทยดังนี้

2.1.1 อักษรเบรลล์อังกฤษ

อักษรเบรลล์อังกฤษมี 3 รูปแบบคือ เบรลล์อังกฤษเกรด 1 เบรลล์อังกฤษเกรด 2 และเบรลล์อังกฤษเกรด 3 สำหรับเบรลล์อังกฤษเกรด 1 ซึ่งวิทยานิพนธ์ฉบับนี้จะเรียกว่าเบรลล์อังกฤษรูปเต็ม หมายถึงอักษรเบรลล์หนึ่งเซลล์ใช้แทนอักษรปกติหนึ่งตัว เบรลล์อังกฤษเกรด 2 ซึ่งวิทยานิพนธ์ฉบับนี้จะเรียกว่าเบรลล์อังกฤษรูปย่อ หมายถึงรหัสอักษรเบรลล์ที่ใช้แทนคำหรือบางส่วนของคำ และเบรลล์อังกฤษเกรด 3 หมายถึงอักษรเบรลล์รูปย่อที่ไม่ใช่รูปมาตรฐาน (non-standardized personal shorthand) ซึ่งใช้ในการจดบันทึกส่วนตัวเฉพาะบุคคลเท่านั้น (Klenk, 2011: 3-4) ในที่นี้จะขอกล่าวถึงเฉพาะเบรลล์อังกฤษรูปเต็ม (Uncontracted English Braille) และเบรลล์อังกฤษรูปย่อ (Contracted English Braille) หรือเบรลล์เกรด 1 และเบรลล์เกรด 2 เท่านั้น

2.1.1.1 อักษรเบรลล์อังกฤษรูปเต็ม

ในอักษรเบรลล์อังกฤษรูปแบบนี้ อักษรเบรลล์มีจำนวน 26 ตัว เท่ากับอักษรปกติ ดังปรากฏในตารางต่อไปนี้ (Petzold, 2000: 18)

ตารางที่ 2.1 ตัวอักษรเบรลล์อังกฤษ

แถว	ตัวอักษรเบรลล์อังกฤษ									
1	a	b	c	d	e	f	g	h	i	j
2	k	l	m	n	o	p	q	r	s	t
3	u	v	x	y	z					
4										w

ตารางข้างต้นแสดงให้เห็นว่า เบรลล์อังกฤษเกิดจากการนำตัวอักษรหลัก 10 ตัว มาใช้เป็นฐานในการสร้างตัวอักษรอื่น ๆ ตัวอักษรหลักได้แก่ ‘⠁’ - ‘⠊’ (=a-j) ดังแสดงในแถวที่1 โดยใช้จุด 1, 2, 4, และ 5 ในการแยกความแตกต่างของตัวอักษร ส่วน ‘⠡’-‘⠪’(=k – t) เกิดจากการใช้จุด 1, 2, 4, 5 ในลักษณะเดียวกับ a-j และเติม ‘⠈’ (จุด 3) เข้าไปในเซลล์ ดังแสดงในแถวที่ 2 ส่วน ‘⠩’-‘⠳’ (=u-z (ยกเว้น w)) เกิดจากการนำตัวอักษรหลัก 5 ตัวแรก คือ a - e มาเติม ‘⠸’ (จุด 3-6) ดังแสดงในแถวที่ 3 และตัวอักษร ‘⠺’ (=w) เกิดจากการเติมเฉพาะ ‘⠨’ (จุด 6) ที่ตัว ‘⠚’ (=j) (Englebretson, 2009: 3; Petzold,. 2000: 16)

ที่กล่าวมานี้เป็นเพียงอักษรเบรลล์อังกฤษตัวพิมพ์เล็กเท่านั้น หากเป็นตัวพิมพ์ใหญ่จะใส่เครื่องหมาย ‘⠠’ (capital sign) นำหน้าตัวอักษรตัวนั้น แต่หากเป็นตัวพิมพ์ใหญ่ทั้งคำก็ให้ใส่เครื่องหมาย ‘⠠⠠’ (double capital sign) นำหน้าคำ

การประสมอักษรในเบรลล์อังกฤษมีลักษณะการเขียนที่เรียงกันไปเหมือนตัวอักษรอังกฤษทั่วไปแบบหนึ่งต่อหนึ่ง โดยแต่ละคำจะมีการเขียนเว้นวรรค (หรือพิมพ์โดยใช้เครื่องหมาย space) ทำให้สามารถแยกคำต่าง ๆ ออกจากกันได้ เช่น "The dog" และ "THE DOG" สามารถเขียนได้ดังนี้

⠠⠮ ⠙⠕⠛ ⠠⠠⠮ ⠙⠕⠛

T h e d o g THE DOG

2.1.1.2 อักษรเบรลล์อังกฤษรูปย่อ

เนื่องจากอักษรเบรลล์เป็นอักษรที่มีขนาดใหญ่ต้องใช้เนื้อที่มากในการปรากฏบนกระดาษ ทำให้การอ่านและการเขียนอักษรระบบนี้ต้องสิ้นเปลืองทรัพยากรและใช้เวลามาก ด้วยข้อจำกัดดังกล่าวจึงมีการประดิษฐ์ตัวย่อเบรลล์เพื่อให้คนตาบอดใช้อักษรเบรลล์ได้สะดวกรวดเร็วยิ่งขึ้น (Tennessee council of the blind, 2008: 3)

ตัวย่อเบรลล์อังกฤษ คือการนำอักขระเบรลล์จำนวนหนึ่งมาใช้แทนคำหรือส่วนของคำภาษาอังกฤษโดยไม่ต้องเขียนทุกตัวอักษรของคำนั้น ซึ่ง Braille Authority of North America (2002: 5) ได้จำแนกตัวย่อเบรลล์อังกฤษออกเป็น 3 กลุ่มได้แก่ ตัวย่อเบรลล์เซลล์เดี่ยว (one-cell contractions) ตัวย่อเบรลล์เซลล์คู่ (two-cell contractions) และคำย่อเบรลล์ (short-form words)

2.1.1.2.1 ตัวย่อเบรลล์เซลล์เดี่ยว

ตัวย่อเบรลล์เซลล์เดี่ยว คือตัวย่อที่อักขระเบรลล์หนึ่งเซลใช้แทนคำหรือส่วนของคำได้ Braille Authority of North America (2002: 5-7) ได้แบ่งตัวย่อเบรลล์เซลล์เดี่ยวออกเป็นประเภทต่าง ๆ ได้แก่ ตัวย่ออักษรโดด ตัวย่อรูปเต็มคงรูป ตัวย่อรูปเต็มกลายรูป ตัวย่อประกอบคำ ตัวย่อกลางคำหรือท้ายคำ และตัวย่อเบรลล์ต่ำ

(1) ตัวย่ออักษรโดด (alphabet contractions)

ในเบรลล์อังกฤษ มีอักขระเบรลล์ใช้แทนตัวอักษร 26 ตัวเช่นเดียวกับอังกฤษทั่วไป และตัวอักษรเหล่านี้สามารถนำมาเรียงต่อกันเป็นคำได้เช่นเดียวกัน เช่น ‘⠃⠥⠞’ สามารถเทียบกับอักษรปกติ “ant” แบบตัวต่อตัวได้ ทำให้อักขระเบรลล์เหล่านี้เมื่อมีอักขระอื่นปรากฏร่วมหน้าและ/หรือหลังภายในคำจะใช้แทนอักษรอังกฤษทั่วไป อย่างไรก็ตาม ในอังกฤษรูปย่ออักขระเบรลล์เหล่านี้สามารถนำมาใช้แทนคำโดดได้ กล่าวคือเมื่ออักขระแต่ละตัวในกลุ่มนี้ปรากฏระหว่างช่องว่างและ/หรือเครื่องหมายแสดงขอบเขตของคำ จะใช้แทนคำโดด เช่น ‘⠸⠃⠸’ ซึ่งเป็นอักขระ b จะใช้แทน “|but|”[2] ‘⠸⠋⠸’ ซึ่งเป็นอักขระ f กับ ? จะใช้แทน “|from?|” เป็นต้น เป็นที่น่าสังเกตว่าอักขระเหล่านี้ไม่สามารถใช้เป็นส่วนหนึ่งของคำได้ เช่น ‘⠃’ ใช้แทน “ but ” แต่คำว่า “butter” ไม่สามารถเขียนเป็น ‘⠃⠥⠞⠻’ (= bter) ได้

ตัวย่ออักษรโดดจะใช้เป็นส่วนหนึ่งของคำได้ ในกรณีต่อไปนี้

o ปรากฏกับ ‘⠤’ (hyphen) เช่น ‘⠍⠑⠗⠗⠽⠤⠤⠤⠛⠕⠤⠤⠤⠗⠤⠳⠝⠙’ (=m-er-r-y---go---r-ound) ซึ่งใช้แทน “merry-go-round” ที่มี ‘⠛⠕’ (= go) ปรากฏระหว่างเครื่องหมาย hypen

---

[2] ในงานวิจัยชิ้นนี้ จะใช้เครื่องหมาย | แทนเครื่องหมายเว้นวรรคในข้อความภาษาอังกฤษ

o ใช้กับคำที่มี ‘⠄’ (apostrophe) เช่น ‘⠠⠉⠄⠞’ (=c’t) เพื่อย่อ “can’t ” เป็นต้น

กล่าวโดยสรุป ตัวย่ออักษรคำโดดจะใช้อักขระเบรลล์หนึ่งตัวแทนคำคงที่หนึ่งคำ ไม่สามารถปรากฏกับอักขระอื่นได้ ยกเว้นการใช้ติดกับอักขระที่แทนเครื่องหมายต่าง ๆ เช่น “-”, “.” , “ ‘ ” , “?” เป็นต้น

การใช้ตัวย่ออักษรโดดที่แสดงใน (Braille Authority of North America, 2002: XI-2) มีดังนี้

ตารางที่ 2.2 การใช้ตัวย่ออักษรโดด

วิธีการใช้	**ตัวอย่าง**		
	รูปย่อในเบรลล์	**แอสกี**	**คำหรือประโยค**
1) ใช้อักษรตามลำพังเพื่อย่อคำโดด	⠽ ⠃ ⠏ ⠇ ⠽	n y b p l y	Not you but people like you.
2) ใช้หน้าหรือหลังเครื่องหมาย hyphen ในกรณีที่เป็นคำประสม	⠎⠤⠉⠁⠇⠇⠫	s-call$	so-called
3) ใช้กับคำวิสามัญนาม	⠠⠍⠗⠲ ⠠⠍	,mr. ,m	Mr. More
4) ใช้กับคำแสดงความเป็นเจ้าของ	⠠⠺⠄⠎	,w’s	Will's
5) ใช้กับคำที่คั่นด้วย apostrophe	⠉⠄⠞	c’t	can’t

ตัวย่ออักษรโดดในกลุ่มเบรลล์อังกฤษ 26 ตัว มีรูปย่อและสิ่งที่ใช้แทนที่เป็นมาตรฐานสำหรับเบรลล์อังกฤษทั่วโลก (Braille Authority of North America, 2002: 2) ดังนี้

ตารางที่ 2.3 รายการตัวย่ออักษรโดด

เบรลล์	แอสกี	รูปเต็ม
⠃	b	but
⠉	c	can
⠙	d	do
⠑	e	every
⠋	f	from
⠛	g	go
⠓	h	have
⠚	j	just
⠅	k	knowledge
⠇	l	like
⠍	m	more
⠝	n	not
⠏	p	people
⠟	q	quite
⠗	r	rather
⠎	s	so
⠞	t	that
⠥	u	us
⠧	v	very
⠺	w	will
⠭	x	it
⠽	y	you
⠵	z	as

(หมายเหตุ: สัญลักษณ์แอสกีที่ปรากฏในตาราง หมายถึงสัญลักษณ์แอสกีที่ใช้แทนตัวย่อโดยการป้อนข้อมูลผ่านคีย์บอร์ด)

เป็นที่น่าสังเกตว่าในการปรากฏของตัวย่ออักษรโดด เบรลล์แต่ละเซลล์จะใช้แทนคำแต่ละคำ ทำให้ไม่สามารถเขียนอักขระเบรลล์เหล่านี้เพื่อหมายถึงอักขระแต่ละตัวตามลำพังได้ ฉะนั้นเมื่อต้องการแสดงตัวอักษรโดด ต้องใช้ ‘⠰’ (เครื่องหมายกำกับตัวอักษร) นำหน้าอักขระเบรลล์เหล่านี้ เช่นเมื่อพบ ‘⠃’ (=b) ปรากฏระหว่างช่องว่างต้องตีความเป็นคำว่า “but” แต่หากหมายถึงตัวอักษร b เท่านั้นจำเป็นต้องใส่เครื่องหมายกำกับอักษรไว้หน้าตัว b ทำให้ได้เป็น ‘⠰⠃’ จึงจะตีความเป็นตัวอักษร b ได้

(2) ตัวย่อรูปเต็มคงรูป (whole-word/part-word contractions type1)

ตัวย่อรูปเต็มคงรูป เป็นตัวย่อที่อักขระเบรลล์ 1 ตัว ใช้แทนกลุ่มตัวอักษรที่ตายตัวไม่ว่าอักษรนั้นจะปรากฏเป็นคำโดดหรือเป็นส่วนของคำก็ตามยังคงใช้แทนกลุ่มตัวอักษรเดิม เช่น ‘⠸⠯’ (=&) ซึ่งใช้แทนคำ “⠸and⠸ ” ที่ปรากฏเป็นคำโดด และ ‘⠛⠗⠯’ (= g-r-and) ซึ่งมี “and” เป็นส่วนหนึ่งของคำ

จะเห็นได้ว่า คำที่ได้จากตัวย่อในกลุ่มนี้มีความถี่ในการปรากฏสูงในภาษาอังกฤษ อีกทั้งคำเหล่านี้มักปรากฏเรียงต่อกัน จึงมีการเขียนตัวย่อเหล่านี้ต่อกันโดยไม่เว้นวรรคเพื่อใช้แทนคำโดดแต่ละคำ (ซึ่งคำเหล่านี้ไม่สามารถเขียนชิดกันโดยไม่เว้นวรรคได้ในอักษรปกติ) เช่น ‘⠯⠮’ (=&!) ใช้ย่อ “and the” ซึ่งเป็นคำโดด 2 คำที่เรียงต่อกัน และ ‘⠯⠯⠁’ (=&)a) ที่ใช้แทนคำโดด 3 คำที่อยู่ชิดกันคือ “and with a” เป็นต้น

ตัวย่อรูปเต็มคงรูป มีตัวอย่างการใช้ซึ่งแสดงได้ดังนี้ (Braille Authority of North America, 2002: XI-3)

ตารางที่ 2.4 วิธีการใช้ตัวย่อรูปเต็มคงรูป

วิธีการใช้	ตัวอย่าง		
	สัญลักษณ์แอสกี	รูปย่อในเบรลล์	คำหรือประโยค
1) เมื่อใช้ตัวย่อกับคำโดดอื่น ๆให้เว้นวรรคระหว่างตัวย่อและคำโดด	,i l hm & h(er).	⠰⠊ ⠰⠇ ⠓⠍ ⠯ ⠓⠻	I like him and her.
2) เมื่อตัวย่ออยู่ติดกันให้เขียนติดต่อกันไม่เว้นวรรค	! cat & ! rat	⠮ ⠉⠁⠞ ⠯ ⠮ ⠗⠁⠞	the cat and the rat.
3) ใช้ย่อส่วนของคำ	! sis	⠮⠎⠊⠎	Thesis

ตัวย่อรูปเต็มคงรูป ประกอบด้วยตัวย่อและกลุ่มตัวอักษรที่ใช้แทน ดังนี้

ตารางที่ 2.5 รายการตัวย่อรูปเต็มคงรูป

เบรลล์	สัญลักษณ์แอสกี	รูปเต็ม
⠯	&	and
⠿	=	for
⠷	(	of
⠮	!	the
⠾	)	with

(3) ตัวย่อรูปเต็มกลายรูป (whole-word/part-word contractions type 2)

ตัวย่อรูปเต็มกลายรูป เป็นตัวย่อที่ใช้แทนกลุ่มตัวอักษรที่แตกต่างกันใน 2 บริบท คือ เมื่อปรากฏเป็นคำโดดจะใช้แทนตัวอักษรกลุ่มหนึ่ง และเมื่อปรากฏเป็นส่วนของคำก็จะใช้แทนตัวอักษรอีกกลุ่มหนึ่ง เช่น ‘⠱’ จะใช้แทน “which” เมื่อปรากฏเป็นคำโดด แต่เมื่อปรากฏภายในคำจะใช้แทน “wh” ตัวย่อกลุ่มนี้ไม่เป็นปัญหาต่อการจดจำเท่าใดนัก เนื่องจากจะใช้แทนกลุ่มตัวอักษรที่มีอักษร 2 ตัวแรกซ้ำกัน เช่น ‘⠩’ ใช้ย่อ “shall” หากเป็นคำโดด และใช้ย่อ “sh” เมื่อปรากฏเป็นส่วนหนึ่งของคำ ซึ่งอักษร 2 กลุ่มนี้มี “sh” ร่วมกัน

ตัวย่อรูปเต็มกลายรูป มีรูปย่อและสิ่งที่ใช้แทนดังนี้ (Braille Authority of North America, 2002: XI-4)

ตารางที่ 2.6 รายการตัวย่อรูปเต็มกลายรูป

เบรลล์	สัญลักษณ์แอสกี	รูปเต็ม	
		คำโดด	ส่วนของคำ
⠡	*	child	ch
⠳	\	out	ou
⠩	%	shall	sh
⠌	/	still	st
⠹	?	this	th
⠱	:	which	wh

(4) ตัวย่อประกอบคำ (part-word contractions)

ตัวย่อประกอบคำ เป็นตัวย่อที่ต้องปรากฏเป็นส่วนหนึ่งของคำ ใช้สำหรับแทนส่วนของคำที่มีตัวอักษรเหล่านี้คือ "ar", "ed", "er", "gh" และ "ow" ตัวย่อประกอบคำมีรูปย่อและสิ่งที่ใช้แทนดังนี้

ตารางที่ 2.7 รายการตัวย่อประกอบคำ

เบรลล์	สัญลักษณ์แอสกี	รูปเต็ม
⠜	>	ar
⠫	$	ed
⠻	]	er
⠣	<	gh
⠪	[	ow

ตัวย่อเหล่านี้มีวิธีการใช้ดังนี้ (Braille Authority of North America, 2002: XII-3)

ตารางที่ 2.8 วิธีการใช้ตัวย่อประกอบคำ

วิธีการใช้	ตัวอย่าง		
	คำหรือประโยค	รูปย่อ	แอสกี
1)ใช้ตัวย่อนี้ที่ตำแหน่งต้น กลางหรือท้ายคำ	error	⠻⠗⠕⠗	]ror
2) ใช้แสดงเสียงของคำ	ar! (หัวเราะ)	⠜⠖	>6

(5) ตัวย่อกลางคำหรือท้ายคำ (middle or end of word contractions)

ตัวย่อกลางคำหรือท้ายคำ หมายถึงตัวย่อที่ปรากฏในตำแหน่งกลางคำหรือท้ายคำได้เท่านั้น ตัวย่อกลุ่มนี้เป็นตัวย่อที่ใช้เบรลล์เซลล์เดียวเพื่อแทนกลุ่มตัวอักษร ประกอบด้วยตัวย่อ 2 ตัวได้แก่ ⠡ (=ble) และ ⠬ (=ing)

เป็นที่น่าสังเกตว่า อักขระเบรลล์ 2 ตัวที่กล่าวมาข้างต้นนี้เมื่อปรากฏต้นสายอักขระจะทำหน้าที่อื่น ๆ ที่ไม่ใช่ตัวย่อ ดังที่แสดงไว้ในหัวข้อที่เกี่ยวกับตัวเลขเบรลล์ (ดูหัวข้อ 2.1.3ประกอบ) ตัวย่อกลางคำหรือท้ายคำ มีรูปย่อและสิ่งที่ใช้แทน ดังนี้

ตารางที่ 2.9 รายการตัวย่อกลางคำหรือท้ายคำ

เบรลล์	สัญลักษณ์แอสกี	รูปเต็ม
⠨⠆	#	ble
⠨⠖	+	ing

ตัวย่อกลางคำหรือท้ายคำมีวิธีการใช้ดังนี้ (Braille Authority of North America, 2002: XII-2)

ตารางที่ 2.10 วิธีการใช้ตัวย่อกลางคำหรือท้ายคำ

วิธีการใช้	ตัวอย่าง		
	คำหรือประโยค	รูปย่อในเบรลล์	สัญลักษณ์แอสกี
1) ใช้ที่ตำแหน่งกลางคำ	Singer	⠎⠬⠻	s+]
2) ใช้ที่ตำแหน่งท้ายคำ	Trouble	⠞⠗⠳⠃⠇	tr\#

(6) ตัวย่อเบรลล์ต่ำ (lower-sign contractions)

หมายถึงตัวย่อเซลล์เดี่ยวที่ไม่มีจุด 1 และ/หรือจุด 4 ปรากฏอยู่[3] ตัวย่อเบรลล์ต่ำจำแนกเป็น 3 ประเภทตามตำแหน่งที่ปรากฏในคำ ได้แก่ ตัวย่อเบรลล์ต่ำย่อคำโดด ตัวย่อเบรลล์ต่ำย่อคำโดดพิเศษ และตัวย่อเบรลล์ต่ำย่อส่วนของคำ

ก. ตัวย่อเบรลล์ต่ำย่อคำโดด (lower-cell word contractions for “be, enough, were, his, in, was”)

ตัวย่อเบรลล์ต่ำย่อคำโดดมีลักษณะคล้ายกับตัวย่ออักษรโดด กล่าวคือจะต้องปรากฏตามลำพัง (stand alone) เท่านั้น ตัวย่อเหล่านี้มีรูปย่อและสิ่งที่ใช้แทนดังนี้ (Braille Authority of North America, 2002: XIII-2)

---

[3] เหตุที่เรียกว่าเบรลล์ต่ำเนื่องจากอักษรเหล่านี้ปรากฏในตำแหน่งที่ต่ำกว่าอักษรเบรลล์กลุ่มอื่น ๆ แม้ว่าดูผิวเผินจะมีรูปเบรลล์เหมือนกันต้องอาศัยบริบทข้างเคียงเพื่อบอกว่าเป็นเบรลล์ต่ำหรือไม่ เช่นอักษร ⠠ จะมีรูปซ้ำกับอักษร ⠨ (ซึ่งอักษรตัวหลังถือว่าเป็นอักษรเบรลล์เซลต่ำ) สัญลักษณ์เบรลล์ต่ำเหล่านี้บางตัวเมื่อปรากฏในตำแหน่งต้นหรือท้ายคำ จะถูกตีความให้เป็นเครื่องหมายวรรคตอน (punctuation) เช่น ⠐ ถ้าปรากฏหน้าคำก็จะหมายถึงวงเล็บเปิด [( ] แต่หากอยู่ท้ายคำจะถูกตีความให้เป็นเครื่องหมายวงเล็บปิด [ )] (ดูตารางที่ 2.13 ประกอบ)

ตารางที่ 2.11 รายการตัวย่อเบรลล์ต่ำย่อคำโดด

เบรลล์	สัญลักษณ์แอสกี	รูปเต็ม
⠃	2	be
⠢	5	enough
⠶	7	were
⠦	8	his
⠔	9	in
⠴	0	was

ข. ตัวย่อเบรลล์ต่ำย่อคำโดดพิเศษ (lower-cell word contractions for “to, into, by”)

ตัวย่อกลุ่มนี้ลักษณะคล้ายกับตัวย่อเบรลล์ต่ำย่อคำโดดในแง่ที่ว่า ต่างก็ใช้ย่อคำโดด แต่ในการใช้ตัวย่อประเภทหลังนี้จำเป็นต้องเขียนติดกับคำต่อไปเลยไม่ต้องมีช่องว่าง เช่น ‘⠾⠕’ใช้เขียน “to go” และ ‘⠔⠾⠊⠞⠻’ใช้เขียน “into water” เป็นต้น

รูปย่อและสิ่งที่ใช้แทนตัวย่อเบรลล์ต่ำย่อคำโดดพิเศษมีดังนี้ (Braille Authority of North America, 2002: XIII-5)

ตารางที่ 2.12 รายการตัวย่อเบรลล์ต่ำที่ย่อคำโดดพิเศษ

เบรลล์	สัญลักษณ์แอสกี	รูปเต็ม
⠮	6	to
⠔⠖	96	into
⠬	0	by

ค. ตัวย่อเบรลล์ต่ำย่อส่วนของคำ (part-word lower-cell contractions)

ตัวย่อเหล่านี้จะต้องใช้ประกอบคำเสมอ การปรากฏของตัวย่อแต่ละตัวในตำแหน่งที่ต่างกัน คือ ต้นคำ กลางคำ หรือท้ายคำ จะทำให้ตีความได้แตกต่างกัน ตัวอย่างเช่น ‘⠒⠉⠑⠽’ (=con-v-e-y) ซึ่งใช้เขียนคำว่า “convey” เป็นต้น

ตัวย่อในตำแหน่งต่าง ๆ และสิ่งที่ใช้แทน มีดังนี้ (Braille Authority of North America, 2002: XIII-7)

ตารางที่ 2.13 รายการตัวย่อเบรลล์ต่ำย่อส่วนของคำ

เบรลล์	สัญลักษณ์แอสกี	รูปเต็ม		
		ต้นคำ	กลางคำ	ท้ายคำ
⠂	1		ea	,
⠆	2	be	bb	;
⠒	3	con	cc	:
⠲	4	dis	dd	.
⠢	5	en	en	en
⠖	6		ff	!
⠶	7	(	gg	)
⠦	8	“		?
⠔	9	in	in	in
⠴	0			”
⠤	-	com	-	-

2.1.1.2.2 ตัวย่อเบรลล์เซลล์คู่ (two-cell contractions)

ตัวย่อเบรลล์เซลล์คู่หมายถึง เบรลล์ 2 เซลล์ ที่ใช้แทนคำโดดหรือส่วนของคำ โดยที่เบรลล์เซลล์แรกจะขึ้นต้นด้วย จุด 5, 4-6, 5-6 หรือจุด 4-5-6

ตัวย่อกลุ่มนี้มี 2 ประเภท (Braille Authority of North America, 2002: XIII-9) ได้แก่

(1) ตัวย่อหน่วยหลัก (initial-letter contractions) เป็นตัวย่อเซลล์คู่ที่ปรากฏตามลำพังและสามารถปรากฏเป็นส่วนหนึ่งของคำได้ ตัวย่อนี้จะขึ้นต้นด้วยจุด 5, 4-5 หรือ 4-5-6 ตัวย่อหน่วยหลักมีรูปย่อและสิ่งที่ใช้แทนดังนี้

ตารางที่ 2.14 รายการตัวย่อหน่วยหลัก

เบรลล์	แอสกี	รูปเต็ม
⠈⠩	_c	cannot
⠈⠡	"*	character
⠈⠙	"d	day
⠈⠑	"e	ever
⠈⠋	"f	father
⠸⠓	_h	had
⠈⠓	"h	here
⠈⠅	"k	know
⠈⠇	"l	lord
⠸⠍	_m	many
⠈⠍	"m	mother
⠈⠕	"o	one
⠈⠏	"p	part
⠈⠟	"q	question
⠈⠗	"r	right
⠈⠳	"\	ought

เบรลล์	แอสกี	รูปเต็ม
⠈⠎	"s	some
⠸⠎	_s	spirit
⠸⠮	_!	their
⠈⠮	"!	there
⠘⠮	^!	these
⠘⠹	^?	those
⠈⠹	"?	through
⠈⠞	"t	time
⠈⠥	"u	under
⠘⠥	^u	upon
⠈⠱	":	where
⠘⠱	^:	whose
⠘⠺	^w	word
⠈⠺	"w	work
⠸⠺	_w	world
⠈⠽	"y	young

ตัวย่อหน่วยหลักมีวิธีการใช้ดังแสดงในตารางต่อไปนี้ (Braille Authority of North America, 2002: XIII-9)

ตารางที่ 2.15 การใช้ตัวย่อหน่วยหลัก

วิธีการใช้	ตัวอย่าง		
	คำหรือประโยค	รูปย่อในเบรลล์	สัญลักษณ์แอสกี
1) ใช้ย่อทั้งคำ	Part	⠈⠏	"p
2) ใช้ย่อส่วนของคำ	apartment	⠁⠈⠏⠞⠍⠑⠝⠞	a"p;t

จากรายการตัวย่อในตารางข้างต้น จะพบว่ามีการใช้ ‘⠈’ (จุด 5) เพื่อนำหน้าเบรลล์เซลล์คู่เป็นหลัก แต่เนื่องจากการใช้ ‘⠈’ เพียงอย่างเดียวทำให้ไม่สามารถจะใช้ย่อคำในภาษาอังกฤษได้ครบถ้วน ดังนั้นจึงมีการใช้ ‘⠸’ (จุด 4-5) และ ‘⠘’ (จุด 4-5-6) เพื่อทำให้สามารถย่อคำอื่น ๆ ได้ เช่น

‘⠺’ (จุด 5 w) ย่อ “work” ‘⠺⠗’ (จุด 4-5 w) ย่อ “word”และ‘⠺⠗⠙’ (จุด 4-5-6 w) ย่อ “world” ดังนั้นการวิจัยครั้งนี้จึงมุ่งที่จะวิเคราะห์ระบบตัวย่อเบรลล์อังกฤษเพื่อนำไปใช้ในการจัดการเรียนการสอนและการประมวลผลทางคอมพิวเตอร์ สำหรับผลการวิเคราะห์จะนำเสนอในหัวข้อที่เกี่ยวข้องต่อไป

(2) ตัวย่อหน่วยเสริม (final-letter contractions) หมายถึง ตัวย่อเบรลล์เซลล์คู่ที่สามารถปรากฏได้ในตำแหน่งกลางคำหรือท้ายคำเท่านั้น โดยเซลล์แรกของตัวย่อเหล่านี้ขึ้นต้นด้วย ‘⠨’ (จุด 6) ‘⠰’ (จุด 5-6) หรือ ‘⠠’ (จุด 4-6)

ตัวย่อหน่วยเสริมและสิ่งที่ใช้แทน แสดงได้ดังนี้ (Braille Authority of North America, 2002: XIII-9)

ตารางที่ 2.16 รายการตัวย่อหน่วยเสริม

เบรลล์	แอสกี	รูปเต็ม
⠨⠇	,l	ful
⠨⠝	,n	ation
⠨⠽	,y	ally
⠸⠙	.d	ound
⠸⠑	.e	ance
⠸⠝	.n	sion
⠸⠎	.s	less

เบรลล์	แอสกี	รูปเต็ม
⠸⠞	.t	ount
⠰⠑	;e	ence
⠰⠛	;g	ong
⠰⠝	;n	tion
⠰⠎	;s	ness
⠰⠞	;t	ment
⠰⠽	;y	ity

ตัวย่อหน่วยเสริมมีวิธีการใช้คือให้ใช้ที่ตำแหน่งกลางคำหรือท้ายคำ เช่น‘⠠⠙⠨⠑’(=d-ance) ใช้แทนคำว่า “dance” และ ‘⠠⠙⠨⠑⠠⠗’ (=d-ance-r) ใช้แทน “dancer”เป็นต้น หากในการย่อคำ ๆ หนึ่งสามารถใช้ตัวย่อได้ 2 กรณี คือ ใช้ตัวย่อหน่วยเสริมเป็นส่วนประกอบของคำได้หรือใช้ตัวย่ออื่นแทนได้ ให้ใช้ตัวย่อหน่วยเสริมแทนที่จะใช้ตัวย่อในกลุ่มอื่น เช่นคำว่า “dancer” ต้องใช้ ‘⠠⠙⠨⠑⠠⠗’ (=d-ance-r) ไม่สามารถเขียนเป็น ‘⠠⠙⠠⠝⠠⠉⠠⠻’ (=d-a-n-c-er) ได้

จากตารางที่ 2.16 จะเห็นว่า ตัวย่อกลุ่มนี้ต้องขึ้นต้นด้วย ‘⠨’ (จุด 6) ‘⠠’ (จุด 4-6) หรือ‘⠰’ (จุด5-6) ตัวย่อประเภทนี้ต้องปรากฏเฉพาะตำแหน่งกลางคำหรือท้ายคำได้เท่านั้น เนื่องจากสัญลักษณ์เหล่านี้ หากปรากฏที่ตำแหน่งต้นคำจะใช้แทนเครื่องหมายอื่นที่ไม่ใช่ตัวย่อเบรลล์ ดังนี้

- ‘⠨’ (จุด 6) แสดงว่าตัวอักษรที่ตามมาตัวแรกเป็นอักษรตัวใหญ่ (capital sign)

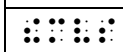
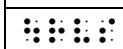
- ‘⠸⠿’ (จุด 4-6) แสดงว่า ตัวอักษรหลังเครื่องหมายนี้จนถึงเว้นวรรคมีการเน้น (เช่นเป็นอักษรตัวเอียง ตัวหนา หรือขีดเส้นใต้) ที่คำนั้น
- ‘⠸⠇’ (จุด5-6) แสดงว่า ตัวอักษรหลังเครื่องหมายนี้จนถึงเว้นวรรคเป็นตัวอักษร ไม่ใช่ตัวย่อหรือคำย่อ

2.1.1.2.3 คำย่อเบรลล์ (shot-form words)

คำย่อเบรลล์หมายถึง คำในอักษรเบรลล์ที่เกิดจากการนำเอาอักขระบางตัวภายในคำมาใช้แทนคำนั้น ๆ เช่น ‘⠗⠉⠧⠛’ (=rcvg) ใช้แทน receiving

คำย่อเบรลล์ที่ปรากฏในภาษาอังกฤษทั้งหมด มีรูปย่อและสิ่งที่ใช้แทนดังนี้

ตารางที่ 2.17 รายการคำย่อเบรลล์

เบรลล์	แอสกี	รูปเต็ม
	!mvs	**them**sel**v**e**s**
	"of	**one**sel**f**
	%d	**sh**oul**d**
	*n	**ch**ildre**n**
	?yf	**th**y**s**el**f**
	\rvs	**our**sel**v**e**s**
	2c	**bec**ause
	2f	**bef**ore
	2h	**beh**ind
	2l	**bel**ow
	2n	**ben**eath
	2s	**bes**ide
	2t	**bet**ween
	2y	**bey**ond
	3cv	**conc**ei**v**e
	3cvg	**conc**ei**v**i**ng**
	ab	**ab**out
	abv	**ab**o**v**e
	ac	**ac**cording
	acr	**acr**oss
	af	**af**ter
	afn	**af**ter**n**oon
	afw	**af**ter**w**ard
	ag	**ag**ain
	ag/	**ag**ain**st**
	al	**al**so
	al?	**alth**ough

เบรลล์	แอสกี	รูปเต็ม
	alm	**alm**ost
	alr	**alr**eady
	alt	**alt**ogether
	alw	**alw**ays
	bl	**bl**ind
	brl	**br**a**ill**e
	cd	**c**oul**d**
	dcl	**decl**are
	dclg	**Decl**arin**g**
	dcv	**d**ece**iv**e
	dcvg	**d**ece**iv**in**g**
	ei	**ei**ther
	f/	**fir**s**t**
	fr	**fr**iend
	gd	**g**oo**d**
	grt	**gr**ea**t**
	h]f	**her**sel**f**
	hm	**him**
	hmf	**him**sel**f**
	imm	**imm**ediate
	ll	**l**itt**l**e
	lr	**l**ette**r**
	m*	**much**
	m/	**must**
	myf	**my**sel**f**
	nec	**nec**essary
	nei	**nei**ther

เบรลล์	แอสกี	รูปเต็ม
⠕⠄⠉	o'c	**o'c**lock
⠏⠻⠉⠧	p]cv	**perc**eive
⠏⠻⠉⠧⠛	p]cvg	**perc**eiving
⠏⠻⠓	p]h	**perh**aps
⠏⠙	pd	**p**ai**d**
⠟⠅	qk	**q**uic**k**
⠗⠉⠧	rcv	**rec**eive
⠗⠉⠧⠛	rcvg	**rec**eivin**g**
⠗⠚⠉	rjc	**r**e**j**oi**c**e
⠗⠚⠉⠛	rjcg	**r**e**j**oi**c**in**g**
⠎⠄	s*	**s**u**ch**

เบรลล์	แอสกี	รูปเต็ม
⠎⠙	sd	sai**d**
⠞⠙	td	to**d**ay
⠞⠛⠗	tgr	to**g**ethe**r**
⠞⠍	tm	to**m**orrow
⠞⠝	tn	to**n**ight
⠺⠙	wd	**w**oul**d**
⠭⠋	xf	**it**sel**f**
⠭⠎	xs	**its**
⠽⠗	yr	**y**ou**r**
⠽⠗⠋	yrf	**y**ou**r**sel**f**
⠽⠗⠧⠎	yrvs	**y**ou**r**sel**ves**

คำย่อเบรลล์ มีวิธีการใช้ดังนี้ (Braille Authority of North America, 2002: XIII-10)

ตารางที่ 2.18 การใช้คำย่ออักษรเบรลล์

วิธีการใช้	ตัวอย่าง		
	คำหรือประโยค	รูปย่อในเบรลล์	แอสกี
1) ใช้ย่อคำโดด	He is my friend.	⠓⠑ ⠊⠎ ⠍⠽ ⠋⠗⠊⠑⠝⠙	,he is my fr4
2) ใช้ย่อส่วนของคำ	accordingly	⠁⠉⠉⠕⠗⠙⠬	acly
3) ใช้กับคำวิสามัญนามที่เป็นคำโดด	Dr. Good	⠠⠙⠗⠄ ⠠⠛⠕⠕⠙	,dr4 ,gd
4) หากคำเต็มใช้รูปเดียวกับคำย่อให้ใส่ ‘⠰’ (เครื่องหมายกำกับตัวอักษร) ที่หน้าคำ	Al Johnson	⠠⠰⠁⠇ ⠠⠰⠚⠕⠓⠝⠎⠕⠝	;,al ,johnson
5) กรณีคำย่อเป็นที่ส่วนหนึ่งของคำ หากแยกคำออกจากกันให้รักษารูปคำย่อไว้	mis- conceived	⠍⠊⠎⠤ ⠉⠕⠝⠉⠑⠊⠧⠙	mis- 3cvd

เป็นที่น่าสังเกตว่าในการปรากฏของคำย่อเบรลล์ กลุ่มตัวอักษรแต่ละกลุ่มจะใช้แทนคำแต่ละคำทำให้ไม่สามารถจะเขียนกลุ่มเบรลล์เหล่านี้ตามลำพังได้ ฉะนั้นเมื่อต้องการแสดงกลุ่มตัวอักษรเหล่านี้ในลักษณะที่ไม่ใช่ตัวย่อ ต้องใช้เครื่องหมายกำกับอักษร (⠰) นำหน้ากลุ่มตัวอักษรเหล่านี้ ตัวอย่างเช่นเมื่อพบ ‘⠸⠃⠰⠁⠸⠃’ (=ab) จะต้องตีความเป็น “about” แต่หากต้องการจะหมายถึงตัวอักษร “ab” จำเป็นต้องใส่เครื่องหมาย ‘⠰⠰’ ไว้ข้างหน้า ซึ่งใช้บอกว่าสิ่งที่ตามมาจะเป็นตัวอักษร เช่น ‘⠰⠰⠁⠃’ จึงจะตีความเป็นตัวอักษร “ab” ได้

จากข้อมูลที่นำเสนอมาข้างต้นพอจะสรุปได้ว่า ในการใช้ตัวย่อเบรลล์อังกฤษนั้นตัวย่อแต่ละตัวจะใช้แทนกลุ่มตัวอักษรที่แตกต่างกันจึงไม่ควรใช้ตัวย่อในตำแหน่งที่ไม่เหมาะสมเพราะจะทำให้เกิดความสับสน นอกจากนี้คำบางคำใช้ตัวย่อได้หลายรูปจึงควรเลือกกฎการใช้ตามลำดับความสำคัญเพื่อให้ได้รูปย่อที่เหมาะสมที่สุด สิ่งที่ควรคำนึงถึงอีกประการคือต้องใช้ตัวย่อกับคำพยางค์เดียวกันโดยเฉพาะในกรณีที่พยางค์นั้นเป็นส่วนหนึ่งของวิภัติ (prefix) หรือปัจจัย (suffix)

2.1.2 อักษรเบรลล์ไทย

มิสเจเนวีฟ คอลฟิลด์[4]ได้ร่วมกับนักวิชาการไทยกำหนดรหัสอักษรเบรลล์ไทยขึ้น (สมทรง พันธ์สุวรรณ, 2538: 19; อาทิตยา บุญมาก และคณะ, 2547: 12) อย่างไรก็ตามในประเทศไทยยังไม่มีเอกสารหรืองานวิจัยชิ้นใดที่อธิบายอักษรเบรลล์ไทยอย่างเป็นระบบ มีเพียงการจัดการเรียนการสอนอักษรเบรลล์โดยการท่องจำรหัสอักษรเบรลล์เท่านั้น สำหรับอักษรเบรลล์ไทยได้มีผู้แสดงรหัสต่าง ๆ ทั้งพยัญชนะ สระ และวรรณยุกต์ ดังจะขอนำเสนอต่อไปนี้

2.1.2.1 พยัญชนะในเบรลล์ไทย

อักษรเบรลล์ไทยประกอบด้วยพยัญชนะภาษาไทย 44 ตัว เช่นเดียวกับตัวอักษรปกติ ซึ่งสามารถแสดงรหัสในอักษรเบรลล์ได้ดังนี้

---

[4] มิสเจเนวีฟ คอลฟิลด์ (Genevive Caulfield) สุภาพสตรีตาบอดชาวอเมริกัน เป็นผู้ริเริ่มการศึกษาของคนตาบอดในประเทศไทย โดยร่วมกับนายแพทย์ ฝน แสงสิงแก้ว จัดตั้งโรงเรียนสอนคนตาบอดกรุงเทพ ขึ้นเมื่อปีพ.ศ.2482

ตารางที่ 2.19 พยัญชนะในเบรลล์ไทย

เบรลล์	แอสกี	อักษรไทย
⠛	g	ก
⠅	k	ข
⠚⠅	0k	ฃ
⠥	u	ค
⠤⠥	-u	ฅ
⠠⠥	,u	ฆ
⠻	]	ง
⠚	j	จ
⠌	/	ฉ
⠖	+	ช
⠮	!	ซ
⠠⠖	,+	ฌ
⠠⠽	,y	ญ
⠠⠙	,d	ฎ
⠠⠜	,\	ฏ
⠠⠞	,t	ฐ
⠠⠣	,)	ฑ
⠤⠣	-)	ฒ
⠠⠝	,n	ณ
⠙	d	ด
⠜	\	ต
⠞	t	ถ

เบรลล์	แอสกี	อักษรไทย
⠣	)	ท
⠚⠣	0)	ธ
⠝	n	น
⠧	v	บ
⠯	&	ป
⠏	p	ผ
⠭	x	ฝ
⠹	?	พ
⠫	$	ฟ
⠠⠹	,?	ภ
⠍	m	ม
⠽	y	ย
⠗	r	ร
⠇	l	ล
⠺	w	ว
⠠⠎	,s	ศ
⠤⠎	-s	ษ
⠎	s	ส
⠓	h	ห
⠠⠇	,l	ฬ
⠕	o	อ
⠿	=	ฮ

สมทรง พันธ์สุวรรณ (2538) ตั้งข้อสังเกตว่า พยัญชนะเบรลล์ไทยบางส่วน เกิดจากการนำเสียงพยัญชนะอังกฤษมาใช้เป็นฐานในการสร้างอักษรเบรลล์ไทย ซึ่งประกอบด้วยตัวอักษรต่อไปนี้

ตารางที่ 2.20 ตัวอักษรเบรลล์ไทยพื้นฐาน 15 ตัว

ไทยปกติ	เบรลล์ไทย	อังกฤษปกติ	เบรลล์อังกฤษ
ก	⠛	g	⠛
ข	⠅	k	⠅
จ	⠚	j	⠚
ด	⠙	d	⠙
ถ	⠞	t	⠞
น	⠝	n	⠝
ผ	⠏	p	⠏
ม	⠍	m	⠍
ย	⠽	y	⠽
ร	⠗	r	⠗
ล	⠇	l	⠇
ว	⠺	w	⠺
ส	⠎	s	⠎
ห	⠓	h	⠓
อ	⠕	o	⠕

จากการทบทวนวรรณกรรม ผู้วิจัยยังไม่พบงานเขียนใด ๆ ที่อธิบายเกี่ยวกับเบรลล์ไทยในแง่มุมอื่น แต่ผู้วิจัยได้ค้นพบว่ามีการนำพยัญชนะเบรลล์ไทยบางตัวมาเป็นฐาน เพื่อสร้างพยัญชนะอื่นที่มีลักษณ์ใกล้เคียงกันใน 2 รูปแบบ ดังนี้

- **รูปแบบที่ 1** การเติม ‘⠂’ (จุด 6) ในพยัญชนะฐานทำให้เกิดพยัญชนะใหม่ที่มีคุณสมบัติใกล้เคียงกันแต่เป็นอักษรที่ต่างหมู่กัน เช่น ‘⠎’ ใช้แทน “ส” ‘⠵’ ใช้แทน “ซ” พยัญชนะทั้งสองตัวนี้ต่างเป็นพยัญชนะเสียงเดียวกัน แต่ต่างกันตรงที่ “ส” เป็นอักษรสูง ในขณะที่ “ซ” เป็นอักษรต่ำ
- **รูปแบบที่ 2** การเพิ่มเซลล์ คือ ‘⠁’ (จุด 6) ‘⠉’ (จุด 3-6) หรือ ‘⠋’ (จุด 3-5-6) หน้าพยัญชนะฐานทำให้เกิดพยัญชนะใหม่ที่มีเสียงใกล้เคียงกันและอยู่ในหมู่เดียวกัน เช่น‘⠵’ใช้แทน“ท” ‘⠁⠵’ ใช้แทน “ฑ” ‘⠉⠵’ใช้แทน “ฒ” และ ‘⠋⠵’แทน “ธ” ซึ่งพยัญชนะเหล่านี้ล้วนอยู่ในหมู่อักษรต่ำ

ข้อสังเกตเหล่านี้ทำให้ผู้วิจัยต้องการหาคำตอบว่าพยัญชนะเบรลล์ไทยมีรูปที่ปรากฏอย่างเป็น

ระบบหรือไม่ และพยัญชนะแต่ละตัวมีความสัมพันธ์กันในลักษณะใด วิทยานิพนธ์ฉบับนี้จึงมุ่งที่จะหาคำตอบให้กับประเด็นเหล่านี้

2.1.2.2 สระและวรรณยุกต์ในเบรลล์ไทย

นอกจากพยัญชนะแล้ว อักษรเบรลล์ไทยยังประกอบด้วยสระและวรรณยุกต์ดังนี้

ตารางที่ 2.21 สระและวรรณยุกต์ในเบรลล์ไทย

เบรลล์	แอสกี	อักษรปกติ
⠁	a	ะ
⠇	>	ั
⠡	*	า
⠵	z	ำ
⠃	b	ิ
⠆	2	ี
⠣	{	ึ
⠱	5	ื
⠉	c	ุ
⠒	3	ู
⠗⠂	r1	ฤ
⠗⠂⠡	r1*	ฤา
⠇⠂	l1	ฦ
⠇⠂⠡	l1*	ฦา
⠋	f	เ
⠣	<	แ
⠊	i	โ
⠇	:	ไ
⠂	1	ๆ
⠇⠂	:1	ใ

เบรลล์	แอสกี	อักษรปกติ
⠄	’	็
⠐	9	่
⠔	4	้
⠶	7	๊
⠦	8	๋
⠴	0	์
⠋⠣	fa	เะ
⠣⠣	<a	แะ
⠊⠣	ia	โะ
⠒⠣	oa	เาะ
⠒	o	อ
⠚⠣	%a	เอะ
⠚	%	เอ
⠦⠣	(a	เียะ
⠦	(	เีย
⠶⠣	qa	เือะ
⠶	q	เือ
⠵⠣	ea	ัวะ
⠵	e	ัว
⠒	6	เา

จากรหัสสระเบรลล์ไทยที่นำเสนอมาข้างต้น ผู้วิจัยพบว่ามีการนำอักขระเบรลล์รูปหนึ่งมาดัดแปลงเพื่อทำให้เป็นอักขระอีกรูปหนึ่งที่มีลักษณ์ใกล้เคียงกัน การดัดแปลงทำได้ 2 ลักษณะคือ ลักษณะแรก นำ*สระเบรลล์รูปหนึ่งมาเติม ‘⠄’ (จุด 6)* เพื่อทำให้เป็นสระอีกรูปหนึ่งที่มีลักษณ์ใกล้เคียงกัน เช่น ‘⠉’ ใช้แทน “ะ” ‘⠍’ ใช้แทน “า” สระทั้ง 2 รูปนี้เป็นสระเสียงเดียวกัน แต่ต่างกันที่ความสั้นยาวของสระเท่านั้น อีกลักษณะหนึ่งคือการนำสระเบรลล์รูปหนึ่งมาเป็นฐานโดยยังคงรูปเดิมไว้แต่มี*การเลื่อนตำแหน่งของจุดให้ต่ำลง* เช่น ‘⠥’ ใช้แทน “ุ” ในขณะที่ ‘⠳’ ใช้แทน “ู” ทั้งสองรูปนี้เป็นสระเสียงเดียวกันแต่แตกต่างกันที่ความสั้นยาวของสระเท่านั้น

วิทยานิพนธ์ฉบับนี้จึงมุ่งที่จะศึกษาว่าสระและวรรณยุกต์เบรลล์ไทย มีความสัมพันธ์กับอักษรเบรลล์ในภาษาอื่น ๆ หรือไม่ และตัวอักษรเบรลล์ไทยแต่ละตัวมีความสัมพันธ์กันหรือไม่ ในลักษณะใด ซึ่งจะขอนำเสนอผลการวิจัยในลำดับต่อไป

2.1.2.3 หลักการประสมอักษรในเบรลล์ไทย

การประสมอักษรในเบรลล์ไทยมีลักษณะคล้ายกับการประสมอักษรในภาษาไทยปกติ แต่แตกต่างตรงที่เบรลล์ไทยไม่มีสระหรือวรรณยุกต์ที่อยู่เหนือหรือใต้พยัญชนะ ฉะนั้นเมื่อพบสระและ/หรือวรรณยุกต์ ก็ให้จัดลำดับอักขระเบรลล์ตามลำดับการจัดเรียงในอักษรปกติ เช่น “ลิ้น” ในเบรลล์ไทยจะปรากฏรูปเป็น ‘⠇⠊⠻⠝’ (ล + ิ + ้ + น)

อย่างไรก็ตาม การประสมอักษรในเบรลล์ไทยยังมีลักษณะแตกต่างจากอักษรปกติบางประการดังนี้ (อาทิตยา บุญมากและคณะ, 2547)

(1) การเขียนสระ ◌ะ ,◌ำ ในอักษรเบรลล์ต้องมีการจัดลำดับคือ พยัญชนะต้น + ◌ะ /◌ำ + (วรรณยุกต์) แต่ในอักษรปกติมีการจัดลำดับเป็น พยัญชนะต้น + วรรณยุกต์ (โดยที่วรรณยุกต์จะปรากฏเหนือพยัญชนะต้น) + ◌ะ /◌ำ ตัวอย่างเช่น ‘⠉⠊⠍’ (จ + ◌ะ + ◌๊) "จ๊ะ" ซึ่งในอักษรปกติจัดลำดับเป็น จ + ◌๊+ ◌ะ

(2) การนำอักขระเบรลล์มาแทนสระที่มีรูปซับซ้อนในอักษรปกติ ได้แก่สระ เ◌ะ แ◌ะ โ◌ะ เ◌อะ เ◌ียะ เ◌ีย เ◌ือะ เ◌ือ เ◌า ◌ัวะ ◌ัว ให้ใช้รหัสอักษรเบรลล์เฉพาะเพื่อแทนสระเหล่านั้น โดยให้เขียนพยัญชนะตามด้วยสระ แล้วจึงตามด้วยวรรณยุกต์ (ถ้ามี) และตัวสะกด (ถ้ามี)

ตัวอย่าง ‘⠗⠊⠻⠝’ ร + เ◌ีย + น ⟹ "เรียน"

‘⠗⠊⠥⠝⠛’ ร + เ◌ือ + ◌่ + ง ⟹ "เรื่อง"

(3) การเขียนไม้ยมก (ๆ) ให้เขียนติดกับคำที่อยู่ข้างหน้า โดยไม่ต้องเว้นวรรค

ตัวอย่าง ‘⠞⠁⠇⠁⠝⠖’ ต + ◌่ + า + ง + ๆ ⟹ "ต่าง ๆ"

2.1.3 ตัวเลขเบรลล์

แม้ว่างานวิจัยนี้มีวัตถุประสงค์ที่จะแปลงข้อความเบรลล์ไทยปนอังกฤษเป็นอักษรปกติเท่านั้นก็ตาม แต่ในความเป็นจริงแล้วข้อความต่าง ๆ ไม่ได้มีเฉพาะตัวอักษรเท่านั้น แต่มักจะมีตัวเลขปนอยู่ด้วย ด้วยเหตุนี้ผู้วิจัยจึงทำการทบทวนวรรณกรรมเกี่ยวกับการเขียนตัวเลขในอักษรเบรลล์ เพื่อใช้เป็นฐานความรู้ในการสร้างระบบการถ่ายถอดอักษรที่รองรับกับการแปลงตัวเลขในข้อความเบรลล์สองภาษาที่ปนกัน

ในการเขียนตัวเลขเบรลล์ ทำได้โดยใช้ตัวอักษร 10 ตัวแรกในภาษาอังกฤษมาใช้แทนตัวเลขตั้งแต่ 1 - 0 ตามลำดับ ดังแสดงในตารางต่อไปนี้ (Braille Authority of North America, 2002: VII-1)

ตารางที่ 2.22 ตัวเลขในอักษรเบรลล์

เบรลล์	⠁	⠃	⠉	⠙	⠑	⠋	⠛	⠓	⠊	⠚
แอสกี	a	b	c	d	e	f	g	h	i	j
ตัวเลข	1	2	3	4	5	6	7	8	9	0

2.1.3.1 วิธีการเขียนตัวเลขในอักษรเบรลล์

สัญลักษณ์ตัวเลขที่ปรากฏในตารางข้างต้น มีวิธีการใช้ดังนี้ (Risjord, 2009: 29-33; California Community Colleges, 2010: 32-36)

(1) การเขียนตัวเลขโดด ให้ใช้เครื่องหมาย ‘⠼’ (=#) หรือเครื่องหมายนำเลข นำหน้าตัวอักษรเหล่านั้น เพื่อแสดงว่าเป็นตัวเลข เช่น

⠼⠁ (= #a) ใช้แทน ‘1’

⠼⠃ (= #b) ใช้แทน ‘2’

⠼⠉ (= #c) ใช้แทน ‘3’

(2) หากเลขจำนวนใดมีตัวเลขมากกว่า 1 ตัวให้ใช้เครื่องหมายนำเลขนำหน้าตัวเลขตัวหน้า

⠼⠁⠚ (= #aj) ใช้แทน ‘10’

⠼⠁⠚⠚ (= #ajj) ใช้แทน ‘100’

⠼⠁⠚⠚⠚ (= #ajjj) ใช้แทน ‘1000’

(3) การเขียนตัวเลขกับคำที่ปรากฏข้างหลัง

ด้วยเหตุที่เบรลล์อังกฤษ 10 ตัวแรกได้นำมาใช้แทนตัวเลขเบรลล์ จึงทำให้จำเป็นต้องเขียนเว้นวรรคระหว่างตัวเลขและตัวอักษรเสมอ เช่น ‘⠼⠁⠚ ⠁⠛⠎’ (= #aj bags) หรือ “10 bags” ใน

อักษรปกติ หากเขียนชิดกันอาจทำให้เกิดการตีความผิดเป็น “10217s” ได้ เนื่องจาก ‘⠰⠁⠛’ (=bag) อยู่ชิดกับตัวเลขทำให้เกิดความสับสนได้ ดังนั้นเครื่องหมายวรรคจึงเป็นตัวบ่งชี้จุดสิ้นสุดของตัวเลขจำนวนใดจำนวนหนึ่ง

(4) การเขียนตัวเลขและตัวอักษรปนกัน

นอกจากการใช้เครื่องหมายเว้นวรรคแล้ว ยังมีการใช้ ‘⠰⠆’ (=;) หลังตัวเลข เพื่อแยกตัวเลขและตัวอักษรออกจากกันในกรณีที่ต้องเขียนตัวอักษรหลังตัวเลขโดยไม่มีเว้นวรรค เช่น ‘⠼⠃⠰⠆⠁’ (=#b⠰⠆a) จะใช้แทน “2a” เป็นต้น ในทางตรงกันข้ามหากต้องการเขียนตัวเลขหลังตัวอักษร ก็ต้องใช้เครื่องหมาย ‘⠼’ (= #) เช่น ‘⠰⠃⠼⠁’ (=b⠼a) ใช้แทน “b1”

(5) การใช้เครื่องหมายที่ปรากฏกับตัวเลข

หากมีเครื่องหมายเหล่านี้ปรากฏในสายอักขระของตัวเลขให้ใช้ ‘⠼’ (= #) ที่ตำแหน่งต้นของตัวเลขเพียงครั้งเดียว เครื่องหมายเหล่านี้แสดงได้ดังนี้

ตารางที่ 2.23 เครื่องหมายที่ปรากฏกับตัวเลข

สัญลักษณ์เบรลล์	สัญลักษณ์แอสกี	ใช้แทนเครื่องหมาย	หมายเหตุ
⠬	+	+	(บวก)
⠤	-	-	(ลบ)
⠐⠔	@*	*	(คูณ)
⠌	/	÷	(หาร)
⠐⠌	./	/	(หาร หรือทับ)
⠨	.	.	(จุดทศนิยม)
⠲	.	.	(จุดท้ายตัวเลข)
⠂	1	,	(คอมม่า)
⠒	3	:	(โคลอน)
⠐⠒⠏	@3p	%	(เปอร์เซ็น)

ตัวอย่างเช่น หากต้องการเขียน 2-3 (สองถึงสาม) ในอักษรเบรลล์สามารถทำได้ดังนี้

ถูก: ‘⠼⠃⠤⠉’ (#b-c)

ผิด: ‘⠼⠃⠤⠼⠉’ (#b-#c)

จะเห็นได้ว่าระบบการเขียนตัวเลขเบรลล์ที่ได้นำเสนอมานี้เป็นเพียงสัญลักษณ์ที่ปรากฏในข้อความทั่วไปในอักษรเบรลล์ อย่างไรก็ตามในศาสตร์ต่าง ๆ ที่เกี่ยวข้องกับตัวเลข เช่น คณิตศาสตร์ วิทยาศาสตร์ ดนตรีฯลฯ จะมีวิธีการเขียนที่แตกต่างออกไป เช่น การใช้เลขต่ำเพื่อแทนตัวเลขเหล่านี้ เนื่องจากวิทยานิพนธ์นี้มุ่งที่จะพัฒนาระบบการแปลงอักษรเบรลล์ที่ปรากฏในข้อความทั่วไป ดังนั้น งานวิจัยฉบับนี้จะไม่ขอกล่าวถึงระบบตัวเลขที่ใช้เฉพาะด้านที่ได้อภิปรายมาข้างต้น

2.2 การจำแนกภาษาด้วยคอมพิวเตอร์

ดังที่ได้กล่าวมาแล้วว่าวัตถุประสงค์สำคัญสำหรับวิทยานิพนธ์ฉบับนี้คือ การถ่ายถอดอักษรเบรลล์ที่มีภาษาไทยและอังกฤษปนกันเป็นอักษรปกติ อย่างไรก็ตามเป็นไปไม่ได้ที่คอมพิวเตอร์จะให้ผลลัพธ์ได้ถูกต้องหากไม่รู้ว่าข้อความเบรลล์ที่ต้องการถ่ายถอดเป็นภาษาใด เนื่องจากเบรลล์อังกฤษและเบรลล์ไทยต่างก็ใช้ตัวอักษรระบบ 6 จุด จึงเกิดการซ้ำรูปในภาษาทั้งสอง ทำให้ยากที่เครื่องคอมพิวเตอร์จะบอกได้ว่ากำลังทำงานกับอักษรเบรลล์ของภาษาใด

ดังนั้นก่อนที่จะทำการถ่ายถอดอักษรเบรลล์เป็นอักษรปกติ คอมพิวเตอร์ต้องจำแนกภาษาในข้อความเบรลล์ให้ได้เป็นอันดับแรก

ในการประมวลผลภาษาธรรมชาติ การจำแนกภาษาด้วยคอมพิวเตอร์ จัดเป็นประเด็นปัญหาหนึ่งในการจัดประเภทข้อความ (text classification) เป็นการระบุว่าข้อความที่คอมพิวเตอร์กำลังประมวลผลอยู่นั้นเป็นภาษาใด (Grothe, De Luca & Nürnberger 2008: 1) การจำแนกภาษาด้วยคอมพิวเตอร์ประกอบด้วย 2 ส่วนหลัก คือ แบบจำลองการจำแนกภาษา (language identification models) และวิธีการในการจำแนกภาษา (language classification method)

ส่วนสำคัญในแบบจำลองการจำแนกภาษาคือ แบบจำลองภาษา (language models) ซึ่งอาจจะเป็นเซ็ทของคำหรือเซ็ทของกลุ่มตัวอักษรซึ่งมีลักษณะพิเศษที่ได้มาจากคลังข้อมูล โดยนำมาใช้เป็นต้นแบบในการเปรียบเทียบกับข้อความที่ต้องการระบุภาษา หากพิจารณาตามลักษณะข้อมูลที่ปรากฏในแบบจำลองภาษาสามารถจะแบ่งวิธีการจำแนกภาษาออกได้เป็น 2 ประเภท คือ การจำแนกภาษาในระดับคำ (word based) และการจำแนกภาษาในระดับอักขระ (character based)

ต่อไปนี้ ผู้วิจัยจะกล่าวถึงการจำแนกภาษาในระดับคำ และการจำแนกภาษาในระดับอักขระ โดยมีรายละเอียดดังนี้

2.2.1 การจำแนกภาษาในระดับคำ

ในการจำแนกภาษาหากใช้คำมาสร้างแบบจำลองภาษา จากการทบทวนวรรณกรรมด้านนี้ พบว่ามีวิธีการทำได้แตกต่างกันอย่างน้อย 2 วิธี วิธีแรกคือการใช้คำที่มีความถี่ในการปรากฏสูงในคลังข้อมูล (frequent word-based approach) มาเป็นตัวแทนภาษานั้นๆ วิธีที่สองคือการใช้คำขนาดสั้น ซึ่งจะมีจำนวนอักขระไม่เกินจำนวนที่กำหนด (short word-based approach) มาเป็นตัวแทนภาษานั้นๆ

2.2.1.1 การใช้คำที่มีความถี่สูง

การจำแนกภาษาด้วยวิธีการใช้คำที่มีความถี่สูง อิงกับแนวคิดที่ว่า ในแต่ละภาษามีคำจำนวนหนึ่งที่มีความถี่ในการปรากฏสูงและมีรูปแตกต่างจากภาษาอื่น คำเหล่านี้สามารถนำมาใช้ในการสร้างแบบจำลองภาษา (language models) เพื่อบ่งชี้แต่ละภาษาได้

ในการจำแนกภาษาด้วยวิธีการนี้ เริ่มต้นจากการแจกแจงรายการคำที่มีความถี่ในการปรากฏสูงที่สุดจำนวนหนึ่งในคลังข้อมูลของแต่ละภาษา แล้วนำคำเหล่านั้นมาสร้างแบบจำลองของแต่ละภาษา จากนั้นจึงนำคำแต่ละคำในข้อความที่ต้องการระบุภาษาไปเปรียบเทียบเพื่อคำนวณคะแนนว่าคำเหล่านั้นใกล้เคียงกับแบบจำลองภาษาใดมากที่สุด หลังจากทำเช่นนี้จนครบทั้งข้อความ ก็จะระบุให้ข้อความนำเข้าเป็นภาษาที่คำนวณแล้วได้คะแนนรวมสูงสุด

ซูเตอร์ (Souter et al., 1994) มาร์ติโน (Martino et al., 2001) และโควีย์ (Cowie et al., 1999) ต่างก็นำแนวคิดนี้มาใช้ในการจำแนกภาษา โดยที่ซูเตอร์และมาร์ติโนต้องการจำแนกข้อความของภาษาที่ใช้ในทวีปยุโรป ในขณะที่โควีย์ต้องการจำแนกภาษาที่ใช้ในทวีปยุโรป เอเชีย และอาฟริกา สำหรับการสร้างแบบจำลองภาษา ซูเตอร์และมาร์ติโนได้ใช้คำที่มีความถี่สูงสุด 100 คำแรกของแต่ละภาษา ในขณะที่โควีย์ใช้คำที่มีความถี่สูงถึง 1,000 คำแรก

สิ่งที่น่าสังเกตคือผู้วิจัยทั้ง 3 คน มีวิธีการคิดคะแนนของแบบจำลองที่แตกต่างกัน ซูเตอร์และโควีย์ใช้วิธีการเพิ่มค่าตัวนับ (counter) ทีละ 1 ให้กับแบบจำลองภาษาที่มีคำตรงกับคำในข้อความนำเข้า ในขณะที่มาร์ติโนใช้วิธีการเพิ่มค่าตัวนับ (ซึ่งมาร์ติโนเรียกว่า accumulator) ให้กับแต่ละแบบจำลองภาษา โดยใช้ค่าความถี่ปรับแบบบรรทัดฐาน (normalized frequency of occurrence หรือ NFO) ซึ่งคำนวณโดย นำความถี่ของคำแต่ละคำในแต่ละแบบจำลองมาหารด้วยจำนวนคำทั้งหมดในคลังข้อมูลภาษานั้น แล้วนำผลลัพธ์ทุกตัวมารวมกันหารด้วยความถี่ของคำที่มีการปรากฏสูงสุดในแบบจำลองนั้น

นอกจากการนำคำที่มีความถี่สูงสุดมาใช้จำแนกภาษาของข้อความได้แล้ว คำขนาดสั้นที่ปรากฏในแต่ละภาษาก็ยังสามารถใช้เป็นตัวบ่งชี้ภาษาได้ ดังจะได้นำเสนอในหัวข้อต่อไป

2.2.1.2 การใช้คำขนาดสั้น

การใช้คำขนาดสั้นเพื่อจำแนกภาษาเป็นการรวบรวมคำที่มีจำนวนอักขระไม่เกินที่กำหนดไว้มาสร้างเป็นแบบจำลองภาษาของแต่ละภาษาโดยไม่คำนึงถึงความถี่ในการปรากฏของคำเหล่านั้น

การจำแนกภาษาด้วยวิธีนี้ใช้แนวคิดที่ว่า คำทั่วไป (common word) ที่ทำให้สามารถระบุภาษาได้ ได้แก่ คำหน้าที่ (function) เช่น คำสันธาน คำบุพบท ส่วนใหญ่เป็นคำที่มีความยาวไม่มากนัก (Grefenstette, G. 1995:3)

ในการจำแนกภาษาด้วยวิธีนี้ เริ่มจากนำข้อความของแต่ละภาษาจากคลังข้อมูลไปแบ่งคำแล้วเลือกเฉพาะคำที่มีจำนวนอักขระไม่เกินจำนวนที่กำหนดไว้ไปสร้างแบบจำลองภาษา หลังจากนั้นนำคำเหล่านี้ไปคำนวนความน่าจะเป็น (probability) โดยคิดจาก ความถี่ในการเกิดของแต่ละคำหารด้วยผลรวมความถี่ของคำทั้งหมดในแบบจำลองภาษา ในการจำแนกภาษาของประโยคที่นำเข้าทำโดยนำคำแต่ละคำในประโยคไปเปรียบเทียบกับคำในแต่ละแบบจำลองภาษา หากตรงกับคำในแบบจำลองใดก็ให้นำค่าความน่าจะเป็นของคำนั้น(probability) ในแบบจำลองไปรวมกับผลรวมเดิมของแบบจำลองภาษานั้น หากไม่ตรงกับคำในแบบจำลองใดก็ให้นำค่าความน่าจะเป็นต่ำสุด (minimum probability) ที่กำหนดขึ้น ไปรวมกับผลรวมเดิมของแบบจำลองภาษานั้นเช่นกัน หลังจากเปรียบเทียบคำทุกคำกับแบบจำลองภาษาทั้งหมด แล้วจึงทำการเปรียบเทียบผลรวมค่าความน่าจะเป็นของประโยคที่ได้ในแต่ละแบบจำลองภาษา เพื่อหาแบบจำลองภาษาที่มีคะแนนรวมสูงสุดอันจะทำให้ระบุได้ว่าเป็นภาษาใด

เกรเฟนส์เตท (Grefenstette, G. 1995) และเพรเกอร์ (Prager, G.M. 1999) ได้ใช้วิธีการดังกล่าวมาจำแนกภาษา โดยเกรเฟนส์เตทเลือกเฉพาะคำที่มีความยาวไม่เกิน 5 ตัวอักษรซึ่งส่วนใหญ่เป็นคำหน้าที่ (function words) มาสร้างแบบจำลองภาษา ในขณะที่เพรเกอร์ใช้ความยาวของคำไม่เกิน 4 ตัวอักษร เนื่องจากเชื่อว่าคำที่มีความยาวเท่าที่ระบุมา ก็สามารถจำแนกภาษาได้ดีเท่ากับการใช้คำที่มีขนาด ตามที่เกรเฟนส์เตทได้นำเสนอ

เป็นที่น่าสังเกตว่า การจำแนกภาษาในระดับคำมีข้อจำกัดบางประการ ประการแรก ในบางกรณีข้อความที่นำมา ไม่มีคำสั้นหรือคำที่มีความถี่สูงปรากฏอยู่เลย เช่น ข้อความพาดหัวข่าว ซึ่งหากพบกรณีนี้คอมพิวเตอร์จะไม่สามารถระบุภาษาของข้อความเหล่านั้นได้ ประการต่อมา ภาษาต่าง ๆ ในโลก เช่น ภาษาไทย จีน เกาหลี ญี่ปุ่น ฯลฯ ไม่มีการเว้นวรรคระหว่างคำ หากจะนำวิธีการดังกล่าวมาใช้กับภาษาเหล่านี้ ก็จำเป็นต้องทำการตัดคำล่วงหน้า ทำให้เกิดความซับซ้อนยิ่งขึ้น จึงไม่เหมาะสมที่จะนำวิธีการจำแนกภาษาในระดับคำมาใช้กับภาษาที่มีข้อจำกัดดังกล่าว

2.2.2 การจำแนกภาษาในระดับอักขระ

ภาษาต่าง ๆ ในโลกอาจใช้ตัวอักษรร่วมกัน เช่นภาษาอังกฤษ เยอรมัน ฝรั่งเศส แต่ภาษาเหล่านี้ก็มีการจัดเรียงตัวอักษรที่แตกต่างกัน เมื่อพิจารณารูปแบบการจัดเรียงตัวอักษรก็อาจระบุได้ว่าเป็นภาษาใด เช่นคำที่ลงท้าย –ck สามารถระบุได้ว่าเป็นภาษาอังกฤษ ในขณะที่คำที่ลงท้าย –ez จะสามารถระบุได้ว่าเป็นภาษาฝรั่งเศส (Grefenstette, G. 1995: 2) ในการจำแนกภาษาในระดับอักขระอาจทำได้ด้วยวิธีการต่าง ๆ คือ การใช้สายอักขระที่มีความเป็นเอกลักษณ์ของแต่ละภาษา (unique character strings) และการใช้แบบจำลองเอ็นแกรม (n-grams model) มาใช้ระบุภาษา

2.2.2.1 การใช้สายอักขระที่เป็นเอกลักษณ์ของแต่ละภาษา

การจำแนกภาษาโดยใช้สายอักขระที่เป็นเอกลักษณ์ของแต่ละภาษา ตั้งอยู่บนสมมติฐานที่ว่า ภาษาต่าง ๆ ในโลกจะมีการเรียงต่อของสายอักขระที่ไม่เหมือนกันแม้จะเป็นภาษาในตระกูลเดียวกันก็ตาม ความแตกต่างในการเรียงตัวของสายอักขระ เป็นลักษณะพื้นฐานของแต่ละภาษาที่สามารถนำมาใช้ในการบ่งชี้เอกลักษณ์ของภาษานั้น ๆ ได้ เช่นคำที่ขึ้นต้นด้วย LL จะเป็นภาษาเวลส์ (Welsh), หรือคำที่มี CZY ปรากฏนั้นจะเป็นภาษาโปแลนด์ (Polish) เป็นต้น

การหาสายอักขระที่เป็นเอกลักษณ์เฉพาะแต่ละภาษาอาจทำได้โดยอาศัยนักภาษาศาสตร์ที่มีความเชี่ยวชาญในภาษาเหล่านั้น หรือนำข้อความที่สุ่มจากภาษาต่าง ๆ มาเปรียบเทียบกันเพื่อหาสายอักขระที่ไม่ซ้ำกัน (Dunning, T. 1994: 3) อย่างไรก็ตาม ในการหาสายอักขระที่ปรากฏเฉพาะแต่ละภาษานั้น อาจจะต้องกำหนดความยาวของสายอักขระที่เหมาะสมจึงจะสามารถจำแนกภาษาได้ถูกต้อง (Souter et al., 1994: 187) ในบางกรณี สายอักขระที่มีความยาว 2 ตัว อาจไม่สามารถนำมาใช้จำแนกภาษาหนึ่งออกจากอีกภาษาหนึ่งได้ เนื่องจากภาษาทั้งสองนั้นมีการเรียงตัวของอักขระ 2 ตัวที่ซ้ำกันเป็นจำนวนมากทำให้จำเป็นต้องใช้อักขระที่มีความยาวมากกว่านั้นจึงจะสามารถระบุภาษาได้

ในการจำแนกภาษาของข้อความนำเข้าด้วยวิธีนี้ ทำได้โดยนำสายอักขระของข้อความนำเข้าไปเปรียบเทียบกับสายอักขระในแต่ละแบบจำลองภาษา แล้วให้ค่าคะแนนสำหรับแต่ละแบบจำลองภาษา แบบจำลองใดได้ค่าคะแนนมากที่สุดก็จะระบุให้ข้อความนำเข้าเป็นภาษานั้น

ซูเตอร์ (1994) ได้ตั้งข้อสังเกตว่า วิธีดังกล่าวให้ความถูกต้องในการจำแนกภาษาได้น้อยกว่าการใช้แบบจำลองเอ็นแกรมและการใช้คำที่มีความถี่สูง เนื่องจากบางครั้งคลังข้อความที่นำมาใช้ในการสร้างแบบจำลองภาษามีข้อมูลที่ไม่ครอบคลุมกับภาษานั้น ๆ เช่นในคลังข้อมูลที่ซูเตอร์นำมาใช้มีการปรากฏของ nx ในคลังข้อมูลภาษาโปรตุเกสแต่ไม่ปรากฏในคลังข้อมูลภาษาอังกฤษ ทำให้ไม่สามารถระบุได้ว่า “anxious” เป็นคำในภาษาอังกฤษได้

2.2.2.2 การใช้แบบจำลองเอ็นแกรม

ในการจำแนกภาษาในระดับอักขระอาจทำได้โดยการใช้แบบจำลองเอ็นแกรม แบบจำลองเอ็นแกรมคือแบบจำลองที่ใช้คำนวณค่าความน่าจะเป็นของสายอักขระ (character sequence) ที่เรียงต่อกันจำนวน n ตัว โดยค่าความน่าจะเป็นของชุดอักขระสามารถคำนวณได้จากคลังข้อมูลที่ได้รวบรวมไว้สำหรับแต่ละภาษา (Grothe, De Luca & Nürnberger, 2008: 2)

สิ่งที่จำเป็นสำหรับการจำแนกภาษาโดยใช้เทคนิคเอ็นแกรม คือแบบจำลองภาษา (language models) ของแต่ละภาษา และแบบจำลองข้อความ (document models) โดยที่แต่ละแบบจำลองภาษา เกิดจากการรวบรวมสายอักขระเอ็นแกรมจากคลังข้อความของแต่ละภาษา ในขณะที่แบบจำลองข้อความ เกิดจากการรวบรวมสายอักขระเอ็นแกรมจากข้อความนำเข้า ในการระบุภาษาของข้อความใดข้อความหนึ่ง ทำได้โดยนำแต่ละสายอักขระในแบบจำลองข้อความไปเปรียบเทียบกับสายอักขระในแต่ละแบบจำลองภาษา เพื่อหาความน่าจะเป็นในการปรากฏของสายอักขระนั้นในแต่ละแบบจำลองภาษา หลังจากนั้นค่าความน่าจะเป็นของสายอักขระทั้งหมดสำหรับแต่ละแบบจำลองภาษาจะถูกนำไปเปรียบเทียบกันเพื่อบ่งชี้ว่าข้อความนำเข้าจัดอยู่ในภาษาใด

เกรเฟนส์เตท(Grefenstette, G. 1995) ได้ใช้ไตรแกรมในการจำแนกภาษาของข้อความโดยสร้างแบบจำลองภาษาที่ประกอบด้วยไตรแกรมที่มีความถี่ในการปรากฏในลำดับแรก ๆ ของคลังข้อความ ECI (European Corpus Initiative) ในการจำแนกภาษา ทำโดยนำแต่ละไตรแกรมของแบบจำลองข้อความ ซึ่งได้มาจากประโยคที่นำเข้าแต่ละประโยคมาเปรียบเทียบกับไตรแกรมในแต่ละแบบจำลองภาษา หากพบไตรแกรมที่เหมือนกันจะนำค่าความน่าจะเป็น (probability) ของไตรแกรมนั้น (ซึ่งคำนวณจากความถี่ของไตรแกรมนั้นหารด้วยความถี่ของไตรแกรมทุกตัวในแบบจำลองภาษา) มาเพิ่มคะแนนให้กับแบบจำลองภาษาที่พบไตรแกรม แต่หากไม่พบก็จะเพิ่มค่าความน่าจะเป็นที่น้อยที่สุด (minimal probability) ที่กำหนดขึ้นให้กับแบบจำลองภาษาที่ไม่พบไตรแกรมเช่นกัน หลังจากที่เปรียบเทียบทุกไตรแกรมในแบบจำลองข้อความกับไตรแกรมในแบบจำลองภาษาทั้งหมดแล้ว จะนำผลรวมค่าความน่าจะเป็นของแต่ละแบบจำลองภาษามาเปรียบเทียบกันเพื่อหาความน่าจะเป็นสูงสุด อันจะทำให้ระบุได้ว่าประโยคที่นำเข้านั้นเป็นภาษาใด

คาฟนาร์และเทรนเคิล (Cavnar, W. B. and Trenkle, J. M. 1994) ได้ใช้วิธีการจำแนกภาษาโดยใช้เอ็นแกรมของสายอักขระเช่นกัน เพียงแต่มีวิธีการนับคะแนนที่ต่างไป โดยไม่ได้นับจากค่าความน่าจะเป็นของสายอักขระแต่ใช้การจัดลำดับ (ranking) ของสายอักขระเอ็นแกรมของแต่ละแบบจำลองภาษา วิธีการระบุภาษาคือการนำแต่ละสายอักขระในแบบจำลองข้อความไปเปรียบเทียบว่าพบสายอักขระดังกล่าวในแต่ละแบบจำลองภาษาหรือไม่ หากพบสายอักขระที่เหมือนกันในแบบจำลองภาษาใดจะนำลำดับของสายอักขระในแบบจำลองข้อความและแบบจำลองภาษานั้นมา

เปรียบเทียบกันเพื่อหาระยะห่างลำดับ (เช่นสายอักขระxxx ปรากฏในลำดับที่ 6 ในแบบจำลองภาษาอังกฤษแต่ปรากฏในลำดับที่ 2 ในแบบจำลองข้อความ ก็จะมีระยะห่างลำดับเป็น 6-2 =4) แล้วนำค่าระยะห่างลำดับที่ได้ไปรวมกับค่าที่มีอยู่เดิมของแบบจำลองภาษานั้น หากไม่พบสายอักขระที่เหมือนกันในภาษาใดก็จะกำหนดค่าระยะห่างลำดับให้เป็นตัวเลขกลางซึ่งมีค่าสูงสุด (default maximum value) แล้วนำค่านี้ไปรวมกับค่าที่มีอยู่เดิมของแบบจำลองภาษานั้นเช่นกัน แบบจำลองภาษาใดที่มีผลรวมค่าระยะห่างลำดับต่ำที่สุดก็จะระบุให้ข้อความนำเข้าเป็นภาษานั้น

จากสิ่งที่ได้อภิปรายข้างต้น ผู้วิจัยพบว่ามีวิธีการหลากหลายในการจำแนกภาษาด้วยคอมพิวเตอร์ บางวิธีก็เหมาะที่จะประยุกต์ใช้กับงานวิจัยที่ผู้วิจัยกำลังทำการศึกษา ดังนั้นผู้วิจัยจะขออภิปรายถึงวิธีการที่เหมาะสมที่จะนำมาใช้ในการระบุภาษาในอักษรเบรลล์ ดังนี้

ดังที่ได้กล่าวมาแล้วว่า วิธีการหนึ่งที่นิยมนำมาใช้ระบุภาษาของข้อความ คือการใช้คำจำนวนหนึ่งที่แตกต่างกันในแต่ละภาษามาจำแนกข้อความของภาษาเหล่านั้นออกจากกัน (ดูหัวข้อ 2.2.2.1) แต่ผู้วิจัยพบว่า ในการจำแนกเบรลล์ไทยออกจากเบรลล์อังกฤษนั้น ไม่สามารถจะใช้คำแต่ละคำมาจำแนกข้อความเบรลล์ของสองภาษานี้ออกจากกันได้ เนื่องจากข้อความภาษาไทยไม่มีการระบุขอบเขตของคำอย่างชัดเจน (กานดา รุณนะพงศา และปโยธร อุราธรรมกุล. 2549: 3) แต่ด้วยเหตุที่ ข้อความแต่ละวรรคในภาษาอังกฤษประกอบด้วยคำเพียงคำเดียว ในขณะที่ภาษาไทยเกิดจากการนำคำหลายคำมาเขียนต่อกัน จึงทำให้ข้อความภาษาอังกฤษมีขนาดสั้นกว่าข้อความภาษาไทยอย่างชัดเจน ดังนั้นจึงสามารถจำแนกภาษาไทยออกจากภาษาอังกฤษได้โดยการพิจารณาความยาวสายอักขระ โดยที่สายอักขระใดมีความยาวมากก็น่าจะระบุได้ว่าเป็นภาษาไทย

นอกจากนี้ ยังสามารถนำวิธีการจำแนกภาษาโดยใช้สายอักขระที่เป็นเอกลักษณ์ของแต่ละภาษา มาใช้ระบุภาษาในอักษรเบรลล์ได้ แม้ว่าเบรลล์ไทยและเบรลล์อังกฤษจะมีรูปตัวอักษรที่เหมือนกัน ทั้งสองภาษานี้ก็มีรูปแบบของการเรียงต่อของตัวอักษรที่แตกต่างกันเป็นจำนวนมาก เช่น ‘⠣⠊⠝’ เป็นภาษาไทย (=กิน) ไม่ใช่ภาษาอังกฤษ(=gbn) ‘⠓⠕⠥⠎⠑’ เป็นภาษาอังกฤษ (= house) ไม่ใช่ภาษาไทย (= หอคสัว) อย่างไรก็ตาม ในบางครั้งอาจมีการเรียงต่อกันของตัวอักษรที่เหมือนกันในสองภาษา เช่น ‘⠙⠕⠛’ ซึ่งอาจใช้แทนคำว่า “dog” หรือ “ดอก” ในภาษาไทยก็ได้ หากเกิดกรณีนี้ขึ้นจะส่งผลให้ไม่สามารถระบุภาษาได้ จึงจำเป็นต้องนำสถิติความน่าจะเป็นมาใช้ร่วมกับการระบุภาษาด้วยสายอักขระที่เป็นเอกลักษณ์ของแต่ละภาษา

สำหรับรายละเอียดเกี่ยวกับแบบจำลองการระบุภาษาในอักษรเบรลล์ ที่จะนำเสนอในวิทยานิพนธ์ฉบับนี้ ผู้วิจัยจะอภิปรายในหัวข้อที่เกี่ยวข้องต่อไป

2.3 การถ่ายถอดอักษร

ดังที่ได้นำเสนอมาแล้วว่า วิทยานิพนธ์ฉบับนี้มีวัตถุประสงค์ที่จะถ่ายถอดอักษรเบรลล์ที่มีการปนกันระหว่างภาษาไทยและภาษาอังกฤษให้เป็นอักษรปกติด้วยคอมพิวเตอร์ ซึ่งกระบวนการนี้มีลักษณะคล้ายคลึงกับการถ่ายถอดข้อความจากระบบตัวเขียนหนึ่งไปสู่อีกระบบตัวเขียนหนึ่ง ด้วยเหตุนี้ผู้วิจัยจึงได้ทำการทบทวนวรรณกรรมที่เกี่ยวข้องเพื่อจะนำมาประยุกต์ใช้ในการถ่ายถอดอักษรเบรลล์เป็นอักษรปกติ โดยหัวข้อที่ 2.4.1 จะนำเสนอ แนวคิดทั่วไปเกี่ยวกับการถ่ายถอดอักษร หลังจากนั้นหัวข้อที่ 2.4.2 จะนำเสนอแนวทางการถ่ายถอดอักษรด้วยคอมพิวเตอร์

2.3.1 แนวคิดทั่วไปเกี่ยวกับการถ่ายถอดอักษร

การถ่ายถอดอักษร (transliteration) เป็นการบันทึกตัวเขียนในภาษาหนึ่ง (ภาษาต้นทาง) ด้วยตัวเขียนที่สอดคล้องกันในอีกภาษาหนึ่ง (ภาษาปลายทาง) โดยที่เสียงของทั้งสองภาษายังคงมีความใกล้เคียงกัน (Knight, K. & Greal, J. 1997: 599 ; Catford, J. Cunnison 1965: 50) ตัวอย่างเช่น คำว่า “ถ่ายเสียง” สามารถถ่ายถอดอักษรในภาษาอังกฤษได้เป็น “taisiang” เป็นต้น

หากจำแนกการถ่ายถอดอักษรโดยยึดความสอดคล้องของภาษาต้นทางและภาษาปลายทาง ก็อาจแบ่งการถ่ายถอดอักษรออกได้เป็น 2 ประเภทคือ การถ่ายถอดอักษรโดยยึดอักษรเป็นสำคัญ และ การถ่ายถอดอักษรโดยยึดเสียงเป็นสำคัญ (Coulmas, Folrian 1992: 241-243)

ในประเภทแรก การถ่ายถอดอักษรโดยยึดอักษรเป็นสำคัญเป็นการถอดรูปอักษรที่ปรากฏในคำนั้น ๆ มาให้ครบทุกรูปในลักษณะอักษร (หรือกลุ่มตัวอักษร) ต่ออักษร (หรือกลุ่มตัวอักษร) โดยไม่คำนึงว่าจะออกเสียงผิดเพี้ยนไปจากคำเดิมหรือไม่ การถ่ายถอดอักษรในลักษณะนี้ใช้ในกรณีที่ต้องการคงรูปศัพท์ในคำดั้งเดิมเอาไว้ เช่น การถ่ายถอดชื่อบุคคล (เช่น บงกช = Bonggoch) และการถ่ายถอดชื่อทางภูมิศาสตร์ต่าง ๆ (เช่น สุวรรณภูมิ = Suvarnabhumi) เป็นต้น จะเห็นได้ว่า การถ่ายถอดอักษรลักษณะนี้มีข้อดีคือ สามารถถอดอักษรเพื่อใช้ในงานค้นคืนข้ามภาษา (cross language information retrieval) ได้ง่าย เนื่องจากเป็นการจับคู่ตัวอักษรหรือกลุ่มตัวอักษรในภาษาหนึ่งกับตัวอักษรหรือกลุ่มตัวอักษรอีกภาษาหนึ่งโดยตรง แต่ก็มีข้อเสียคือ คำที่ได้จากการถ่ายถอดอักษรในภาษาปลายทางจะออกเสียงผิดเพี้ยนไปจากคำในภาษาต้นทาง เช่น ชื่อ Bonggoch ถ้ายึดตามการออกเสียงแบบภาษาอังกฤษจะต้องออกเสียงว่า “บนกอช” ซึ่งมีเสียงแตกต่างจากคำว่า “บงกช” ที่เป็นคำดั้งเดิมในภาษาไทย

ในอีกประเภทหนึ่ง การถ่ายถอดอักษรโดยยึดเสียงเป็นสำคัญ เป็นการถ่ายถอดที่ไม่จำเป็นต้องถอดครบทุกตัวอักษร แต่ต้องปรับให้คำที่เกิดจากการถ่ายถอดอักษร ออกเสียงได้

ใกล้เคียงกับคำดั้งเดิมมากที่สุด การถ่ายถอดอักษรประเภทนี้จะคำนึงถึงคุณสมบัติทางเสียงมากกว่าตัวอักษร ลักษณะเช่นนี้จึงคล้ายกับการถ่ายถอดอักษรแบบถ่ายเสียง ตัวอย่างเช่น คำว่า “สุวรรณภูมิ” เมื่อทำการถ่ายถอดอักษรตามแนวทางนี้จะได้เป็น Suwannaphum เป็นต้น

เป็นที่น่าสังเกตว่า ข้อดีของการถ่ายถอดอักษรแบบยึดเสียงเป็นสำคัญคือ ทำให้คำที่ถ่ายถอดอักษรมีเสียงใกล้เคียงกับคำดั้งเดิมซึ่งจะก่อให้เกิดประโยชน์ต่าง ๆ เช่น การเรียนภาษาต่างประเทศ แต่การถ่ายถอดอักษรวิธีการนี้ จะส่งผลให้การค้นคืนข้ามภาษาทำได้ยาก เนื่องจากกฎในการออกเสียงของแต่ละภาษามีความแตกต่างกัน การจะถ่ายถอดอักษรได้ถูกต้องนั้นต้องอาศัยกฎที่จำเป็นสำหรับการถ่ายถอดอักษรจากภาษาหนึ่งไปยังอีกภาษาหนึ่งที่เพียงพอทำให้กระบวนการในการทำงานเกิดความซับซ้อนมากยิ่งขึ้น

อย่างไรก็ตาม หากพิจารณาทิศทาง (direction) ในการถ่ายถอดอักษรระหว่างภาษาต้นทางและภาษาปลายทาง ก็สามารถแบ่งออกได้เป็น 2 ประเภท คือการถ่ายถอดอักษรแบบปกติและการถ่ายถอดอักษรแบบย้อนกลับ (Lee, C.-J. and Chang, J. S. 2003) การถ่ายถอดอักษรแบบปกติหรือแบบไปข้างหน้า (forward-direction transliteration) เป็นกระบวนการในการแปลงคำดั้งเดิม (original word) ในภาษาต้นทางให้เป็นคำใหม่ในภาษาปลายทางโดยให้มีเสียงใกล้เคียงกับภาษาต้นทาง เช่นคำว่า “ต้มยำกุ้ง” ซึ่งเป็นคำเดิมในภาษาไทยสามารถถอดอักษรได้เป็น “tom yum kung” ในภาษาอังกฤษ อีกประเภทหนึ่งคือการถ่ายถอดอักษรแบบย้อนกลับ (backward-direction transliteration) ซึ่งเป็นการถ่ายถอดอักษรจากคำยืมหรือคำทับศัพท์จากภาษาหนึ่ง กลับเป็นตัว อักษรดั้งเดิมในอีกภาษาหนึ่งเช่น คำว่า “tom yum kung” ซึ่งเป็นคำภาษาอังกฤษที่เกิดจากการถ่ายถอดมาจากภาษาไทยถูกแปลงกลับให้เป็น “ต้มยำกุ้ง” ซึ่งเป็นคำดั้งเดิมที่ปรากฏในภาษาไทยอยู่แล้ว

2.3.2 แนวทางการถ่ายถอดอักษรด้วยคอมพิวเตอร์

Oh, Choi และ Isahara (2006: 186) ได้แบ่งวิธีการในการถ่ายถอดอักษรด้วยคอมพิวเตอร์ออกเป็น 3 ประเภท ได้แก่ (1) การถ่ายถอดอักษรโดยการเทียบตัวอักษร (grapheme based approach หรือ direct method) (2) การถ่ายถอดอักษรโดยอาศัยตัวกลางทางเสียง (phoneme based approach หรือ pivot method) และ (3) การถ่ายถอดอักษรแบบผสมผสาน (hybrid approach) ซึ่งเป็นการผสมผสานระหว่างการอาศัยตัวกลางทางเสียงและการเทียบตัวอักษรเข้าด้วยกัน ซึ่งจะขอนำเสนอรายละเอียดดังนี้

2.3.2.1 การถ่ายถอดอักษรโดยการเทียบตัวอักษร

การถ่ายถอดอักษรโดยการเทียบตัวอักษร (Grapheme based approach) เป็นการจับคู่ระหว่างตัวเขียนภาษาต้นทางกับตัวเขียนของภาษาปลายทางโดยไม่ผ่านตัวกลางทางเสียง ซึ่งตัวอักษรที่จับคู่นี้อาจเป็นความสัมพันธ์แบบหนึ่งต่อหนึ่งหรือไม่ใช่หนึ่งต่อหนึ่งก็ได้ (Oh, J.-H. Choi, K.-S. and Isahara, H. 2006: 188)

ในการถ่ายถอดอักษรแบบเทียบตัวอักษรสามารถใช้วิธีการต่างๆ ได้แก่ การใช้กฎทางภาษา (Rule-based model) และ การใช้แบบจำลองทางสถิติ (Statistical machine transliteration model หรือ SMT) ซึ่งมีรายละเอียดดังนี้

2.3.2.1.1 การใช้กฎทางภาษา

การใช้กฎทางภาษา (Rule-based model) คือ การนำกฎทางภาษามาใช้ในการถ่ายถอดอักษรจากภาษาหนึ่งไปสู่อีกภาษาหนึ่ง ซึ่งกฎต่าง ๆ ในการถ่ายถอดอักษรอาจได้มาจาก การนำกฎที่มนุษย์สร้างขึ้นไปใช้กับเครื่องคอมพิวเตอร์โดยตรง (manual construction) หรือการหากฎด้วยคอมพิวเตอร์โดยอาศัยคลังข้อความ (corpus-based) (วิโรจน์ อรุณมานะกุล 2548:37)

สำหรับกฎที่กำหนดขึ้นโดยมนุษย์นั้น อาจจำแนกได้เป็น (1) กฎที่มีผู้กำหนดไว้แล้วเช่นกฎที่ปรากฏในตำราต่าง ๆ รวมทั้งกฎที่ตราขึ้นเป็นมาตรฐานโดยภาครัฐ เช่น เกณฑ์ศัพท์บัญญัติของราชบัณฑิตยสถาน หรือ (2) กฎที่เกิดจากการศึกษาเพิ่มเติมในภายหลัง ไม่ว่าจะเป็นกฎที่ได้จากผู้เชี่ยวชาญ หรือการทำงานวิจัยเพิ่มเติม ฯลฯ ส่วนการหากฎด้วยคอมพิวเตอร์โดยอาศัยคลังข้อความนั้นทำได้โดยเก็บรวบรวมคลังข้อความ (corpus) ที่ต้องการจะหากฎให้ได้จำนวนมากพอสมควร จากนั้นเขียนโปรแกรมให้คอมพิวเตอร์ได้เรียนรู้จากคลังข้อความนั้นๆภายใต้บริบทต่าง ๆ แล้วนำกฎที่เครื่องคอมพิวเตอร์หาได้มาใช้ในการประมวลผลต่อไป

ในการถ่ายถอดอักษรด้วยเครื่องคอมพิวเตอร์โดยใช้กฎทางภาษา อาจทำได้โดย การใช้ผังต้นไม้ช่วยตัดสินใจ (decision trees) การใช้ตารางกฎ (table lookup) และการจับคู่อักขระ (character mapping) เป็นต้น

ก. ผังต้นไม้ช่วยตัดสินใจ

โดยทั่วไป ในการใช้ผังต้นไม้ช่วยตัดสินใจ (decision trees) จะมีการกำหนดผังต้นไม้ขึ้นซึ่งมีลักษณะคล้ายต้นไม้จริงกลับหัวที่มีโหนดรากอยู่ด้านบนสุดและโหนดใบอยู่ล่างสุด โดยที่แต่ละกิ่งแทนแต่ละขั้นตอนของการตัดสินใจ ในการประมวลผลจะเริ่มจากด้านบนสุดของผังต้นไม้ แล้วไล่ไปตามกิ่งต่าง ๆ เป็นลำดับ ตามเงื่อนไขที่ระบุไว้ในแต่ละกิ่งจนสิ้นสุดผังต้นไม้แล้วใช้ผลลัพธ์สุดท้ายเป็นผลลัพธ์ในการตัดสินใจ ในการถ่ายถอดโดยวิธีนี้จึงเป็นการพิจารณาตัวอักษรพร้อมกับ

บริบทข้างเคียงเพื่อตัดสินใจทีละขั้นตามลำดับจนถึงขั้นสุดท้ายคือปลายกิ่งจนได้ตัวอักษรเป้าหมาย (สุขชาตรี ประสมสุข. 2555)

คังและชอย (Kang, I.-H. and Choi, K.-S. 2000) ได้เสนอแบบจำลองการถ่ายถอดจากภาษาอังกฤษเป็นภาษาเกาหลีโดยการใช้ผังต้นไม้ช่วยตัดสินใจ โดยที่ในขั้นแรก คังและชอยได้ทำการปรับแนว (align) ตัวอักษรของข้อความในภาษาต้นทางกับตัวอักษรในภาษาปลายทาง แล้วจึงใช้ผังต้นไม้มาช่วยตัดสินว่าสายอักขระใดเหมาะสมที่สุดที่จะนำมาใช้เป็นผลลัพธ์ของการถ่ายถอดอักษรในภาษาปลายทางได้

ข. การใช้ตารางกฎ

การใช้ตารางกฎ (table lookup) เป็นวิธีการที่ต้องอาศัยตารางที่ประกอบด้วยกฎต่าง ๆ ที่สร้างขึ้น โดยที่ตารางกฎนี้จะประกอบด้วยตัวอักษรแต่ละชุด (โดยที่แต่ละชุดจะประกอบด้วยอักษร 1 ตัว 2 ตัวหรือ 3 ตัวก็ได้ แล้วแต่กรณี) โดยแต่ละชุดอักษรนี้ จะมีการระบุถึงตัวอักษรทางขวาและตัวอักษรทางซ้ายที่ปรากฏร่วมด้วย และระบุว่าหากพบบริบทตัวอักษรทางซ้ายและทางขวานี้แล้วจะถ่ายถอดอักษรเป็นตัวอักษรใดในภาษาปลายทาง

ในการถ่ายถอดอักษร คอมพิวเตอร์จะนำตัวอักษรแต่ละชุดพร้อมกับบริบทซ้าย-ขวาของคำที่ต้องการถ่ายถอดมาเปรียบเทียบกับชุดอักษรในตารางทีละชุด เพื่อตรวจสอบว่าชุดอักษรและบริบทซ้ายและขวาของคำที่ต้องการถ่ายถอดนั้นตรงกับชุดอักษรและบริบทใดในตาราง แล้วนำผลลัพธ์ของกรณีนั้นมาใช้เป็นผลลัพธ์ของการถอดชุดอักษรนั้น แต่หากพบว่าชุดอักษรนั้นไม่ตรงกับกรณีใด หรือพบว่าชุดอักษรนั้นตรงกับกรณีต่าง ๆ มากกว่า 1 กรณีในตาราง คอมพิวเตอร์ก็จะทำการตรวจสอบกับคลังข้อมูลฝึกระบบที่ได้เตรียมไว้แล้ว เพื่อหาว่าชุดอักษรที่มีปัญหานี้มีโอกาสที่จะถอดเป็นอักษรใดได้มากที่สุด แล้วจึงสร้างตารางคาดเดาขึ้นเพื่อนำมาใช้ในการประมวลผลชุดอักษรที่มีปัญหานั้น (วิโรจน์ อรุณมานะกุล 2548)

วัลย์วรา ไชยฤกษ์ (2547) ได้ศึกษาการถ่ายถอดคำทับศัพท์ภาษาอังกฤษเป็นภาษาไทย โดยเริ่มจากสร้างตารางกฎซึ่งได้มาจากข้อกำหนดของราชบัณฑิตยสถาน โดยตารางกฎที่ใช้มี 2 ประเภท ได้แก่ ตารางกฎที่มีบริบทซ้าย-ขวาอย่างละ 1 ตำแหน่ง 2 ตำแหน่งและ 3 ตำแหน่ง และตารางคาดเดา ซึ่งประกอบด้วยตารางคาดเดาที่มีบริบทซ้าย-ขวา 2 ตำแหน่ง 1 ตำแหน่ง และตารางคาดเดาที่ไม่มีในบริบท

ค. Character Mapping

Character Mapping เป็นการจับคู่ตัวอักษรของภาษาหนึ่งกับอักษรของอีกภาษาหนึ่งโดยตรง (Min, Z. Haizhou, L. and Jian, S. 2004)

Deep, Kamal and Goyal, Vishal (2011) ได้ถ่ายถอดอักษรชื่อบุคคลและชื่อสถานที่ภูมิศาสตร์จากภาษาปัญจาบ (Punjabi) เป็นภาษาอังกฤษ โดยใช้ Character Mapping ในการจับคู่ตัวอักษรระหว่าง Gurumakhi script กับ English letter โดยในการถ่ายถอดอักษรได้จำแนกกฎออกเป็น 2 ประเภท คือ กฎที่ใช้ในการจับคู่ตัวอักษรระหว่างภาษาทั้งสอง และ กฎเกี่ยวกับตำแหน่งของการปรากฏของตัวอักษรของคำในภาษาปัญจาบ เช่นปรากฏที่ตำแหน่งต้นคำ ท้ายคำ หลังสระ หรือหลังพยัญชนะ เป็นต้น

2.3.2.1.2 การใช้แบบจำลองทางสถิติ

การถ่ายถอดอักษรโดยใช้แบบจำลองทางสถิติ (Statistical machine transliteration model หรือ SMT) เป็น แบบจำลองที่ใช้ข้อมูลสถิติที่พบจากการถ่ายถอดอักษรจริงมาช่วยในการพัฒนาระบบ แบบจำลองนี้มองว่า การถ่ายถอดอักษรมีความเป็นไปได้ที่จะให้ผลลัพธ์ในหลาย ๆ รูปแบบ โดยที่แบบจำลองทางสถิติจะเป็นตัวเลือกรูปแบบที่เหมาะสมมากที่สุดมาใช้เป็นผลลัพธ์

การพัฒนาระบบที่ใช้แบบจำลองทางสถิติ กระทำได้ในหลายลักษณะ ดังนี้

ก. Noisy Channel Model

แบบจำลองนี้ใช้ในการปรับข้อมูลนำเข้าให้เป็นผลลัพธ์ที่ถูกต้อง เนื่องจากอิงอยู่กับสมมติฐานว่าข้อมูลที่นำเข้ามานี้มีข้อผิด จึงต้องให้คอมพิวเตอร์ทำกระบวนการต่าง ๆ เพื่อให้ได้ผลลัพธ์เป็นข้อมูลที่ถูกต้อง เช่น หากใช้แบบจำลองนี้มาแก้คำผิดก็จะเชื่อว่าคำที่นำเข้าสะกดไม่ถูกต้อง เนื่องจากอาจมีอักขระบางตัวหายไป (เช่น leter แทนที่จะเป็น letter) มีอักขระเกินมา (เช่น misstake แทนที่จะเป็น mistake) มีการสลับที่อักขระ (เช่น receiving แทนที่จะเป็น receiving) หรือมีการแทนที่อักขระตัวหนึ่งด้วยอักขระอีกตัวหนึ่ง (เช่น fimite แทนที่จะเป็น finite) เป็นต้น

ในแบบจำลองนี้ประกอบด้วยแบบจำลองย่อยอีก 2 แบบจำลองคือ (1) Noisy Channel model เป็นแบบจำลองที่จะสร้างเซ็ตของคำเป้าหมายที่เป็นไปได้ทั้งหมดจากคำนำเข้า และ(2) Language model เป็นแบบจำลองที่หาความเป็นไปได้สูงสุดจากเซ็ตของคำที่เครื่องคอมพิวเตอร์สร้างขึ้นจากแบบจำลองที่ผ่านมา โดยที่รูปแบบที่ใช้อาจเป็น maximum likelihood, maximum a posteriori หรือ minimum distance เป็นต้น

Jia, Zhu and Yu (2009) ได้ใช้ noisy channel model ในการถ่ายถอดภาษาอังกฤษเป็นภาษาจีนแบบไปข้างหน้าและย้อนกลับเพื่อใช้ในงานค้นคืนข้ามภาษาแนวคิดคือการมองว่าจากคำอังกฤษที่กำหนดให้ (E) ซึ่งอาจเป็นคำจีน (C) ได้หลายแบบ โดยขั้นแรกเมื่อนำเข้าข้อมูลที่เป็นคำอังกฤษแบบจำลองนี้จะสร้างคำจีน (C) ที่เป็นไปได้ทั้งหมดขึ้นมา จากนั้นใช้กฎของ Bayes Rule เพื่อจะหาคำจีน ($\hat{C}$) ซึ่งเป็นคำที่สอดคล้องกับคำอังกฤษมากที่สุด

ข. ระบบที่ใช้แบบจำลองจากคู่ตัวอักษรสองภาษา (Joint source-channel model)

ระบบดังกล่าวจะใช้ข้อมูลจากทั้งภาษาต้นทางและภาษาปลายทางไปพร้อมๆกัน ระบบนี้มีแนวคิดหลักที่ว่า ในการตัดสินใจเลือกตัวอักษรตัวต่อไปในภาษาปลายทางนั้น ขึ้นอยู่กับข้อมูลของทั้งในภาษาต้นทางและในภาษาปลายทางที่มาก่อนหน้า วิธีการนี้มีข้อดีคือ ไม่จำเป็นต้องใช้ข้อมูลคำอ่านของทั้งสองภาษา เนื่องจากบริบทของทั้งสองภาษาจะช่วยให้เลือกตัวอักษรในภาษาปลายทางที่เหมาะสมได้อยู่แล้ว

วิโรจน์ อรุณมานะกุล (2548: 37-38) ได้นำวิธีการนี้มาใช้เป็นส่วนหนึ่งในการถ่ายถอดคำภาษาอังกฤษเป็นภาษาไทย โดยใช้ไตรแกรมในการดูค่าความน่าจะเป็นสูงสุดของการปรากฏของตัวอักษรไทยและอังกฤษไปพร้อม ๆ กัน ในการถ่ายถอดอักษรนั้นจะยึดหลักการที่ว่า คำอังกฤษหนึ่ง คำ สามารถถ่ายถอดให้เป็นคำไทยได้หลายรูปแบบ ส่วนรูปแบบของคำไทยที่เหมาะสมที่สุดนั้นขึ้นอยู่กับการปรากฏของตัวอักษร 3 ตัวในภาษาอังกฤษที่ปรากฏร่วมกันตั้งแต่ต้นคำถึงท้ายคำ

ค. ระบบที่ใช้แบบจำลองจากคู่ตัวอักษรสองภาษาและมองเป็นกลุ่มก้อนได้(Chunk-based)

วิธีการนี้มีแนวคิดว่า ในการถ่ายถอดอักษรไม่จำเป็นต้องเทียบคู่ตัวอักษรระหว่างภาษาต้นทางและภาษาปลายทาง แต่มองความสัมพันธ์ระหว่างตัวอักษรของภาษาต้นทางและภาษาปลายทางนั้นเป็นแบบกลุ่มตัวอักษรต่อกลุ่มตัวอักษร โดยที่กลุ่มของตัวอักษรสามารถมีความยาวตั้งแต่หนึ่งตัวอักษรขึ้นไปได้

ตัวอย่างเช่น locarno = ล/โ/ค/า/ร์/โ/น นอกจากจะมองว่า ‘l’ ถ่ายถอดเป็น ‘ล’, ‘o’ ถ่ายถอดเป็น ‘โ’, ‘ c’ ถ่ายถอดเป็น ‘ค’ แล้ว ยังสามารถมองการถ่ายถอดอักษรเป็นกลุ่มของตัวอักษรที่มีขนาดความยาวต่างกันไม่ว่าจะเป็น 2, 3, 4 จนถึง n ตัวอักษร เช่น ‘lo’ เป็น ‘ลโ’, ‘oc’ เป็น ‘โค’, ‘ca’ เป็น ‘คา’, ‘ar’ เป็น ‘าร์’ , no เป็น ‘นโ’, ‘loc’ เป็น ‘ลโค’ ‘locarno’ เป็น ‘ลโคาร์นโ’ เป็นต้น (วิโรจน์ อรุณมานะกุล 2548: 39) ดังนั้น จากข้อมูลฝึกสอนระบบซึ่งได้จับคู่ตัวอักษรระหว่างภาษาต้นทางและภาษาปลายทางแล้ว โปรแกรมจะคำนวณค่าสถิติของการถ่ายถอดกลุ่มอักษรต่างๆ หลังจากนั้นจะนำค่าสถิติที่ได้ไปใช้ในการตัดสินใจถ่ายถอดอักษรอื่นๆ ต่อไป

เท่าที่ได้อภิปรายมาพอจะกล่าวได้ว่า การถ่ายถอดอักษรโดยการเทียบตัวอักษร เป็นระบบที่ง่ายต่อการพัฒนา ระบบนี้มีข้อดีคือใช้เพียงข้อมูลของตัวเขียนทั้งสองภาษาซึ่งเป็นข้อมูลพื้นฐานที่จำเป็นต้องจัดเตรียมไว้แล้ว แต่การถ่ายถอดระหว่างตัวเขียนของภาษาหนึ่งไปยังอีกภาษาหนึ่งโดยตรงมีข้อจำกัดคืออาจได้ผลไม่ถูกต้องในกรณีที่ตัวเขียนและเสียงในภาษานั้นมีความสัมพันธ์ที่ซับซ้อน ซึ่งในกรณีนี้ควรเลือกใช้การถ่ายถอดโดยอาศัยตัวกลางทางเสียงซึ่งจะได้อภิปรายในหัวข้อต่อไป

2.3.2.2 การถ่ายถอดอักษรโดยอาศัยตัวกลางทางเสียง

การถ่ายถอดโดยอาศัยตัวกลางทางเสียง (phoneme based approach) เป็นการสะกดคำในภาษาปลายทางที่ทำให้ออกเสียงใกล้เคียงกับเสียงพูดของคำดั้งเดิมในภาษาต้นทางมากสุด

วิโรจน์ อรุณมานะกุล (2548: 30) กล่าวว่าการถ่ายถอดอักษรโดยใช้เสียงเป็นตัวกลางสามารถทำได้หลายลักษณะคือ

ก. รูปแบบที่ถ่ายอักษรภาษาต้นทางเป็นเสียงภาษาต้นทางแล้วเทียบเป็นเสียงในภาษาปลายทาง ก่อนจะถอดออกมาเป็นตัวอักษรในภาษาปลายทาง (Source Language Texts->Source Language Sound ->Target Language Sound-> Target Language Texts) การถ่ายถอดอักษรรูปแบบนี้ ทำโดยผ่านตัวกลางคือเสียงพูดของทั้งภาษาต้นทางและปลายทาง ดังที่ปรากฏในงานวิจัยของ ลี (Lee, J. S. 1999) ไนท์และเกรล (Knight, K. and Graehl, J. 1997)

ข. รูปแบบที่ถ่ายเสียงตัวอักษรภาษาต้นทางเป็นเสียงในภาษาต้นทางแล้วแปลงเป็นตัวอักษรในภาษาปลายทางโดยลดขั้นตอนการเทียบหน่วยเสียงจากภาษาต้นทางเป็นหน่วยเสียงภาษาปลายทาง (Source Language Texts -> Source Language Sounds -> Target Language Texts) ดังปรากฏในงานของ Glover and Knight (1998) ที่ถ่ายถอดจากภาษาอังกฤษเป็นอารบิค

ค. รูปแบบที่ถ่ายเสียงตัวอักษรภาษาต้นทางไปเป็นเสียงในภาษาปลายทางโดยตรงก่อนจะถอดเป็นตัวอักษรในภาษาปลายทาง (Source Language Texts -> Target Language Sounds -> Target Language Texts) ดังปรากฏในงานของ Kuo and Yang (2004) ที่ถ่ายถอดจากภาษาอังกฤษเป็นภาษาจีน

ในการถ่ายถอดอักษรโดยอาศัยตัวกลางทางเสียง สามารถทำได้โดย การใช้กฎทางภาษา การใช้แบบจำลองทางสถิติ การใช้กฎร่วมกับแบบจำลองทางสถิติ และการใช้เครือข่ายประสาทเทียม ซึ่งมีรายละเอียดดังนี้

2.3.2.2.1 การใช้กฎทางภาษา

วิธีการใช้กฎทางภาษาทำได้หลายวิธี วิธีการหนึ่งคือการใช้รหัสตัวเลขหรือที่เรียกว่า Soundex Technique ซึ่งริเริ่มโดย M. K. Odell และ R. C. Russell ในปี พ.ศ.2461

ซาวด์เด็กซ์คือรหัสทางเสียง (phonetic code) ที่ได้มาจากการเปลี่ยนตัวอักษรต่าง ๆให้เป็นตัวเลข (numeric code) โดยมีหลักการคือ ต้องมีการเก็บรวบรวมข้อมูลคำในภาษาต้นทางและภาษาปลายทาง แล้วนำคำในภาษาทั้งสองที่อยู่ในคลังข้อมูลมาเปลี่ยนเป็นรหัสซาวด์เด็กซ์ของแต่ละภาษา หลังจากนั้นเมื่อต้องการถอดคำในภาษาใดภาษาหนึ่งก็จะนำคำนั้นมาเปลี่ยนเป็นรหัสซาวด์เด็กซ์

แล้วนำรหัสนี้ไปใช้เป็น key ในการค้นว่าตรงกับรหัสซาวด์เด็กซ์ของอีกภาษาหนึ่งหรือไม่ ถ้าตรงกันก็ให้นำคำที่ตรงกับรหัสซาวด์เด็กซ์ในภาษาปลายทางมาเป็นผลลัพธ์

Suwanvisat and Prasitjutrakul (1998) ได้นำวิธีการนี้มาใช้ในการถ่ายถอดอักษรระหว่างภาษาไทยและอังกฤษ เพื่อใช้ในการสืบค้นเอกสารอิเล็กทรอนิกส์ (retrieval of document) โดยได้ปรับตารางรหัสและอัลกอริทึมของโอเด็ลและรัสเซลเพื่อให้คอมพิวเตอร์ประมวลผลได้ดีขึ้น

2.3.2.2.2 การใช้แบบจำลองทางสถิติ

การถ่ายถอดอักษรที่อาศัยตัวกลางทางเสียงโดยใช้แบบจำลองทางสถิติ สามารถกระทำได้ในหลายลักษณะ ดังนี้

ก. การใช้แบบจำลองเอ็นแกรม

แบบจำลองเอ็นแกรม (N-gram model) คือแบบจำลองที่ใช้คำนวณค่าความน่าจะเป็นของสายอักขระที่เรียงต่อกันจำนวน n ตัว แบบจำลองเอ็นแกรมจะตั้งอยู่บนสมมติฐานมาร์คอฟ (Markov assumption) ที่ว่า “การปรากฏของตัวอักษรตัวหนึ่งขึ้นกับตัวอักษรก่อนหน้าเพียง n-1 ตัว” นอกจากนี้เราสามารถนำสมมติฐานมาร์คอฟมาใช้ประมาณค่าความเป็นไปได้ที่ 2-gram 3-gram 4-gram ฯลฯ ได้

ลี (Lee, J. S. 1999) ได้ศึกษาการถ่ายถอดภาษาอังกฤษเป็นภาษาเกาหลีโดยใช้แบบจำลองเอ็นแกรมเพื่อหาความน่าจะเป็นสูงสุดของชุดตัวอักษรต่าง ๆ ในภาษาเกาหลีที่ได้มาจากคำภาษาอังกฤษ ในการถ่ายถอด ลีจะใช้การจับคู่ระหว่างหน่วยเสียงพูดของภาษาอังกฤษและภาษาเกาหลี วิธีการเริ่มจากนำคำอังกฤษมาแยกออกเป็นหน่วยเสียงพูด (PU’s) เช่น ‘board’ อาจแบ่งได้เป็น b:o:a:r:d, b:oa:r:d, b:o:ar:d b:oar:d ฯลฯ (ซึ่ง ลี ใช้สัญลักษณ์แทนแต่ละรูปแบบด้วย E_i) หลังจากนั้นนำแต่ละ E_i ไปสร้างรูปแบบหน่วยเสียงพูดในภาษาเกาหลีให้ได้จำนวนมากที่สุด แล้วคำนวณหารูปแบบหน่วยเสียงพูดในภาษาเกาหลีที่มีความเป็นไปได้สูงสุด แล้วนำรูปแบบที่ได้นี้ไปจับคู่กับตัวเขียนในภาษาเกาหลีเพื่อใช้เป็นผลลัพธ์

ค่าความน่าจะเป็น P(K) ถูกกำหนดโดยโมเดลของภาษาและสามารถประมาณได้โดยการใช้แบบจำลองไบแกรม ในการคำนวณหาหน่วยเสียงพูดที่เป็นไปได้สูงสุดในภาษาเกาหลีทำโดยสมการดังนี้

$$\arg\max_k P(K \mid E) = \arg\max_k P(K)P(E \mid K)$$

$$P(K) \cong p(kpu_1)\prod_{i=2}^{n} p(kpu_i \mid kpu_{i-1})$$

$$P(K \mid E) \cong \prod_{i=1}^{n} p(epu_i \mid kpu_i)$$

ในงานวิจัย ลี ได้พบปัญหาจากการถ่ายถอดอักษรด้วยวิธีดังกล่าว 2 ประการ ประการแรก คือ ในการแบ่งตัวอักษรออกเป็นหน่วยเสียงพูดในภาษาต้นทางและภาษาปลายทาง จึงมีโอกาสสูงที่จะเกิดความคลาดเคลื่อนทำให้การถอดตัวอักษรเกิดความผิดพลาด ประการต่อมาคือในการสร้างรูปแบบหน่วยเสียงพูดของทั้งสองภาษาต้องใช้เวลามากเนื่องจากวิธีการนี้จะทำให้เกิดสายอักขระที่คอมพิวเตอร์สร้างขึ้นในปริมาณมากเป็นทวีคูณ ทำให้ต้องใช้เวลามากในการประมวลผล จึงไม่เหมาะที่จะนำมาถอดตัวอักษรกับคำหลายๆคำในข้อความเดียวกัน

ข. การใช้แบบจำลอง Weighted finite-state transducer (WFST)

Weighted finite-state transducer หรือเรียกว่าเครือข่ายการเปลี่ยนตำแหน่ง (transition network) มีแนวคิดที่ว่า เมื่อมีข้อมูลเข้ามาจะทำให้เกิดการเปลี่ยนสถานะ โดยในการทำงานจะเริ่มจากการรับข้อมูลเข้าเป็นตัวอักษร เข้าสู่สถานะเริ่มต้น (start state) หลังจากนั้นจะรับข้อมูลเข้าและพิจารณาสภาพไปพร้อม ๆ กัน โดยที่สถานะต่าง ๆ จะขึ้นกับข้อมูลที่นำเข้า คอมพิวเตอร์จะทำเช่นนี้ไปเรื่อย ๆ จนถึงเงื่อนไขที่จะทำให้สู่สถานะสิ้นสุดสภาพซึ่งอาจจะไปจบที่สถานะสุดท้าย (final state) หรืออาจจะไปไม่ถึงสถานะสุดท้ายก็ได้ (โดยที่การเปลี่ยนสถานะจะขึ้นกับสถานะปัจจุบันและค่าของ input) โดยวิธีการนี้จะมีการกำหนดน้ำหนักของคำเพื่อช่วยให้การถ่ายถอดอักษรมีความถูกต้องและแม่นยำมากขึ้น

วิธีการถ่ายถอดอักษรดังกล่าว จะมีลักษณะดังนี้

ภาพที่ 2.2 วิธีการ Weighted finite-state transducer (ที่มา Knight and Graehl,1997)

การถ่ายถอดอักษรด้วยวิธีนี้จะได้แบบจำลองทางสถิติดังนี้

$$P_s(w \mid a) \cong \sum_{\forall e} P(w)P(e \mid w)P(a \mid e)$$

โดยที่ P_p (w|a) คือความน่าจะเป็นของการถ่ายถอดอักษรคำในภาษา a

ไนท์และเกรล (Knight and Graehl ,1997) ได้ทำการถ่ายถอดคำทับศัพท์ในตัวเขียนคาตาคานะในภาษาญี่ปุ่นกลับเป็นภาษาอังกฤษโดยใช้แบบจำลองแบบ weighted finite-state transducer (WFSTs) ในการถ่ายถอดอักษรเริ่มจากการนำเข้าคำในรูปตัวเขียนคาตาคานะ แล้วระบบจะทำการถ่ายถอดอักษรคำนั้น โดยการเลือกคำในภาษาอังกฤษที่มีความเป็นไปได้สูงสุดจากรายการคำที่

เป็นไปได้ทั้งหมดที่คอมพิวเตอร์สร้างขึ้น โดยใช้สมการ P(w). P(e|w) . P(j|e) . P(k|j) . P(o|k) โดยที่ประกอบด้วย 5 พารามิเตอร์คือ (1) P(w) แสดงความเป็นไปได้ของลำดับคำอังกฤษจากคลังข้อมูลฝึกระบบ (2) P(e|w) ซึ่งใช้ระบุคำอังกฤษที่ได้ถอดเสียงให้เป็นเสียงพูด (3) P(j|e) ใช้แสดงเสียงในภาษาอังกฤษที่ทำการปรับให้เข้ากับระบบเสียงในภาษาญี่ปุ่น (4) P(k|j) ใช้บอกเสียงของภาษาญี่ปุ่นที่แปลงให้อยู่ในรูปอักษรคาตาคานะ และ (5) P(o|k) ใช้แทนตัวเขียนคาตาคานะที่ได้จากการแปลงอักษรภาษาญี่ปุ่นจากโปรแกรม OCR

2.3.2.2.3 การใช้กฎร่วมกับแบบจำลองทางสถิติ

การถ่ายถอดอักษรในลักษณะนี้จำเป็นต้องอาศัยกฎต่าง ๆ ร่วมกับแบบจำลองทางสถิติ เพื่อช่วยให้การทำงานมีประสิทธิภาพยิ่งขึ้น ในการใช้แบบจำลองนั้นจำเป็นจะต้องอาศัยคลังข้อความของทั้งสองภาษา เพื่อหาความน่าจะเป็นทางสถิติในการถ่ายถอดอักษรเป็นอีกภาษาหนึ่ง และในการสร้างกฎสามารถทำได้ 2 รูปแบบคือ การกำหนดกฎเองโดยนักภาษาศาสตร์ และการสร้างโปรแกรมให้เครื่องคอมพิวเตอร์เรียนรู้และสร้างกฎเอง

สุรนา และสิงห์ (Surana, H. and Singh, A.K. 2008) ศึกษาการถ่ายถอดอักษรแบบไป-กลับระหว่างภาษาอังกฤษและภาษาในอินเดีย ได้แก่ Hindi, Telugu โดยใช้วิธีการทางสถิติและกฎร่วมกัน ขั้นแรกแบบจำลองที่สร้างขึ้นจะทำการจำแนกภาษาต้นทางว่าเป็นภาษาใด โดยวิธีการ symmetric cross entropy หลังจากจำแนกภาษาแล้ว หากพบว่าเป็นภาษาอังกฤษจะนำคำอังกฤษมาถ่ายเป็นเสียงอ่าน จากนั้นนำเสียงอ่านแต่ละเสียงในคำนั้นไปจับคู่กับตัวอักษรในภาษาอินเดีย ในทำนองเดียวกันหากพบว่าคำนั้นเป็นภาษาอินเดียก็จะถอดคำนั้นเป็นตัวเขียนอังกฤษ โดยเริ่มจากนำคำอินเดียมาสร้างอักขระคำในรูปแบบต่าง ๆ ในภาษาอังกฤษให้มากเท่าที่จะเป็นไปได้ แล้วนำตัวอักษรเหล่านี้มาผ่านการกรองและจัดลำดับในขั้นสุดท้ายจะได้ผลลัพธ์ที่ถ่ายถอดเป็นคำอังกฤษ

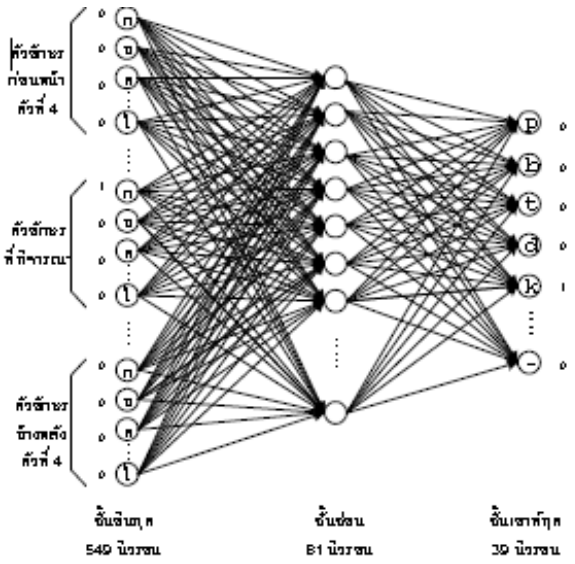
2.3.2.2.4 การใช้เครือข่ายประสาทเทียม

เครือข่ายประสาทเทียม (Neural network) คือการคำนวณที่เลียนแบบการทำงานของระบบสมองมนุษย์เพื่อใช้ประโยชน์ในการคาดคะเนเหตุการณ์จากข้อมูลที่มีอยู่

เครือข่ายประสาทเทียมจะจำลองการทำงานของระบบสมองโดยใช้ฟังก์ชั่นถ่ายโอน (transfer function) ค่าถ่วงน้ำหนัก (weight) และค่าไบแอส (bias) เป็นเครื่องมือจำลองคุณสมบัติของเซลล์ประสาท โดยในการนำมาใช้ถ่ายถอดอักษรนั้นมีการสร้างเครือข่ายต่าง ๆ ให้โปรแกรมเรียนรู้การถ่ายถอดอักษรจากข้อมูลที่เตรียมไว้ให้ จากนั้นนำค่าน้ำหนักที่ได้จากการฝึกและเครือข่ายที่ให้ผลดีที่สุดไปใช้เป็นผลลัพธ์ของรูปแทนเสียงหรือรูปอักษรที่รับเข้ามา (Kaur, K. and Singh, P.

2014) เครือข่ายประสาทเทียมจึงไม่จำเป็นต้องอาศัยกฎหรือความรู้ทางภาษา

ทัศนวรรณ ศูนย์กลาง (2543) ศึกษาการเข้ารหัสคำทับศัพท์ภาษาไทยและภาษาอังกฤษเพื่อการค้นคืนข้ามภาษา โดยใช้ Backpropagation neural networks ในการทดลองทัศนวรรณได้แบ่งข้อมูลเป็นข้อมูลฝึก (train set) และข้อมูลทดสอบ (test set) นำข้อมูลทดสอบไปรับการเข้ารหัสนิวรอลเน็ตเวิร์กที่ได้จากข้อมูลฝึก ซึ่งแสดงในภาพต่อไปนี้



ภาพที่ 2.3 โครงสร้างแบ็กพรอพาเกชันนิวรอลเน็ตเวิร์กที่ใช้ในการเรียนรู้คำไทย

ผลที่ได้จากขั้นตอนนี้ คือนิวรอลเน็ตเวิร์กที่ใช้เข้ารหัสที่ได้กับคำอังกฤษทับศัพท์ และคำไทยทับศัพท์ โดยคำอังกฤษทับศัพท์จะมี 35 นิวรอล และคำไทยทับศัพท์จะมี 39 นิวรอล สำหรับขั้นตอนการค้นคืนคำข้ามภาษาจะอาศัยการเปรียบเทียบรหัสของคำแบบประมาณ โดยอนุญาตให้มีความแตกต่างของรหัสที่ได้ไม่เกิน 1 ซึ่งทำให้ได้คำที่ค้นคืนข้ามภาษาที่มีความถูกต้องมากยิ่งขึ้น

เป็นที่น่าสังเกตว่า วิธีการถ่ายถอดอักษรโดยการใช้หน่วยเสียงมีข้อจำกัดคือในแต่ละภาษามีหน่วยเสียงที่แตกต่างกัน จึงยากที่จะถ่ายถอดอักษรจากภาษาหนึ่งไปเป็นอีกภาษาหนึ่งได้ เนื่องจากในบางกรณีหน่วยเสียงในภาษาต้นทางอาจไม่มีหน่วยเสียงของภาษาปลายทางมารองรับ เช่น ภาษาจีนมีหลายหน่วยเสียงที่ไม่ปรากฏในภาษาอังกฤษ นอกจากนี้ยังเกิดข้อจำกัดในกรณีที่คู่ภาษาเป็นภาษาที่ใช้ตัวอักษรแทนความหมาย (ideographic language) เช่น ภาษาจีน เพราะตัวอักษรแต่ละตัวใช้แทน แต่ละความคิด(concept)ไม่ใช่แทนเสียง ข้อจำกัดอีกประการคือ วิธีการนี้มีขั้นตอนการทำงานที่ซับซ้อนต้องสร้างรายการหน่วยเสียงของแต่ละภาษาซึ่งจำเป็นต้องอาศัยผู้เชี่ยวชาญจึงทำให้ต้องใช้เวลามากและมีค่าใช้จ่ายสูง ด้วยเหตุนี้จึงทำให้เกิดการถ่ายถอดอักษรโดยใช้ตัวอักษรของภาษาต้นทางไปสู่ภาษาปลายทางโดยตรงโดยไม่ผ่านตัวกลางเสียง ซึ่งจะขอนำเสนอในหัวข้อต่อไป

2.3.2.3 การถ่ายถอดอักษรแบบผสมผสาน (Hybrid approach)

การถ่ายถอดอักษรแบบผสมผสาน หมายถึงการผสมผสานวิธีการเทียบหน่วยเสียงและวิธีการเทียบตัวอักษรเข้าด้วยกัน กล่าวคือ กระทำทั้งสองวิธีแล้วนำทั้งสองวิธีนี้มาผ่านกระบวนการเชื่อมเข้าด้วยกัน เนื่องจาก ผู้วิจัยที่ใช้วิธีการนี้เชื่อว่าจะให้ผลดีกว่าการใช้การถ่ายถอดอักษรแบบใดแบบหนึ่ง

นักวิจัยหลายคนได้เสนอแบบจำลองการถ่ายถอดอักษรแบบผสมผสาน โดยได้สร้างแบบจำลองสำหรับ grapheme และแบบจำลองสำหรับ phoneme โดยวิธีการต่าง ๆ เช่น Weighted finite-state transducer (WFST) (Bilac and Tanake. 2004) และ Source- channel model (LeeJ.-S. 1999, Al-Onaizan and Knight. 1998)

ในการประมวลผลโดยวิธีการผสมผสานนั้น จะประกอบด้วยแบบจำลอง 2 ประเภทคือ (1) แบบจำลองสำหรับ phoneme ซึ่งใช้ พารามิเตอร์ต่าง ๆ เช่น ความเป็นไปได้ที่ source grapheme จะกลายเป็น source phoneme ความเป็นไปได้ที่ source phoneme จะกลายเป็น target grapheme และความเป็นไปได้ที่จะเกิดคำที่ถูกต้องในภาษาเป้าหมาย (2) ส่วนแบบจำลองสำหรับ grapheme นั้นจะพิจารณาความเป็นไปได้ที่ source grapheme จะกลายเป็น target grapheme เป็นหลัก หลังจากที่ได้แบบจำลองสำหรับ grapheme และแบบจำลองสำหรับ phoneme แล้ว จะเชื่อมแบบจำลองทั้งสองประเภทนี้ เข้าด้วยกันโดยใช้ linear interpolation (Oh, J.-H. and Choi, K.-S. 2005: 450-461)

ข้อดีของแบบจำลองการถ่ายถอดอักษรแบบผสมผสาน คือ ในกระบวนการเชื่อมส่วนที่เป็นแบบจำลองของ grapheme และ phoneme เข้าด้วยกัน โดยที่ไม่คำนึงว่า source grapheme และ source phoneme มีการพึ่งพาอาศัยกันหรือไม่

อย่างไรก็ตามแบบจำลองดังกล่าวมีข้อจำกัดในด้านความถูกต้องของการถ่ายถอดอักษร เนื่องจากรูปแบบนี้เพียงแค่นำแบบจำลอง grapheme และแบบจำลอง phoneme มาเชื่อมต่อกันโดยใช้ linear interpolation โดยที่ไม่สนใจความสอดคล้องระหว่าง source grapheme และphoneme ความสอดคล้องของ source grapheme และphoneme ก็มีความสำคัญต่อการถ่ายถอดอักษรให้ถูกต้อง เช่น เสียง /AH/ ในภาษาอังกฤษสามารถถ่ายถอดอักษรเป็นสระได้เกือบทุกตัวในภาษาเกาหลี การที่จะรู้ว่าต้องถ่ายถอดเป็นสระใดในภาษาเกาหลีนั้น ต้องอาศัย source phoneme จึงจะให้ผลลัพธที่ถูกต้องได้

จากการทบทวนวรรณกรรมที่ได้อภิปรายมาข้างต้น ผู้วิจัยพบว่าการถ่ายถอดอักษรโดยอาศัยตัวกลางทางเสียงเป็นกระบวนการที่มีความซับซ้อน และการถ่ายถอดอักษรเบรลล์นั้นไม่มี

ความเกี่ยวข้องกับเสียงพูดโดยตรง แต่เป็นการจับคู่ตัวอักษรในระบบการเขียนหนึ่งกับตัวอักษรอีกระบบการเขียนหนึ่งภายใต้ภาษาเดียวกัน การถ่ายถอดโดยนำอักษรเบรลล์ไปผ่านตัวกลางทางเสียงแล้วนำเสียงที่ได้มาแปลงให้เป็นอักษรปกติ จะทำให้กระบวนการถ่ายถอดเบรลล์มีความซับซ้อนมากขึ้นเป็นทวีคูณ การใช้วิธีการนี้จึงไม่เหมาะสมเป็นอย่างยิ่ง

เนื่องจากอักษรเบรลล์และอักษรปกติต่างก็เป็นระบบตัวเขียน 2 ระบบ ที่อยู่ในภาษาเดียวกัน จึงมีแนวโน้มว่าอักษรทั้ง 2 ระบบนี้จะมีอักขรวิธีที่ใกล้เคียงกันอยู่จำนวนมากพอสมควร การถ่ายถอดจึงทำได้อย่างสะดวกโดยการจับคู่อักขระเบรลล์และอักษรปกติโดยตรง ไม่ว่าจะเป็นการจับคู่แบบหนึ่งต่อหนึ่ง (‘⠅’ กับ “ก”) หรือแบบสองต่อหนึ่ง (‘⠩⠱’ กับ “ศ”) อย่างไรก็ตาม แม้ว่าเบรลล์และอักษรปกติจะมีรูปแบบการประสมอักษรที่แตกต่างกันบ้าง แต่ในการถ่ายถอดเบรลล์สามารถใช้กฎจำนวนหนึ่งมาช่วยในการจัดเรียงข้อความอักษรปกติให้เป็นไปตามอักขรวิธีของภาษา ทำให้ได้ข้อความปกติที่สอดคล้องกับข้อความเบรลล์ในแต่ละภาษาได้

สำหรับวิธีการที่จะนำมาใช้ในการถ่ายถอดอักษรเบรลล์เป็นอักษรปกตินั้น มีอยู่หลายวิธีการแต่บางวิธีการก็ไม่เหมาะสมที่จะนำมาใช้ โดยเฉพาะอย่างยิ่งวิธีการที่ต้องอาศัยคลังข้อความที่ใช้คู่เทียบคำระหว่างอักษรเบรลล์และอักษรปกติ เช่นวิธี Joint source-channel model และ Chunk-based approach เนื่องจากการจัดทำคลังข้อความของทั้งสองระบบที่มีเนื้อหาเดียวกันนั้น (comparable corpora) จะทำให้สิ้นเปลืองเวลาและมีค่าใช้จ่ายสูง เพราะต้องอาศัยผู้เชี่ยวชาญอักษรเบรลล์ในการตรวจสอบความถูกต้องและความสอดคล้องของคลังข้อความทั้งสองประเภทนี้

ดังนั้น ผู้วิจัยจึงขอเสนอว่า ควรจะนำวิธีการใช้กฎทางภาษา (Rule-based) มาใช้ร่วมกับแบบจำลองทางสถิติในการถ่ายถอดอักษรเบรลล์ให้เป็นอักษรปกติ โดยที่ใช้กฎกับกรณีที่สามารถจับคู่อักษรเบรลล์และอักษรปกติได้โดยตรง และใช้แบบจำลองเอ็นแกรมซึ่งผ่านการฝึกระบบด้วยข้อความปกติมาใช้ในการเลือกรูปแบบที่ถูกต้องที่สุดในกรณีที่สามารถถ่ายถอดเบรลล์ให้เป็นอักษรปกติได้มากกว่า 1 กรณี เนื่องจากวิธีการต่าง ๆ ที่กล่าวมานี้ไม่ต้องอาศัยคลังข้อความเบรลล์ซึ่งมีความยุ่งยากและต้องใช้เวลามากในการจัดทำ สำหรับรายละเอียดของแบบจำลองที่จะนำเสนอนี้ผู้วิจัยจะได้อภิปรายในหัวข้อที่เกี่ยวข้องต่อไป

2.4 งานวิจัยที่เกี่ยวข้องกับการถ่ายถอดอักษรเบรลล์ไทยและเบรลล์อังกฤษ

จากการทบทวนวรรณกรรม ผู้วิจัยพบว่าในประเทศไทยมีการทำวิจัยเกี่ยวกับการถ่ายถอดอักษรเบรลล์น้อยมาก โดยพบว่ามีงานวิจัยที่ศึกษาการถ่ายถอดเบรลล์ไทยอังกฤษให้เป็นอักษรปกติเพียงสองชิ้น โดยงานชิ้นแรก (รวิศม์ วงษ์สมาน, 2551) ศึกษาการถ่ายถอดระหว่างเบรลล์ไทยปนอังกฤษกับอักษรปกติ โดยเน้นที่การระบุภาษาให้ได้ว่าเป็นเบรลล์ไทยหรืออังกฤษก่อนถ่ายถอดอักษรและแก้ปัญหาความกำกวมในการถ่ายถอดอักษร ส่วนงานชิ้นที่สอง (ทศวัฒน์ ชุณหวิทยะธีระ, 2555) ศึกษาการถ่ายถอดเบรลล์ไทยให้เป็นอักษรปกติโดยเน้นที่การแก้ปัญหาความกำกวมที่เกิดจากคำควบกล้ำในภาษาไทยโดยใช้แบบจำลองเอ็นแกรม ดังมีรายละเอียดดังนี้

2.4.1 การออกแบบและพัฒนาระบบถ่ายถอดอักษรไปกลับ อังกฤษผสมไทยเป็นเบรลล์ (รวิศม์ วงษ์สมาน, 2551)

รวิศม์ วงษ์สมาน (2551) ได้ศึกษาการถ่ายถอดอักษรเบรลล์ที่มีภาษาไทยและอังกฤษปนกัน โดยมีวัตถุประสงค์ที่จะสร้างระบบถ่ายถอดอักษรเบรลล์ไทยปนอังกฤษทั้งไปและกลับ

ในการทำงานกับข้อมูลที่มีทั้งสองภาษาปนกัน คอมพิวเตอร์จำเป็นที่จะต้องระบุให้ได้ว่าส่วนของข้อความนั้นเป็นของภาษาใดจึงจะทำงานได้ถูกต้อง โดยทั่วไปคอมพิวเตอร์สามารถจะจำแนกข้อความอักษรปกติที่มีภาษาปนกันได้โดยอัตโนมัติ เนื่องจากทั้งสองภาษาแสดงด้วยชุดตัวอักขระที่แตกต่างกันทำให้ง่ายต่อการระบุภาษา แต่ในกรณีของอักษรเบรลล์เป็นการยากที่จะจำแนกเบรลล์ที่มีการปนกันของสองภาษา เนื่องจากเบรลล์ไทยและเบรลล์อังกฤษต่างก็ใช้ระบบ 6 จุดเพื่อแสดงตัวอักษรแต่ละภาษาจึงทำให้อักษรทั้ง 2 ภาษามีรูปเบรลล์ที่ซ้ำกันได้

นอกจากนี้ในการถ่ายถอดเบรลล์ไทยเป็นอักษรปกติ คอมพิวเตอร์จำเป็นต้องแก้ความกำกวมที่เกิดจากการที่เบรลล์ 1 ตัว ใช้แทนอักษรปกติได้มากกว่า 1 กรณี การที่จะบ่งชี้ได้ว่าหมายถึงอักษรปกติตัวใดจำเป็นต้องอาศัยบริบทในการตีความเบรลล์ตัวนั้น ตัวอย่างเช่น ‘⠷’ หากปรากฏหน้าหรือหลัง ‘⠡’ (◌า) หรือ ‘⠽’ (◌อ) จะตีความเป็นไม้โท เช่น ‘⠷⠡⠷⠽’ พบว่า ‘⠷’ปรากฏหน้า ‘⠡’ (◌า) จึงต้องตีความให้เป็น “◌้” แต่หาก ‘⠷’ปรากฏกับพยัญชนะแล้วไม่สามารถอ่านเป็นคำได้จะตีความเป็น “.” (จุด) เช่น ‘⠷⠤⠷⠷’ ต้องตีความ ‘⠷’ ให้เป็นเครื่องหมายจุด เพราะคำว่า “ม้ค้” ไม่สามารถตีความได้ เป็นต้น

เนื่องจากผู้วิจัยสนใจศึกษาเฉพาะการถ่ายถอดอักษรเบรลล์เป็นอักษรปกติเท่านั้น ฉะนั้นผู้วิจัยจึงจะขอนำเสนองานวิจัยของรวิศม์ เฉพาะส่วนของการถ่ายถอดเบรลล์ไทยปนอังกฤษเป็นอักษรปกติ และจะไม่กล่าวถึงงานของรวิศม์ในส่วนที่เป็นการถ่ายถอดอักษรปกติเป็นอักษรเบรลล์

ในการถ่ายถอดเบรลล์ไทยปนอังกฤษ รวิศม์ได้เสนอกระบวนการต่าง ๆ คือ เริ่มจากการนำเข้าข้อมูลเบรลล์แต่ละบรรทัด แล้วทำการตัดแบ่งออกเป็นสายอักขระโดยใช้ช่องว่าง (space) เป็นตัวแบ่ง (delimiter) ก่อนที่จะนำแต่ละสายอักขระไปตรวจสอบว่าเป็นภาษาไทยหรืออังกฤษ ในการตัดสินว่าอักขระแต่ละสายเป็นภาษาใดทำได้โดยใช้แบบจำลองเอ็นแกรมที่ผ่านการฝึกฝนด้วยคลังข้อมูลอักษรเบรลล์ทั้งไทยและอังกฤษ (ซึ่งรวิศม์พบว่าค่าที่ 4 แกรมเหมาะสมที่สุด) โดยเปรียบเทียบคะแนนที่ได้จากแบบจำลอง หากภาษาใดมีระดับคะแนนสูงกว่าคอมพิวเตอร์จะระบุว่าเป็นภาษานั้น

สมการคำนวณระดับคะแนนของการเลือกภาษาโดยใช้แบบจำลองเอ็นแกรม

$$\text{ThScore} = \frac{w1 * S1th + w2 * S2th + w3 * S3th + w4 * S4th}{\sum_{i=1}^{4} wi}$$

$$\text{EnScore} = \frac{w1 * S1en + w2 * S2en + w3 * S3en + w4 * S4en}{\sum_{i=1}^{4} wi}$$

<u>หมายเหตุ</u>

w1, w2, w3, w4 คือ ค่าน้ำหนักสำหรับแบบจำลองที่ 1, 2, 3 และ 4 แกรมตามลำดับ

S1th, S2th, S3th, S4th คือความน่าจะเป็นของการคิดคะแนนเบรลล์ไทยด้วยแบบจำลองที่ 1, 2, 3 และ 4 แกรมตามลำดับ

S1en, S2en, S3en, S4en คือความน่าจะเป็นของการคิดคะแนนเบรลล์อังกฤษด้วยแบบจำลอง 1, 2, 3 และ 4 แกรมตามลำดับ

จากการศึกษา รวิศม์พบว่าค่าน้ำหนักที่ทำให้แยกเบรลล์ไทยและเบรลล์อังกฤษออกจากกันได้มากที่สุดคือ w1 = 1, w2 = 10^{10}, w3 = 10^{20}, w4 = 10^{30}

หลังจากที่ทำการเลือกภาษาแล้วก็จะนำแต่ละสายอักขระไปถ่ายถอดให้ตรงกับภาษาที่เลือก โดยแบ่งเป็น 2 สายงานคือการถ่ายถอดเบรลล์อังกฤษ และการถ่ายถอดเบรลล์ไทย

ในสายงานแรก การถ่ายถอดอักษรไม่มีความซับซ้อนมากนักสามารถถ่ายถอดแบบตัวต่อตัวได้ รวิศม์จะทำการถ่ายถอดโดยเทียบอักขระต่ออักขระทันที ในสายงานการถ่ายถอดเบรลล์ไทยแบ่งการทำงานออกเป็น 2ส่วน คือการถ่ายถอดโดยการใช้กฎ และการถ่ายถอดเบรลล์ไทยที่กำกวมโดยอาศัยแบบจำลองภาษา

ในการถ่ายถอดเบรลล์ไทยด้วยกฎ รวิศม์ได้เสนอให้ตรวจสอบค่าอักขระแต่ละตัวจากซ้ายไปขวาทีละอักขระ ซึ่งสามารถนำมาสรุปไว้เป็นกฎต่าง ๆ ได้ดังนี้

(1) หากพบ ‘◌ร’ (ร) หรือ ‘◌ล’ (ล) ให้ตรวจสอบอักขระถัดไป ว่าเป็น ‘◌ั’ หรือไม่

ก. ถ้าใช่ ให้ตรวจสอบตัวถัดไปว่าเป็น ‘◌ิ’ หรือไม่

ถ้าใช่ ก็ให้ถ่ายถอด ‘◌ร◌ั◌ิ’ หรือ ‘◌ล◌ั◌ิ’ เป็น “ฤา” หรือ “ฦา” ตามลำดับ

ถ้าไม่ใช่ ให้ถ่ายถอด ‘◌ร◌ั’หรือ ‘◌ล◌ั’ เป็น “ฤ” หรือ “ฦ” ตามลำดับ

ข. ถ้าไม่ใช่ ให้ถ่ายถอด ‘◌ร’ หรือ ‘◌ล’ เป็น “ร” หรือ “ล” ตามลำดับ

(2) หากพบ ‘◌ิ’ (ไ) ให้ตรวจสอบอักขระถัดไป ว่าเป็น ‘◌ั’ หรือไม่

ก. ถ้าใช่ ให้ถ่ายถอด ‘◌ิ◌ั’ เป็น “ไ แล้วจึงตรวจสอบอักขระต่อไป

ข. ถ้าไม่ใช่ ให้ถ่ายถอด ‘◌ิ’ เป็น “ไ

(3) หากพบ ‘◌ุ’ (เ◌), ‘◌ู’ (แ◌) , ‘◌ั’ (โ◌) ให้ตรวจสอบอักขระถัดไปว่าเป็น‘◌ิ’ (◌ะ) หรือไม่

ก. ถ้าใช่ ให้ถ่ายถอด ‘◌ุ◌ิ’ , ‘◌ู◌ิ’หรือ ‘◌ั◌ิ’ เป็น “เ◌ะ”, “แ◌ะ” หรือ “โ◌ะ” โดยดึงอักขระก่อนหน้ามาเป็นพยัญชนะต้น

ข. ถ้าไม่ใช่ ให้ถ่ายถอด ‘◌ุ’, ‘◌ู’ หรือ ‘◌ั’ เป็น “เ◌”, “แ◌” หรือ “โ◌” ตามลำดับ

(4) หากพบ ‘◌ิ’ (ะ) หรือ ‘◌ำ’ (◌ำ) ให้ตรวจสอบอักขระถัดไปว่าเป็นวรรณยุกต์หรือไม่

ก. ถ้าใช่ ให้ถ่ายถอด ‘◌ิ’หรือ ‘◌ำ’ ที่ตามด้วยวรรณยุกต์ ให้เป็นวรรณยุกต์แล้วตามด้วย “◌ะ” หรือ “◌ำ”

ข. ถ้าไม่ใช่ ให้ถ่ายถอด ‘◌ิ’ หรือ ‘◌ำ’ เป็น “◌ะ” หรือ “◌ำ” ตามลำดับ

(5) หากพบ ‘◌ี’ (เ◌ีย) , ‘◌ื’ (เ◌ือ) , ‘◌ั’ (◌ัว) , ‘◌ื ’ (เ◌อ) หรือ ‘◌ี’ (◌อ) ให้ตรวจสอบอักขระถัดไป ว่าเป็น ‘◌ิ’ (◌ะ) หรือไม่

ก. ถ้าใช่ ให้ถ่ายถอด ‘◌ี◌ิ, ‘◌ื◌ิ, ‘◌ั◌ิ, ‘◌ื◌ิ หรือ ‘◌ี◌ิ เป็น “เ◌ียะ” , “เ◌ือะ”, “◌ัวะ”, “เ◌อะ” หรือ “◌อะ” ตามลำดับแล้วโดยดึงอักขระก่อนหน้ามาเป็นพยัญชนะต้น

ข. ถ้าไม่ใช่ ให้ถ่ายถอด‘◌ี’ , ‘◌ื’ , ‘◌ั’ , ‘◌ื’ หรือ ‘◌ี’ เป็น “เ◌ีย, “เ◌ือ, “ ◌ัว”, “เ◌อ” หรือ “◌อ” ตามลำดับ

(6) ถ้าพบ ‘◌า’ (เ◌า)ให้ถ่ายถอดเป็นสระ “เ◌า”โดยดึงอักขระก่อนหน้ามาเป็นพยัญชนะต้น

(7) หากพบพยัญชนะ (ยกเว้นพยัญชนะ อ) จะเก็บค่าไว้ 1 ตัว จากนั้นตรวจสอบอักขระถัดไปว่าเป็นพยัญชนะหรือไม่ ถ้าใช่จะเก็บค่าไว้อีก 1 ตัว (รวมเป็น 2 ตัว) แล้วตรวจสอบตัวถัดไป หากพบพยัญชนะอีกให้ถ่ายถอดตัวที่ 1 ที่เก็บเอาไว้เป็นอักษรปกติ พร้อมเก็บค่าพยัญชนะตัวใหม่ไว้ โดยจะต้องเก็บค่าพยัญชนะที่อยู่ติดกันได้สูงสุดไม่เกิน 2 ตัว จากนั้นตรวจสอบตัวถัดไปเรื่อยๆ จนกว่าพบอักขระเบรลล์ที่ไม่ใช่พยัญชนะจึงถ่ายถอดอักขระที่เก็บไว้

อย่างไรก็ตาม หากพบสระตามข้อ (3), (5) หรือ(6) ตามหลัง ให้ตรวจสอบพยัญชะ 2 ตัวที่เก็บค่าไว้ ว่าอยู่ในรายการอักษรควบกล้ำหรือไม่

ก. ถ้าใช่ ให้ถ่ายถอดพยัญชนะ 2 ตัวนั้นร่วมกับสระเหล่านี้ตามกฎการประสมอักษรไทย

ข. ถ้าไม่ใช่ ให้ถ่ายถอดพยัญชนะตัวแรกก่อน แล้วถ่ายถอดพยัญชนะตัวที่ 2 ร่วมกับสระ

(8) หากพบสระ หรือวรรณยุกต์อื่น ๆ ที่ไม่เป็นไปตามกฎที่กล่าวมาข้างต้น ให้ถ่ายถอดเป็นอักษรปกติได้ทันที โดยยึดจากพจนานุกรมอักษรเบรลล์

อย่างไรก็ตามรวิศม์ได้ตั้งข้อสังเกตว่ากฎที่เสนอมานี้ไม่สามารถใช้ถ่ายถอดได้ถูกต้องในหลายกรณี เช่น ในกรณีที่อักขระเบรลล์บางตัวสามารถแทนอักขระปกติได้มากกว่า 1 ตัว รวิศม์ได้รวบรวมอักขระที่กำกวมเหล่านี้ไว้ 8 อักขระ ดังนี้

ตารางที่ 2.24 เบรลล์เซลล์เดี่ยวที่กำกวม

สัญลักษณ์	กรณีที่ 1	กรณีที่ 2	กรณีที่ 3
⠦	้ (ไม้โท)	. (จุด)	
⠒	ู (สระอู)	: (โคลอน)	
⠔	ี (สระอี)	; (เซมิโคลอน)	
⠢	์ (การันต์)	” (ปิดคำพูด)	
⠐	ๆ (ไม้ยมก)	, (คอมม่า)	
⠶	๊ (ไม้ตรี)	((วงเล็บเปิด)	) (วงเล็บปิด)
⠖	๋ (ไม้จัตวา)	“ (เปิดคำพูด)	
⠤	็ (ไม้ไต่คู้)	ฺ (ฝนทอง)	

ในการแก้ปัญหาอักขระเบรลล์ไทยที่ถ่ายถอดกลับได้ 2 กรณีขึ้นไปเหล่านี้ รวิศม์ได้เสนอให้ใช้ทฤษฎีแบบจำลองเอ็นแกรมตั้งแต่ 1-7 แกรม โดยใช้ฐานข้อมูลภาษาไทยในการฝึกฝนแบบจำลองการประมาณค่าความน่าจะเป็นของการที่เบรลล์ 1 เซลล์จะตีความให้เป็นกรณีต่าง ๆ ได้ตามสมการ

สมการความน่าจะเป็นในการถ่ายถอดกลับแต่ละกรณีมีดังนี้

$$\text{Case 1} = \frac{W1}{\sum_{i=1}^{7} wi} * (w1 * S1_1 + w2 * S2_1 + w3 * S3_1 ...)$$

$$\text{Case 2} = \frac{W2}{\sum_{i=1}^{7} wi} * (w1 * S1_2 + w2 * S2_2 + w3 * S3_2...)$$

$$\text{Case 3} = \frac{W3}{\sum_{i=1}^{7} wi} * (w1 * S1_3 + w2 * S2_3 + w3 * S3_3...)$$

จากสมการข้างต้น ในการคำนวณความเป็นไปได้ที่จะเกิดอักขระในแต่ละกรณีต้องใช้ผลรวมของแต่ละแกรม ในการคำนวณความเป็นไปได้ของแต่ละแกรมต้องคำนวณจากจุดเริ่มต้นของสายอักขระปัจจุบันที่ได้รับการถ่ายถอดเป็นอักษรปกติแล้วจนถึงอักษรที่ตีความได้มากกว่าหนึ่งกรณี โดยใช้สมการ

$$P(C) = P(c_1, c_2 ... c_N) = P(c_1) \prod_{i=2}^{N} P(c_i \mid c_{i-1}, ..., c_1)$$

นอกจากนี้ในสมการถ่ายถอดกลับแต่ละกรณียังมีการปรากฏของตัวแปรต่าง ๆ คือ ค่า w1-w7 (ตัวเลขหลัง w เป็นค่าที่ระบุจำนวนแกรม) เป็นค่าน้ำหนักที่ทำให้ผลการถ่ายถอดกลับถูกต้องมากที่สุด และ W1 - W3 (ตัวเลขหลัง W เป็นค่าบอกลำดับของกรณีต่าง ๆในการถ่ายถอดกลับ) เป็นค่าน้ำหนักในความเป็นไปได้ที่จะเกิดอักขระในแต่ละกรณี

จากการศึกษาวิทยานิพนธ์ฉบับนี้ ผู้วิจัยพบว่าข้อมูลอักษรเบรลล์บางส่วนที่รวิศม์นำมาใช้ในการวิจัยมีความคลาดเคลื่อนโดยเฉพาะอย่างยิ่งเบรลล์ที่กำกวม หลายตัวไม่ได้มีการใช้จริงในเบรลล์ไทย เช่น ‘⠒’ ที่รวิศม์อ้างว่าเป็นได้ทั้ง “ู” และ “ : ” นั้น ในความเป็นจริงเครื่องหมาย “:” ในเบรลล์ไทยคือ ‘⠐⠒’ ไม่ใช่ ‘⠒’ เหมือนในเบรลล์อังกฤษ ทำให้รหัสดังกล่าวไม่เป็นอักษรที่มีความกำกวมตามที่อ้างไว้

นอกจากอักขระที่ไม่ได้ปรากฏจริงแล้ว กฎที่รวิศม์เสนอยังไม่สามารถแก้ปัญหาความกำกวมที่พยัญชนะ 2 ตัวติดกันมีรูปคล้ายควบกล้ำแต่ไม่ใช่พยัญชนะควบกล้ำ ในกรณีนี้เมื่อผ่านกฎที่รวิศม์ได้เสนอไว้จะส่งผลให้ถ่ายถอดเป็นพยัญชนะควบกล้ำทันที เช่น ‘⠃⠕⠐⠓⠔⠔⠕⠅⠇⠜⠎’ (บ-่อ-ห-ิน-น-อ-ก-ล-แ-ะ) ถ่ายถอดเป็น “บ่อหินนอแกละ” แทนที่จะเป็น “บ่อหินนอกและ” นอกจากนี้ยังขาดกฎที่จำเป็นในการถ่ายถอดสระที่เบรลล์ตัวหนึ่งสามารถแทนอักษรปกติได้ 2 รูป รวิศม์ระบุว่า ‘⠐’ ใช้แทน “เ◌อ” ได้เพียงกรณีเดียวเท่านั้น แต่ในเบรลล์ไทย สัญลักษณ์ดังกล่าวใช้แทน “เ◌ิ” ได้เช่นกัน ทำให้ไม่สามารถถ่ายถอดคำว่า เดิน เหิน เงิน ฯลฯ ได้ถูกต้อง

นอกจากความคลาดเคลื่อนเกี่ยวกับเบรลล์และกฎที่ใช้แล้ว เป็นที่น่าสังเกตว่าคลังข้อความที่รวิศม์นำมาใช้ก่อให้เกิดความไม่สะดวกหลายประการ ประการแรก การใช้ข้อมูลอักษรเบรลล์มาฝึกระบบทำให้สิ้นเปลืองเวลาและทรัพยากร เพราะในการใช้คลังข้อความเบรลล์ต้องมีผู้เชี่ยวชาญ

ตรวจสอบความถูกต้องเสียก่อน มิฉะนั้นจะทำให้คอมพิวเตอร์ประมวลผลได้ไม่ถูกต้อง นอกจากนี้การใช้ข้อมูลเบรลล์มาฝึกระบบ ในบางกรณีจะส่งผลให้ไม่สามารถจะจำแนกภาษาไทยและอังกฤษออกจากกันได้ เพราะสายอักขระบางส่วนในคลังข้อความภาษาไทยอาจตรงกับสายอักขระในคลังข้อความภาษาอังกฤษ เช่น ‘⠇⠁⠉⠗’ ที่ปรากฏในคลังข้อความเบรลล์ทั้งสองภาษา คือคำว่า "ละคร" หรือ "laur" เป็นต้น

2.4.2 การใช้เอ็นแกรมช่วยในการเพิ่มประสิทธิภาพการตัดสินใจถ่ายถอดอักษรเบรลล์ ในกรณีที่มีคำควบกล้ำ สระประสมและอักษรเบรลล์ 2 เซลล์ (ทศวัฒน์ ชุณหวิริยะธีระ, 2555)

ในการถ่ายถอดเบรลล์ไทยโดยใช้กฎทางภาษานั้น ทศวัฒน์พบว่ามีปัญหาความกำกวมที่เบรลล์บางคู่อาจตีความเป็นอักษรปกติได้หลายกรณี ได้แก่ ความกำกวมที่เกิดจากพยัญชนะควบกล้ำที่ปรากฏร่วมกับสระประสม ดังที่ได้นำเสนอในหัวข้อ 2.4.1 มาแล้ว นอกจากนี้ยังเกิดความกำกวมจากเบรลล์ 2 เซลล์ซึ่งอยู่ชิดกัน ที่อาจตีความให้เป็นอักขระเดียวหรือสองอักขระในอักษรปกติก็ได้ ขึ้นอยู่กับบริบทที่เบรลล์คู่นั้นปรากฏอยู่ เช่น ‘⠠⠞’ อาจตีความให้เป็น “ธ” หรือ “ ์ท” ได้ เช่น ‘⠠⠞⠠⠞⠁⠝⠁⠀⠠⠞⠠⠞⠁’ “พุทธศาสนา” ต้องตีความ ‘⠠⠞’ให้เป็น “ธ” แต่ ‘⠀⠠⠞⠠⠞⠠⠞⠁⠀⠠⠞⠁⠀’ (ศุกร์ที่) ต้องตีความ‘⠠⠞’ให้เป็น “ ์ท” ซึ่งในกรณีหลังนี้ ทศวัฒน์ได้ระบุว่ามีเบรลล์ 2 เซลล์ที่กำกวมทั้งหมด 5 คู่ ในภาษาไทย ดังนี้

ตารางที่ 2.25 เบรลล์เซลล์คู่ที่กำกวม

เบรลล์ 2 เซลล์ที่กำกวม	กรณีที่ 1	กรณีที่ 2
⠠⠗	ฤ	รๆ
⠠⠇	ฦ	ลๆ
⠠⠁	เ ะ	ะ
⠠⠞	ธ	์ท
⠠⠃	ฺ	์ข

ความกำกวมที่กล่าวมาข้างต้นก่อให้เกิดปัญหาในการถ่ายถอดเบรลล์ไทย เนื่องจากคอมพิวเตอร์ไม่สามารถถอดรหัสเบรลล์ที่กำกวมออกเป็นอักษรปกติได้อย่างถูกต้อง อย่างไรก็ตาม ทศวัฒน์เชื่อว่าสามารถแก้ปัญหาความกำกวมเหล่านี้ได้ โดยเสนอให้มีการตัดคำในเบรลล์ไทยด้วยเทคนิคเอ็นแกรมก่อนที่จะทำการแปลงเป็นอักษรปกติ

ก่อนที่จะทำการตัดคำ ทศวัฒน์ได้สร้างแบบจำลองเอ็นแกรมของเบรลล์ไทยระดับ 1 (ดูหัวข้อ 2.1.2) ที่จำนวนแกรมเท่ากับ 3 แกรม, 4 แกรม และ 5 แกรมตามลำดับ โดยเริ่มจากสกัดประโยคที่ประกอบด้วยคำที่มีอักษรควบกล้ำ สระประสม และประโยคที่มีการปรากฏของเบรลล์ 2 เซลล์ที่กำกวมจากคลังข้อมูล BEST (คลังข้อความภาษาไทยที่พัฒนาโดยศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ) แล้วนำประโยคทั้งหมดมาตัดคำด้วยโปรแกรม SWATH โดยคั่นระหว่างคำด้วยเครื่องหมาย | (vertical bar) จากนั้นนำประโยคที่ตัดแบ่งคำแล้วไปถ่ายถอดเป็นอักษรเบรลล์ ผลที่ได้คือคลังข้อความเบรลล์ที่มีการแยกคำตามอักขระการแยกคำภาษาไทยในคลังข้อความ BEST หลังจากนั้นนำคลังข้อความเบรลล์นี้ไปฝึกฝนเพื่อสร้างแบบจำลองเอ็นแกรม

หลังจากที่ได้สร้างแบบจำลองเอ็นแกรมแล้ว ทศวัฒน์ได้นำแบบจำลองนี้ไปใช้ในการแบ่งคำอักษรเบรลล์ ก่อนที่จะถ่ายถอดเบรลล์ไทยเป็นไทยปกติ โดยมีสายงานดังนี้

ขั้นตอนที่ 1 pre_ngrams_generate เป็นการรับเข้าข้อความอักษรเบรลล์ภาษาไทย เริ่มจากนำเข้าข้อความเบรลล์ทีละสายอักขระ ซึ่งตัดแบ่งด้วยเครื่องหมายเว้นวรรค (space) ตัดออกเป็นส่วน ๆ ตามขนาดแกรมที่กำหนดไว้ ใส่ในตัวแปรชนิด array

ขั้นตอนที่ 2 generate_ngram_pattern มีขั้นตอนคือ (1) สร้างตัวแปรชนิด array of String อีกชุดหนึ่งที่มีจำนวน gram+(gram-1) (2) นำอินพุตทีละชุดจากขั้นตอนแรกมาเติมอักขระช่องว่าง เป็นจำนวนแกรมลบหนึ่งทั้งข้างหน้าและข้างหลัง เช่นหากข้อความอินพุตเป็นสี่แกรม (เช่น “เพลา”) จะได้สายอักขระใหม่ที่ประกอบด้วย 3 ช่องว่างตามด้วยข้อความ4แกรมแล้วตามด้วยช่องว่าง (3) นำผลลัพธ์มาคัดลอกอักขระครั้งละเป็นจำนวนแกรมที่กำหนดใส่ในตัวแปรarrayที่สร้างขึ้นใหม่ โดยหลังคัดลอกแต่ละครั้งจะเลื่อนไปทางขวา 1 อักขระ ทำเช่นนี้จนสิ้นสุดอักขระนำเข้า เช่นในคำว่า___เพลา___ ถ้าทดลองที่ 4 แกรม จะได้เป็น

array [1] = _ _ _เ
array [2] = _ _เพ
array [3] = _เพล
array [4] = เพลา
array [5] = พลา_
array [6] = ลา_ _
array [7] = า_ _ _

ขั้นตอนที่ 3 select_output_ pattern นำข้อมูลใน array ขั้นตอนที่ 2 มาเปรียบเทียบกับคลังข้อความเบรลล์ซึ่งมีการแยกคำไว้แล้ว เพื่อตัดสินว่าควรจะแบ่งคำในข้อมูลเบรลล์อย่างไร โดยเลือกแบบที่ให้ค่าความน่าจะเป็นสูงสุดเมื่อคำนวณจากแบบจำลองเอ็นแกรมที่สร้างขึ้น โดยมี

การใส่เครื่องหมาย “<pipe>” เพื่อเป็นสัญลักษณ์ในการแบ่งคำแต่ละคำออกจากกัน แล้วส่งค่าความยาวของประโยคที่เหลือกลับไปยังขั้นตอนที่ 1 เพื่อทำซ้ำทั้งกระบวนการ จนจบสายอักขระที่นำเข้า

ขั้นตอนที่ 4 translate_to_thai ทำการถ่ายถอดเบรลล์ที่ได้จากขั้นตอนที่ 3 เป็นอักษรปกติ โดยการจับคู่คำในเบรลล์ไทย (ซึ่งเบรลล์แต่ละคำจะถูกแบ่งด้วย “<pipe>”) กับคำไทยปกติ ซึ่งอ้างอิงจากรายการคู่คำเบรลล์ไทย-ไทยปกติที่ได้สร้างไว้ก่อนหน้านี้ โดยคำไทยปกติที่เกิดขึ้นจากการจับคู่คำนี้จะถูกนำมาเรียงต่อกันตามลำดับ(sequence)ทำให้ได้ผลลัพธ์เป็นข้อความไทยปกติ ที่ตรงกับข้อความเบรลล์ที่นำเข้า

เพื่อทดสอบประสิทธิภาพของแบบจำลองดังกล่าว ทศวัฒน์ได้ทำการเปรียบเทียบการทำงานของโปรแกรมที่พัฒนาขึ้นภายใต้แบบจำลองที่ได้เสนอกับ โปรแกรมถ่ายถอดเบรลล์เป็นแอสกีภาษาไทยแบบไฟล์ต่อไฟล์ซึ่ง วรพล (2551) ได้เสนอไว้ โดยทศวัฒน์ได้เปรียบเทียบใน 3 ลักษณะ คือ ความถูกต้องในการถ่ายถอดอักษร ปริมาณการใช้หน่วยความจำ และเวลาที่ใช้

ผลการทดสอบทศวัฒน์พบว่า วิธีการที่เสนอใหม่สามารถแปลงคำควบกล้ำ สระประสม และเบรลล์สองเซลล์ได้ถูกต้องมากกว่าวิธีการเดิมคิดเป็น 13.94% 14.94% และ 40.51% ตามลำดับ ในด้านหน่วยความจำพบว่าวิธีการใหม่ใช้หน่วยความจำน้อยกว่าวิธีเดิม 21.81 เท่า แต่วิธีการใหม่ใช้เวลามากกว่าวิธีการเดิม 272.71 เท่า อย่างไรก็ตามเพื่อให้วิธีการใหม่นี้ทำงานได้ผลลัพธ์ตามที่ได้กล่าวมาข้างต้น จะต้องใช้แบบจำลองเอ็นแกรมที่มีขนาด 4 แกรมซึ่งเป็นผลลัพธ์ที่ดีที่สุด

ในงานวิจัยชิ้นนี้ทศวัฒน์ตั้งข้อสังเกตว่า ระบบที่ได้เสนอไม่ควรนำไปใช้แต่เพียงลำพัง เนื่องจากต้องใช้เวลามากในการทำงาน นอกจากนี้ยังพบว่ามีการถ่ายถอดอักษรผิดพลาดในบางกรณี เพราะคลังข้อความอักษรเบรลล์ที่นำมาใช้ฝึกระบบขาดการตรวจสอบจากผู้เชี่ยวชาญอักษรเบรลล์

ผู้วิจัยมีความเห็นสอดคล้องกับทศวัฒน์ว่า ระบบที่เสนอมาข้างต้นไม่สามารถนำไปใช้งานโดยลำพังได้ โดยได้พบเพิ่มเติมว่า นอกจากระบบดังกล่าวจะใช้เวลามากแล้ว ยังถ่ายถอดได้เฉพาะคำที่ปรากฏในรายการคู่คำเบรลล์ไทย-ไทยปกติที่ทศวัฒน์ได้รวบรวมไว้เท่านั้น แต่ในบางกรณีอักษรเบรลล์ที่นำมาไม่ใช่คำที่สมบูรณ์ หรือเป็นคำที่ไม่ได้ปรากฏในรายการคำที่รวบรวมไว้ ระบบจะไม่สามารถจับคู่คำอักษรเบรลล์กับคำอักษรปกติได้ นอกจากนี้ผู้วิจัยยังพบว่าคำที่ปรากฏในรายการคู่คำเบรลล์ไทยและไทยปกตินั้นยังจำกัดอยู่แค่เพียงภาษาเดียวคือภาษาไทยเท่านั้น ทำให้ไม่รองรับกับการถ่ายถอดคำในภาษาอังกฤษซึ่งใช้ในการสื่อสารของคนตาบอดเช่นกัน

เพื่อแก้ปัญหาต่าง ๆ ดังที่ได้อภิปรายมาข้างต้น ผู้วิจัยจึงต้องการที่จะศึกษาการถ่ายถอดอักษรเบรลล์ โดยการใช้กฎทางภาษาร่วมกับแบบจำลองเอ็นแกรมที่ฝึกฝนระบบด้วยอักษรไทยปกติ

โดยจะถ่ายถอดเบรลล์ไทยผ่านการใช้กฎทางภาษา แล้วนำอักษรปกติที่ได้จากการถ่ายถอดมาสร้างรูปแบบการจัดเรียงตัวอักษรในรูปแบบต่าง ๆ เท่าที่จะเป็นไปได้ หลังจากนั้นนำรูปแบบเหล่านี้ไปเทียบกับแบบจำลองเอ็นแกรม เพื่อเลือกรูปแบบที่เหมาะสมที่สุดที่จะใช้เป็นผลลัพธ์ในการถ่ายถอดอักษร ซึ่งจะขอเสนอรายละเอียดเกี่ยวกับระบบที่สร้างขึ้นในหัวข้อที่เกี่ยวข้องต่อไป

2.5 โปรแกรม Duxbury Barille Translator

ในหัวข้อนี้ ผู้วิจัยจะขออภิปรายเกี่ยวกับโปรแกรม Duxbury Braille Translator ที่ได้ทำการทบทวนวรรณกรรมมาแล้ว เนื่องจากในงานวิจัยนี้ ผู้วิจัยได้นำโปรแกรมดังกล่าวมาใช้เป็นเครื่องมือในการสร้างคลังข้อความอักษรเบรลล์เพื่อฝึกฝนระบบ

โปรแกรม Duxbury Braille Translator (DBT) เป็นโปรแกรมถ่ายถอดอักษรเบรลล์ที่ผลิตโดยบริษัท Duxbury Systems, Inc. ประเทศสหรัฐอเมริกา ซึ่งโปรแกรมดังกล่าวรองรับการถ่ายถอดอักษรปกติและอักษรเบรลล์แบบไป-กลับ (Duxbury systems. 2014: 1) โดยที่ถ่ายถอดได้ทั้งเบรลล์เกรด 1 และเบรลล์เกรด 2 และรองรับภาษาได้มากกว่า 130 ภาษา รวมทั้งภาษาไทย นอกจากนี้โปรแกรมนี้ยังรองรับการถ่ายถอดสัญลักษณ์ทางคณิตศาสตร์เป็นอักษรเบรลล์ ด้วยเหตุนี้โปรแกรมดังกล่าวจึงเป็นที่นิยมใช้กันในกลุ่มคนตาบอดทั่วโลก โปรแกรม Duxbury มีการพัฒนามาอย่างต่อเนื่องตั้งแต่ปีค.ศ. 1975 จนถึงปัจจุบันพัฒนามาถึง DBT 12.1 ซึ่งเวอร์ชันดังกล่าวรองรับการถ่ายถอดเบรลล์ไทยรูปเต็มทั้งไปและกลับ แม้ว่าโปรแกรมนี้จะสามารถถ่ายถอดเบรลล์ไทยเป็นอักษรปกติได้แต่ก็ไม่สามารถถ่ายถอดได้ถูกต้องสมบูรณ์ ดังแสดงในตัวอย่างต่อไปนี้

ตารางที่ 2.26 ตัวอย่างข้อความไทยที่โปรแกรม Duxbury Braille Translator ถ่ายถอดผิด

ข้อความเบรลล์	การถอดรหัสเบรลล์	อักษรปกติ	ผลลัพธ์จากโปรแกรม
⠝⠑⠗⠊⠽⠝	น-ั-ก-ร-เ◌ีย-น	นักเรียน	นัเกรียน
⠏⠗⠞⠕⠝⠅⠞⠗⠝	พ-ร-ะ-อ-ง-ค-์-ท-ร-ง	พระองค์ทรง	พระองคธรง (์ท = ธ)
⠓⠇⠊⠅⠑⠇⠽⠝	ห-ล-ี-ก-ล-เ◌ีย -่-ง	หลีกเลี่ยง	หลีเกลี่ยง
⠉⠓⠝⠓⠁	ค- ้- น- ห-า	ค้นหา	ค. นหา (้ = .)
⠃⠁⠝⠉⠁⠇⠑⠎⠁⠇⠽⠝	ข-้-า-ง-น-อ- ก- ล- แ ะ	ข้างนอกและ	ข้างนอแกละ

จากตารางข้างต้นจะพบว่าโปรแกรม Duxbury ยังไม่รองรับการถ่ายถอดสระที่พยัญชนะสองตัวหน้าไม่ใช่พยัญชนะควบกล้ำแท้ และยังไม่รองรับกับความกำกวมที่เกิดจากเบรลล์สองเซลล์ซึ่งสามารถตีความให้เป็นอักษรปกติตัวเดียวหรือสองตัวได้ ซึ่งผู้วิจัยได้พบว่าปัญหาดังกล่าวเกิดขึ้นกับงานวิจัยของรวิศ วงษ์สมาน (2551) เช่นเดียวกัน

นอกจากนี้ผู้วิจัยยังพบว่า หากข้อความภาษาไทยในวรรคใดลงท้ายด้วยรหัสเบรลล์ที่คล้ายกับเครื่องหมายวรรคตอนในภาษาอังกฤษ (punctuation) เช่น ‘⠲’ (เป็น “.” (จุด)ในภาษาอังกฤษ หรือ “◌้” ในภาษาไทยได้) หรือ ‘⠆’ (เป็น “;” หรือ “ ◌ี” ได้) โปรแกรม Duxbury จะทำการถ่ายถอดเครื่องหมายเหล่านี้ให้เป็นเครื่องหมายวรรคตอนในภาษาอังกฤษทันทีเช่นข้อความไทยที่จบวรรคด้วย‘⠠⠊⠅⠲’ คือคำว่า “เช่นนี้” แต่โปรแกรมดังกล่าวจะให้ผลลัพธ์เป็น “เช่นน;.”

นอกจากปัญหาที่กล่าวมาข้างต้น โปรแกรมนี้ยังไม่รองรับการถ่ายถอดข้อความเบรลล์ไทยปนอังกฤษ เนื่องจากเบรลล์ไทยและเบรลล์อังกฤษต่างก็เกิดจากอักษรระบบ 6 จุดทำให้อักขระเบรลล์ในภาษาหนึ่งอาจถูกตีความให้เป็นอีกภาษาหนึ่งได้ อีกทั้งในการเขียนเบรลลล์ไทยปนอังกฤษนั้นไม่มีการใช้เครื่องหมายกำกับว่าเป็นภาษาใด ในการถ่ายถอดข้อความเบรลล์โปรแกรมนี้จะทำการถ่ายถอดตามสถานะภาษาปัจจุบันที่ตั้งไว้ หากตั้งสถานะไม่ตรงกับภาษาก็จะทำให้โปรแกรมถ่ายถอดผิดพลาด ดังตัวอย่างต่อไปนี้

ตารางที่ 2.27 ตัวอย่างข้อความไทยปนอังกฤษที่โปรแกรม Duxbury Braille Translator ถ่ายถอดผิด

ข้อความเบรลล์	การถอดรหัสเบรลล์	อักษรปกติ	ผลลัพธ์จากโปรแกรม
⠓⠁⠗⠙⠺⠁⠗⠑ ⠝⠤◌ี ⠤◌้	h-a-r-d-w-a-r-e น-◌ี -◌้	hardware นี้	หะรดวะรัว นี้

จากตัวอย่าง ข้อความประกอบด้วย 2 สายอักขระ สายอักขระแรกเป็นภาษาอังกฤษ และสายอักขระต่อมาเป็นภาษาไทย ในการถ่ายถอดครั้งนี้ได้มีการตั้งสถานะให้เป็นภาษาไทย เมื่อสายอักขระที่พบเป็นภาษาอังกฤษ โปรแกรมจึงทำการถ่ายถอดสายอักขระที่พบนี้เป็นภาษาไทยทันที ทำให้ได้ตัวอักษรภาษาไทยที่อ่านแล้วไม่เข้าใจ แสดงให้เห็นว่า โปรแกรมดังกล่าวยังไม่รองรับกับการสลับภาษาแบบอัตโนมัติ หากต้องการถ่ายถอดข้อความเบรลล์ไทยปนอังกฤษให้ถูกต้อง จำเป็นต้องทำการกำกับภาษาของข้อความที่ต้องการถ่ายถอดทีละส่วน ซึ่งทำได้เฉพาะผู้ที่มีความเชี่ยวชาญอักษรเบรลล์ ทำให้ต้องใช้เวลาและค่าใช้จ่ายอย่างมากในการจัดการกับปัญหาดังกล่าว

บทที่ 3

ระบบหน่วยอักขระเบรลล์อังกฤษและเบรลล์ไทย

อักษรเบรลล์เกิดจากการนำจุดในจำนวนที่จำกัดมาเรียงต่อกันในรูปแบบตายตัว เพื่อใช้แสดงตัวอักษรในภาษาต่าง ๆ ด้วยเหตุนี้อักขระเบรลล์แต่ละตัวจึงสามารถแสดงตัวอักษรที่ต่างกันระหว่างภาษาได้ แต่เนื่องจากแต่ละภาษามีตัวเขียนและอักขรวิธีที่ไม่เหมือนกัน จึงทำให้รูปของอักขระที่ซ้ำกันเหล่านี้ไม่เป็นปัญหาสำหรับการอ่านอักษรเบรลล์ เพราะผู้อ่านสามารถเข้าใจความต่างของระบบอักขรวิธีทำให้เข้าใจได้ว่ากำลังอ่านเบรลล์ภาษาใดอยู่ อย่างไรก็ตาม เนื่องจากการพัฒนาระบบ อักษรเบรลล์มักอาศัยต้นแบบจากภาษาอื่นที่พัฒนามาก่อน จึงเป็นไปได้ว่าจะมีการสร้างอักษรเบรลล์ภาษาใหม่เพิ่มเติมโดยอิงอักขระในอักษรเบรลล์ของภาษาที่มีมาก่อน

ฉะนั้น ในบทนี้ผู้วิจัยจะได้นำเสนอระบบอักขระเบรลล์ไทยที่ผู้วิจัยได้ทำการศึกษา ว่าเกิดจากการนำอักขระภาษาอื่น ๆ มาใช้เป็นอักขระพื้นฐานในเบรลล์ไทยแล้วสร้างอักขระขึ้นเพิ่มเติมเพื่อให้การอภิปรายในประเด็นต่าง ๆ ชัดเจนยิ่งขึ้น ผู้วิจัยจะเริ่มจาก การนำเสนอระบบหน่วยอักขระเบรลล์อังกฤษ หลังจากนั้นจะได้กล่าวถึง กรอบการวิเคราะห์หน่วยอักขระเบรลล์ไทย และระบบหน่วยอักขระเบรลล์ไทย ที่ผู้วิจัยได้ค้นพบในการวิจัยครั้งนี้

3.1 ระบบหน่วยอักขระเบรลล์อังกฤษ

สำหรับภาษาที่ใช้อักษรโรมันนั้น มีการศึกษาวิจัยตัวอักษรเบรลล์มาแล้วอย่างกว้างขวางทำให้มีคำอธิบายอย่างเป็นระบบ และเนื่องจากอักษรเบรลล์มีต้นกำเนิดมาจากประเทศฝรั่งเศส ซึ่งมีตัวอักษร 26 ตัวเหมือนกับที่ใช้ในภาษาอังกฤษ ดังนั้นอักษรเบรลล์ที่พัฒนาสำหรับภาษาฝรั่งเศสจึงใช้กับภาษาอังกฤษได้โดยตรง ซึ่งอักษรทั้ง 26 ตัวที่ใช้ในภาษาอังกฤษดังมีรายละเอียดดังนี้

3.1.1 รูปแบบของหน่วยอักขระเบรลล์อังกฤษ

มีเอกสารและตำราจำนวนหนึ่ง ที่อธิบายความสัมพันธ์ของอักขระเบรลล์อังกฤษแต่ละตัวว่าอักขระเบรลล์อังกฤษเกิดจากการสร้างอักขระเบรลล์พื้นฐาน 10 ตัวขึ้นมาก่อน แล้วนำอักขระเหล่านี้มาดัดแปลงโดยการเติมจุดเบรลล์ในบางตำแหน่ง ทำให้เกิดอักขระเบรลล์อังกฤษเพิ่มขึ้นจนครบ 26 ตัว (Petzold, 2000: 16; Englebreston, 2009: 3) โดยสามารถแบ่งอักขระแต่ละกลุ่มออกได้ดังนี้

3.1.1.1 หน่วยอักขระเบรลล์อังกฤษพื้นฐาน

หน่วยอักขระพื้นฐานในเบรลล์อังกฤษประกอบด้วยอักขระ 10 ตัว โดยที่อักขระเบรลล์แต่ละตัวมีจุดเบรลล์ปรากฏในตำแหน่งต่างกัน ซึ่งไม่รวมจุด 3 และ/หรือจุด 6 ดังนี้

ตารางที่ 3.1 หน่วยอักขระเบรลล์อังกฤษพื้นฐาน

หน่วยอักขระเบรลล์								
⠁ a	⠃ b	⠉ c	⠙ d	⠑ e	⠋ f	⠛ g	⠓ h	⠊ i	⠚ j

ด้วยเหตุที่หน่วยอักขระเบรลล์พื้นฐานเหล่านี้ใช้เนื้อที่เฉพาะส่วนบนและส่วนกลางของเซลล์จึงมีการนำจุดเบรลล์ในส่วนล่างของเซลล์คือ จุด 3 และ 6 มาสร้างอักขระเบรลล์ขึ้นเพิ่มเติมอีกโดยการเติมจุดเหล่านี้ภายในเซลล์

3.1.1.2 หน่วยอักขระเบรลล์อังกฤษดัดแปลง

หน่วยอักขระเบรลล์อังกฤษดัดแปลงเป็นอักขระที่เกิดขึ้นเพิ่มอีก 16 รูป จากการปรับหน่วยอักขระเบรลล์อังกฤษพื้นฐาน โดยมีรูปแบบในการสร้างหน่วยอักขระเบรลล์อังกฤษดัดแปลงดังนี้

ก. การเติมจุด 3 ภายในเซลล์

หน่วยอักขระเบรลล์ที่สร้างจากการนำหน่วยอักขระเบรลล์อังกฤษพื้นฐานมาเติมจุด 3 ภายในเซลล์ ทำให้ได้หน่วยอักขระเบรลล์ที่ใช้แทนหน่วยอักขระฉัดไปอีก 10 ตัว ดังแสดงในตารางต่อไปนี้

ตารางที่ 3.2 หน่วยอักขระเบรลล์อังกฤษดัดแปลงที่ได้จากการเติม '⠄' ภายในเซลล์

หน่วยอักขระเบรลล์										
หน่วยอักขระพื้นฐาน	⠁ a	⠃ b	⠉ c	⠙ d	⠑ e	⠋ f	⠛ g	⠓ h	⠊ i	⠚ j
หน่วยอักขระที่สร้างใหม่	⠅ k	⠇ l	⠍ m	⠝ n	⠕ o	⠏ p	⠟ q	⠗ r	⠎ s	⠞ t

ตารางข้างต้นแสดงให้เห็นว่า หน่วยอักขระที่เกิดขึ้นใหม่อีก 10 ตัว จากการเติมจุด 3 ภายในเซลล์เดียวกับหน่วยอักขระเบรลล์พื้นฐานได้แก่ '⠅' – '⠞' (= k ถึง t) ทำให้ได้หน่วยอักขระรวม 20 ตัวในระบบเบรลล์อังกฤษ แต่ยังขาดหน่วยอักขระอีก 6 ตัวซึ่งมีการเติมจุดในรูปแบบอื่น ๆ

ข. การเติมจุด 3-6 ภายในเซลล์

หน่วยอักขระเบรลล์อีก 5 ตัวต่อมาเกิดจากการนำหน่วยอักขระพื้นฐาน 5 ตัวแรกคือ ‘⠁’- ‘⠑’ (= a - e) มาเติมจุด 3-6 ภายในเซลล์เดียวกับอักขระฐาน ดังแสดงในตารางต่อไปนี้

ตารางที่ 3.3 หน่วยอักขระเบรลล์อังกฤษดัดแปลงที่ได้จากการเติม‘⠒’ ภายในเซลล์

หน่วยอักขระเบรลล์					
หน่วยอักขระพื้นฐาน	⠁ a	⠃ b	⠉ c	⠙ d	⠑ e
หน่วยอักขระที่สร้างใหม่	⠥ u	⠧ v	⠭ x	⠽ y	⠵ z

ค. การเติมจุด 6 ภายในเซลล์

อย่างไรก็ตาม ยังเหลือหน่วยอักขระอีก 1 ตัวที่ไม่เป็นไปตามรูปแบบที่แสดงไว้ข้างต้นคือ “w” หน่วยอักขระนี้เกิดจากการเติมจุด 6 ภายในเซลล์ “j” ทำให้ได้รูปเป็น “w”

‘⠚’ (j) ⟹ ‘⠺’ (w)

เหตุที่เป็นเช่นนี้เนื่องจากอักษรเบรลล์เกิดขึ้นในประเทศฝรั่งเศสที่มีการใช้ “w” กันน้อย ทำให้หน่วยอักขระนี้เกิดขึ้นในภายหลัง และมีรูปแบบแตกต่างจากอักขระอื่น ๆ ที่กล่าวมาข้างต้น

3.1.2 รูปแบบของตัวเลขเบรลล์

นอกจากการนำหน่วยอักขระพื้นฐานมาใช้สร้างตัวอักษรเพิ่มเติมแล้ว ยังมีการนำหน่วยอักขระเหล่านี้มาแทนตัวเลขอารบิกจาก 1 ถึง 0 โดยการเติมเครื่องหมายนำเลข ‘⠼’ (= #) ข้างหน้าหน่วยอักขระพื้นฐาน ซึ่งแสดงได้ดังนี้

ตารางที่ 3.4 ตัวเลขเบรลล์ 1 - 0

หน่วยอักขระเบรลล์										
หน่วยอักขระพื้นฐาน	⠁ a	⠃ b	⠉ c	⠙ d	⠑ e	⠋ f	⠛ g	⠓ h	⠊ i	⠚ j
หน่วยอักขระที่สร้างใหม่	⠼⠁ 1	⠼⠃ 2	⠼⠉ 3	⠼⠙ 4	⠼⠑ 5	⠼⠋ 6	⠼⠛ 7	⠼⠓ 8	⠼⠊ 9	⠼⠚ 0

นอกจากการนำเบรลอังกฤษตัว a - j มาใช้แทนตัวเลขแล้ว หน่วยอักขระเบรลล์เหล่านี้ยังสามารถนำไปปรับจุดให้ต่ำลงเพื่อใช้แทนเครื่องหมายต่าง ๆ ได้

3.1.3 รูปแบบของเครื่องหมายในเบรลล์อังกฤษ

ในเบรลล์อังกฤษมีการสร้างเครื่องหมายขึ้นเพิ่มเติม โดยการเลื่อนตำแหน่งจุดของอักขระพื้นฐาน ให้ต่ำลงอีกหนึ่งตำแหน่งเท่า ๆ กันทั้งตัวอักษร นั่นคือ ปรับตำแหน่งจากจุด 1 เป็นจุด 2, จากจุด 2 เป็นจุด 3, จากจุด 4 เป็นจุด 5 และจากจุด 5 เป็นจุด 6 เป็นผลให้มีเครื่องหมายที่เกิดจากอักขระพื้นฐาน ดังนี้

ตารางที่ 3.5 เครื่องหมายในเบรลล์อังกฤษ

หน่วยอักขระเบรลล์										
หน่วยอักขระพื้นฐาน	a	b	c	d	e	f	g	h	i	j
หน่วยอักขระที่สร้างใหม่	,	;	:	.	-	!	( )	“ ?	-	”

จากตารางจะพบว่า ไม่มีเครื่องหมายใดที่สอดคล้องกับตัว ‘⠑’ (= e) และ‘⠊’ (= i) เนื่องจากมีการนำหน่วยอักขระ 2 ตัวนี้ไปใช้เป็นตัวย่อเบรลล์ต่ำย่อส่วนของคำแล้ว (ดูตารางที่ 2.14 ประกอบ)

3.1.4 อักขรวิธีในเบรลล์อังกฤษ

ดังที่ได้กล่าวไว้ในส่วนของการทบทวนวรรณกรรมในบทที่ 2 ว่า เบรลล์อังกฤษที่ใช้กันทั่วโลกประกอบด้วยเบรลล์ 2 ประเภท คือเบรลล์อังกฤษรูปเต็มและเบรลล์อังกฤษรูปย่อ โดยที่ตัวอักษรประเภทต่างๆที่ได้อภิปรายมาแล้วทั้งหมด สามารถปรากฏได้ในเบรลล์ทั้งสองประเภทนี้

สำหรับเบรลล์อังกฤษรูปเต็มและรูปย่อมีวิธีการผสมอักษรที่แตกต่างกัน ซึ่งผู้วิจัยได้นำเสนอมาแล้วในบทที่ 2 หัวข้อที่ 2.1.1.1 และ 2.1.1.2 จึงไม่ขออภิปรายประเด็นเหล่านี้ซ้ำอีก ณ ที่นี้

3.2 ระบบหน่วยอักขระเบรลล์ไทย

ด้วยเหตุที่อักษรเบรลล์เกิดจากการนำจุดที่มีการเรียงต่อแบบตายตัวมาใช้แทนตัวอักษรในภาษาต่าง ๆ จึงทำให้ภาษาที่นำอักษรระบบนี้ไปใช้มีอักษรซ้ำรูปกันระหว่างภาษาเป็นจำนวนมาก แต่คงจะไม่ใช่ความบังเอิญ ที่เบรลล์สองภาษาจะใช้อักษรรูปเดียวกันเพื่อแทนเสียงที่ใกล้เคียงกันระหว่างภาษา ดังที่สระในเบรลล์ไทยมีรูปและเสียงใกล้เคียงกับสระในเบรลล์ญี่ปุ่นและพยัญชนะส่วนหนึ่งในเบรลล์ไทยมีรูปและคุณลักษณะทางเสียงใกล้เคียงกับพยัญชนะในเบรลล์อังกฤษ

เหตุที่เป็นเช่นนั้น เพราะผู้ที่กำหนดอักษรเบรลล์ไทยคือ มิสเจเนวีฟ คอลฟิลด์ (Genevieve Caulfield) สุภาพสตรีสายตาพิการชาวอเมริกัน เคยทำงานด้านการศึกษาที่ประเทศญี่ปุ่นเป็นระยะเวลาหนึ่งก่อนเข้ามาบุกเบิกงานการศึกษาสำหรับคนตาบอดในประเทศไทย (Caulfield, G. 1960: 81) ผู้วิจัยจึงสันนิษฐานว่า มิสเจเนวีฟ คอลฟิลด์ ได้นำสระในเบรลล์ญี่ปุ่นที่มีเสียงใกล้เคียงกับภาษาไทยมาใช้แทนสระในเบรลล์ไทยด้วย นอกจากนี้ สมทรง พันธ์สุวรรณ (2538: 19) ได้ตั้งข้อสังเกตว่าเบรลล์ไทย 15 รูป เกิดจากการนำพยัญชนะเบรลล์อังกฤษที่มีเสียงตรงกับพยัญชนะไทยมาใช้แทนอักษรเบรลล์ไทย (ซึ่งจะได้กล่าวต่อไปในหัวข้อพยัญชนะในเบรลล์ไทย) เนื่องจากอักษรเบรลล์ในภาษาอื่นโดยเฉพาะภาษาอังกฤษได้มีผู้วิเคราะห์ไว้แล้วอย่างเป็นระบบ ดังที่ได้นำเสนอมาแล้วในหัวข้อที่ 3.1 งานวิจัยนี้จึงมุ่งที่จะวิเคราะห์ระบบอักษรเบรลล์ไทย โดยผู้วิจัยจะนำเสนอกรอบการวิเคราะห์ และผลที่ได้จากการวิจัย ดังนี้

3.2.1 กรอบการวิเคราะห์ระบบหน่วยอักขระเบรลล์ไทย

ในการศึกษาครั้งนี้ผู้วิจัยต้องการหาคำตอบว่า มีการนำหน่วยอักขระในภาษาอังกฤษและภาษาอื่นมาใช้ในเบรลล์ไทยหรือไม่ และในเบรลล์ไทยมีรูปแบบในการสร้างหน่วยอักขระอย่างไร เพื่อตอบคำถามเหล่านี้ ผู้วิจัยได้นำหน่วยอักขระเบรลล์ไทยแต่ละตัวไปเปรียบเทียบกับเบรลล์อังกฤษและเบรลล์ญี่ปุ่น โดยมีขั้นตอนในการวิเคราะห์ ดังนี้

(1) หารูปในอักขระเบรลล์ที่ซ้ำรูปกันระหว่างภาษาไทยกับภาษาอังกฤษหรือภาษาญี่ปุ่น

ในขั้นแรก ผู้วิจัยได้นำอักษรเบรลล์ใน 3 ภาษาคือ ไทย ญี่ปุ่น และอังกฤษ มาเปรียบเทียบกันเพื่อหารูปที่เหมือนกันระหว่างภาษาไทยกับภาษาอังกฤษหรือภาษาญี่ปุ่น ทั้งนี้เพื่อยืนยันข้อสันนิษฐานว่า มิสเจเนวีฟ คอลฟิลด์ ได้รับอิทธิพลจากเบรลล์อังกฤษและญี่ปุ่นในการกำหนดตัวอักษรเบรลล์ไทย

ต่อมา นำหน่วยอักขระเบรลล์ที่พบว่ามีการใช้รูปซ้ำกันระหว่างภาษามาเปรียบเทียบคุณลักษณะทางเสียง เพื่อยืนยันว่าหน่วยอักขระเบรลล์ไทยที่ซ้ำรูปกับเบรลล์อังกฤษหรือญี่ปุ่นนั้น

เป็นเพราะยืมมาใช้ด้วยเหตุที่มีคุณลักษณะทางเสียงคล้ายกัน เช่น ‘ ’ (= k) ในเบรลล์อังกฤษมีรูปเหมือน ‘ ’ (= ข) ในเบรลล์ไทย และใช้แทนเสียง /kh/ ในทั้งสองภาษา หน่วยอักขระเบรลล์ไทยที่ได้ในขั้นตอนนี้จะเรียกว่าเบรลล์ไทยพื้นฐานซึ่งจะนำไปใช้ในการวิเคราะห์ความสัมพันธ์กับหน่วยอักขระเบรลล์ไทยอื่น ๆ ต่อไป

(2) วิเคราะห์หาระบบหน่วยอักขระเบรลล์ไทย

ในขั้นนี้ ผู้วิจัยนำเบรลล์ไทยแต่ละรูปมาเปรียบเทียบกันเพื่อหาระบบอักษรเบรลล์ไทยว่าหน่วยอักขระเบรลล์ไทยตัวอื่น ๆ มีความสัมพันธ์กับหน่วยอักขระพื้นฐานที่พบในข้อ (1) หรือไม่อย่างไร โดยพิจารณาทั้งส่วนที่เป็นสระ พยัญชนะ และวรรณยุกต์

จากกรอบการวิเคราะห์ที่นำเสนอข้างต้น ทำให้ผู้วิจัยพบว่า มีการนำหน่วยอักขระเบรลล์ส่วนหนึ่งจากภาษาญี่ปุ่นและภาษาอังกฤษมาใช้เป็นหน่วยอักขระในเบรลล์ไทย ส่วนหน่วยอักขระในเบรลล์ไทยอื่น ๆ ที่เหลืออยู่ เกิดจากการดัดแปลงหน่วยอักขระในเบรลล์ไทย และการสร้างหน่วยอักขระขึ้นเพิ่มเติม ซึ่งจะขออภิปรายรายละเอียดดังนี้

3.2.2 รูปแบบของหน่วยอักขระเบรลล์ไทย

หน่วยอักขระเบรลล์ไทยพื้นฐานที่พบว่ามาจากสองส่วน ส่วนแรกเป็นเบรลล์ไทยที่แทนรูปสระ 5 รูปที่มีเสียงคล้ายกับสระของเบรลล์ภาษาญี่ปุ่น ส่วนที่สองเป็นเบรลล์ไทยที่แทนพยัญชนะไทย15 รูปที่มีเสียงคล้ายกับเบรลล์อังกฤษ ดังมีรายละเอียดดังนี้

3.2.2.1 หน่วยอักขระเบรลล์ไทยที่มีรูปซ้ำกับหน่วยอักขระเบรลล์อังกฤษหรือญี่ปุ่น

จากการเปรียบเทียบอักขระที่ปรากฏในเบรลล์ไทย[5] เบรลล์อังกฤษ และเบรลล์ญี่ปุ่น ผู้วิจัยพบว่า อักขระในเบรลล์ไทยมีรูปซ้ำกับอักขระในเบรลล์ญี่ปุ่นหรือเบรลล์อังกฤษ ดังนี้

---

[5] ตารางที่ 3.6 แสดงเฉพาะเบรลล์ไทยที่แทนด้วยเบรลล์หนึ่งเซลล์เท่านั้น ส่วนเบรลล์ไทยที่แทนด้วยเบรลล์มากกว่าหนึ่งเซลล์จะกล่าวถึงในส่วนของการสร้างหน่วยอักขระเพิ่มเติมจากเบรลล์พื้นฐานต่อไป

ตารางที่ 3.6 อักขระเบรลล์ที่มีรูปซ้ำกันระหว่างภาษาไทยและญี่ปุ่น หรือไทยและอังกฤษ

เบรลล์	ภาษาไทย		ภาษาญี่ปุ่น		ภาษาอังกฤษ	
	ตัวเขียน	เสียง	ตัวเขียน	เสียง	ตัวเขียน	เสียง
⠛	ก	/k/	れ	/re/	g	/g/
⠅	ข	/kh/	な	/na/	k	/k/
⠇	ค	/kh/	は	/ha/	u	/u:/
⠻	ง	/ŋ /	せ	/se/		
⠩	จ	/tɕ/	ろ	/ro/	j	/dʒ/
⠍	ฉ	/tʃʰ/	や	/ya/		
⠝	ช	/tʃʰ/	ゆ	/yu/		
⠎	ซ	/s/	ほ	/ho/		
⠙	ด	/d/	る	/ru/	d	/d/
⠞	ต	/t/	し	/shi/		
⠹	ถ	/tʰ /	と	/to/	t	/t/
⠾	ท	/tʰ /	も	/mo/		
⠝	น	/n/	つ	/tsu/	n	/n/
⠃	บ	/b/	ひ	/hi/	v	/v/
⠏	ป	/p/	へ	/he/		
⠙	ผ	/pʰ /	ね	/ne/	p	/pʰ/
⠋	ฝ	/f/	ふ	/fu/	x	/z/
⠗	พ	/pʰ /	す	/su/		
⠋	ฟ	/f/	け	/ke/		
⠍	ม	/m/	ぬ	/nu/	m	/m/
⠽	ย	/j/	む	/mu/	y	/j/
⠗	ร	/r/	ち	/chi/	r	/r/
⠇	ล	/l/	に	/ni/	l	/l/
⠺	ว	/w/	そ	/so/	w	/w/
⠎	ส	/s/	の	/no/	s	/s/
⠓	ห	/h/	り	/ri/	h	/h/
⠕	อ	/ɔ /	た	/ta/	o	/ɔ /
⠓	ฮ	/h/	め	/me/		
⠁	ะ	/a/	あ	/a/	a	/a/
⠎	ั	/a/	よ	/yo/		
⠧	ัว	/ua/	ら	/ra/	e	/e/
⠜	า	/a:/	か	/ka/		
⠎	ำ	/am/	ま	/ma/	z	/z/

เบรลล์	ภาษาไทย		ภาษาญี่ปุ่น		ภาษาอังกฤษ	
	ตัวเขียน	เสียง	ตัวเขียน	เสียง	ตัวเขียน	เสียง
⠁	ิ	/i/	い	/i/	b	/b/
⠂	ี	/i:/	わ	/wi/	;	
⠃	เ	/e:/	え	/e/	f	/f/
⠄	เา	/aw/	ゑ	/we/	!	
⠅	เื	/ɯa/	て	/te/	q	/k^h/
⠆	เีย	/ia/	み	/mi/		
⠇	เิ หรือ เอ	/ɤ:/	く	/ku/		
⠈	แ	/ɛ:/	き	/ki/		
⠉	โ	/o:/	お	/o/	i	/i/
⠊	ไ	/aj/	さ	/sa/		
⠋	่		を	/wo/		
⠌	้				.	
⠍	๊				(หรือ)	
⠎	๋				“ หรือ ?	
⠏	็		わ	/wa/	‘	
⠐	์				”	

จากตารางข้างต้น มีหน่วยอักขระเบรลล์ที่ซ้ำรูปกันระหว่างภาษาอยู่ 2 ประเภท ประเภทแรกเป็นหน่วยอักขระที่ซ้ำรูปกันแต่มีเสียงแตกต่างกัน เช่น ‘⠿’ ใช้แทน “ฝ” ในภาษาไทยและแทน “x” ในภาษาอังกฤษ อีกประเภทหนึ่งเป็นหน่วยอักขระเบรลล์ที่ซ้ำรูปกันและมีเสียงใกล้เคียงกันเช่น ‘⠹’ใช้แทน “ถ” ในภาษาไทย และแทน “t” ในภาษาอังกฤษ หน่วยอักขระที่ซ้ำรูปในประเภทที่สองนี้เองที่น่าจะเป็นหน่วยอักขระเบรลล์ไทยพื้นฐานที่สร้างขึ้นโดยการเทียบจากหน่วยอักขระภาษาอังกฤษหรือญี่ปุ่น และเป็นหน่วยอักขระพื้นฐานในการสร้างหน่วยอักขระตัวอื่น ๆ ในเบรลล์ไทยขึ้นเพิ่มเติม

3.2.2.2 ความสัมพันธ์ระหว่างหน่วยอักขระเบรลล์ไทยและเบรลล์อังกฤษหรือเบรลล์ญี่ปุ่นที่มีรูปซ้ำกัน

หากพิจารณาหน่วยอักขระเบรลล์ไทยที่ซ้ำรูปกับเบรลล์อังกฤษหรือเบรลล์ญี่ปุ่น (ตามตารางที่ 3.6) จะพบว่าสระในเบรลล์ไทยมีเสียงสอดคล้องกับสระในเบรลล์ญี่ปุ่น 5 รูป และพยัญชนะในเบรลล์ไทยมีเสียงเหมือนกับพยัญชนะเบรลล์อังกฤษ 15 รูป โดยมีหน่วยอักขระซ้ำรูปที่อยู่ในกลุ่มเบรลล์อังกฤษพื้นฐาน 10 ตัวแรก ดังนี้

ตารางที่ 3.7 อักขระในเบรลล์ญี่ปุ่นและเบรลล์ไทยที่อยู่ในกลุ่มเบรลล์อังกฤษพื้นฐาน

เบรลล์	อังกฤษ	ญี่ปุ่น	ไทย
⠁	a	あ /a/	ะ
⠃	b	い /i/	ิ
⠉	c	う /u/	ุ
⠙	d		ด
⠑	e		
⠋	f	え /e/	เ
⠛	g		ก
⠓	h		ห
⠊	i	お /0/	โ
⠚	j		จ

จากตารางข้างต้นจะเห็นได้ว่า มีการนำเบรลล์พื้นฐานเหล่านี้มาใช้แทนหน่วยอักขระชุดแรกในเบรลล์ไทย โดยที่ได้นำรูปอักษรที่ใช้แทนเสียงสระในภาษาญี่ปุ่นมาใช้แทนสระในเบรลล์ไทย จากตารางจะสังเกตได้ว่าไม่มีการนำอักขระ '⠑' (=e) มาใช้เป็นเบรลล์พื้นฐานในเบรลล์ญี่ปุ่นและเบรลล์ไทยเลย

สำหรับสระในเบรลล์ไทยที่มีรูปและเสียงใกล้เคียงกับเบรลล์ญี่ปุ่น สามารถแสดงได้ดังนี้

ตารางที่ 3.8 รูปของสระในเบรลล์ไทยและเบรลล์ญี่ปุ่นที่มีเสียงใกล้เคียงกัน

ภาษาญี่ปุ่น			ภาษาไทย		
เบรลล์	อักษรปกติ	ใช้แทนเสียง	เบรลล์	อักษรปกติ	ใช้แทนเสียง
⠁	あ	/a/	⠁	ะ	/a/
⠃	い	/i/	⠃	ิ	/i/
⠉	う	/u/	⠉	ุ	/u/
⠋	え	/e/	⠋	เ	/e:/
⠊	お	/o/	⠊	โ	/o:/

ตารางข้างต้นแสดงให้เห็นว่า สระที่ซ้ำรูปกัน 5 รูประหว่างเบรลล์ไทยและเบรลล์ญี่ปุ่นใช้แทนเสียงที่สอดคล้องกัน โดยที่สระ 3 รูปคือ ‘⠥’ ‘⠕’ และ ‘⠽’ ล้วนเป็นสระเสียงสั้นที่ใช้แทนเสียงที่เหมือนกัน ในขณะที่อีก 2 รูปคือ ‘⠷’ และ ‘⠾’ เป็นสระใช้แทนเสียงที่เหมือนกันแต่มีความยาวแตกต่างกันระหว่างภาษา คือในเบรลล์ไทยใช้แทนสระยาวแต่ในเบรลล์ญี่ปุ่นใช้แทนสระสั้น เหตุที่ไม่ใช้รูป ‘⠷’ และ ‘⠾’ แทนสระ เ-ะ และ โ-ะ ในเบรลล์ไทย สันนิษฐานว่าเพราะสระดังกล่าวในภาษาไทยใช้สองตัวอักษรจึงไม่เหมาะที่จะนำมาใช้เป็นสระพื้นฐานในเบรลล์ไทย

นอกจากนี้ เมื่อเปรียบเทียบรูปที่ซ้ำกันระหว่างเบรลล์ไทยและเบรลล์อังกฤษ ผู้วิจัยพบว่า มีพยัญชนะ 15 ตัวที่ใช้แทนเสียงที่ใกล้เคียงกันระหว่างภาษา ซึ่งสอดคล้องกับสิ่งที่ สมทรง พันธ์สุวรรณ (2538) ได้ตั้งข้อสังเกตไว้ดังนี้

ตารางที่ 3.9 พยัญชนะเบรลล์ไทยที่มีรูปและเสียงตรงกับเบรลล์อังกฤษ

เบรลล์	⠅	⠞	⠏	⠎	⠓	⠛	⠚	⠙	⠕	⠝	⠍	⠽	⠗	⠇	⠺
อังกฤษ	k	t	p	s	h	g	J	d	o	n	m	y	r	l	w
ไทย	ข	ถ	ผ	ส	ห	ก	จ	ด	อ	น	ม	ย	ร	ล	ว

แม้ว่าพยัญชนะทั้ง 15 รูปในเบรลล์ไทยจะมีรูปซ้ำกับเบรลล์อังกฤษและมีเสียงเหมือนกันก็ตาม แต่จากการเปรียบเทียบหน่วยอักขระเบรลล์ไทยและเบรลล์อังกฤษตามตารางที่ 3.6 ผู้วิจัยได้พบเพิ่มเติมว่ามีพยัญชนะที่ใช้แทนเสียงที่เหมือนกันระหว่างภาษาอยู่อีก 2 คู่ คือ ‘⠃’ (=b) และ ‘⠋’ (=f) ที่มีเสียงเหมือน “บ” และ “ฟ”

เนื่องจากหน่วยอักขระเบรลล์ทั้งคู่มีรูปตรงกับ ‘⠃’ (=◌ิ) และ ‘⠋’ (=เ-) หรือ“เ◌” และ“ะ” ในภาษาญี่ปุ่น ทำให้ไม่สามารถนำเบรลล์อังกฤษสองตัวนี้มาใช้แทนพยัญชนะได้ จึงมีการดัดแปลงโดยการเติมจุด 6 ให้กับอักษร f ได้เป็น ‘⠯’ เพื่อแทนตัว “ฟ” ในเบรลล์ไทย แต่สำหรับอักษร bไม่สามารถเติมจุด 6 โดยลำพังได้เพราะไปซ้ำกับ ‘⠂’ (= แ-) จึงต้องเติม จุด 3-6 ได้เป็น ‘⠳’ เพื่อใช้แทน “บ” ในเบรลล์ไทย อักขระเบรลล์พื้นฐานสำหรับพยัญชนะไทยจึงมีรูปตรงกับเบรลล์อังกฤษที่มีเสียงตรงกัน 15 รูป และมีรูปที่ขยายเพิ่มเติมสำหรับพยัญชนะที่มีเสียงตรงกันระหว่างภาษาอีก 2 รูป

3.2.2.3 อักขรวิธีของเบรลล์ไทย

ในเบรลล์อังกฤษ หน่วยอักขระแต่ละตัวมีรูปแบบการเรียงต่อของจุดเบรลล์ที่สัมพันธ์กันระหว่างหน่วยอักขระอย่างเป็นระบบ โดยที่มีหน่วยอักขระพื้นฐานได้แก่ '⠁' – '⠑' (=a-j) ส่วนหน่วยอักขระที่เหลือ เกิดจากการเติม '⠂'(จุด 3) '⠄' (จุด 3-6) และ '⠈' (จุด 6) เข้าไปในเซลล์เดียวกับหน่วยอักขระพื้นฐาน นอกจากนี้มีการแยกความแตกต่างระหว่างอักขระตัวพิมพ์เล็กและตัวพิมพ์ใหญ่ ด้วยการเติมเซลล์ '⠈' (จุด 6) หน้าอักขระเบรลล์เพื่อแสดงว่าเป็นตัวพิมพ์ใหญ่

สำหรับการสร้างตัวอักษรในเบรลล์ไทยขึ้นเพิ่มเติมนั้น มีการนำสระเบรลล์ไทยที่มีรูปและเสียงสอดคล้องกับภาษาญี่ปุ่นตามตารางที่ 3.8 มาใช้เป็นสระพื้นฐานเพื่อสร้างสระตัวอื่น ในทำนองเดียวกันมีการนำพยัญชนะในเบรลล์ไทยและเบรลล์อังกฤษที่มีรูปและเสียงใกล้เคียงกันตามตารางที่ 3.9 มาใช้เป็นพยัญชนะพื้นฐานเพื่อสร้างพยัญชนะอื่น ๆ ดังนี้

3.2.2.3.1 สระและอักขระอื่น ๆ ในเบรลล์ไทย

ในภาษาไทย มีเสียงสระเดี่ยว 18 เสียง ได้แก่ อะ อา อิ อี อึ อื อุ อู เอะ เอ แอะ แอ โอะ โอ เอาะ ออ เออะ เออ เสียงสระประสม 6 เสียง ได้แก่ เอียะ เอีย เอือะ เอือ อัวะ อัว และสระเกินที่เป็นเสียงสระประสมกับเสียงพยัญชนะอีก 8 เสียง ได้แก่ ฤ ฤๅ ฦ ฦๅ อำ ใอ ไอ เอา ในขณะที่รูปตัวอักษรสระมีเพียง 15 รูป (ไม่รวมสระเกิน ฤ ฤๅ ฦ ฦๅ) (วิโรจน์ อรุณมานะกุล. 2551.) โดยที่เสียงสระบางเสียงต้องใช้มากกว่าหนึ่งรูปอักษรมาผสมกัน เช่น เอะ ใช้ทั้งรูป เ และ ะ เป็นต้น ดังนั้น การใช้อักษรเบรลล์แทนสระในภาษาไทยจึงเกี่ยวข้องกับทั้งรูปและเสียงของสระเหล่านี้

ดังที่ได้กล่าวมาแล้วว่าสระพื้นฐานในเบรลล์ไทยมีอยู่ 5 รูป (สำหรับ 5 เสียงที่คล้ายคลึงทางเสียงกับสระภาษาญี่ปุ่น และเสียงสระ /ɔ/ ใช้รูปเดียวกับพยัญชนะ "อ") ในการสร้างอักขระขึ้นเพิ่มเติมสำหรับรูปและเสียงสระที่เหลือทำได้โดยการปรับเปลี่ยนจุดเบรลล์ในสระพื้นฐาน การสร้างอักขระเฉพาะขึ้นใหม่และการเติมเซลล์ที่มีจุดเบรลล์รูปแบบเฉพาะ โดยที่แต่ละรูปแบบมีรายละเอียดดังนี้

1) การปรับเปลี่ยนจุดเบรลล์ภายในเซลล์

ในการสร้างสระขึ้นเพิ่มเติม มีการนำสระพื้นฐานในเบรลล์ไทย 5 รูป มาปรับเปลี่ยนจุดภายในเซลล์ ซึ่งมีอยู่ 4 รูปแบบ ดังนี้

ก. การเติมจุด 6 ภายในเซลล์

จากการศึกษาเบรลล์ไทย ผู้วิจัยพบว่ามีสระที่เกิดจากการเติมจุด 6 เข้าไปในเซลล์เดียวกับสระพื้นฐาน ดังนี้

ตารางที่ 3.10 การสร้างสระขึ้นเพิ่มเติมโดยการเติม ‘⠈’ (จุด 6)

สระฐาน		สระที่ได้ใหม่	
เบรลล์	อักษรปกติ	เบรลล์	อักษรปกติ
⠐	ะ	⠘	า
⠒	ิ	⠸	แ
⠖	ุ	⠶	เ◌อ, เ◌ิ
⠦	เ	-	-
⠴	โ	⠾	◌ึ

ตารางที่ 3.8 แสดงให้เห็นว่า แม้จะมีสระพื้นฐาน 5 รูปก็ตาม แต่มีสระเพียง 4 รูปเท่านั้นที่เกิดขึ้นใหม่จากสระพื้นฐานเหล่านี้ ยกเว้น ‘⠦’ (=เ) ที่ไม่สามารถเติม จุด 6 ให้เป็นสระอื่นได้ เนื่องจาก ‘‘⠦’ ถูกใช้เป็นพยัญชนะ “ฟ” ในเบรลล์ไทยแล้ว

อย่างไรก็ตามการเติม จุด 6 นี้ไม่ได้ถูกใช้อย่างเป็นระบบ แม้ว่าการเติมจุด 6 จะเปลี่ยนสระสั้นเป็นยาวคือจาก “ะ” เป็น “า” แต่ยังมีการเปลี่ยนสระสั้นเป็นยาวจาก “ ิ” เป็น “ ี” ที่ใช้วิธีการอื่นคือการเลื่อนจุดให้ต่ำลง โดยสรุป การเติมจุด 6 ภายในเซลรูปเบรลล์สระพื้นฐานทำให้ได้รูปสำหรับแทนเสียงสระเพิ่มอีก 4 เสียง

ข. การเลื่อนตำแหน่งจุดเบรลล์ให้ต่ำลง

การรักษารูปของอักษรเบรลล์ที่ใช้เป็นฐานโดยการเลื่อนจุดเบรลล์ให้อยู่ในตำแหน่งที่ต่ำลง เป็นกลวิธีหนึ่งที่ทำให้เกิดอักษรใหม่ในเบรลล์ไทย ดังปรากฏในตารางต่อไปนี้

ตารางที่ 3.11 การสร้างอักษรโดยการเลื่อนตำแหน่งจุดเบรลล์ให้ต่ำลง

สระฐาน		สระที่ได้ใหม่	
เบรลล์	อักษรปกติ	เบรลล์	อักษรปกติ
⠐	ะ	⠒	ๆ
⠒	ิ	⠔	ี
⠖	ุ	⠦	ู
⠦	เ	⠤	เ◌า
⠴	โ	⠲	◌่

เมื่อพิจารณาอักขระเบรลล์ที่เกิดขึ้นใหม่ในตาราง จะพบว่ามีสระเบรลล์ 3 รูป ที่เกิดจากการเลื่อนตำแหน่งจุดเบรลล์ลงได้แก่ ‘⠽’ (=ี) ‘⠾’ (=ู) และ ‘⠿’ (=เา) นอกจากนี้อักขระอีก 2 ตัวคือ ‘⠷’ (=ๆ) และ ‘⠵’ (=่) ก็ได้มาจากการเลื่อนจุดเบรลล์ให้ต่ำลงเช่นกัน เนื่องจากการขยายจาก ‘⠇’ (=ะ) เป็น ‘⠧’ (=า) ใช้วิธีการเติม จุด 6 ไปแล้ว การเลื่อน ‘⠇’ ให้ต่ำลงจึงใช้แทน “ๆ” ในขณะที่ “โ” ไม่จำเป็นต้องขยายเป็นสระอื่นที่ใกล้กัน การเลื่อนตำแหน่งจุดเบรลล์ของ “โ” จึงถูกนำไปใช้แทน “่” ในภายหลัง วิธีการนี้ทำให้ได้รูปเบรลล์สำหรับแทนเสียงสระเพิ่มอีก 3 รูป

ค. การเติมจุดเบรลล์ในเซลล์เดียวกับสระ “ะ”

นอกจากการสร้างสระโดยการเติมจุด 6 ในเซลล์เดียวกับสระพื้นฐานแล้ว ยังมีการเติมจุด 3-5-6, จุด 5-6 และจุด 5 ในเซลล์เดียวกับ ‘⠇’ (=ะ) ทำให้เกิดสระขึ้นอีก 3 รูป 3 เสียง โดยที่สระทั้งสามรูปนี้ต่างก็มีเสียง /a/ เป็นส่วนประกอบภายในสระประสม ดังนี้

ตารางที่ 3.12 การสร้างสระจาก ⠇ (= ะ)

สระฐาน	จุดเบรลล์ที่เติมเข้ามา	ทำให้ได้รูปเป็น		
		เบรลล์	ไทยปกติ	สัทอักษร
⠇ (/a/)	⠍	⠷	ำ	/am/
	⠊	⠔	ไ	/aj/
	⠥	⠳	ัว	/u:a/
	⠫	⠵	เือ	/ɯa/
	⠽	⠽	เีย	/ia/

สระที่ปรากฏในตารางล้วนเป็นสระที่มีเสียง /a/ ประกอบทั้งสิ้น โดยที่ ‘⠷’ (=ำ) และ ‘⠔’ (=ไ) เป็นสระที่มีรูปแทนในลักษณะหนึ่งเสียงต่อหนึ่งตัวอักษรเช่นเดียวกับอักษรปกติ สระที่เหลือเช่น ‘⠳’ (=ัว) เป็นสระประสม ซึ่งในเบรลล์ไทยจะใช้เบรลล์เพียงเซลล์เดียวเพื่อแสดงรูปของสระตัวนี้ ในขณะที่ภาษาไทยจะใช้อักขระถึง 2 ตัวคือ" ั" และ "ว"

ง. การปรับเปลี่ยนจุดเบรลล์แบบอื่น ๆ

ในเบรลล์ไทย มีอักขระอยู่ส่วนหนึ่งที่ไม่มีความสัมพันธ์ทางรูปและเสียงกับเบรลล์ไทยพื้นฐานโดยตรง และไม่ได้ใช้วิธีการขยายอักษรตามวิธีปรับจุดเบรลล์แบบที่กล่าวมา หรือไม่มีการใช้อย่างเป็นระบบเหมือนอักษรหลาย ๆ ตัว ได้แก่ ‘⠵’ (= ี) แม้จะเป็นรูปสระเสียงยาวของ ‘⠽’ (= ึ) ที่มีลักษณะจุดคล้ายกัน แต่การเติมจุด 4 หรือลบจุด 4 ก็ไม่ใช่วิธีที่พบในการสร้างอักขระ

ในเบรลล์ไทย ส่วนอักขระแทนรูปวรรณยุกต์ 4 รูปในภาษาไทย พบว่า มีลักษณะของการใช้ตำแหน่ง จุด 2-3-5-6 เป็นหลัก โดยหากวรรณยุกต์เอกเกิดจากการเลื่อนจุดลงของ ‘⠕’ (=โ) เป็น ‘⠳’ (=่) วรรณยุกต์โท ตรี และจัตวา ก็เปลี่ยนตำแหน่งจุดของเบรลล์อังกฤษพื้นฐานที่เหลือ ได้เป็น ‘⠽’ (=้) ‘⠷’ (=๊) และ ‘⠟’ (=๋) ตามลำดับ เหตุที่ไม่สามารถใช้ ‘⠵’ (จุด 2-3-5) ได้เพราะใช้สำหรับสระ เ-า ไปแล้ว และเหลือ ‘⠻’ (จุด 3-5-6) ที่ว่างอยู่จึงใช้กับอักษรการันต์

สำหรับสระอื่น ๆ ก็กำหนดโดยใช้ตัวอักขระที่ว่างอยู่เป็นหลัก แต่มีข้อสังเกตว่า สระในเบรลล์ไทยที่แสดงด้วยเบรลล์เพียงเซลล์เดียวมักแสดงด้วยจุดเบรลล์จำนวน 2 หรือ 3 จุด แต่มีสระบางรูปที่แสดงด้วยจุดเบรลล์ 4 หรือ 5 จุด ได้แก่ ‘⠩’ (=ำ) ‘⠹’ (=เีย) และ ‘⠿’ (=เือ) ซึ่งครอบคลุมเสียงสระเพิ่มอีก 3 เสียง คือ /am/, /ia/ และ /ɯa/

ตารางที่ 3.13 อักขระที่เกิดจากการปรับเปลี่ยนจุดเบรลล์แบบอื่น ๆ

อักขระฐาน		**อักขระที่ได้ใหม่**	
เบรลล์	**อักษรปกติ**	**เบรลล์**	**อักษรปกติ**
⠕	e	⠫	ึ
⠙	d	⠽	้
⠛	g	⠷	๊
⠓	h	⠮	็
⠚	j	⠪	์

นอกจากนี้ยังมีอักขระที่ได้จากการปรับเปลี่ยนจุดเบรลล์แบบอื่น ๆ สำหรับอักษรไทยที่ยังไม่ได้ถูกกำหนดรหัสอักษรเบรลล์ คือ ‘⠣’ (=ึ) และ ‘⠭’ (= ั)

2) การเพิ่มอักขระเบรลล์ที่มีรูปแบบเฉพาะ

ด้วยเหตุที่อักขระเบรลล์แต่ละเซลล์มีจุดเบรลล์อยู่อย่างจำกัด จึงไม่สามารถใช้แทนอักษรไทยได้ทั้งหมด ฉะนั้น จึงจำเป็นต้องนำเบรลล์ที่มีการปรากฏของจุดที่มีรูปแบบเฉพาะมาวางไว้หลังอักษรที่ใช้เป็นฐาน ทำให้อักษรฐานและอักษรที่เกิดขึ้นใหม่มีลักษณะบางประการร่วมกัน ดังที่ปรากฏกับสระในกลุ่มต่าง ๆ ต่อไปนี้

ก. การเพิ่มอักขระเบรลล์ '⠐' (จุด 1) เพื่อทำให้สระยาวเป็นสระสั้น

โดยทั่วไปในเบรลล์ไทยจะแสดงสระยาวด้วยเบรลล์เพียงเซลล์เดียว โดยที่สระยาวเหล่านี้ส่วนหนึ่งมีคู่ที่เป็นสระสั้น ที่เกิดจากการเติมเซลล์ '⠐' ที่ใช้แทนรูป "ะ" มาไว้ข้างหลังสระเหล่านี้ ทำให้เกิดคู่สระสั้น-ยาว และรูปแทนเสียงสระเพิ่มอีก 8 เสียง ตามที่ปรากฏในตารางต่อไปนี้

ตารางที่ 3.14 การสร้างสระสั้นโดยการเพิ่มอักขระ '⠐' (จุด 1)

อักขระฐาน		สระที่ได้ใหม่	
เบรลล์	**อักษรปกติ**	**เบรลล์**	**อักษรปกติ**
⠣	อ	⠣⠐	เ◌าะ
⠭	เ	⠭⠐	เ◌ะ
⠕	โ	⠕⠐	โ◌ะ
⠡	แ	⠡⠐	แ◌ะ
⠝	เ◌อ	⠝⠐	เ◌อะ
⠧	◌ัว	⠧⠐	◌ัวะ
⠽	เ◌ีย	⠽⠐	เ◌ียะ
⠿	เ◌ื่อ	⠿⠐	เ◌ื่อะ

การเพิ่มอักขระเบรลล์วิธีนี้ ทำให้เกิดความกำกวมในบางกรณีขึ้นได้ เช่น '⠣⠐' นอกจากจะใช้แทน "เ◌าะ" แล้ว ยังใช้แทน "อะ" ตามรูปอักขระเบรลล์แต่ละตัวได้ด้วย ปัญหาความกำกวมในการถ่ายถอดอักษรนี้จะได้กล่าวถึงในบทต่อไป

ข. การเพิ่มอักขระเบรลล์ '⠠' (จุด 2) เพื่อสร้างสระเพิ่มเติม

มีการสร้างสระขึ้นเพิ่มเติมในเบรลล์ไทยด้วยการเพิ่มเซลล์ '⠠' (จุด 2) หลังสระหรือพยัญชนะที่มีเสียงใกล้เคียงกับอักขระฐาน ดังแสดงในตารางต่อไปนี้

ตารางที่ 3.15 การสร้างสระเพิ่มเติมโดยการเพิ่มอักขระ '⠠' (จุด 2)

อักขระฐาน		สระที่ได้ใหม่	
เบรลล์	**อักษรปกติ**	**เบรลล์**	**อักษรปกติ**
⠗	ร	⠗⠠	ฤ
⠇	ล	⠇⠠	ฦ
⠌	ไ	⠌⠠	ใ

ตารางนี้แสดงให้เห็นว่า มีการสร้างสระขึ้น 3 รูปจากอักขระฐาน โดยสระที่เพิ่มขึ้น 2 รูปแรกคือ ‘⠿⠽’(=ฤ) และ‘⠿⠺’(=ฦ) นั้นเกิดจากการใช้พยัญชนะเป็นฐาน เนื่องจากชื่อของพยัญชนะสอดคล้องกับชื่อของสระที่ได้ใหม่ การเพิ่มเซลล์ ‘⠽’ เข้าไปหลังพยัญชนะจะทำให้จดจำได้ง่ายขึ้นอีกทั้ง‘⠽’ยังแสดงถึงความเป็นสระอีกด้วย ส่วนสระที่ได้อีกรูปคือ‘⠿⠽’(=ใ) ซึ่งเกิดจากการเติมเซลล์‘⠽’หลัง ‘⠿’(=ไ) เหตุที่มีการสร้างสระนี้ด้วยวิธีการดังกล่าว เนื่องจากทั้ง “ไ” และ “ใ” ต่างเป็นสระที่ปรากฏหน้าพยัญชนะและใช้แทนเสียงเดียวกันในภาษาไทย

3.2.2.3.2 พยัญชนะในเบรลล์ไทย

สมทรง พันธ์สุวรรณ (2538) พบว่า ในเบรลล์ไทยมีการนำเบรลล์อังกฤษที่มีหน่วยเสียงตรงกับพยัญชนะไทยมาใช้แทนอักษรไทย 15 ตัว นอกจากนี้ผู้วิจัยยังพบเพิ่มเติมว่า พยัญชนะไทยเหล่านี้เป็นพยัญชนะที่พบได้ในภาษาไทยปัจจุบันและปรากฏในอักษรหมู่ต่าง ๆ สิ่งที่น่าสังเกตคือเสียงซึ่งตรงกับอังกฤษ ในกรณีที่อักษรไทยเป็นได้ทั้งอักษรต่ำและอักษรสูง ในเบรลล์ไทยจะเลือกใช้อักษรสูงเป็นฐาน

ตารางที่ 3.16 พยัญชนะเบรลล์ไทยที่มีรูปและเสียงตรงกับเบรลล์อังกฤษจำแนกตามหมู่อักษร

หมู่อักษร	เบรลล์ไทย	ไทยปกติ	เบรลล์อังกฤษ	อังกฤษปกติ
สูง	⠅	ข	⠅	k
	⠞	ถ	⠞	t
	⠏	ผ	⠏	p
	⠎	ส	⠎	s
	⠓	ห	⠓	h
กลาง	⠛	ก	⠛	g
	⠚	จ	⠚	j
	⠙	ด	⠙	d
	⠕	อ	⠕	o
ต่ำ	⠝	น	⠝	n
	⠍	ม	⠍	m
	⠽	ย	⠽	y
	⠗	ร	⠗	r
	⠇	ล	⠇	l
	⠺	ว	⠺	w

นอกจากการนำเบรลล์อังกฤษที่มีเสียงและรูปตรงกับอักษรไทย มาใช้แทนพยัญชนะเบรลล์ไทยโดยตรงแล้ว ยังมีเบรลล์อังกฤษที่เสียงตรงกับพยัญชนะไทยอีกสองตัวแต่ไม่สามารถนำมาใช้ได้โดยตรงเนื่องจากเบรลล์สองตัวนี้ถูกนำมาใช้แทนสระพื้นฐานแล้ว จึงต้องดัดแปลงด้วยการเติมจุด 6 และ จุด 3-6

ตารางที่ 3.17 พยัญชนะเบรลล์ไทยที่เกิดจากการดัดแปลงพยัญชนะเบรลล์อังกฤษ

เบรลล์ไทยที่มาจากเบรลล์ญี่ปุ่น				**เบรลล์ไทยที่ดัดแปลงจากเบรลล์อังกฤษ**			
ภาษาญี่ปุ่น		**ภาษาไทย**		**ภาษาญี่ปุ่น**		**ภาษาไทย**	
เบรลล์	**ปกติ**	**เบรลล์**	**ปกติ**	**เบรลล์**	**ปกติ**	**เบรลล์**	**ปกติ**
⠊	เ	⠊	ิ	⠃	b	⠫	บ
⠋	ะ	⠋	เ	⠋	f	⠯	ฟ

เบรลล์อังกฤษ ‘⠃’ (=b) จะใช้แทนเสียงที่ตรงกับเสียง “บ” ในภาษาไทย และ ‘⠋’ (=f) จะมีเสียงตรงกับเสียง “ฟ” ก็ตาม แต่ไม่สามารถนำเบรลล์อังกฤษดังกล่าวมาใช้ในเบรลล์ไทยได้โดยตรง เนื่องจากพยัญชนะเบรลล์2ตัวนี้มีรูปเหมือนสระในภาษาญี่ปุ่นซึ่งได้นำไปใช้กับสระเบรลล์ไทยแล้ว

เพื่อให้สามารถนำเบรลล์อังกฤษ 2 ตัวข้างต้น มาใช้แทนพยัญชนะไทยที่มีเสียงตรงกันได้ จึงต้องดัดแปลงเบรลล์ ‘⠃’(=b) และ ‘⠋’ (=f) โดยการเติมจุด 6 เข้าไปในเซลล์ทำให้ได้รูปเป็น ‘⠷’ และ ‘⠯’ ตามลำดับ อย่างไรก็ตามการดัดแปลงในลักษณะนี้เกิดขึ้นได้เฉพาะกับ ‘⠋’(=f) เท่านั้น ทำให้ได้ ‘⠯’ ซึ่งใช้แทน “ฟ” แต่ไม่สามารถนำ ‘⠃’ มาเติมจุด 6 เป็น ‘⠷’ เพื่อใช้แทน “บ” ได้ เนื่องจากมีรูปตรงกับสระ “แ” ในเบรลล์ไทย ดังนั้นจำเป็นต้องเติมจุด 3-6 แทน ทำให้ได้รูป ‘⠫’ เพื่อใช้แทน “บ”

จากจำนวนอักษรพยัญชนะไทย 44 ตัว จะเห็นได้ว่าเบรลล์พื้นฐานจำนวน 15 ตัวและส่วนขยายอีก 2 ตัวได้ถูกกำหนดขึ้นตามที่กล่าวมา แต่ยังเหลืออักษรพยัญชนะอีก 27 ตัว ซึ่งต้องกำหนดรหัสตัวอักษรเบรลล์ไทยเพิ่มเติม ผู้วิจัยพบว่ามีการนำพยัญชนะเหล่านี้มาใช้เป็นฐานแล้วเติมจุดเบรลล์ต่างเซลล์หรือภายในเซลล์เพื่อสร้างอักษรใหม่ ดังนี้

1) การเติมจุดเบรลล์ภายในเซลล์

ในการสร้างพยัญชนะใหม่โดยการเติมจุด 6 ให้กับพยัญชนะฐาน ซึ่งพบว่ามีอักษรพื้นฐาน 8 ตัวที่ขยายเพิ่มเติมได้ ดังแสดงในตารางต่อไปนี้

ตารางที่ 3.18 การสร้างพยัญชนะใหม่ที่มีหมู่อักษรต่างจากพยัญชนะฐาน

พยัญชนะฐาน		พยัญชนะที่ได้ใหม่	
เบรลล์	อักษรปกติ	เบรลล์	อักษรปกติ
⠅	ข	⠇	ค
⠋	ถ	⠓	ท
⠎	ส	⠉	ช
⠛	ก	⠝	ง
⠙	ด	⠏	พ
⠍	ม	⠟	ฝ
⠗	ผ	⠕	ป
⠚	ห	⠞	ต
⠡	ฉ	⠩	ช

ตารางข้างต้นแสดงอักขระเบรลล์ที่เกิดขึ้นด้วยการเติมจุด 6 เข้าไปในเซลล์พยัญชนะฐาน ซึ่งพยัญชนะฐานเกือบทั้งหมดอยู่ในกลุ่มอักษรพื้นฐาน 15 ตัวตามตารางที่ 3.9 ยกเว้น ‘⠡’ (=ฉ) สำหรับ‘⠡’(=ฉ) นั้นผู้วิจัยสันนิษฐานว่าเป็นพยัญชนะที่สร้างเพิ่มเติมจากพยัญชนะพื้นฐานด้วยเหตุที่ “ฉ” ใช้แทนเสียง /c^h/ ซึ่งไม่สามารถใช้อักษรอังกฤษตัวเดียวมาแทนเสียงนี้ได้ จึงต้องกำหนดรหัสเบรลล์ชุดใหม่มาแทนเสียงนี้ ฉะนั้น การเติมจุด 6 จึงทำให้เกิดพยัญชนะเบรลล์ขึ้นอีก 9 ตัว

หากพิจารณาพยัญชนะที่นำมาใช้เป็นฐานในอีกแง่มุมหนึ่ง จะพบว่าการเติม จุด 6 ช่วยเพิ่มอักขระที่เป็นหมู่อักษรต่ำเสียงเดียวกับอักษรสูง ดังจะเห็นได้จากการนำ ‘⠋’ (=ถ) มาเติม จุด 6 ได้เป็น ‘⠓’ (=ท) ดังนั้นจึงมีพยัญชนะฐานในอักษรสูง 4 ตัว ได้แก่ ‘⠅’ (=ข) ‘⠡’ (=ฉ) ‘⠋’ (=ถ) และ ‘⠎’ (=ส) ที่เมื่อเติมจุด 6 จะได้พยัญชนะใหม่เป็นอักษรต่ำ คือ‘⠇’ (=ค)‘⠩‘ (=ช)‘⠓‘ (=ท) และ ‘⠉’ (=ช) ตามลำดับ

ด้วยเหตุนี้จึงดูเหมือนว่าการเติมจุด 6 ในเซลล์พยัญชนะฐานในอักษรสูงจะทำให้ได้อักษรต่ำที่มีเสียงเหมือนกับพยัญชนะฐานเสมอไป แต่ในความเป็นจริงปรากฏว่า มีพยัญชนะฐานในอักษรสูง 2 ตัวคือ ‘⠟’ (=ผ) และ ‘⠭’ (=ห) ที่หลังจากเติมจุด 6 แล้วได้เป็น ‘⠝’ และ ‘⠽’ ซึ่งน่าจะใช้แทน “พ” และ “ฮ” แต่กลับนำมาใช้แทน “ป” และ “ต” ซึ่งเป็นอักษรกลาง นอกจากนี้การเติมจุด 6 ก็ไม่ได้จำกัดอยู่เพียงการเพิ่มอักษรเสียงเดียวกัน เช่นการเติมจุด 6 เข้าไปในเซลล์ ‘⠛’ (=ก) ซึ่งเป็นอักษรกลางได้เป็น ‘⠿’ (=ง) ซึ่งเป็นอักษรต่ำ เป็นต้น การเติมจุด 6 ในที่นี้จึงดูเป็นระบบเพียงบางส่วน

เท่านั้น ตารางที่ 3.16 จึงแสดงส่วนขยายเพิ่มเติมเบรลล์พยัญชนะไทยอีก 10 ตัว จึงยังมีพยัญชนะอีก 17 ตัวที่ต้องกำหนดรหัสอักษรเบรลล์ขึ้นเพิ่มเติม ซึ่งผู้วิจัยพบว่ามีการสร้างขึ้นจากพยัญชนะฐาน โดยมีความสัมพันธ์กันทางเสียงอย่างเป็นระบบ แต่มีกลวิธีในการเติมจุดเบรลล์แตกต่างจากที่กล่าวมาข้างต้น ซึ่งจะได้นำเสนอต่อไป

2) การเติมอักขระเบรลล์

ในการสร้างพยัญชนะเพิ่มเติม มีการนำเบรลล์พื้นฐานที่มีเสียงตรงกับเบรลล์อังกฤษบางส่วน และพยัญชนะเบรลล์ที่เกิดจากพยัญชนะเบรลล์พื้นฐานเหล่านี้ส่วนหนึ่ง มาใช้เป็นฐาน แล้วเติมอักขระเบรลล์ข้างหน้าพยัญชนะทั้งสองกลุ่มนี้ ทำให้ได้พยัญชนะขึ้นเพิ่มเติมอีก 16 ตัว ดังนี้

ตารางที่ 3.19 การสร้างพยัญชนะใหม่ที่มีหมู่ตัวอักษรและเสียงตรงกับพยัญชนะฐาน

พยัญชนะฐาน		**พยัญชนะที่สร้างขึ้นใหม่**					
		เติมอักขระ จุด 6		**เติมอักขระ จุด 3-6**		**เติมอักขระ จุด 3-5-6**	
พยัญชนะพื้นฐาน	**พยัญชนะเบรลล์**	**เบรลล์**	**อักษรปกติ**	**เบรลล์**	**อักษรปกติ**	**เบรลล์**	**อักษรปกติ**
⠅ (ข)						⠸⠅	ฃ
	⠇ (ค)	⠄⠇	ฆ	⠸⠇	ฅ		
	⠉ (ช)	⠄⠉	ฉ				
⠙ (ด)		⠄⠙	ฎ				
	⠞ (ต)	⠄⠞	ฏ				
⠋ (ถ)		⠄⠋	ฐ				
	⠓ (ท)	⠄⠓	ฑ	⠸⠓	ฒ	⠸⠓	ธ
⠝ (น)		⠄⠝	ณ				
	⠏ (พ)	⠄⠏	ภ				
⠽ (ย)		⠄⠽	ญ				
⠇ (ล)		⠄⠇	ฬ				
⠎ (ส)		⠄⠎	ศ	⠸⠎	ษ		

ในเบรลล์ไทยมีการสร้างพยัญชนะเพิ่มเติม โดยการนำเบรลล์หนึ่งเซลล์ที่มีจุดเบรลล์รูปแบบเฉพาะคือ‘⠄’ (จุด 6) ‘⠸’ (จุด 3-6) และ ‘⠸⠅’ (จุด 3-5-6) มาวางไว้ข้างหน้าพยัญชนะฐาน ทำ

ให้อักษรที่เกิดขึ้นใหม่มีเสียงและหมู่อักษรเหมือนกับพยัญชนะฐาน เช่น ‘ ’(=ท),‘ ’(=ฒ) และ‘ ’ (=ธ) เกิดจากการเพิ่มอักขระข้างหน้า ‘ ’ (=ท) ซึ่งเป็นพยัญชนะฐาน

เป็นที่น่าสังเกตว่าในตารางมีพยัญชนะ 2 ตัวที่ไม่ปรากฏในภาษาไทยปัจจุบันแล้ว คือ “ฆ” และ “ฅ” ซึ่งเกิดจากการเพิ่มอักขระรูปแบบต่างกัน ในกรณีแรกมีการเพิ่มอักขระจุด 3-5-6 หน้า ‘ ’ (=ข) ได้เป็น ‘ ’ (=ฆ) และยังมีอักษรอีกเพียงหนึ่งตัวที่เพิ่มอักขระรูปแบบดังกล่าว คือ ‘ ’ (=ธ) ในขณะที่อีกกรณีหนึ่งมีการเพิ่มอักขระจุด 3-6 หน้า ‘ ’ (=ค) ได้เป็น ‘ ’ (=ฅ) และมีอักษรอีก 2 ตัวที่เพิ่มอักขระรูปแบบนี้คือ ‘ ’ (=ฒ) และ ‘ ’ (=ษ)

จึงพอสรุปได้ว่า ในการสร้างพยัญชนะเบรลล์ไทยในหมู่เดียวกันจะเริ่มจากการเติมอักขระ ‘ ’ (จุด 6) ข้างหน้าพยัญชนะซึ่งมาจากเบรลล์พื้นฐาน 15 ตัวและพยัญชนะที่ได้จากเบรลล์พื้นฐานเหล่านี้เป็นอันดับแรก แต่หากยังคงมีอักษรในหมู่เดียวกันที่ยังไม่ปรากฏในเบรลล์ไทยก็จะเติมอักขระ ‘ ’ (จุด 3-6) และอักขระ ‘ ’ (จุด 3-5-6) ข้างหน้าพยัญชนะฐานตามลำดับ

อย่างไรก็ตาม ยังเหลือพยัญชนะอีกหนึ่งตัวคือ “ฮ” ที่ยังไม่ได้ถูกกำหนดเป็นเบรลล์ ดังนั้นจึงมีการนำทั้งหกจุด คือ ‘ ’ มาใช้แทน “ฮ” ในภายหลัง แม้กระนั้นก็ตามการสร้างพยัญชนะใหม่ด้วยวิธีเติมอักขระหน้านี้ ทำให้เกิดความกำกวมขึ้นได้ในบางกรณี คือ ‘ ’ (=ธ) แต่ในบางครั้งจะแทน “◌์ท” ได้ด้วย เพราะ‘ ’ ตรงกับรูปที่ใช้แทน “◌์” ความกำกวมของอักขระนี้จะกล่าวถึงอีกครั้งในบทต่อไป

3.2.2.4 อักขรวิธีในเบรลล์ไทยและภาษาไทยทั่วไป

ในภาษาไทย ตัวอักษรแต่ละตัวไม่จำเป็นต้องใช้แทนเสียงหนึ่งเสียงเสมอไป ดังจะพบได้ว่าสระเดี่ยวอาจใช้มากกว่าหนึ่งอักขระ เช่น เสียง /e/ ใช้ 2 อักขระคือ "เ◌ะ" และเสียง /ɔ/ ใช้ 3 อักขระคือ "เ◌าะ" นอกจากนี้แม้ว่าภาษาไทยจะมีคู่เสียงสระสั้น-ยาวแต่ไม่ได้แสดงความแตกต่างนี้ในอักขรวิธีโดยตรง แต่ใช้วิธีสร้างรูปต่างกันไปเพื่อแทนสระสั้น-ยาว เช่น "ะ" คู่กับ "า" และ "◌ิ" คู่กับ "◌ี" และมีบางกรณีที่ใช้รูปสระที่มี "ะ" เพื่อแทนกลุ่มสระสั้น "เ◌ะ" "แ◌ะ" "เ◌าะ" "เ◌ียะ" เป็นต้น ดังนั้น ในการเรียนภาษาไทยจึงต้องเรียนรู้ทั้งระบบเสียงและระบบอักขรวิธีควบคู่กันไป

ลักษณะสำคัญประการหนึ่งที่เบรลล์ไทยแตกต่างจากไทยปกติคือ ในไทยปกติสระหนึ่งเสียงในบางกรณีต้องแทนด้วยอักขระหลายตัว ในขณะที่ในเบรลล์ไทยใช้เบรลล์เพียงตัวเดียวเพื่อแทนสระนั้น เช่น เสียง /ia/ ใช้สามอักขระคือ "เ◌ีย" ในอักษรปกติ แต่ใช้อักขระเบรลล์ตัวเดียวคือ ‘ ’ ซึ่งแสดงให้เห็นว่าเบรลล์ไทยส่วนหนึ่งอิงอยู่กับเสียง

อย่างไรก็ตาม วิธีการเขียนเบรลล์ไทยโดยปกติจะอิงลำดับการเขียนแบบไทยปกติ ยกเว้นกรณีสระที่มีรูปซับซ้อนจะเขียนเป็นเบรลล์ตัวเดียวไว้หลังเบรลล์ที่เป็นพยัญชนะต้น ส่วนกรณีสระ

ที่ใช้ตัวอักษรได้มากกว่าหนึ่งแบบ เช่นเสียง /a/ ในไทยปกติเขียนอักษรได้ 2 แบบ ก็จะใช้อักขระเบรลล์ต่างกันตามตัวอักษรคือ ‘⠣’ (=ั) และ‘⠡’ (=ะ) นั่นเอง

ในเบรลล์ไทย มีการประสมอักษรโดยยึดตามลำดับการจัดเรียงก่อนหลังเช่นเดียวกับอักษรปกติ ในอักษรปกติมีอักขระที่ปรากฏข้างบนหรือข้างล่างพยัญชนะต้น แต่เบรลล์ไทยจะจัดเรียงอักขระจากซ้ายไปขวา ดังแสดงในตัวอย่างต่อไปนี้

⠇⠊⠝	ล + ิ + ้ + น	“ลิ้น”
⠗⠢⠝	ร + เ◌ีย + น	“เรียน”
⠧⠢⠁	ว + ่ + า	"ว่า"
⠅⠍⠝	แ + ก + ้ + ม	"แก้ม"

จากตัวอย่างข้างต้นจะเห็นได้ว่า เบรลล์ไทยมีการประสมอักษรโดยจัดเรียงลำดับก่อนหลังเช่นเดียวกับอักษรปกติโดยที่ครอบคลุมสระส่วนใหญ่ในภาษาไทย อย่างไรก็ตาม หากยึดตามอักขรวิธีในภาษาไทยแล้ว “ะ” และ “◌ำ” จะต้องอยู่ภายใต้กฎเกณฑ์เดียวกับการจัดเรียงสระตำแหน่งกลาง (คือ “า” “อ” หรือ “ว”) แต่ในความเป็นจริง เบรลล์ไทยจะวางวรรณยุกต์ไว้หลัง “ะ” และ “◌ำ” ซึ่งผู้วิจัยสันนิษฐานว่าเหตุที่เป็นเช่นนี้เนื่องจากเกิดการเทียบแบบในการจัดเรียงกับสระต่าง ๆ ในเบรลล์ไทย ทำให้เกิดความลักลั่นในการจัดเรียงระหว่างเบรลล์ไทยและไทยปกติเช่น ‘⠴⠝⠍’ จัดเรียงเป็น น+◌ำ+้ ในขณะที่อักษรปกติจัดเรียงเป็น น +้+◌ำ เป็นต้น

สำหรับวรรณยุกต์ในเบรลล์ไทย มีการผันวรรณยุกต์กับอักษรสามหมู่ในลักษณะเดียวกับไทยปกติ เนื่องจากเบรลล์ไทยเขียนตามรูปอักษรจึงต้องออกเสียงวรรณยุกต์โดยอาศัยหลักการผันวรรณยุกต์เช่นเดียวกับไทยปกติซึ่งผู้เรียนอักษรเบรลล์ต้องจดจำอักษรสามหมู่ให้ได้ จึงจะผันวรรณยุกต์ได้ถูกต้อง

กล่าวโดยสรุป เบรลล์ไทยมีหน่วยอักขระที่คล้ายคลึงทางรูปและเสียงกับเบรลล์อังกฤษและเบรลล์ญี่ปุ่น โดยที่สระส่วนหนึ่งมีเสียงคล้ายคลึงกับเบรลล์ญี่ปุ่นและพยัญชนะส่วนหนึ่งมีเสียงคล้ายคลึงกับเบรลล์อังกฤษ จึงได้มีการนำหน่วยอักขระเหล่านี้มาใช้เป็นหน่วยอักขระพื้นฐานในการสร้างหน่วยอักขระเบรลล์อื่น ๆ ขึ้นเพิ่มเติม โดยวิธีการสร้างหน่วยอักขระนั้นมีการนำหน่วยอักขระในเบรลล์ไทย ซึ่งรวมถึงอักขระเบรลล์พื้นฐานมาดัดแปลงโดยการปรับเปลี่ยนจุดเบรลล์ภายในเซลล์ หรือการเพิ่มอักขระที่มีจุดเบรลล์รูปแบบเฉพาะข้างหน้าหรือข้างหลังอักษรที่ใช้เป็นฐาน อย่างไรก็ตามมีอักขระเบรลล์ไทยอยู่จำนวนหนึ่งที่ไม่สามารถอธิบายถึงความสัมพันธ์ที่มีต่ออักขระเบรลล์ตัวอื่น ๆ ได้ ฉะนั้นผู้ที่ศึกษาเบรลล์ไทยจึงจำเป็นต้องจดจำรหัสอักษรเบรลล์ในกลุ่มนี้ให้ได้ จึงจะเรียนรู้หน่วยอักขระในเบรลล์ไทยได้ครบทั้งหมด

บทที่ 4

การออกแบบและพัฒนาระบบ
การถ่ายถอดเบรลล์ไทยปนอังกฤษโดยใช้กฎ

อักษรเบรลล์เกิดจากการนำจุดไม่เกิน 6 จุดมาเรียงต่อกันโดยที่มีรูปแบบที่จำกัด ทำให้อักษรเบรลล์มีรูปซ้ำกันระหว่างภาษาได้ นอกจากนี้ด้วยเหตุที่แต่ละภาษาไม่สามารถใช้ 6 จุดมาแสดงตัวอักษรได้ครบทุกตัว เป็นผลให้ต้องใช้เบรลล์มากกว่าหนึ่งเซลล์แทนอักษรปกติตัวเดียว และยังมีความพยายามที่จะลดเนื้อที่แสดงตัวอักษรโดยกำหนดให้มีเบรลล์รูปย่อ ทำให้สามารถใช้เบรลล์บางตัวแทนตัวอักษรปกติได้มากกว่าหนึ่งตัว ปัจจัยเหล่านี้ล้วนเป็นความท้าทายในการถ่ายถอดเบรลล์ไทยปนอังกฤษทั้งสิ้น ดังนั้นในบทนี้ผู้วิจัยจะได้กล่าวถึงแนวคิดที่นำมาใช้ในการถ่ายถอดเบรลล์ไทยปนอังกฤษ โดยจะได้กล่าวถึงแนวคิดในการออกแบบการถ่ายถอดเบรลล์ไทย แนวคิดในการออกแบบการถ่ายถอดเบรลล์อังกฤษ และแนวคิดเกี่ยวกับการระบุภาษาในอักษรเบรลล์

4.1 แนวคิดที่ใช้ในการออกแบบและพัฒนาระบบการถ่ายถอดเบรลล์ไทย

หากมองผิวเผิน อาจดูเหมือนว่าการถ่ายถอดเบรลล์ไทยให้เป็นไทยปกตินั้นทำได้ไม่ยาก เพียงแค่นำเบรลล์แต่ละตัวไปจับคู่กับอักษรปกติเท่านั้น[6] แต่การถ่ายถอดโดยการจับคู่ตัวอักษรโดยตรง จะทำให้ได้ผลลัพธ์ของการถ่ายถอดถูกต้องเพียงส่วนหนึ่งเท่านั้น เนื่องจากเบรลล์ไทยและไทยปกติมีอักขรวิธีในการประสมอักษรที่แตกต่างกัน ดังที่ได้นำเสนอมาแล้วในบทที่ 3

ในหัวข้อนี้ ผู้วิจัยจะอภิปรายเกี่ยวกับแนวคิดในการออกแบบระบบการถ่ายถอดเบรลล์ไทยให้เป็นไทยปกติ ภายใต้ความแตกต่างระหว่างอักษรเบรลล์และอักษรปกติ ซึ่งความแตกต่างเหล่านี้จะช่วยให้เข้าใจปัญหาสามารถออกแบบและพัฒนาระบบที่เหมาะสมสำหรับถ่ายถอดเบรลล์ไทยให้เป็นไทยปกติได้

4.1.1 ความไม่สอดคล้องระหว่างเบรลล์ไทยและไทยปกติ

ความไม่สอดคล้องระหว่างเบรลล์ไทยและไทยปกติ มีอยู่ 2 ลักษณะคือ ความไม่สอดคล้องด้านการจัดเรียงตัวอักษร และ ความไม่สอดคล้องที่เกิดจากการใช้จำนวนตัวเขียนไม่เท่ากันเพื่อแสดงอักษรรูปเดียวกันในเบรลล์ไทยและไทยปกติ

---

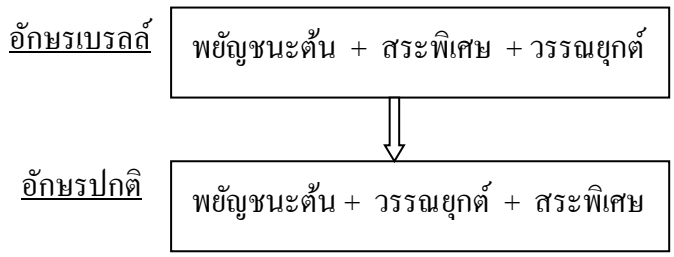
[6] เช่น ‘⠯’ จับคู่กับ “ก” ได้ ดังนั้นเมื่อระบบพบ ‘⠯’ ก็จะถ่ายถอดเป็น “ก” ได้ทันที

4.1.1.1 การจัดเรียงตัวอักษรที่แตกต่างกัน

จากการศึกษาระบบเบรลล์ไทย ผู้วิจัยพบว่า อักษรเบรลล์เซลล์เดี่ยวจำนวนหนึ่งมีการจัดเรียงลำดับตัวอักษรเหมือนกับอักษรไทยปกติจึงไม่เป็นปัญหาต่อการถ่ายถอดอักษร แต่ยังมีอักษรเบรลล์อีกจำนวนหนึ่ง ที่เมื่อทำการถ่ายถอดแล้วจำเป็นต้องสลับตำแหน่งตัวอักษรเพื่อให้ถูกต้องตามอักขรวิธีไทย ได้แก่กลุ่มตัวอักษรเบรลล์ต่อไปนี้

ก. สระพิเศษ

ในอักษรเบรลล์มีสระ 2 ตัว ที่เมื่อสระเหล่านี้ปรากฏร่วมกับวรรณยุกต์จะมีรูปแบบการจัดเรียงแตกต่างกับอักษรปกติ สระเหล่านี้ได้แก่ ‘⠪’ (=ะ) และ ‘⠣’ (= ำ) ซึ่งในงานวิจัยฉบับนี้ผู้วิจัยจะขอเรียกสระ 2 ตัวนี้ว่า “สระพิเศษ” สระพิเศษเมื่อปรากฏร่วมกับวรรณยุกต์จะมีรูปแบบการจัดเรียงแตกต่างกันระหว่างอักษรเบรลล์และอักษรปกติ ดังแสดงในแผนภาพต่อไปนี้



ภาพที่ 4.1 การจัดเรียงสระพิเศษในอักษรเบรลล์และในอักษรปกติ

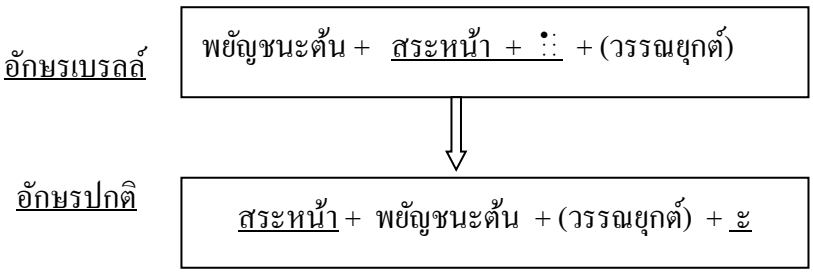
ในการถ่ายถอดสระพิเศษ จะมีการจัดเรียงในอักษรปกติดังที่แสดงในตัวอย่าง เช่น ‘⠝⠪’ (=น-ะ) จัดเรียงเป็น “นะ” (=น- ะ) และ ‘⠝⠣⠊’ (=น- ำ- ้) จัดเรียงเป็น “น้ำ” (=น- ้- ำ) ดังนั้นหากระบบพบว่า สระพิเศษปรากฏร่วมกับวรรณยุกต์ในเบรลล์ไทย เมื่อถ่ายถอดอักษรระบบจะทำการสลับที่ระหว่างสระพิเศษกับวรรณยุกต์ในอักษรปกติ ซึ่งในกรณีนี้คือสลับที่ระหว่าง ‘⠣’ (=ำ) และ ‘⠊’ (=้) จึงจะทำให้การถ่ายถอดอักษรครบถ้วนสมบูรณ์

ข. สระหน้า

จากการวิเคราะห์การปรากฏของสระหน้าในเบรลล์ไทย ผู้วิจัยพบว่า สระหน้า 2 ตัวคือ ‘⠣’ (= ไ) และ ‘⠣⠒’ (= ใ) มีการจัดเรียงอักษรในเบรลล์ไทยและไทยปกติเหมือนกัน กล่าวคือสระสอง ตัวนี้จะปรากฏหน้าพยัญชนะต้นพยางค์ จึงไม่เป็นอุปสรรคต่อการถ่ายถอดอักษร

อย่างไรก็ตามสระหน้าอีก 3 ตัวได้แก่ ‘⠪’ (=เ) ‘⠡’ (=แ) และ ‘⠒’ (= โ) มีการจัดเรียงอักษรทั้งเหมือนและแตกต่างจากอักษรปกติ ดังนั้นในการถ่ายถอดสระหน้าจำเป็นต้องจัดเรียงสระเหล่านี้ให้ถูกต้องตามอักขรวิธีไทย

ดังจะเห็นได้ว่า รูปสระกลุ่มนี้มีการจัดเรียงในอักษรเบรลล์และอักษรปกติเหมือนกันในกรณีที่เป็นสระยาว เพราะเขียนลำดับเป็น สระหน้า + พยัญชนะ ต้น +... แต่จะมีการจัดเรียงต่างกันเมื่อรูปสระเหล่านี้ปรากฏร่วมกับ '⠡' (= ะ) คือใช้เป็นสระสั้น ตามแผนภาพต่อไปนี้



ภาพที่ 4.2 การจัดเรียงสระหน้าในอักษรเบรลล์และในอักษรปกติ

ภาพที่ 4.2 แสดงให้เห็นว่าเมื่อสระหน้าปรากฏร่วมกับ '⠡' (= ะ) สระ 2 ตัวนี้จะต้องอยู่ชิดกันเสมอ โดยที่จะมีวรรณยุกต์ตามหลังหรือไม่ก็ได้ ซึ่งแตกต่างจากอักษรปกติ ที่สระ 2 ตัวนี้จะปรากฏห่างกัน ดังแสดงในตัวอย่างต่อไปนี้

{1} ก. '⠇⠕⠇⠡' (= โ-ล-ะ-ล) ⟹ โลเล

ข. '⠇⠡⠇⠡' (= ล - ะ – ล) ⟹ ละ

ค. '⠇⠿⠇⠡' (=ต -โ – ะ- ◌๊) ⟹ โต๊ะ

จากตัวอย่างข้อ {1} ก. มีการจัดเรียงอักษรระหว่างเบรลล์ไทยและไทยปกติเหมือนกัน คือ การจัดเรียงแบบ สระหน้าตามด้วยพยัญชนะต้นพยางค์ ในขณะที่ข้อ {1} ข. และ {1} ค. มีการจัดเรียงสระหน้าที่แตกต่างกันในระบบการเขียนทั้ง 2 ระบบ ดังนั้นในการถ่ายถอด ระบบจึงต้องจัดเรียงอักษรให้ถูกต้องตามรูปแบบที่กล่าวมาข้างต้น

4.1.1.2 การใช้จำนวนตัวเขียนที่แตกต่างกันเพื่อแสดงอักษรรูปเดียวกัน

เมื่อพิจารณาความสัมพันธ์ของรูปอักษรเบรลล์และอักษรปกติ จะพบว่า เบรลล์ 1 เซลล์ ล้วนแทนอักษรปกติ 1 ตัวทั้งสิ้นซึ่งมีเบรลล์ไทยเพียงบางส่วนเท่านั้นที่สัมพันธ์กันในลักษณะนี้

นอกเหนือจากนี้ในเบรลล์ไทยมีการใช้เบรลล์ 2 เซลล์เพื่อแทนอักษรปกติ 1 ตัว และในทางกลับกันมีการใช้เบรลล์ 1 เซลล์เพื่อแทนอักษรปกติหลายตัว ความไม่สอดคล้องระหว่างจำนวนตัวอักษรเบรลล์และจำนวนตัวอักษรปกติที่กล่าวมานี้ ล้วนก่อให้เกิดปัญหาต่อการถ่ายถอดอักษร ซึ่งจะขออภิปรายดังนี้

4.1.1.2.1 เบรลล์สองเซลล์แทนอักษรปกติหนึ่งตัว

ในเบรลล์ไทย เบรลล์เซลล์คู่จำนวนหนึ่งเกิดจากการเติมอักขระหน้าเบรลล์เซลล์เดี่ยว (หรือบางกรณีเติมข้างหลัง) เพื่อสร้างอักขระใหม่ซึ่งมีลักษณะบางประการโดยเฉพาะลักษณะทางเสียงที่สอดคล้องกับเบรลล์เซลเดี่ยวนั้นๆ ดังที่ได้กล่าวมาแล้วในบทที่ 3 และเบรลล์เซลล์คู่อีกส่วนหนึ่งเป็นเครื่องหมายต่างๆ

ดังจะพบได้ว่าเบรลล์เซลล์คู่หลายคู่เกิดจากการนำเบรลล์เซลล์เดี่ยวมาปรับเปลี่ยน (modify) เพื่อทำให้เกิดตัวอักษรใหม่ เช่น ‘⠨⠍’ (=ฒ) เกิดจากการเติม ‘⠨’ ข้างหน้า ‘⠍’(=ท) เป็นต้น

ดังนั้นในการถ่ายถอดอักษร ระบบจำเป็นต้องถ่ายถอดเบรลล์เซลล์คู่ก่อนเบรลล์เซลล์เดี่ยว จึงจะทำให้การถ่ายถอดได้ผลลัพธ์ที่ถูกต้อง หากถ่ายถอดเบรลล์เซลล์เดี่ยวก่อนอาจทำให้ได้ผลลัพธ์เป็นตัวอักษรสองตัว แทนที่จะเป็นตัวเดียวเช่น ได้เป็น “-ท” แทนที่จะเป็น “ฒ” เป็นต้น

4.1.1.2.2 เบรลล์หนึ่งเซลล์แทนอักษรปกติมากกว่าหนึ่งตัว

นอกจากอักษรปกติบางตัวต้องแทนด้วยเบรลล์ 2 เซลล์แล้ว ยังมีอักษรเบรลล์อีกประเภทหนึ่งที่เบรลล์ 1 เซลล์ใช้แทนอักษรปกติหลายตัวดังที่ปรากฏกับ สระประสม

ดังที่ได้แสดงมาแล้วในบทที่ 3 ว่า สระประสมรูปเดียวกันในเบรลล์ไทยและไทยปกติจะใช้จำนวนตัวอักษรแตกต่างกัน ด้วยเหตุนี้จึงทำให้มีการจัดเรียงอักษรตัวอื่นที่ปรากฏร่วมกับสระประสมในระบบตัวเขียนทั้งสองระบบมีความแตกต่างกัน ดังนั้นในการถ่ายถอดอักษรเบรลล์จึงต้องจัดเรียงสระประสมให้เป็นตามอักขรวิธีไทย ดังแสดงในตัวอย่างต่อไปนี้

{2} ก. ‘⠃⠥’ (= บ - ◌ัว) ⟹ บัว

ข. ‘⠎⠜’ (= ส - เ◌ีย) ⟹ เสีย

ค. ‘⠎⠬’ (= ส- เ◌ือ) ⟹ เสือ

จากตัวอย่าง มีการใช้เบรลล์เพียง 1 เซลล์แทนสระประสม และเมื่อเปรียบเทียบการจัดเรียงสระประสม จะพบว่าระบบการเขียนทั้ง 2 ระบบนี้มีการจัดเรียงสระที่เหมือนกันและแตกต่างกันนั่นคือ ในข้อ {2}ก.สระประสมในเบรลล์ยังคงรักษาตำแหน่งเช่นเดียวกับในอักษรปกติ (คือปรากฏหลังพยัญชนะต้น) แต่ในข้อ {2}ข. และ {2}ค. สระประสมจะตามหลังพยัญชนะต้นในอักษรเบรลล์ในขณะที่ในอักษรปกติมีการสลับที่ระหว่าง “เ” (ซึ่งเป็นส่วนหนึ่งของสระประสม) กับพยัญชนะต้น ความแตกต่างในการจัดเรียงสระประสมในเบรลล์ไทยและไทยปกตินี้เองที่เป็นปัญหาต่อการถ่ายถอดเนื่องจากก่อให้เกิดความกำกวมในเบรลล์ไทยซึ่งจะได้กล่าวในหัวข้อที่เกี่ยวข้องต่อไป

นอกจากนี้รูปของสระที่ใช้จำนวนตัวอักษรไม่เท่ากันในอักษรเบรลล์และอักษรปกติ ยังส่งผลให้วรรณยุกต์ที่ปรากฏร่วมกับสระเหล่านี้ มีการจัดเรียงที่แตกต่างกัน ดังแสดงในตัวอย่าง

{3} ก. ‘⠎⠗⠠⠧’ (= ร – ั ว -่) ⟹ รั่ว (=ร- ั- ่-ว)

ข. ‘⠃⠊⠽’ (= บ – เ ี ย - ้) ⟹ เบี้ย

ค. ‘⠏⠢⠪’ (= พ - เ อ - ้) ⟹ เพ้อ

ง. ‘⠏⠢⠬⠛’ (= พ-เ ิ-้-ง) ⟹ เพิ้ง

จากตัวอย่างจะพบว่าเบรลล์ไทยมีการจัดเรียงของสระที่ปรากฏร่วมกับวรรณยุกต์ แตกต่างจากการจัดเรียงในอักษรปกติเป็นส่วนใหญ่ โดยในข้อ {3} ก. {3} ข. และ {3} ค. วรรณยุกต์จะปรากฏหลังสระประสมในอักษรเบรลล์ ในขณะที่วรรณยุกต์จะปรากฏในตำแหน่งก่อนตัวอักษรสุดท้ายของสระประสมในอักษรปกติ อย่างไรก็ตาม ยังมีบางกรณีที่ระบบการเขียนทั้งสองระบบมีการปรากฏของวรรณยุกต์ตามหลังสระเช่นเดียวกัน ดังที่ปรากฏในข้อ {3} ง.

จากที่ได้กล่าวมาในหัวข้อนี้พอจะสรุปได้ว่า เบรลล์ไทยและไทยปกติใช้จำนวนตัวอักษรในการแสดงตัวอักษรบางประเภทแตกต่างกัน สำหรับวิธีการจัดเรียงตัวอักษรที่เป็นปัญหานี้มีวิธีการแก้ไขที่แตกต่างกันในระบบการถ่ายถอดอักษรด้วยกฎ ด้วยแบบจำลองเอ็นแกรม และด้วยวิธีแบบผสม ซึ่งจะได้กล่าวในหัวข้อที่เกี่ยวข้องต่อไป

4.1.2 ความกำกวมของเบรลล์เซลล์เดี่ยว

หากพิจารณาถึงความกำกวมของเบรลล์เซลล์เดี่ยวในภาษาไทย ผู้วิจัยพบว่า สามารถแบ่งความกำกวมออกเป็น 2 ประเภท คือ (1) ความกำกวมที่เกิดจากอักขระเบรลล์หนึ่งตัวใช้แทนอักษรปกติได้มากกว่าหนึ่งกรณี และ (2) ความกำกวมจากการจัดเรียงพยัญชนะสองตัวที่ปรากฏหน้าสระประสม ซึ่งความกำกวมแต่ละประเภทล้วนเป็นปัญหาที่ต้องอาศัยกระบวนการที่เหมาะสม เพื่อให้ถ่ายถอดอักษรได้ถูกต้อง

4.1.2.1 อักขระเบรลล์หนึ่งตัวใช้แทนอักษรปกติได้หลายกรณี

ในเบรลล์ไทย เบรลล์เซลล์เดี่ยวบางตัวสามารถถ่ายถอดเป็นอักษรปกติได้มากกว่าหนึ่งกรณี โดยที่ตัวอักษรเบรลล์ที่กำกวมนั้นอาจปรากฏอยู่เฉพาะท้ายสายอักขระ หรืออยู่ส่วนใดของสายอักขระก็ได้

4.1.2.1.1 อักขระเบรลล์กำกวมที่ปรากฏในส่วนใดของสายอักขระก็ได้

อักขระเบรลล์ในกลุ่มนี้ประกอบด้วยอักขระ 2 ตัว ได้ ‘⠪’ และ ‘⠣’

ก. ความกำกวมของ ‘⠪’ ("เ ิ" หรือ "เ ือ")

‘⠪’ มีรูปแทนในอักษรปกติ 2 รูปคือ "เ ิ" หรือ "เ ือ" ดังแสดงในตัวอย่างต่อไปนี้

{4} ก. ‘⠛⠢⠝’ (= ง – เ◌ิ - น) ⟹ เงิน

ข. ‘⠉⠢’ (= จ – เ◌อ) ⟹ เจอ

จากตัวอย่าง ‘⠢’ มีรูปในอักษรปกติ 2 รูปคือ "เ◌ิ" ในข้อ {4}ก. และ"เ◌อ" ในข้อ {4} ข. ซึ่ง ‘⠢’ จะมีรูปใดนั้นสามารถทำนายได้จากบริบทที่เบรลล์ตัวนี้ปรากฏอยู่ กล่าวคือจะเป็น “เ◌ิ” ก็ต่อเมื่อ ‘⠢’ ตามหลังด้วยพยัญชนะตัวสะกดท้ายพยางค์ และจะมีรูปเป็น “เ◌อ” ในกรณีตรงกันข้าม ดังนั้นในการถ่ายถอดอักขระเบรลล์ตัวนี้ ระบบต้องอาศัยบริบทางขวาเพื่อกำหนดว่าควรจะถ่ายถอด ‘⠢’ โดยใช้รูปแทนใดในอักษรปกติ

อย่างไรก็ตาม มีคำ 3 คำ ที่ไม่เป็นไปตามเกณฑ์ที่กล่าวมาข้างต้น แม้คำเหล่านี้จะมีตัวสะกดท้ายพยางค์ก็ตาม แต่ก็มีรูปเป็น เ◌อ ได้แก่ “เทอญ” “เทอม” และ “เยอว” ในอักษรเบรลล์จึงต้องเขียนคำเหล่านี้โดยใช้หลักเกณฑ์ในการสะกดคำเช่นเดียวกับอักษรปกติ เช่น ‘⠞⠓⠢⠍’ (เ-ท-อ-ญ) ซึ่งไม่เป็นปัญหาต่อการถ่ายถอดอักษร เพราะหากระบบพบคำเหล่านี้ระบบก็จะทำการถ่ายถอดอักษรแบบตัวต่อตัว

อย่างไรก็ตาม การพิจารณาอักขระที่ปรากฏหลัง ‘⠢’ เพียงตัวเดียว ยังไม่เพียงพอที่จะระบุได้ว่าอักขระเบรลล์นั้นเป็นตัวสะกดท้ายพยางค์หรือไม่ เนื่องจากเบรลล์ตัวนั้นอาจมีรูปเหมือนพยัญชนะตัวสะกดท้ายพยางค์ แต่อาจเป็นพยัญชนะต้นของพยางค์ถัดไปก็ได้ ดังตัวอย่าง

{5} ก. ‘⠞⠓⠢⠍⠊’ (= ธ-เ อ-<u>ม</u>- ◌ี) ⟹ เธอมี

ข. ‘⠉⠢⠝⠐⠕⠛’ (= จ-เ อ-<u>น</u>- ◌้-อ-ง) ⟹ เจอน้อง

ค. ‘⠎⠢⠝⠐⠎⠦’ (= ส-น-เ อ-<u>น</u>-แ-ะ) ⟹ เสนอแนะ

ในตัวอย่าง ล้วนมีพยัญชนะที่มีรูปคล้ายตัวสะกดท้ายพยางค์ ปรากฏหลัง ‘⠢’ ทั้งสิ้น ระบบจึงจำเป็นต้องอาศัยบริบททางขวาเพิ่มเติม ในข้อ {5} ก. ระบบพบ ‘⠊’ (= ◌ี) ปรากฏหลัง ‘⠍’ (= ม) เนื่องจาก ‘⠊’ (= ◌ี) เป็นสระทั่วไปที่ต้องปรากฏหลังพยัญชนะต้นพยางค์เสมอ จึงเป็นผลให้‘⠍’เป็นพยัญชนะต้นของพยางค์ ระบบจึงต้องเลือกรูปในอักษรปกติตามที่แสดงไว้ข้างต้น

ในข้อ {5} ข. ‘⠐’ (=◌้) ปรากฏหลัง ‘⠝’ (=น) ทำให้ ‘⠝’ (=น) ที่มีรูปคล้ายตัวสะกดท้ายพยางค์ ไม่ใช่ตัวสะกดท้ายพยางค์ ดังนั้น ‘⠢’ จึงต้องมีรูปเป็น "เ-อ"

ในข้อ {5} ค. ‘⠦’ (=แ) เป็นสระหน้าซึ่งปรากฏหลัง ‘⠝’ (=น) หากพิจารณาเพียง ‘⠦’ เท่านั้น จะต้องจัดให้ ‘⠝’ (=น) เป็นตัวสะกดท้ายพยางค์เนื่องจาก ‘⠦’ เป็นสระหน้าที่ใช้กำกับจุดเริ่มต้นของพยางค์ได้ (เช่นใน “เงินแท้” ที่ “น” เป็นตัวสะกดท้ายพยางค์) แต่ในอักษรเบรลล์ หากสระหน้าปรากฏร่วมกับ ‘⠡’ (= ะ) อักษร 2 ตัวนี้ต้องอยู่ชิดกันเสมอ และต้องตามหลังพยัญชนะต้นด้วย ดังนั้นเมื่อพบสระหน้า ระบบจำเป็นต้องตรวจสอบอักษรที่ตามหลังสระหน้าว่าเป็น ‘⠡’ (= ะ)

หรือไม่ หากเป็นสระตัวนี้ พยัญชนะที่นำหน้า ‘◌’ ก็ไม่ใช่ตัวสะกดท้ายพยางค์ เป็นผลให้ในกรณีนี้ต้องถ่ายถอด ‘◌’ ให้มีรูปเป็น "เ◌อ"

ข. ความกำกวมของ ‘◌’ ("◌้" หรือ ".")

‘◌’ สามารถตีความให้เป็น “◌้” หรือ “.” ได้ โดยจะแปรไปตามบริบทที่เบรลล์ตัวนี้ปรากฏอยู่ ซึ่งสามารถแสดงได้ในตัวอย่างต่อไปนี้

{6} ก. ‘◌ ◌ ◌ ◌ ◌’ (=บ- ◌้-า-น-น-า) ⟹ บ้านนา

ข. ‘◌ ◌ ◌ ◌ ◌’(=พ-.-ศ-.) ⟹ พ.ศ.

จากตัวอย่างจะเห็นว่า ‘◌’ ในข้อ {6} ก. ต้องตีความให้เป็น “◌้” แต่ในข้อ{6} ข. ต้องตีความให้เป็น “.” การที่จะตีความให้เป็นกรณีใดนั้นจะต้องพิจารณาอักขระตัวนี้พร้อมบริบททางซ้ายและบริบททางขวาซึ่งทำได้โดยการใช้กฎและแบบจำลองเอ็นแกรม ซึ่งจะได้กล่าวในหัวข้อที่เกี่ยวข้องต่อไป

4.1.2.1.2 อักขระเบรลล์ที่กำกวมเมื่อปรากฏท้ายสายอักขระ

นอกจากความกำกวมของอักษรเบรลล์ที่ปรากฏในตำแหน่งใดของสายอักขระตามที่ได้อภิปรายในหัวข้อที่ผ่านมาแล้ว ยังมีเบรลล์เซลล์เดี่ยวอีกกลุ่มหนึ่งที่มีความกำกวมเฉพาะเมื่ออักขระเบรลล์เหล่านี้ปรากฏท้ายสายอักขระเท่านั้น จากการศึกษาเบรลล์ไทย ผู้วิจัยพบว่าเบรลล์ที่กำกวมท้ายสายอักขระลักษณะนี้ มีอยู่ 2 ตัว ได้แก่ ‘◌’ และ ‘◌’

ก. ความกำกวมของ ‘◌’ (“◌์” หรือ “ ” ”)

‘◌’ เมื่อปรากฏท้ายสายอักขระสามารถตีความให้เป็น “◌์” หรือ “ ” ” ได้ ดังนี้

{7} ก. ‘◌ ◌ ◌ ◌ ◌’ (= ศ- ◌ุ-ก-ร์) ⟹ ศุกร์

ข. ‘◌ ◌ ◌ ◌ ◌’ (= น-ะ-ค-ะ-”) ⟹ นะคะ”

จากตัวอย่าง เมื่อ ‘◌’ ปรากฏท้ายสายอักขระ สามารถตีความให้เป็น “◌์” หรือเป็น “ ” ” (คำพูดปิด) ได้ ขึ้นอยู่กับบริบทที่อักษรตัวนี้ปรากฏอยู่ การที่จะถ่ายถอดอักขระนี้ให้ถูกต้องได้ ต้องพิจารณาบริบททางซ้ายของอักขระตัวนี้ ดังจะเห็นได้ว่าข้อ {7} ก. บริบทซ้ายของ‘◌’ คือ “ศุกร” และข้อ {7} ข. บริบทซ้ายของ‘◌’ คือ “นะคะ” ส่วนจะดำเนินการตรวจสอบบริบททางซ้ายอย่างไรนั้นขึ้นอยู่กับวิธีการถ่ายถอดในแต่ละวิธี ซึ่งจะได้อภิปรายในบทที่ 5 และ 6 ต่อไป

ข. ความกำกวมของ ‘◌’ (“◌๊” หรือ “) ”)

‘◌’ สามารถตีความให้เป็น “◌๊” หรือ “)” เมื่อปรากฏท้ายสายอักขระ ดังตัวอย่างต่อไปนี้

{8} ก. ‘⠎⠥⠅⠓⠁⠏⠒’ (= ส-ุ-ข-ภ-า-พ-)) ⟹ สุขภาพ)

ข. ‘⠃⠝⠞⠕⠒’ (บ-น-ต-โ-ะ- ้) ⟹ บนโต๊ะ

ในตัวอย่างนี้ต้องตีความ ‘⠒’ ต่างกันในอักษรปกติ กล่าวคือ ในข้อ{8} ก. ต้องตีความให้เป็น “)” ส่วนข้อ{8}ข. ต้องตีความให้เป็น “้” ในการถ่ายถอด ‘⠒’ให้ได้ผลลัพธ์ที่ถูกต้องนั้นต้องอาศัยบริบททางซ้ายของอักขระเบรลล์ตัวนี้เช่นเดียวกับการถ่ายถอด ‘⠒’

4.1.2.2 ความกำกวมในการจัดเรียงพยัญชนะสองตัวหน้าสระประสม

ความกำกวมลักษณะนี้เกิดจากความต่างของอักขรวิธีภาษาไทยและอักขรวิธีเบรลล์ไทย กล่าวคือ ตามหลักอักขรวิธีไทยสระประสมที่ประกอบด้วยรูปตัวอักษรมากกว่าหนึ่งอาจเขียนในลักษณะที่มีพยัญชนะ 1-2 ตัวปรากฏระหว่างสระ เช่น เรือ เรียน เรื่อง หรือเขียนในลักษณะที่ติดกัน เช่น หัว ทั่ว แต่ในอักขรวิธีเบรลล์ไทยสระสองกลุ่มนี้มีการจัดเรียงอักษรในลักษณะเดียวกันคือเขียนติดกันโดยไม่มีการแทรกของรูปพยัญชนะ เช่น ‘⠏⠒⠝’ (= ต-เ◌ีย- ้) และ ‘⠞⠒⠺’ (=ท- ◌ัว- ่)

การถ่ายถอดเบรลล์ไทยในกรณีสระประสมกลุ่มแรกจึงมีความกำกวมเกิดขึ้น จากการนำรูปพยัญชนะมาแทรกให้ถูกต้อง เพราะในเบรลล์ไทย บ่อยครั้งที่มีการปรากฏของพยัญชนะ 2 ตัวหน้าสระประสมโดยที่พยัญชนะคู่นี้มีรูปคล้ายพยัญชนะควบกล้ำแต่อาจเป็น (1) พยัญชนะควบกล้ำจริง หรือ (2) อักษรตัวหน้าเป็นพยัญชนะท้ายของพยางค์หน้าและตัวต่อมาเป็นพยัญชนะต้นของพยางค์เดียวกับสระประสม ดังตัวอย่างต่อไปนี้

{9} ก. ‘⠏⠗⠁⠙⠏⠗⠑⠞⠒’ (= ป-ร-า-ด-<u>ป-ร</u>-เ◌ือ- ่-ง) ⟹ ปราดเปรื่อง

ข. ‘⠃⠁⠏⠗⠑⠞⠒⠝⠥⠊’ (= บ-า-<u>ป-ร</u>-เ◌ือ- ่-ง-น- ◌ี-้) ⟹ บาปเรื่องนี้

หากเปรียบเทียบการจัดเรียงอักษรเบรลล์และอักษรปกติข้างต้น จะพบว่าข้อ {9} ก.และข้อ {9} ข. ต่างก็มีพยัญชนะ ‘⠏⠗’ (= ปร) ซึ่งอยู่ชิดกันในอักษรเบรลล์ แต่ในอักษรปกติ ข้อ {9} ก. “ปร” เป็นพยัญชนะควบกล้ำ ในขณะที่ข้อ {9} ข. “ป” เป็นพยัญชนะท้ายของพยางค์หน้าและ “ร” เป็นพยัญชนะต้นของพยางค์เดียวกับสระประสม

ผู้วิจัยพบว่า ความกำกวมของสระประสม เกิดขึ้นใน 2 ลักษณะ คือ

(1) <u>สระประสมในอักษรเบรลล์แสดงด้วยเบรลล์เพียงเซลล์เดียว</u> ทำให้ยากที่จะระบุได้ว่าพยัญชนะ 2 ตัวที่ปรากฏหน้าสระประสมเป็นพยัญชนะควบกล้ำอย่างแท้จริงหรือไม่ เช่น

{10} ก. ‘⠇⠕⠛⠇⠊⠜⠝’ (= ล-อ-<u>ก-ล</u>-เ◌ีย-น) ⟹ ลอกเลียน

* ข. ‘⠇⠕⠛⠇⠊⠜⠝’ (= ล-อ -<u>ก-ล</u>-เ◌ีย-น) ⟹ ลอเกลียน

จากตัวอย่าง ‘⠛⠣’ (=กล) เนื่องจากเป็นพยัญชนะที่อยู่ชิดกัน ทำให้ยากที่จะระบุได้ว่า ‘⠛’ (=ก) เป็นส่วนหนึ่งของพยางค์หน้าดังในข้อ{10} ก. หรือเป็นส่วนหนึ่งของพยางค์เดียวกับ ‘⠣‘ (=เ◌ีย) ในข้อ {10} ข.

(2) <u>สระประสมในอักษรเบรลล์แสดงด้วยเบรลล์สองเซล</u> โดยที่เซลล์หลังเป็น ‘⠱’ (=ะ) และเบรลล์เซลหน้าเป็น ‘⠛’(=เ), ‘⠩’ (=แ) หรือ ‘⠕’ (โ) โดยจะเกิดความกำกวมในการจัดเรียงอักษรไทยปกติได้ ดังตัวอย่างต่อไปนี้

{11} ก.‘⠝⠕⠛⠇⠩⠝’ (=น-อ-ก-ล-<u>แ-ะ</u>-ไ-น) ⟹ นอก<u>และ</u>ใน

* ข. ‘⠝⠕⠛⠇⠣⠝’ (=น-อ-ก-ล-<u>เ-ีย</u>-ไ-น) ⟹ นอ<u>เกลีย</u>ใน

จากตัวอย่าง ‘⠛⠣’ (=กล) มีรูปคล้ายพยัญชนะควบกล้ำ แต่ก็ยากที่ระบบจะระบุได้ว่า ‘⠛⠣’เป็นพยัญชนะควบกล้ำในอักษรเบรลล์จึงเกิดความกำกวมดังที่แสดงไว้ใน{11}ก.และ{11} ข.

ฉะนั้นในการถ่ายถอดสระประสม ระบบจึงจำเป็นต้องตรวจสอบสระประสมพร้อมบริบทซ้ายและบริบทขวา แล้วตัดสินว่าพยัญชนะคู่กำกวมที่แสดงไว้ข้างต้นควรจะปรากฏในพยางค์เดียวกันหรือต่างพยางค์กัน

4.1.3 ความกำกวมของเบรลล์เซลล์คู่

ความกำกวมนี้เกิดจากการที่เบรลล์ไทยสองตัวซึ่งอยู่ชิดกันบางคู่สามารถใช้แทนอักษรปกติได้มากกว่า 1 กรณี โดยเกิดขึ้นกับเบรลล์เซลล์คู่ 4 ชุด ได้แก่ ‘⠕⠱’ , ‘⠍⠍’ , ‘⠕⠕’ และ ‘⠍⠕’

4.1.3.1 ความกำกวมของ ‘⠕⠱’

‘⠕⠱’ มีรูปในอักษรปกติ 2 รูป คือ “เาะ” หรือ“อะ” ดังแสดงในตัวอย่างต่อไปนี้

{12} ก. ‘⠃⠕⠛⠕⠱⠗’ (=บ-อ-ก-อ-ะ-ไ-ร) ⟹ บอกอะไร

ข. ‘⠛⠕⠱’ (=ก-เาะ) ⟹ เกาะ

จากข้อ {12} ก. ‘⠕’ ใช้แทน “อ” และ‘⠱’ ใช้แทน “ะ” ส่วนในข้อ {12} ข ‘⠕⠱’ มีรูปเป็น ‘เาะ’ ซึ่งกรณีนี้ ‘⠕⠱’จัดเป็น 1 หน่วย เนื่องจากไม่สามารถตีความแต่ละเซลล์ให้แยกจากกันได้ หากระบบเลือกถ่ายถอด ‘⠕⠱’ ให้เป็นรูปใดรูปหนึ่งจะทำให้ถ่ายถอดถูกต้องเพียงบางส่วนเท่านั้นไม่สามารถครอบคลุมได้ทุกกรณี เช่นหากเลือกถ่ายถอด ‘⠕⠱’ ให้เป็น “เาะ” เสมอ จะได้ผลลัพธ์ของ ‘⠃⠕⠛⠕⠱⠗’ เป็น “*บอ<u>เกาะ</u>ไร*” แต่หากเลือกถ่ายถอด ‘⠕⠱’ ให้เป็น “อะ” ในทุกกรณี จะได้ผลลัพธ์การถ่ายถอด ‘⠛⠕⠱’ เป็น “*ก<u>อะ</u>*”

ดังนั้นหากระบุได้ว่า อักขระเบรลล์หน้า ‘⠅⠕’ นั้นเป็นอักษรสุดท้ายของพยางค์ แสดงว่า ‘⠅⠕’ ปรากฏในตำแหน่งต้นของพยางค์ต่อมาจึงต้องถ่ายถอดให้เป็น “อะ” แต่ในทางตรงกันข้าม หากไม่สามารถบ่งชี้ได้ว่าอักขระข้างหน้า ‘⠅⠕’ เป็นพยัญชนะสุดท้ายของพยางค์ แสดงว่า ‘⠅⠕’ เป็นส่วนหนึ่งของพยางค์ จึงต้องถ่ายถอดให้เป็น "เาะ" ดังแสดงในตัวอย่างต่อไปนี้

{13} ก. ‘⠃⠕⠈⠅⠕⠗⠁⠊’ (= บ-อ-ก-อ-ะ-ไ-ร) ⟹ บอกอะไร

ข. ‘⠓⠧⠺⠗⠁⠴’ (= ห- ัว-ร-เาะ) ⟹ หัวเราะ

จากตัวอย่างที่ {13} ก. ‘⠅⠕’ ปรากฏต้นพยางค์เนื่องจากบริบทซ้ายของเบรลล์สองตัวนี้เป็นพยางค์ที่สมบูรณ์คือ ‘⠃⠕⠈’ (=บอก) ปรากฏอยู่ จึงต้องถ่ายถอดเบรลล์คู่นี้ให้เป็น “อะ” แต่อีกนัยหนึ่ง ในข้อ {13} ข. ‘⠓⠧⠺⠗’ (=หัวร) ไม่ใช่พยางค์ที่สมบูรณ์ ระบบจึงต้องนำ "เาะ" มาเป็นผลลัพธ์ของการถ่ายถอดอักษร ส่วนการตรวจสอบว่าบริบทซ้ายเป็นพยางค์ที่สมบูรณ์หรือไม่ต้องอาศัยการตรวจสอบด้วยกฎหรือแบบจำลองเอ็นแกรม

4.1.3.2 ความกำกวมของ ‘⠈⠈’

‘⠈⠈’ สามารถตีความเป็น “์ท” หรือ “ธ” ได้ จะเห็นว่า ‘⠈⠈’ จะใช้แทน“์ท”ได้ ก็ต่อเมื่อ “์”ปรากฏที่ท้ายพยางค์หน้า และ “ท” ปรากฏในตำแหน่งต้นของพยางค์ถัดไป แต่หากไม่เป็นดังที่กล่าวมา ก็ต้องตีความ เช่นนั้น ‘⠈⠈’ ให้เป็น “ธ” ดังแสดงในตัวอย่างต่อไปนี้

{14} ก. ⠉⠁⠝⠞⠺⠊⠈⠈⠙⠺⠊ (= จ- ั -น-ท-ร- ์-ท- ี-่) ⟹ จันทร์ที่

ข. ⠏⠺⠜⠈⠈⠁⠞⠥ (= พ-ร-ะ-ธ-า-ต- ุ) ⟹ พระธาตุ

จากตัวอย่างข้อ{14}ก. ต้องตีความ ‘⠈⠈’ให้เป็น “์ท” เนื่องจาก “์” ปรากฏเป็นอักษรตัวสุดท้ายของพยางค์ “จันทร์” และ “ท” ปรากฏเป็นพยัญชนะต้นของพยางค์ “ที่” แต่ในทางตรงกันข้าม ในข้อ {14}ข. “์” ไม่ใช่อักษรสุดท้ายของ “พระ” และ “ท” ไม่ใช่อักษรตัวแรกของพยางค์ “ธาตุ” จึงต้องถ่ายถอดเบรลล์คู่นี้ให้เป็น “ธ”

4.1.3.3 ความกำกวมของ ‘⠹⠜’

‘⠹⠜’ สามารถตีความเป็น “รๆ” หรือ “ฤ” ได้ การที่จะตีความให้เป็นกรณีใดนั้น จะต้องพิจารณาบริบทซ้ายที่ปรากฏร่วมกับอักขระเบรลล์สองตัวนี้ ดังตัวอย่างต่อไปนี้

{15} ก. ‘⠉⠗⠹⠜’ (=ค-ไ-ร-ๆ) ⟹ ใคร ๆ

ข. ‘⠹⠜⠚⠁’ (=ฤ-ช-า) ⟹ ฤชา

จากตัวอย่างข้อ {15}ก. เมื่อตีความ ‘⠐⠗’ เป็น “รๆ” แล้วนำบริบททางซ้ายคือ ‘⠐⠗ ⠐⠗ ⠐⠗ มาพิจารณาร่วมด้วย จะพบว่า “ใคร” เป็นพยางค์ที่สมบูรณ์ จึงต้องยึด “รๆ” เป็นผลลัพธ์ของการถ่ายถอด แต่ในทางตรงกันข้ามในข้อ {15} ข. หากตีความ ‘⠐⠗’ ให้เป็น “รๆ” แล้วนำ “ร” พร้อมบริบททางซ้ายมาพิจารณาว่าเป็นพยางค์ที่สมบูรณ์หรือไม่ ทำให้ต้องถ่ายถอด ‘⠐⠗’ให้เป็น “ฤ” เนื่องจาก “ร”ไม่ใช่พยางค์ที่สมบูรณ์

4.1.3.3 ความกำกวมของ ‘⠨⠝’

‘⠨⠝’ สามารถตีความได้ 2 กรณีคือ “ษ” หรือ “-ส” การที่จะตีความ ‘⠨⠝’ ให้เป็นกรณีใดนั้นจะต้องอาศัยบริบทที่ปรากฏร่วมกับอักขระเบรลล์คู่นี้ ซึ่งต้องดำเนินการต่างกันในการถ่ายถอดโดยการใช้กฎ และการใช้แบบจำลองเอ็นแกรม ซึ่งจะได้อภิปรายในบทที่ 5 ต่อไป

ในหัวข้อนี้ ผู้วิจัยได้กล่าวถึงแนวคิดที่ใช้สำหรับการออกแบบระบบการถ่ายถอดเบรลล์ไทยเท่านั้น ซึ่งแนวคิดเหล่านี้ไม่สามารถนำไปใช้กับการถ่ายถอดเบรลล์อังกฤษได้ เนื่องจากทั้งสองภาษามีลักษณะทั่วไปแตกต่างกัน แม้ว่าแต่ละภาษาจะใช้ระบบ 6 จุดมาแทนตัวอักษรเช่นเดียวกันก็ตาม เพื่อให้ระบบสามารถถ่ายถอดเบรลล์อังกฤษได้ ซึ่งเป็นวัตถุประสงค์หลักข้อหนึ่งของงานวิจัยฉบับนี้ ผู้วิจัยจึงได้สร้างระบบการถ่ายถอดเบรลล์อังกฤษขึ้นโดยอาศัยแนวคิดต่าง ๆ ดังจะได้นำเสนอในหัวข้อต่อไป

4.2 แนวคิดที่ใช้ในการออกแบบและพัฒนาระบบการถ่ายถอดเบรลล์อังกฤษ

ในการออกแบบระบบการถ่ายถอดเบรลล์อังกฤษ ผู้วิจัยได้นำแนวคิดที่สำคัญในอักษรเบรลล์ มาใช้ในการพัฒนากระบวนการต่าง ๆ เพื่อใช้ในการถ่ายถอดเบรลล์อังกฤษรูปเต็มและเบรลล์อังกฤษรูปย่อ ซึ่งจะขออภิปรายดังนี้

4.2.1 เบรลล์อังกฤษรูปเต็ม เบรลล์อังกฤษรูปย่อ และคอมพิวเตอร์เบรลล์

ดังที่ได้กล่าวมาในส่วนของการทบทวนวรรณกรรมว่า เบรลล์อังกฤษที่คนตาบอดทั่วโลกนิยมใช้มีอยู่ 2 ระบบ คือ เบรลล์อังกฤษรูปเต็ม (English Braille Grade 1) และเบรลล์อังกฤษรูปย่อ (English Braille Grade 2)

เบรลล์อังกฤษรูปเต็ม เป็นระบบที่ตัวอักษรเบรลล์และอักษรปกติสอดคล้องกันแบบหนึ่งต่อหนึ่งเป็นส่วนใหญ่ ในขณะที่*เบรลล์อังกฤษรูปย่อ*จะใช้เบรลล์ 1 เซลล์ หรือ 1 กลุ่ม แทนอักษรปกติในจำนวนที่มากกว่า ซึ่งแสดงดังตัวอย่างต่อไปนี้

{16} ‘⠽ ⠉ ⠙ ⠞’ อังกฤษรูปเต็ม ⟹ y c d t

อังกฤษรูปย่อ ⟹ you can do that

จากตัวอย่างจะพบว่า ตัวเขียนทั้ง 2 ระบบนี้ เป็นตัวเขียนที่แยกจากกันอย่างเห็นได้ชัด เมื่ออักขระเบรลล์ตัวหนึ่งปรากฏในระบบตัวเขียนเบรลล์ต่างกัน ก็จะตีความให้แตกต่างกัน

อย่างไรก็ตาม ใช่ว่าตัวเขียนทั้ง 2 ระบบนี้จะแตกต่างกันโดยสิ้นเชิง จากการวิเคราะห์ข้อความที่เขียนด้วยเบรลล์รูปเต็มและเบรลล์รูปย่อ ผู้วิจัยพบว่าเบรลล์ 2 ประเภทนี้มีความเหมือนและความแตกต่างกัน

สิ่งที่เบรลล์รูปเต็มและเบรลล์รูปย่อมีลักษณะเหมือนกัน ประการที่สำคัญที่สุด คือ ตัวเขียนทั้ง 2 ระบบนี้ต่างก็มีตัวอักษรครบทั้ง 26 ตัวเช่นกัน และตัวอักษรเหล่านี้สามารถจะนำมาเรียงต่อกันเพื่อให้เป็นคำในภาษาอังกฤษได้ หากเขียนคำบางคำ เช่น ‘⠃ ⠊ ⠗ ⠙’ (= bird) ‘⠉ ⠁ ⠞’ (= cat) ‘⠙ ⠕ ⠛’ (=dog) โดยไม่มีคำอื่น ๆ ปรากฏร่วมด้วย ก็ยากที่จะระบุได้ว่าใช้ตัวเขียนเบรลล์ระบบใด เพราะคำเหล่านี้สามารถปรากฏได้ทั้งในข้อความเบรลล์รูปเต็มและข้อความเบรลล์รูปย่อ เนื่องจากคำเหล่านี้ไม่สามารถนำอักขระเบรลล์รูปย่อมาเขียนแทนได้

ในเบรลล์อังกฤษ ไม่สามารถเขียนอักษรตัวพิมพ์ใหญ่โดยใช้อักขระเบรลล์เพียงตัวเดียวได้ หากต้องการเขียนอักษรตัวพิมพ์ใหญ่ต้องใช้เครื่องหมาย ‘⠠’ นำหน้าตัวอักษรเพื่อแสดงว่าอักษรตัวแรกเป็นอักษรตัวพิมพ์ใหญ่ และ ‘⠠ ⠠’ เพื่อแสดงว่า อักษรที่ตามหลังเบรลล์คู่นี้เป็นตัวพิมพ์ใหญ่ทั้งหมด การใช้เครื่องหมายแสดงตัวพิมพ์ใหญ่นี้จะเกิดขึ้นในทั้งเบรลล์อังกฤษรูปเต็มและเบรลล์อังกฤษรูปย่อ

ลักษณะร่วมที่เบรลล์อังกฤษรูปเต็มและเบรลล์อังกฤษรูปย่อมีร่วมกันอีกประการหนึ่ง คือ ตัวเขียนเบรลล์ทั้ง 2 ประเภทนี้ต่างก็ใช้เครื่องหมาย (punctuations) กลุ่มเดียวกันและเครื่องหมายเหล่านี้ต้องตีความจากตำแหน่งการปรากฏ จึงจะระบุได้ว่าเป็นเครื่องหมายใด เช่น ‘⠶’ เมื่อปรากฏหน้าคำจะใช้แทน “(” แต่หากปรากฏหลังคำจะใช้แทน “)” ‘⠦’เมื่อปรากฏหน้าคำจะใช้แทน “ “ ” (เครื่องหมายคำพูดเปิด) และเมื่อปรากฏหลังคำจะใช้แทนเครื่องหมาย “?” เป็นต้น เครื่องหมายเหล่านี้ได้แสดงไว้ในตารางคู่เทียบเบรลล์ต้นสายอักขระ (ตารางที่ 5.15 เฉพาะส่วนที่เป็นเครื่องหมายเท่านั้น) และตารางคู่เทียบเครื่องหมายหลังคำ (ตารางที่ 5.16)

นอกจากนี้ ในการเขียนตัวเลขในเบรลล์อังกฤษรูปเต็มและเบรลล์อังกฤษรูปย่อก็ใช้อักขรวิธีเดียวกัน คือ นำอักษรภาษาอังกฤษตั้งแต่ ‘⠁-⠚’ (a-j) มาใช้แทนตัวเลข 1-0 โดยเขียน

เครื่องหมาย‘⠼’ เพื่อกำกับตัวเลขไว้ข้างหน้าอักษรเหล่านี้เช่นเดียวกัน เช่น ‘⠼⠙’ (เครื่องหมายนำเลข + d) เพื่อใช้แทนเลข “4” และ ‘⠼⠃⠊⠛’ (เครื่องหมายนำเลข + big) เพื่อแสดงเลข “297”

หากพิจารณาลักษณะที่คล้ายคลึงของระบบตัวเขียนเบรลล์ทั้ง 2 ประเภทที่กล่าวมานี้ ก็อาจดูเหมือนว่าเบรลล์ 2 ประเภทนี้น่าจะเป็นตัวเขียนประเภทเดียวกัน แต่เมื่อนำตัวเขียนทั้ง 2 ประเภทมาเปรียบเทียบกันจะพบว่า เบรลล์อังกฤษรูปเต็มเป็นระบบที่มีตัวอักษรอยู่จำกัด คือมีตัวอักษรอังกฤษเพียง 26 ตัวและเครื่องหมายต่าง ๆ ในปริมาณที่จำกัด นอกจากนี้ อักขระเบรลล์แต่ละตัวยังสามารถใช้แทนอักษรปกติได้แบบตัวต่อตัวเสียเป็นส่วนใหญ่ เช่น ‘⠃⠊⠗⠙’ (=bird) เทียบกับอักษรปกติได้ กล่าวคือ ‘⠃’ เทียบได้กับ “b” ‘⠊’ เทียบได้กับ “i” และ ‘⠗’ เทียบได้กับ “r” เป็นต้น สำหรับอังกฤษรูปย่อนั้น ต้องมีอักขระเบรลล์จำนวนหนึ่งที่ใช้แทนอักษรปกติในปริมาณที่มากกว่านั้น เช่น อักขระเบรลล์ 1 ตัวใช้แทนอักษรปกติ 2 ตัว (เช่น ‘⠻’ ใช้แทน “er”) อักขระเบรลล์ 2 ตัวใช้แทนคำ 5 ตัวอักษร (เช่น ‘⠰⠝’ ใช้แทน “ation”) เป็นต้น

อย่างไรก็ตาม ข้อสรุปเกี่ยวกับเบรลล์รูปเต็มและเบรลล์รูปย่อที่กล่าวมาข้างต้นอาจไม่เป็นจริงเสมอไป หากพิจารณาตัวย่อเบรลล์ 2 ชุดคือ ‘⠰⠝’ และ ‘⠰⠽’ จะพบว่า อักขระ 2 ชุดนี้ปรากฏได้ทั้งในเบรลล์รูปเต็มและเบรลล์รูปย่อ เมื่ออักขระชุดเหล่านี้ปรากฏในตำแหน่งต้นคำต้องตีความให้เป็น “N” และ “Y” ในตัวเขียนทั้ง 2 ระบบตามลำดับ แต่เมื่อปรากฏในตำแหน่งอื่นจะต้องตีความให้แตกต่างกัน กล่าวคือ ในเบรลล์รูปเต็มต้องตีความให้เป็น “N” และ “Y” ในขณะที่ในเบรลล์รูปย่อต้องตีความให้เป็น “ation” และ “ally” เนื่องจากอักขระเบรลล์ 2 ชุดนี้เป็นตัวย่อหน่วยเสริม (ดูตารางที่ 2.16 ประกอบ) ดังตัวอย่างต่อไปนี้

{17} ก. ⠝⠰⠝ เบรลล์รูปย่อ = nation
*เบรลล์รูปเต็ม = nN
ข. ⠰⠎⠰⠽ เบรลล์รูปย่อ = Sally
*เบรลล์รูปเต็ม = SY

จากตัวอย่าง จำเป็นต้องตีความ 2 คำนี้ให้ปรากฏในเบรลล์อังกฤษรูปย่อจึงจะสื่อความได้ ดังนั้นในการถ่ายถอดอักษร จึงต้องมีระบบ 2 ระบบ คือระบบหนึ่งใช้ถ่ายถอดเบรลล์รูปเต็ม และอีกระบบหนึ่งใช้ถ่ายถอดเบรลล์รูปย่อ

หากพิจารณากระบวนการถ่ายถอดเบรลล์รูปย่อที่ผู้วิจัยได้พัฒนาขึ้นซึ่งจะได้นำเสนอในบทต่อไป จะพบว่า ต้องมีกระบวนการถ่ายถอดเบรลล์รูปเต็มเป็นองค์ประกอบของกระบวนการถ่ายถอดเบรลล์รูปย่อด้วยเช่นกัน เนื่องจากคำบางคำไม่สามารถตีความให้เป็นเบรลล์รูปย่อได้ แต่จำเป็นต้องตีความเสมือนว่าคำนั้นเป็นเบรลล์รูปเต็ม จึงจะสื่อความได้ ดังแสดงในตัวอย่างต่อไปนี้

{18} ⠏⠍ ⠏⠙⠍⠼⠙ เบรลล์รูปเต็ม = p.m.
เบรลล์รูปย่อ = pddm4

จากตัวอย่าง เมื่อ ‘⠏⠙’ ปรากฏที่ตำแหน่งกลางคำ ในเบรลล์รูปย่อต้องตีความเป็น “dd” ทำให้ได้คำว่า “pddm.” ซึ่งเป็นคำที่เป็นไปไม่ได้ในภาษาอังกฤษ ในทางตรงกันข้ามต้องตีความคำนี้ให้ปรากฏเป็นเบรลล์รูปเต็ม เนื่องจาก ‘⠏⠙’ ในเบรลล์รูปเต็มใช้แทน “.” ไม่ว่าจะอยู่ในบริบทใดก็ตาม ด้วยเหตุนี้จึงต้องมีส่วนของการถ่ายถอดเบรลล์รูปเต็มในกระบวนการถ่ายถอดเบรลล์รูปย่อด้วยเช่นกัน

นอกจากนี้ ยังมีอักขระเบรลล์อีกตัวหนึ่งที่เป็นปัญหาต่อการถ่ายถอดในเบรลล์รูปย่ออย่างยิ่งคือ ‘⠻’ เมื่ออักขระตัวนี้ปรากฏในอังกฤษรูปย่อและไม่ใช่ต้นวรรคอาจตีความให้ใช้แทน “ble” หรือตีความเป็นเครื่องหมายนำเลข ได้ (นั่นคือ ต้องตีความอักขระเบรลล์ “a-j” ที่ตามหลังเครื่องหมายนี้ให้เป็นตัวเลข) ดังตัวอย่างต่อไปนี้

{19} ก. ⠁⠻⠙ = a4
= * abled
ข. ⠙⠁⠻⠙ = disabled
= * disa4

จากตัวอย่างในข้อ{19} ก. ต้องตีความให้ ‘⠻’ ใช้กำกับตัวเลขและต้องตีความให้‘⠻’ ในข้อ {19} ข. ใช้ย่อ “ble”

เพื่อความสะดวกในการถ่ายถอดอักษร ผู้วิจัยได้นำกระบวนการถ่ายถอดเบรลล์รูปเต็มไปใช้กับการถ่ายถอดตัวเลขเบรลล์ด้วยเช่นกัน เนื่องจากการถ่ายถอดตัวเลขเบรลล์มีลักษณะคล้ายคลึงกับการถ่ายถอดเบรลล์รูปเต็ม คือมีการจับคู่ตัวอักษรแบบหนึ่งต่อหนึ่งเป็นส่วนใหญ่ แต่ต่างกันที่ต้องแปลงจากอักษรอังกฤษ “a-j” เป็นตัวเลข “1-0”

เท่าที่ได้อภิปรายมาแล้ว จะพบว่า การที่จะตีความอักขระเบรลล์ให้ใช้แทนอักขระใดในอักษรปกตินั้นต้องพิจารณาว่าอักขระเบรลล์นั้นปรากฏในเบรลล์อังกฤษรูปเต็มหรือเบรลล์อังกฤษรูปย่อ ซึ่งในบางกรณีทำให้เกิดความสับสนต่อผู้ใช้อักษรเบรลล์ เช่น ‘⠻’ ที่สามารถตีความให้เป็นเครื่องหมายนำเลขหรือใช้แทน “ble” ในอักษรรูปย่อได้ ขึ้นอยู่กับว่าอักขระเบรลล์ตัวนี้จะปรากฏในบริบทใด

เพื่อให้เกิดความสะดวกในการใช้อักษรเบรลล์โดยไม่ต้องพึ่งพาบริบท จึงได้มีการนำอักขระเบรลล์แต่ละตัวมาใช้แทนอักขระแอสกีนั้นโดยตรงแบบหนึ่งต่อหนึ่ง ซึ่งเป็นที่รู้จักกันในกลุ่มผู้ใช้อักษรเบรลล์ทั่วโลกว่าเป็น “*คอมพิวเตอร์เบรลล์*” เช่น ‘⠻’ ซึ่งแทนด้วยอักขระแอสกีคือ

“#” ในคอมพิวเตอร์ หากผู้ใช้อักษรเบรลล์ตระหนักว่ากำลังอ่านอักษรเบรลล์ที่เป็นรหัสคอมพิวเตอร์เบรลล์ก็จะต้องตีความให้ ‘⠼’ เป็น “#” ในอักษรปกติ จะตีความแตกต่างออกไปจากนี้ไม่ได้

ในข้อความเบรลล์ทั่วไป รหัสคอมพิวเตอร์เบรลล์มักจะใช้ในการเขียนที่อยู่ของเว็บไซต์หรือที่อยู่อีเมล์ โดยมักจะเขียนครอบด้วย ‘⠸⠶’ (เครื่องหมายเปิดรหัสคอมพิวเตอร์เบรลล์) และ ‘⠸⠴’ (เครื่องหมายปิดรหัสคอมพิวเตอร์เบรลล์) เช่น‘⠸⠶⠓⠞⠞⠏⠒⠌⠌⠁⠉⠂⠚⠂⠙⠂⠉⠕⠍⠌⠸⠴’ (=เครื่องหมายเปิดรหัสคอมพิวเตอร์เบรลล์-h-t-t-p-:-/-/-a-c-1-0-4-.-c-o-m-เครื่องหมายปิดรหัสคอมพิวเตอร์เบรลล์) นอกจากนี้ รหัสคอมพิวเตอร์เบรลล์ยังเป็นประโยชน์โดยตรงต่อนักเขียนโปรแกรมผู้พิการทางสายตาที่ต้องใช้อักษรเบรลล์ในการพัฒนาโปรแกรมต่าง ๆ เพราะจะทำให้สามารถอ่านคำสั่งคอมพิวเตอร์แบบตัวต่อตัวได้โดยไม่ต้องตีความจากอักขระเบรลล์แวดล้อม

4.2.2 แนวคิดที่ใช้ในการถ่ายถอดเบรลล์อังกฤษรูปย่อ

ในการพัฒนากระบวนการถ่ายถอดเบรลล์อังกฤษรูปย่อนั้น ผู้วิจัยได้อาศัยแนวคิดต่าง ๆ ดังนี้

4.2.2.1 การตีความตัวย่อจากตำแหน่งที่ปรากฏ

ในการเขียนเบรลล์รูปย่อ เบรลล์เซลล์เดี่ยวบางประเภทอาจใช้แทนอักษรปกติได้หลายกรณี การที่จะระบุได้ว่าเบรลล์แต่ละเซลล์ ใช้แทนตัวอักษรหรือกลุ่มตัวอักษรใด จำเป็นต้องตรวจสอบตำแหน่งที่อักขระเบรลล์ตัวนั้นปรากฏอยู่ เช่น ‘⠸⠸’ หากปรากฏหน้าคำจะใช้แทน “(” ที่ตำแหน่งกลางคำจะใช้ย่อ “gg” ที่ตำแหน่งหลังคำจะหมายถึง ‘)’ หรือหากปรากฏเป็นคำโดดจะใช้ย่อ ‘were’ เป็นต้น

ด้วยเหตุที่ตัวย่อเบรลล์อังกฤษในกลุ่มนี้ สามารถตีความในบริบทต่างกันได้ ผู้วิจัยจึงได้ใช้แนวคิดนี้ มาพัฒนาระบบเพื่อให้รองรับกับการถ่ายถอดอักขระเบรลล์กลุ่มนี้ในกรณีต่าง ๆ ดังนี้

4.2.2.1.1 อักษรเบรลล์ต้นสายอักขระ

จากการวิเคราะห์การใช้ตัวย่อเบรลล์อังกฤษ ผู้วิจัยพบว่ามีอักขระเบรลล์ 2 กลุ่ม ที่ต้องปรากฏในตำแหน่งหน้าคำและมีลักษณะบางประการคล้ายคลึงกันจึงจำเป็นต้องจัดให้อยู่ในประเภทเดียวกันเพื่อความสะดวกต่อการถ่ายถอดอักษร อักขระเบรลล์ 2 กลุ่มนี้คือ *ตัวย่อเบรลล์ต่ำย่อคำโดดพิเศษ* และ *เครื่องหมายหน้าคำ*

(1) ตัวย่อเบรลล์ต่ำย่อคำโดดพิเศษ ได้แก่ ‘⠐’ (=to) ‘⠐⠔’ (=into) ‘⠆’ (=by) ซึ่งตัวย่อเหล่านี้จะเขียนติดกับคำต่อไปโดยไม่มีเครื่องหมายเว้นวรรค (ดูหัวข้อ 2.1.1.2.1 (6) ข. ประกอบ)

(2) เครื่องหมายหน้าคำประกอบด้วย ‘⠣’ (= ( ) ‘⠦’ (= “) ‘⠠⠦’ (= ‘) ‘⠠⠤’ (= _) และ ‘⠠⠣’ (= [ )

อักขระเบรลล์ทั้ง 2 กลุ่มที่กล่าวมานี้มีลักษณะร่วมกัน คือ อักขระเบรลล์ที่ตามมาจะถือว่าเป็น จุดเริ่มต้นของคำใหม่โดยอัตโนมัติ ตัวอย่างเช่น

{20} ก. ‘⠣⠿’ ⟹ **to** go

ข. ‘⠦⠞’ ⟹ **“**but

จากตัวอย่างจะต้องตีความ ‘⠣⠿’ ใน{20} ก. ให้เป็น “to go” โดยที่ ‘⠣’แทน “to” และ ‘⠿’ แทน “go” นั่นคือต้องตีความเสมือนว่า ‘⠿’ เป็นเบรลล์รูปย่อที่ปรากฏเป็นคำใหม่ ไม่สามารถตีความเป็น “tog” หรือ “to g” ได้[7] ในทำนองเดียวกันต้องตีความ ‘⠦⠞’ ใน {20} ข. ให้เป็น “ “but” โดยที่ ‘⠦’ ใช้แทน “ “ ” และ ‘⠞’ ใช้แทน “but”

ผู้วิจัยยังพบเพิ่มเติมว่า *เครื่องหมายหน้าคำ* อาจปรากฏเพียงตัวเดียว ปรากฏเรียงต่อกัน หรือปรากฏร่วมกับตัวย่อเบรลล์ต่ำย่อคำโดดพิเศษก็ได้ ดังตัวอย่างต่อไปนี้

{21} ก. ‘⠣⠣⠣’ ⟹ **(“**the

ข. ‘⠦⠦⠶’ ⟹ **“**to do

จากตัวอย่าง มี ‘⠣’ = ((วงเล็บเปิด) และ‘⠦’ = “ (คำพูดเปิด) เรียงต่อกันในข้อ {21} ก. นอกจากนี้มี‘⠦’(เครื่องหมายคำพูดเปิด) ตามด้วย ‘⠶’ (to ;ตัวย่อเบรลล์ต่ำย่อคำโดดพิเศษ)ในข้อ {21} ข. ซึ่งในการถ่ายถอดอักษรเบรลล์ต้นสายอักขระนั้นมีวิธีดำเนินการแตกต่างกัน ในวิธีการถ่ายถอดด้วยกฎและวิธีการถ่ายถอดด้วยแบบจำลองเอ็นแกรม

4.2.2.1.2 อักษรเบรลล์ท้ายสายอักขระ

อักษรเบรลล์ท้ายสายอักขระ เช่น ‘⠠⠂’ (=,)‘⠰⠆’(=;) ‘⠒⠒’(=:) ‘⠲⠲’(=.) ‘⠖⠖’(=!) ‘⠶⠶’(=)) (ดูตารางที่ 5.16 ประกอบ)

อักษรเบรลล์ในตัวอย่างข้อ {20} และ{21} ข้างต้น สามารถปรากฏในตำแหน่งท้ายสายอักขระ ได้เช่นกัน ซึ่งจะต้องตีความให้แตกต่างออกไป (ดูหัวข้อ 2.1.1.2.1 (6) ค.) ดังตัวอย่างต่อไปนี้

{22} ก. ‘⠠⠒⠒’ ⟹ about**?**

ข. ‘⠠⠒⠶’ ⟹ about**)**

[7] ในการถ่ายถอดตัวย่อเบรลล์ต่ำย่อคำโดดพิเศษ ต้องมีช่องว่างตามหลังอักษรปกติที่ถ่ายถอดแล้วเสมอ

จากตัวอย่างข้อ {22} ต้องตีความให้ ‘⠦’ เป็น “?” และ ‘⠴’ เป็น “)” เนื่องจากเครื่องหมายเหล่านี้ ต่างปรากฏในตำแหน่งท้ายคำ นอกจากเครื่องหมายกลุ่มนี้จะปรากฏในสายอักขระเพียงตัวเดียวดังปรากฏในตัวอย่างแล้ว เครื่องหมายประเภทนี้ยังสามารถปรากฏเรียงต่อกันได้เช่นเดียวกัน ดังตัวอย่างต่อไปนี้

{23} ก. ‘⠠⠦⠴’ ⟹ more**?)**

ข. ‘⠠⠠⠦⠦’ ⟹ have**?”**

เมื่อเปรียบเทียบ เครื่องหมายหลังคำในข้อ {22}และ {23} จะพบว่าเครื่องหมายที่แสดงไว้ในตัวอย่างเหล่านี้มีรูปซ้ำกับข้อ{21} แต่ปรากฏในตำแหน่งต่างกันจึงใช้แทนเครื่องหมายที่ต่างกัน คือข้อ {21} ‘⠦’ เมื่อปรากฏที่หน้าคำใช้แทน “**“**” แต่เมื่อปรากฏหลังคำจะใช้แทน “**?**” ดังข้อ{23}

4.2.2.1.3 อักษรเบรลล์กลางสายอักขระ

อักษรเบรลล์กลางสายอักขระ คือ อักษรที่ปรากฏระหว่างอักษรเบรลล์ในหัวข้อ 4.2.2.1.2 และ/หรือ 4.2.2.1.3. หรือปรากฏระหว่างช่องว่าง โดยอักษรที่ปรากฏเบรลล์กลางสายอักขระได้ประกอบด้วยตัวย่อ 2 ประเภท คือ (1) ตัวย่อทั่วไป (ซึ่งจะได้อภิปรายในหัวข้อที่ 4.2.3 ต่อไป) และ (2) ตัวย่อที่ต้องตีความตามตำแหน่งที่ปรากฏ

สำหรับตัวย่อในประเภทที่ (2) แบ่งออกเป็น 2 กลุ่มย่อยคือ ตัวย่อต้นคำ และตัวย่อกลางคำ (ดูหัวข้อที่ 2.1.1.2.1 (6) ค) โดยในการถ่ายถอดอักษร ระบบที่พัฒนาขึ้นจะต้องรองรับกับตัวย่อเหล่านี้ ดังรายละเอียดต่อไปนี้

ก. ตัวย่อต้นคำ

ตัวย่อต้นคำ ประกอบด้วย ‘⠆’ (=be) ‘⠒’(=con) และ ‘⠲’(=dis) ซึ่งหากมองโดยผิวเผินดูเหมือนว่าตัวย่อเบรลล์ต่ำย่อคำโดดพิเศษ (ได้แก่ตัวย่อที่ใช้ย่อ to, into, by ตามที่แสดงไว้ใน 4.2.2.1.1 (1) หน้า 97) จะมีความคล้ายคลึงกับตัวย่อต้นคำ เนื่องจากตัวย่อทั้ง 2 กลุ่มนี้ปรากฏต้นสายอักขระได้เช่นเดียวกัน แต่ในความเป็นจริง ตัวย่อ 2 กลุ่มนี้มีความแตกต่างกันคือ อักษรที่ตามหลังตัวย่อเบรลล์ต่ำย่อคำโดดพิเศษจะถือว่าเป็นจุดเริ่มต้นของคำใหม่ ในขณะที่ตัวย่อต้นคำเป็นอักษรเบรลล์ที่ใช้แทนหน่วยคำประเภทอุปสรรค (prefix) ซึ่งเป็นส่วนหนึ่งของคำ นอกจากนี้อักษรทั้ง 2 กลุ่มนี้ยังมีรหัสอักษรเบรลล์ที่แตกต่างกัน

ข. ตัวย่อกลางคำ

ตัวย่อกลางคำ ได้แก่ตัวย่อต่อไปนี้ ‘⠂’ (=ea)‘⠆’(=bb) ‘⠒’(=cc) ‘⠶’(=dd ‘⠖’(=ff) ‘⠤’(=gg) ตัวย่อกลุ่มนี้มีรหัสอักษรเบรลล์ตรงกับเครื่องหมายส่วนหนึ่งที่ปรากฏหลังคำตามหัวข้อ

4.2.2.1.2 อย่างไรก็ตาม จากการศึกษาครั้งนี้ผู้วิจัยพบว่าอักขระเบรลล์ที่ปรากฏกลางสายอักขระบางตัว สามารถตีความได้สองกรณีคือตีความเป็นเครื่องหมาย หรือตัวย่อกลางสายอักขระ ซึ่งพบในอักขระเบรลล์ 2 ตัว ดังนี้

(1) ความกำกวมของ‘⠸⠸’

‘⠸⠸’ ที่ปรากฏกลางสายอักขระ อาจตีความให้เป็น ตัวย่อกลางคำ หรือ เครื่องหมายวงเล็บ ดังตัวอย่างต่อไปนี้

{24} ก. ‘⠋⠊⠇⠑⠤⠎⠤⠸⠸’ (= f-i-l-e-(-s-)) ⇨ file(s)

*ข. ‘⠋⠊⠇⠑⠸⠸⠎⠸⠸’ ⇨ fileggs)

ค. ⠎⠥⠸⠸⠑⠎⠞⠊⠕⠝⠤⠸⠸ (= s-u-gg-e-s-tion-)) ⇨ suggestion)

จากตัวอย่างข้อ {24} หากยึดตามตำแหน่งการปรากฏ จะต้องถ่ายถอด ‘⠸⠸’ ที่กลางสายอักขระ ให้เป็น “*gg*” ทำให้ได้ผลลัพธ์ดังที่ปรากฏในข้อ {24} ข. “fileggs)” ซึ่งไม่ถูกต้อง

(2) ความกำกวมของ ‘⠌’

‘⠌’ ในตำแหน่งกลางสายอักขระที่อาจตีความได้เป็น “st”หรือ “/” จากการวิเคราะห์การปรากฏของ ‘⠌’ ผู้วิจัยพบว่า เมื่อ ‘⠌’ ปรากฏในตำแหน่งต้นคำหรือท้ายคำ จะต้องตีความให้เป็น “st” เสมอ จึงไม่เป็นปัญหาต่อการถ่ายถอด ดังตัวอย่างต่อไปนี้

{25} ก. ‘⠌⠕⠏’ (= st-o-p) ⟹ stop

ข. ‘⠏⠕⠌’ (= p-o-st) ⟹ post

อย่างไรก็ตามเมื่อ ‘⠌’ ปรากฏกลางสายอักขระ (หรือวรรค) อาจตีความเป็น “st” หรือ “/” ด้วยเหตุนี้จึงไม่สามารถเลือกถ่ายถอดให้เป็นกรณีใดกรณีหนึ่งได้โดยเฉพาะ ดังตัวอย่างข้อ {26}

{26} ก. ‘⠔⠌⠁⠇⠇’ (= in-st-a-l-l) ⟹ install

ข. ‘⠓⠊⠍⠤⠸⠓⠑⠗’(=him-/-h-er) ⟹ him/her

จากตัวอย่างข้อ{26}ก. จะต้องถ่ายถอด ‘⠌’ให้เป็น “st” ในขณะที่ต้องถ่ายถอด ‘⠌’ในข้อ {26}ข. ให้เป็น “/” ซึ่งการที่จะถ่ายถอด ‘⠌’ ให้ถูกต้องได้นั้นต้องอาศัยบริบทซ้าย-ขวา ที่ปรากฏร่วมกับอักขระเบรลล์ตัวนี้

จากที่ได้กล่าวมาในหัวข้อนี้ พอสรุปได้ว่า อักขระเบรลล์แต่ละตัวในบางประเภทต้องอาศัยการตีความตามตำแหน่งที่ปรากฏ ผู้วิจัยได้นำแนวคิดนี้มาใช้ในการพัฒนาระบบการถ่ายถอดเบรลล์ไทยปนอังกฤษด้วยวิธีการต่าง ๆ ระบบที่ผู้วิจัยได้พัฒนาขึ้นนี้ สามารถรองรับกับอักขระเบรลล์ที่ได้เสนอมาในหัวข้อนี้ครบทุกกรณี ดังจะได้นำเสนอในบทต่อ ๆ ไป

4.2.2.2 การตีความเบรลล์อังกฤษรูปย่อจากส่วนใหญ่มาหาส่วนย่อย

โดยทั่วไปในการตีความเบรลล์อังกฤษรูปย่อนั้น ต้องนำเบรลล์หลายเซลล์มาประกอบกันแล้วพิจารณาว่า สามารถตีความเป็นเบรลล์รูปย่อได้หรือไม่ หากไม่ได้ ต้องทำการลดจำนวนเซลล์ลงทีละหนึ่ง จนกว่าจะตีความได้ หากเริ่มตีความจากหนึ่งเซลล์ก็จะทำให้การตีความไม่ถูกต้อง ดังตัวอย่างต่อไปนี้

{27} ก. ‘⠿⠗⠎’ (=f-r-s) ⟹ friends

* ข. ‘⠿⠗⠎’ (=f-r-s) ⟹ frs

ในการตีความ ‘⠿⠗⠎’ ในข้อ {27} ก. นั้น ต้องเริ่มจากการตีความอักขระทั้ง 3 เซลล์ ซึ่งพบว่า ไม่ตรงกับรูปย่อมาตรฐานใดในเบรลล์อังกฤษ ดังนั้นจึงต้องลดเซลล์ทางขวาลงอีก 1 เซลล์ จึงเหลือ 2 เซลล์คือ ‘⠿⠗’ (=fr) ซึ่งใช้เป็นคำย่อของ “friend” ยังเหลือ ‘⠎’ ที่ยังไม่ได้ตีความ ซึ่งต้องตีความให้เป็น “s” เนื่องจากเป็นอักษรที่ปรากฏเป็นส่วนหนึ่งของคำ ดังนั้น ‘⠿⠗⠎’ จึงใช้ย่อ “friends” การตีความคำนี้จะไม่ถูกต้องหากตีความเริ่มจากหนึ่งเซลล์ แล้วเพิ่มจำนวนเซลล์ไปเรื่อย ๆ เพราะจะทำให้ได้ผลลัพธ์ที่ผิดตามที่ปรากฏในข้อ {27} ข.

ในการถ่ายถอดตัวย่อเบรลล์อังกฤษ ผู้วิจัยจึงอาศัยแนวคิดการตีความตัวย่อจากส่วนใหญ่ลงมาหาส่วนย่อยมาใช้ในการพัฒนาระบบ โดยจะเริ่มจาก การนำสายอักขระเบรลล์อังกฤษ มาพิจารณาทั้งสายอักขระ หากไม่ตรงกับตัวย่อมาตรฐาน ก็จะลดจำนวนเซลล์ทางขวาลงเรื่อย ๆ จนกว่าจะตรงกับ ตัวย่อระดับคำ ตัวย่อเบรลล์เซลล์คู่ หรือตัวย่อเบรลล์เซลล์เดี่ยว กลุ่มใดกลุ่มหนึ่งเป็นอันดับแรก จึงจะทำการถ่ายถอดเบรลล์ที่พบว่าตรงกับตัวย่อมาตรฐานทั้งสามกลุ่มนี้ แล้วดำเนินการกับเซลล์อื่น ๆ ถัดไป

การตีความตัวย่อจากส่วนใหญ่มาหาส่วนย่อย แสดงตามลำดับได้ดังนี้

ตัวย่อระดับคำ ⟶ ตัวย่อเบรลล์เซลล์คู่ ⟶ ตัวย่อเบรลล์เซลล์เดี่ยว

4.2.2.2.1 ตัวย่อระดับคำ

บางครั้ง ในการตีความว่าตัวย่อใดใช้แทนอักษรปกติตัวใดนั้นจะต้องพิจารณาตัวอักษรที่ปรากฏทั้งคำ กล่าวคือ กลุ่มตัวอักษรนั้นต้องปรากฏ ระหว่างช่องว่าง เครื่องหมายหน้าคำ และ/หรือ เครื่องหมายหลังคำ ดังตัวอย่างต่อไปนี้

{28} ก. ‘ ⠁⠃ ’ (= ab) ⟹ about

ข. ‘⠦⠁⠃ ’ ⟹ “about

ค. ‘ ⠠⠁⠃⠥⠞’ ⟹ about”

ง. ‘⠠⠁⠃⠥⠞’ ⟹ “about”

เมื่อพิจารณาตัวอย่างในข้อ{28} ซึ่งมี ‘⠠⠁⠃’ ปรากฏในทุกสายอักขระ จะเห็นว่าตัวอักษรนี้เป็นตัวย่อระดับคำทั้งสิ้น แม้ว่าจะปรากฏในบริบทต่างกันก็ตาม นั่นคือใน{28}ก. ‘⠠⠁⠃’ ปรากฏระหว่างช่องว่าง ข้อ{28}ข. ปรากฏร่วมกับเครื่องหมายหน้าคำและช่องว่าง ข้อ{28}ค. ปรากฏร่วมกับช่องว่างและเครื่องหมายหลังคำ และในข้อ{28}ง. ปรากฏกับเครื่องหมายหน้าคำและเครื่องหมายหลังคำ

ในงานวิจัยฉบับนี้ผู้วิจัยได้กำหนดให้ตัวย่อระดับคำประกอบด้วยตัวย่อ 4 ประเภท (ดังที่ปรากฏในภาคผนวก ข.1) ได้แก่ (1) คำย่อเบรลล์ เช่น ‘ ⠠⠙⠠⠙ ’ มีคู่เทียบเป็น “could” (2) ตัวย่ออักษรโดด เช่น ‘⠠⠅’ มีคู่เทียบเป็น “knowledge” (3) ตัวย่อรูปเต็มกลายรูป เฉพาะส่วนที่เป็นคำโดด เช่น ‘⠠⠉’ มีคู่เทียบเป็น “child” และ (4) ตัวย่อเบรลล์ต่ำย่อคำโดด เช่น ‘⠠⠺’ มีคู่เทียบคือ “were”

เป็นที่น่าสังเกตว่า การถ่ายถอดตัวย่อระดับคำจะไม่เกิดปัญหาใด ๆ หากส่วนของสายอักขระที่กำลังถ่ายถอดตรงกับคำย่อมาตรฐาน เช่น

{29} ก. ‘⠠⠋⠗’ (= **fr**iend) ⟹ “friend”

ข. ‘⠠⠊⠍⠍’ (= **imm**ediate) ⟹ “immediate”

อย่างไรก็ตาม หากระบบต้องถ่ายถอดตัวย่อระดับคำที่มีความซับซ้อนยิ่งขึ้น ซึ่งเกิดจากการนำอักษรเบรลล์อื่น ๆ มาใช้ร่วมกับตัวย่อระดับคำ จึงเป็นปัญหาต่อการถ่ายถอดอักษร ดังตัวอย่าง

{30} ก. ‘⠠⠋⠗⠠⠎’ (=frs; **fr**iend-**s**) ⟹ “friends”

ข. ‘⠠⠊⠍⠍⠇⠠⠽’ (=immly; immediate-**ly**) ⟹ “immediately”

ในตัวอย่างเหล่านี้ ระบบจะนำคำย่อมาตรฐานไปใช้ร่วมกับอักษรเบรลล์อื่น ๆ ซึ่งจะทำให้เกิดข้อผิดในการถ่ายถอดอักษร เนื่องจากระบบมองคำย่อเหล่านี้ว่าเป็นอักขระที่นำมาเรียงต่อกันแทนที่จะเป็นคำย่อ เช่นในตัวอย่างข้อ {30} ข. ‘⠠⠊⠍⠍⠇⠠⠽’ อาจตีความเป็น “immly” แทนที่จะเป็น “immediately”

นอกจากนี้ จากการวิเคราะห์การใช้ตัวย่อกลุ่มนี้ ผู้วิจัยพบว่าบางครั้งในสายอักขระเดียวกันอาจมีตัวย่อระดับคำปรากฏได้มากกว่าหนึ่งตำแหน่ง ดังแสดงในตัวอย่าง

{31} ก. ‘⠠⠃⠑⠠⠯⠠⠁⠋⠠⠻’ (= **bef**ore-**and**-**af**ter) ⟹ before-and-after

* ข. ‘⠠⠃⠑⠠⠯⠠⠁⠋⠠⠻’ (=bef-and-af) ⟹ bef-and-af

จากตัวอย่างข้างต้น อาจเกิดปัญหาในการตีความ‘⠃⠗⠁⠙⠋⠁⠞⠓⠻’ ผิดเป็น“bef-and-af” ดังที่แสดงไว้ใน {31} ข. ได้ หากนำทั้งสายอักขระไปทำการถ่ายถอดโดยไม่มีกระบวนการใด ๆ เพิ่มเติม

4.2.2.2.2 ตัวย่อเบรลล์เซลล์คู่

ดังที่ได้กล่าวมาแล้วว่า ในอักษรเบรลล์มีอักษรภาษาอังกฤษเฉพาะตัวพิมพ์เล็กเท่านั้น หากต้องการเขียนอักษรตัวพิมพ์ใหญ่ ต้องเขียน ‘⠠’ (จุด6) นำหน้า แต่อย่างไรก็ตาม มีตัวย่อเบรลล์เซลล์คู่ที่ใช้ ‘⠠’ นำหน้าอักขระ อยู่ 2 คู่ คือ ‘⠠⠝’ ที่ตีความเป็น “ation” และ “N” และ ‘⠠⠽’ ที่ตีความเป็น “ally”และ “Y” ในภาษาอังกฤษตามลำดับ จึงทำให้เกิดปัญหาต่อการถ่ายถอดอักษร หากไม่ทำการตรวจสอบบริบทแวดล้อมของอักขระเบรลล์ 2 คู่นี้

4.2.2.2.3 ตัวย่อเบรลล์เซลล์เดี่ยวทั่วไป

ตัวย่อกลุ่มนี้เป็นตัวย่อที่ใช้แทนอักษรปกติได้เพียงกรณีเดียวไม่ว่าจะปรากฏในตำแหน่งใดของสายอักขระก็ตาม จึงไม่เป็นปัญหาต่อการถ่ายถอดเนื่องจากสามารถจับคู่เบรลล์กับอักษรปกติได้โดยตรง

4.2.2.3 การเว้นวรรคระหว่างคำที่แตกต่างกันของเบรลล์อังกฤษและอังกฤษปกติ

ตัวย่อเบรลล์มีขึ้นเพื่อประหยัดเนื้อที่ในการเขียน ซึ่งวิธีการหนึ่งที่ทำได้คือ ขจัดช่องว่างบางตำแหน่งในอักษรเบรลล์ ดังนั้นเบรลล์อังกฤษและอังกฤษปกติจึงมีการเว้นวรรคที่ต่างกันใน 2 ลักษณะคือ (1) การเว้นวรรคเมื่อมีการปรากฏของตัวย่อเบรลล์ต่ำ และ(2) การเว้นวรรคในตัวย่อรูปเต็มคงรูป

4.2.2.3.1 การเว้นวรรคในตัวย่อเบรลล์ต่ำย่อคำโดด

ตัวย่อเบรลล์ต่ำ ประกอบด้วย*ตัวย่อของ to, into และ by* ในการใช้ตัวย่อกลุ่มนี้จะต้องเขียนชิดกับคำที่ตามมาเสมอ (ดังที่ได้อภิปรายมาแล้วในหัวข้อ 4.2.2.1.1 (1))

ในการถ่ายถอดตัวย่อกลุ่มนี้ ระบบจะต้องเพิ่มช่องว่างท้ายคำในอักษรปกติที่ได้จากตัวย่อเหล่านี้ เพื่อให้ตรงตามหลักการเว้นวรรคในอังกฤษปกติ

4.2.2.3.2 การเว้นวรรคในตัวย่อรูปเต็มคงรูป

ตัวย่อกลุ่มนี้ประกอบด้วย*รูปย่อของ* ‘⠯’ (and), ‘⠿’ (for), ‘⠷’ (of), ‘⠮’ (the), ‘⠾’ (with) และ ‘⠁’ (a) ตัวย่อเหล่านี้สามารถใช้ย่อคำโดด หรือใช้ย่อส่วนของคำก็ได้ (ดูหัวข้อ 2.1.1.2.1(2)) ในการใช้เพื่อย่อคำโดดที่เรียงต่อกัน จะต้องเขียนตัวย่อแต่ละตัวต่อกันโดยไม่มีช่องว่าง ดังนี้

{32} ก. ‘⠯⠿⠮’ ⟹ “and for the”

ข. ‘⠾⠁’ ⟹ “with a”

ฉะนั้นในการถ่ายถอดอักษร หากระบบพบตัวย่อเหล่านี้เรียงต่อกัน จะต้องทำการแทรกช่องว่างหลังจากที่ถ่ายถอดอักษรแต่ละตัวแล้ว ดังที่ได้แสดงไว้ในตัวอย่าง

4.2.2.4 การใช้สัญลักษณ์ที่ปรากฏเฉพาะอักษรเบรลล์

ในภาษาอังกฤษมีสัญลักษณ์ที่ปรากฏเฉพาะในอักษรเบรลล์เพื่อกำกับสถานะหรือประเภทของอักษรเบรลล์ ในงานวิจัยชิ้นนี้ผู้วิจัยได้แบ่งสัญลักษณ์เหล่านี้ออกเป็น 3 กลุ่ม ได้แก่ (1) เครื่องหมายประกอบคำ (special composition signs) (2) เครื่องหมายนำเลข (number sign) และ (3) เครื่องหมายกำกับคอมพิวเตอร์เบรลล์

4.2.2.4.1 เครื่องหมายประกอบคำ

เครื่องหมายประกอบคำประกอบด้วย 3 สัญลักษณ์ ได้แก่ เครื่องหมายแสดงตัวพิมพ์ใหญ่เดี่ยว (single capital sign) เครื่องหมายแสดงตัวพิมพ์ใหญ่คู่ (double capital sign) และเครื่องหมายกำกับตัวอักษร (letter sign)

(1) เครื่องหมายแสดงตัวพิมพ์ใหญ่เดี่ยว

ในระบบอักษรเบรลล์หกจุดมาตรฐาน ไม่สามารถแสดงอักษรตัวพิมพ์ใหญ่โดยใช้เบรลล์เพียง 1 เซลล์ได้ เนื่องจากมีจำนวนจุดที่ใช้อยู่อย่างจำกัด ดังนั้น หากต้องการแสดงว่าอักษรปัจจุบันเป็นตัวพิมพ์ใหญ่ ต้องเขียน ‘⠠’ (เครื่องหมายแสดงตัวพิมพ์ใหญ่เดี่ยว) ไว้หน้าตัวอักษรนั้น เช่น

{33}ก.‘⠠⠃⠊⠗⠙’ (=CAPITAL SIGN-b-i-r-d) ⟹ “Bird”

ข. ‘⠠⠁⠗⠑’ (=CAPITAL SIGN-ar-e) ⟹ “Are”

ในการถ่ายถอดอักษร หากระบบพบ ‘⠠’ ระบบจะทำการลบเครื่องหมายนี้ออกจากสายอักขระแล้วทำการถ่ายถอดตัวต่อไป หลังจากนั้นแปลงเฉพาะอักษรตัวแรกที่ถ่ายถอดได้ให้เป็นอักษรตัวพิมพ์ใหญ่

(2) เครื่องหมายแสดงตัวพิมพ์ใหญ่คู่ ‘⠠⠠’

ในเบรลล์อังกฤษ มีการใช้ ‘⠠⠠’ (เครื่องหมายแสดงตัวพิมพ์ใหญ่คู่) เพื่อระบุว่า อักษรที่อยู่ตามหลังเครื่องหมายนี้ต้องเป็นอักษรตัวพิมพ์ใหญ่ทั้งหมดจนจบวรรค เครื่องหมายนี้มักปรากฏในตำแหน่งต้นสายอักขระเช่น ‘⠠⠠⠁⠙⠧⠊⠉⠑’ (=DOUBLE CAPITAL SIGN–a-d-v-i-c-e) ต้องตีความให้เป็น ‘ADVICE’ ในการถ่ายถอดอักษร เมื่อระบบพบ ‘⠠⠠’ ระบบจะลบเครื่องหมายนี้ออกจากสายอักขระ แล้วทำการถ่ายถอดอักษรตัวต่อไปพร้อมกับแปลงอักษรทุกตัวที่ถ่ายถอดแล้วหลังเครื่องหมายนี้ ให้เป็นอักษรตัวพิมพ์ใหญ่ทั้งหมด จนสิ้นสุดสายอักขระปัจจุบัน

(3) เครื่องหมายกำกับตัวอักษร ‘⠰’

ในเบรลล์อังกฤษ ‘⠰’ (เครื่องหมายกำกับตัวอักษร) เป็นสัญลักษณ์ที่ใช้นำหน้าตัวอักษรหรือกลุ่มตัวอักษรเพื่อแสดงว่า สิ่งที่ตามหลังเครื่องหมายนี้ไม่ใช่ตัวย่อระดับคำ ตัวอย่างเช่น

{34} ก. ‘⠨’ (= b) ⟹ “but”

ข. ‘⠰⠃’ (=LETTER SIGN-b) ⟹ “b”

ค. ‘⠁⠃’(= ab) ⟹ “about”

ง. ‘⠰⠁⠃’(=LETTER SIGN-ab) ⟹ “ab”

จะเห็นได้ว่าโดยทั่วไป หาก ‘⠨’ ปรากฏโดด ๆ ในข้อความดังที่แสดงในข้อ {34} ก. ต้องตีความตัวย่อนี้ให้ใช้แทน“but” เนื่องจากตัวย่อนี้เป็นตัวย่อระดับคำ (ดู 2.1.1.2.1 ตัวย่ออักษรโดด)

แต่หากต้องการเขียนเฉพาะตัว “b” ต้องใช้ ‘⠰’ นำหน้าอักษร ดังที่แสดงในข้อ {34} ข. ดังนั้นหากระบบพบ ‘⠰’ ระบบจะทำการลบเครื่องหมายนี้ออกจากสายอักขระ แล้วนำอักษรที่ปรากฏหลังเครื่องหมายนี้มาถ่ายถอดทีละตัว จนจบสายอักขระ

4.2.2.4.2 เครื่องหมายนำเลข

เครื่องหมาย ‘⠼’ (= #; number sign) เป็นเครื่องหมายที่เขียนไว้ข้างหน้าอักษรอังกฤษจาก ‘⠁’ –‘⠊’ (= a-j) เพื่อระบุว่าอักขระเบรลล์นั้น ๆ ใช้แทนตัวเลข ในการเขียนตัวเลขที่เรียงต่อกัน จะเขียนเครื่องหมายนี้ไว้ข้างหน้าตัวเลขเฉพาะตัวแรกเท่านั้น (ดูหัวข้อ 2.1.3 ประกอบ)

ฉะนั้นในการถ่ายถอดอักษรเมื่อระบบพบเครื่องหมายนำเลข จะทำการลบเครื่องหมายนี้ออกจากสายอักขระ แล้วถ่ายถอดอักขระเบรลล์หลังเครื่องหมายนี้ให้เป็นตัวเลข

อนึ่ง หน้าเครื่องหมาย ‘⠼’ (number sign) อาจมีเครื่องหมายอื่น ๆ กำกับอยู่ ดังนี้

- ‘⠈⠼’ (=$+ตัวเลข) แสดงว่ามีเครื่องหมาย $ ปรากฏหน้าตัวเลข เช่น ‘⠈⠼⠓’(= “$8”)
- ‘⠘⠼’ (=£+ตัวเลข) แสดงว่ามีเครื่องหมาย £ ปรากฏหน้าตัวเลข เช่น ‘⠘⠼⠓’ (= “£8”)

ในการถ่ายถอดตัวเลขที่มีเครื่องหมายเหล่านี้กำกับอยู่ ต้องลบเครื่องหมายออกจากสายอักขระ แล้วถ่ายถอดอักขระเบรลล์หลังเครื่องหมายเหล่านี้ให้เป็นตัวเลข

นอกจากนี้โดยทั่วไปแล้วเมื่อเขียนตัวเลขและตัวอักษรปนกัน จำเป็นต้องใช้เครื่องหมายกำกับอักษร ‘⠰’ (จุด 5-6) เพื่อบ่งชี้ว่าอักขระเบรลล์ที่ตามหลังเครื่องหมายนี้เป็นตัวอักษร เช่น ‘⠼⠃⠰⠁’ (= เครื่องหมายนำเลข + 2 + เครื่องหมายกำกับอักษร + a) คือ “2a” อย่างไรก็ตาม ในการเขียนลำดับที่ (ordinal number) จะเขียนอักษรที่ปรากฏหลังตัวเลขทันทีโดยที่ไม่ต้องใช้ ‘⠰’นำหน้าตัวอักษร เช่น ‘⠼⠃⠝⠙’ (=เครื่องหมายนำเลข + 2 +n + d) ใช้แทน “2nd” เป็นผลให้ในเบรลล์รูป

ย่อ สามารถใช้เครื่องหมาย ‘⠌’ (=st) และ ‘⠹’(=th) หลังตัวเลขได้ทันที เช่น ‘⠼⠋⠹’(=เครื่องหมายนำเลข+6+th) เพื่อใช้แทน “6th”

4.2.2.4.3 เครื่องหมายกำกับคอมพิวเตอร์เบรลล์

จากการศึกษาเอกสารเบรลล์ ผู้วิจัยพบว่าเอกสารส่วนใหญ่มักจะเขียนที่อยู่ของเว็บไซต์ด้วยรหัสคอมพิวเตอร์เบรลล์ (รหัสASCII) [8] โดยจะกำกับด้วยเครื่องหมายต่อไปนี้

- ‘⠈⠼’ (เครื่องหมายเปิดรหัส) ใช้กำกับจุดเริ่มต้นของคอมพิวเตอร์เบรลล์
- ‘⠈⠼’ เขียนไว้ท้ายบรรทัดเพื่อระบุว่า ข้อความในบรรทัดต่อไปยังคงเป็นส่วนหนึ่งของคอมพิวเตอร์เบรลล์
- ‘⠈⠼’ (เครื่องหมายปิดรหัส) ใช้กำกับจุดสิ้นสุดของคอมพิวเตอร์เบรลล์

ในการถ่ายถอดที่อยู่ของเว็บไซต์ ระบบต้องทำการลบเครื่องหมายข้างต้นออกจากสายอักขระ (เนื่องจากเครื่องหมายเหล่านี้ไม่สอดคล้องกับอักษรใดในอักษรปกติ) หลังจากทำการลบแล้ว นำสายอักขระที่ได้ไปใช้เป็นผลลัพธ์ของการถ่ายถอดอักษรโดยตรง ด้วยเหตุที่สายอักขระนี้เป็นรหัสคอมพิวเตอร์เบรลล์อยู่แล้ว จึงไม่ต้องผ่านกระบวนการถ่ายถอดใด ๆ

เท่าที่ได้อภิปรายมาข้างต้น จะพบว่า การถ่ายถอดเบรลล์รูปย่อ ต้องอาศัยแนวคิดต่าง ๆ ได้แก่ การแบ่งสายอักขระในการตีความตัวย่อ การตีความตัวย่อจากตำแหน่งที่ปรากฏ การเว้นวรรคที่แตกต่างกันระหว่างอักษรเบรลล์และอักษรปกติ รวมถึงการใช้สัญลักษณ์พิเศษที่ปรากฏเฉพาะในอักษรเบรลล์ซึ่งแนวคิดเหล่านี้ผู้วิจัยได้นำมาใช้ในการลดความซับซ้อนของการถ่ายถอดเบรลล์รูปย่อ

อย่างไรก็ตาม กระบวนการถ่ายถอดเบรลล์ไทยและกระบวนการถ่ายถอดเบรลล์อังกฤษจะทำงานร่วมกันได้ ก็ต้องอาศัยกระบวนการระบุภาษา ซึ่งจะใช้เป็นตัวกลางในการส่งข้อความนำเข้าไปดำเนินการถ่ายถอดอักษรในภาษาที่เหมาะสม ซึ่งจะได้อภิปรายในหัวข้อต่อไป

4.3 แนวคิดที่ใช้ในการระบุภาษาในอักษรเบรลล์

เบรลล์ไทยและเบรลล์อังกฤษต่างก็ใช้จุดไม่เกินหกจุดแทนตัวอักษร ทำให้ทั้งสองภาษานี้มีตัวอักษรที่ซ้ำกันระหว่างภาษาเป็นจำนวนมาก แต่ยังมีอักษรอยู่จำนวนหนึ่งที่มีรูปแตกต่างกัน ซึ่งจัดว่าเป็นเอกลักษณ์ของแต่ละภาษา ที่นำมาใช้จำแนกภาษาทั้งสองออกจากกันได้ นอกจากนี้เบรลล์

---

[8] รหัสคอมพิวเตอร์เบรลล์ เป็นรหัสอักษรเบรลล์ที่สอดคล้องกับตัวอักษรแต่ละตัวที่พิมพ์จากแป้นพิมพ์โดยตรง เช่น พิมพ์ “.” บนแป้นพิมพ์ จะได้ ‘⠲’ (จุด 4-6) ในขณะที่อักษรเบรลล์มาตรฐาน “.” จะแทนด้วย ‘⠨⠄’ (จุด 2-5-6)

ไทยและเบรลล์อังกฤษมีการเรียงตัวอักษรในแต่ละภาษาที่ไม่เหมือนกัน จึงสามารถนำอักษรที่เรียงต่อกันมาใช้ระบุภาษาได้

4.3.1 การจำแนกภาษาโดยใช้สายอักขระที่เป็นเอกลักษณ์ของแต่ละภาษา

หากเปรียบเทียบอักษรเบรลล์ไทยกับเบรลล์อังกฤษ จะพบว่าตัวเขียนของทั้งสองภาษานี้มีรูปซ้ำกันจำนวนมาก เช่น ‘⠙’(=ด) ซ้ำกับ ‘⠙’(=d), ‘⠕’(=อ) ซ้ำกับ‘⠕’(=o) และ‘⠛’(=ก) ซ้ำกับ ‘⠛’(=g) เป็นต้น แม้ว่าตัวเขียนของเบรลล์จะมีรูปซ้ำกันแต่ยังมีตัวเขียนอีกจำนวนหนึ่งที่มีรูปแตกต่างกันระหว่างภาษา ซึ่งสามารถนำมาใช้ในการระบุภาษาได้ เช่น ‘⠞’ (=ต) สามารถใช้ระบุว่าเป็นภาษาไทยได้ หากข้อความนั้นมีการปนกันระหว่างเบรลล์ไทยและเบรลล์อังกฤษรูปเต็ม

นอกจากนี้ แม้อักษรจะซ้ำกันระหว่างภาษาแต่รูปแบบในการจัดเรียงตัวอักษรในแต่ละภาษาก็มีความแตกต่างกันทำให้เกิดสายอักขระที่เป็นเอกลักษณ์ของแต่ละภาษาขึ้น ดังจะเห็นได้จากตัวอย่างต่อไปนี้

{35} ก.‘⠃⠥⠞’ ⟹ “but” (อังกฤษ) หรือ “ิคถ” (ไทย)

ข. ‘⠛⠃⠝’ ⟹ “กิน” (ไทย) หรือ “gbn” (อังกฤษ)

จากตัวอย่างจะพบว่า แม้อักษรจะปรากฏได้ในทั้งสองภาษาแต่การเรียงต่อของอักษรเหล่านี้สามารถเกิดขึ้นได้เฉพาะภาษาใดภาษาหนึ่งเท่านั้น ดังนั้นตัวอักษรที่จัดเรียงต่างกันระหว่างภาษาจึงสามารถใช้ระบุว่าเป็นภาษาใดภาษาหนึ่งได้

4.3.2 การจำแนกภาษาโดยอาศัยบริบทแวดล้อม

ดังที่ได้กล่าวมาแล้วว่า การระบุภาษา โดยใช้ตัวเขียนที่แตกต่างกันระหว่างภาษา สามารถใช้ระบุภาษาของสายอักขระปัจจุบันได้ แต่การระบุภาษาด้วยวิธีการดังกล่าว ไม่สามารถทำได้กับทุกสายอักขระ เนื่องจากเบรลล์ไทยและเบรลล์อังกฤษมีตัวอักษรและการเรียงต่อของตัวอักษรที่คล้ายคลึงกันจำนวนหนึ่ง เช่น ‘⠙⠕⠅’ สามารถตีความเป็น “ดอก” ในภาษาไทย หรือ “dog” ในภาษาอังกฤษได้ ดังนั้นจึงสามารถระบุสถานะภาษาของสายอักขระปัจจุบันโดยอาศัยสถานะภาษาของสายอักขระก่อนหน้าได้

อย่างไรก็ตาม การพึ่งพาบริบทในการถ่ายถอดอักษรก็ยังไม่ครอบคลุมในบางกรณี โดยเฉพาะอย่างยิ่ง กรณีที่ระบบไม่สามารถระบุภาษาสำหรับสายอักขระแรกของข้อความเบรลล์ที่กำลังถ่ายถอดได้ จึงจำเป็นต้องนำความยาวของสายอักขระมาช่วยในการระบุภาษา

4.3.3 การจำแนกภาษาโดยใช้ความยาวของสายอักขระ

ในการเขียนข้อความอักษรปกติ การเว้นวรรคของข้อความในภาษาไทยและภาษาอังกฤษมีความแตกต่างกัน ในข้อความไทย แต่ละวรรคจะประกอบด้วยคำต่าง ๆ ที่นำมาเรียงต่อกัน ขณะที่ในข้อความอังกฤษ วรรคหนึ่งจะประกอบด้วยคำเพียงคำเดียว ซึ่งความแตกต่างนี้ก็เป็นไปในทำนองเดียวกันสำหรับเบรลล์ไทยและเบรลล์อังกฤษ ด้วยเหตุนี้จึงทำให้เบรลล์ไทยแต่ละวรรคมีความยาวมากกว่าเบรลล์อังกฤษอย่างมีนัยสำคัญ

ในการระบุภาษาโดยใช้ความยาวของสายอักขระ จะเกิดขึ้นเมื่อระบบไม่สามารถระบุภาษาด้วยสองวิธีการที่กล่าวมาข้างต้นได้ โดยการระบุภาษาด้วยวิธีนี้จะอาศัยความยาวที่เหมาะสมเพื่อระบุสายอักขระปัจจุบันให้เป็นภาษาไทย ซึ่งจะขอนำเสนอในบทต่อไป

ในบทนี้ ผู้วิจัยได้กล่าวถึงแนวคิดทั่วไปที่ใช้ในการออกแบบระบบการถ่ายถอดเบรลล์ไทยปนอังกฤษ ไม่ว่าจะเป็น ระบบการถ่ายถอดอักษรโดยใช้กฎ ระบบการถ่ายถอดอักษรโดยใช้แบบจำลองเอ็นแกรม และระบบการถ่ายถอดอักษรโดยใช้วิธีการแบบผสม ซึ่งแนวคิดเหล่านี้เป็นที่มาของการทำงานในส่วนต่าง ๆ ของระบบการถ่ายถอดเบรลล์ไทยปนอังกฤษ ที่จะได้นำเสนอในบทต่อไป

บทที่ 5

ระบบการถ่ายถอดเบรลล์ไทยปนอังกฤษโดยใช้กฎ

ในบทที่ผ่านมาผู้วิจัยได้ชี้ให้เห็นว่า อักษรเบรลล์เป็นอักษรที่ใช้จุดในปริมาณที่จำกัดทำให้อักษรในภาษาหนึ่งซ้ำรูปกับอีกภาษาหนึ่งอย่างหลีกเลี่ยงไม่ได้ จึงเป็นปัญหาต่อการระบุว่า ข้อความเบรลล์นั้นเป็นข้อความในภาษาใด หากพิจารณาการถ่ายถอดอักษรเฉพาะภาษา ในไทยเบรลล์เองก็ยังมีข้อจำกัดในการถ่ายถอดให้เป็นอักษรปกติ เนื่องจากอักขระเบรลล์บางตัวอาจตีความได้มากกว่าหนึ่งกรณีและอักษรเบรลล์ไทยมีอักขรวิธีที่แตกต่างจากอักษรปกติบางประการ ในขณะเดียวกันการถ่ายถอดเบรลล์อังกฤษให้เป็นอักษรปกติก็ทำได้ไม่ง่ายเลย เพราะเบรลล์อังกฤษมีระบบการเขียนอยู่ 2 ระบบ คือเบรลล์อังกฤษรูปเต็ม และเบรลล์อังกฤษรูปย่อ ทั้งสองระบบนี้มีปริมาณตัวอักษรและอักขระวิธีที่แตกต่างกัน เป็นผลให้อักขระเบรลล์แต่ละตัวเมื่อปรากฏในข้อความเบรลล์ต่างระบบตัวเขียนอาจตีความได้แตกต่างกัน ข้อจำกัดเหล่านี้ทำให้การพัฒนาระบบการถ่ายถอดเบรลล์ไทยปนอังกฤษด้วยกฎในงานวิจัยนี้ได้แยกเป็นสองระบบ คือ *ระบบที่ใช้ถ่ายถอดเบรลล์ไทยปนอังกฤษรูปเต็ม* และ*ระบบที่ใช้ถ่ายถอดเบรลล์ไทยปนอังกฤษรูปย่อ*

ในบทนี้ผู้วิจัยจะนำเสนอระบบการถ่ายถอดเบรลล์ไทยปนอังกฤษที่ผู้วิจัยได้พัฒนาขึ้น โดยในหัวข้อที่ 5.1 จะกล่าวถึงการเตรียมข้อมูลและข้อมูลที่ใช้ในงานวิจัย หัวข้อที่ 5.2 จะกล่าวถึงระบบการถ่ายถอดอักษรเบรลล์ที่ผู้วิจัยได้พัฒนาขึ้น และเพื่อวัดประสิทธิภาพของระบบอย่างเป็นรูปธรรม ผู้วิจัยจะได้นำเสนอผลการทดสอบระบบในหัวข้อที่ 5.3 และอภิปรายผลการทดสอบระบบในหัวข้อที่ 5.4 ซึ่งแต่ละหัวข้อมีรายละเอียดดังนี้ต่อไป

5.1 ข้อมูลที่ใช้ในงานวิจัย

ในการวิจัยการถ่ายถอดอักษรเบรลล์ให้เป็นอักษรปกติโดยอาศัยกฎทางภาษานั้น ผู้วิจัยได้ใช้ข้อมูล 2 ประเภท ได้แก่ ข้อมูลที่ใช้สำหรับวิเคราะห์เพื่อนำผลมาใช้ในระบบถ่ายถอดอักษร และข้อมูลที่นำมาใช้ทดสอบระบบ ซึ่งแต่ละประเภทประกอบด้วยข้อมูลต่าง ๆ ดังนี้

5.1.1 ข้อมูลที่ใช้วิเคราะห์เพื่อสร้างระบบการถ่ายถอดอักษร

ข้อมูลที่ใช้ในการวิเคราะห์และนำผลมาใช้ในระบบการถ่ายถอดอักษรเบรลล์แบ่งออกได้เป็น 3 ประเภทได้แก่ ข้อมูลที่ใช้ในการระบุภาษา ข้อมูลที่ใช้ในการถ่ายถอดเบรลล์ไทย และข้อมูลที่ใช้ในการถ่ายถอดเบรลล์อังกฤษ

5.1.1.1 ข้อมูลที่ใช้ในการระบุภาษา

เนื่องจากข้อมูลเบรลล์ที่จะเป็นข้อมูลนำเข้าในระบบ สามารถเป็นสายอักขระของเบรลล์ไทย เบรลล์อังกฤษรูปเต็ม หรือเบรลล์อังกฤษรูปย่อได้ การระบุภาษาจึงเป็นได้อย่างใดอย่างหนึ่งในสามทางเลือกนี้

อย่างไรก็ตาม ข้อความเบรลล์ที่มีการปนกันของภาษาไทยและอังกฤษต้องเป็นไปในรูปแบบใดรูปแบบหนึ่งคือ ไทยปนอังกฤษรูปเต็มหรือไทยปนอังกฤษรูปย่อเท่านั้น เนื่องจากหากผิดประเภทแล้วจะทำให้ได้ผลลัพธ์ไม่ถูกต้อง เช่น หากนำข้อความอังกฤษรูปเต็มไปถ่ายถอดด้วยระบบอังกฤษรูปย่อจะทำให้อักษรโดดที่ปรากฏในอังกฤษรูปเต็มถูกถ่ายถอดเป็นคำย่อเบรลล์ (เช่น “vitamin C” ถูกถ่ายถอดเป็น “vitamin Can”) และอักษรที่เรียงต่อกันบางชุดในอังกฤษรูปเต็มจะถูกถ่ายถอดเป็นคำย่อเบรลล์ (เช่น “cd” ถูกถ่ายถอดเป็น “could”) ในทางตรงกันข้าม หากนำเบรลล์รูปย่อไปถ่ายถอดด้วยเบรลล์รูปเต็มจะทำให้คำย่อถูกถ่ายถอดเป็นตัวอักษร (เช่น ‘⠗⠉⠧’ ซึ่งแทนคำว่า “receive” ถูกถ่ายถอดเป็น *“rcv”) และอักขระเบรลล์บางตัวที่ไม่ปรากฏในเบรลล์รูปเต็มจะถูกถ่ายถอดให้เป็นอักษรปกติที่ไม่สามารถเป็นคำได้เช่น ‘⠏⠗⠑⠎⠑⠝⠞⠁⠞⠊⠕⠝’(p-r-e-s-e-n-t-ation) ซึ่งใช้แทน“presentation” ถูกถ่ายถอดเป็น *“pres5tN”

ฉะนั้น ในการระบุภาษาโดยใช้เซ็ตของสายอักขระเฉพาะของแต่ละภาษาตามระบบที่พัฒนาออกมาเป็นสองระบบ จึงต้องมีชุดสายอักขระ 2 ชุด ชุดแรกประกอบด้วยเซ็ตกลุ่มอักษรเบรลล์ไทยและเซ็ตกลุ่มอักษรเบรลล์อังกฤษรูปเต็ม ส่วนอีกชุดหนึ่งประกอบด้วยเซ็ตกลุ่มอักษรเบรลล์ไทยและเซ็ตกลุ่มอักษรเบรลล์อังกฤษรูปย่อ ซึ่งสมาชิกของแต่ละเซ็ตประกอบด้วย อักษรเบรลล์ 1 ตัว อักษรเบรลล์ 2 ตัว และอักษรเบรลล์ 3 ตัว ที่แตกต่างจากอีกเซ็ตหนึ่งในชุด

ในการสร้างเซ็ตสายอักขระเบรลล์ของแต่ละภาษา ผู้วิจัยเริ่มจากการสร้างคลังข้อความเบรลล์ 3 ประเภท ได้แก่ คลังข้อความเบรลล์ไทย คลังข้อความเบรลล์อังกฤษรูปเต็ม และคลังข้อความเบรลล์อังกฤษรูปย่อ โดยที่ข้อความเบรลล์ไทยและข้อความเบรลล์อังกฤษรูปย่อนั้น ได้มาจากเอกสารเบรลล์ ที่รวบรวมจากศูนย์เทคโนโลยีการศึกษาเพื่อคนตาบอด ประเภทละ 500 แฟ้มข้อมูล (ซึ่งเป็นข้อความที่แตกต่างจากข้อความที่นำมาใช้ในการทดสอบประสิทธิภาพของระบบ) แต่เนื่องจากโดยทั่วไป ข้อความเบรลล์อังกฤษ มักจะเขียนโดยใช้เบรลล์อังกฤษรูปย่อจึงทำให้หาข้อความที่เป็นเบรลล์อังกฤษรูปเต็มได้น้อยมาก ผู้วิจัยจึงได้สร้างคลังข้อความเบรลล์อังกฤษรูปเต็ม โดยการสกัดคำทุกคำจาก British National Corpus แล้วนำมาแปลงเป็นเบรลล์อังกฤษรูปเต็มด้วยโปรแกรม Duxbury Braille Translator เวอร์ชั่น 12.1

หลังจากได้คลังข้อความเบรลล์ทั้ง 3 ประเภทแล้ว ผู้วิจัยได้ทำการสกัดอักษร 1 ตัวแรก 2 ตัวแรก และ 3 ตัวแรก ของแต่ละวรรค ทีละคลังข้อความ ทำให้ได้เซ็ตของตัวอักษร 3 ประเภท ได้แก่ เบรลล์ไทย เบรลล์อังกฤษรูปเต็ม และเบรลล์อังกฤษรูปย่อ ที่มีความยาวต่างกัน

เนื่องจากข้อมูลนำเข้าเป็นข้อความที่เป็นการผสมของเบรลล์ไทยกับเบรลล์อังกฤษรูปเต็ม หรือเป็นข้อความผสมของเบรลล์ไทยกับเบรลล์อังกฤษรูปย่อ ไม่มีข้อความใดที่เป็นทั้งเบรลล์อังกฤษรูปเต็มและเบรลล์อังกฤษรูปย่อปนกัน ผู้วิจัยจึงได้นำเซ็ตกลุ่มตัวอักษรเบรลล์อังกฤษรูปเต็มมาเปรียบเทียบกับเซ็ตกลุ่มตัวอักษรเบรลล์ไทย เพื่อสร้างสายอักขระที่เป็นเอกลักษณ์ของเบรลล์ไทย และสายอักขระที่เป็นเอกลักษณ์ของเบรลล์อังกฤษรูปเต็ม ซึ่งจะนำไปใช้ในการระบุภาษาของข้อความที่มีการปนกันระหว่างไทยและอังกฤษรูปเต็ม ตัวอย่างเช่น ‘⠓’(=ฮ) ‘⠓⠊’ (=เฮ) ‘⠓⠊⠙’ (=ปิด) เป็นสายอักขระที่เป็นเอกลักษณ์ของเบรลล์ไทย และ ‘⠑’(=e) ‘⠅⠁’(=K) ‘⠞⠁⠇’(=tal) เป็นสายอักขระที่เป็นเอกลักษณ์เบรลล์อังกฤษ (ดูรายละเอียดเกี่ยวกับข้อมูลในภาคผนวก ค.1 และ ค.2)

นอกจากนี้ ผู้วิจัยได้นำเซ็ตกลุ่มตัวอักษรเบรลล์อังกฤษรูปย่อมาเปรียบเทียบกับเซ็ตกลุ่มตัวอักษรเบรลล์ไทย เพื่อสร้างสายอักขระที่เป็นเอกลักษณ์ของเบรลล์ไทย และเบรลล์อังกฤษรูปย่อ ซึ่งจะนำไปใช้ในการระบุภาษาของข้อความที่มีการปนกันระหว่างเบรลล์ไทยและเบรลล์อังกฤษรูปย่อ ตัวอย่างเช่น ‘⠁⠏’ (=แป) ‘⠛⠊⠃’ (กลบ) เป็นสายอักขระที่เป็นเอกลักษณ์ของเบรลล์ไทย[9] ‘⠁’ (=a) ‘⠰⠁’(= and the) ‘⠝⠔’ (=Nin) เป็นสายอักขระที่เป็นเอกลักษณ์ของเบรลล์อังกฤษ ที่ใช้ในการถ่ายถอดเบรลล์อังกฤษรูปย่อ (ดูรายละเอียดเกี่ยวกับข้อมูลทั้งหมดได้ในภาคผนวก ค.3 และ ค.4)

หลังจากดำเนินการทุกขั้นตอนแล้ว จะได้*เซ็ตของสายอักขระที่เป็นเอกลักษณ์ของแต่ละภาษา* ซึ่งสามารถนำมาใช้ระบุภาษาของข้อความเบรลล์ไทยและเบรลล์อังกฤษประเภทต่าง ๆ ได้ โดยที่สายอักขระที่เป็นเอกลักษณ์ของแต่ละภาษาที่มีความยาว 1 , 2 และ 3 ตัวอักษร มีจำนวนต่างกันในแต่ละภาษา ดังนี้

[9] จะสังเกตได้ว่าไม่มีสายอักขระที่เป็นเอกลักษณ์ที่มีความยาว 1 อักขระ ซึ่งสามารถใช้ระบุภาษาไทยในอักษรเบรลล์ได้

ตารางที่ 5.1 จำนวนสายอักขระที่เป็นเอกลักษณ์ของแต่ละภาษา

ขนาดของสายอักขระที่สกัด	จำนวนสายอักขระสำหรับการถ่ายถอดเบรลล์ไทยปนอังกฤษรูปเต็ม		จำนวนสายอักขระสำหรับการถ่ายถอดเบรลล์ไทยปนอังกฤษรูปย่อ	
	เบรลล์ไทย	เบรลล์อังกฤษรูปเต็ม	เบรลล์ไทย	เบรลล์อังกฤษรูปย่อ
หนึ่งอักขระ	9	7	0	21
สองอักขระ	659	116	317	760
สามอักขระ	1,464	839	5,183	7,678

อย่างไรก็ตาม บางครั้งการระบุภาษาด้วยเซ็ตของสายอักขระที่เป็นเอกลักษณ์ของแต่ละภาษาไม่สามารถทำได้โดยตรง เนื่องจากมีอักขระเบรลล์อยู่จำนวนหนึ่งที่มีรูปซ้ำกันในตำแหน่งต้นสายอักขระระหว่างภาษาไทยและภาษาอังกฤษ ดังแสดงในตารางต่อไปนี้

ตารางที่ 5.2 เครื่องหมายเบรลล์ต้นสายอักขระ

เบรลล์	แอสกี	อักษรปกติ
⠨⠼	,7	[
⠨⠬	,8	‘
⠤	-	-
⠨	.	_
⠼	7	(
⠬	8	“

เพื่อให้สามารถระบุภาษาได้ ระบบจึงต้องหาจุดเริ่มต้นของการระบุภาษา โดยตรวจสอบว่ามีอักขระเบรลล์เหล่านี้ปรากฏต้นสายอักขระหรือไม่

5.1.1.2 ข้อมูลที่ใช้ในการถ่ายถอดเบรลล์ไทย

ข้อมูลที่ใช้ในการถ่ายถอดเบรลล์ไทยให้เป็นอักษรปกติ เป็นข้อมูลที่ผู้วิจัยวิเคราะห์และสร้างขึ้นในรูปแบบตารางคู่เทียบตัวอักษร ซึ่งประกอบด้วยข้อมูล 2 ประเภทได้แก่ คู่เทียบเบรลล์ไทย และ เซ็ตของอักษรเบรลล์ที่ใช้แก้ปัญหาความกำกวมในการถ่ายถอดอักษร

5.1.1.2.1 คู่เทียบเบรลล์ไทย

ในการถ่ายถอดอักษรระบบต้องนำอักษรเบรลล์มาจับคู่กับอักษรปกติ จึงต้องมีคู่เทียบตัวอักษรระหว่างสองระบบที่ตรงกัน ซึ่งคู่เทียบเหล่านี้ผู้วิจัยได้กำหนดขึ้นจากการศึกษาอักษรเบรลล์ และการปรับเปลี่ยนแก้ไขในช่วงพัฒนาระบบและเพื่อให้นำมาใช้กับระบบถ่ายถอดอักษรได้ ผู้วิจัยได้จัดเก็บคู่เทียบเหล่านี้ในรูปของตารางซึ่งประกอบด้วยอักษรเบรลล์และอักษรปกติที่จับคู่กัน

คู่เทียบในอักษรเบรลล์ไทยประกอบด้วยคู่เทียบประเภทต่าง ๆ ตามลักษณะความสัมพันธ์ระหว่างอักษรและการนำไปใช้เพื่อถ่ายถอดอักษรในแต่ละกรณี ดังนี้

(1) คู่เทียบเบรลล์เซลล์เดี่ยวทั่วไป

คู่เทียบเบรลล์เซลล์เดี่ยวทั่วไป ทำหน้าที่หลักคือ ใช้จับคู่อักษรเบรลล์และอักษรปกติแบบหนึ่งต่อหนึ่งเท่านั้น ซึ่งประกอบด้วยอักษรเบรลล์และอักษรปกติที่ตรงกันดังนี้

ตารางที่ 5.3 คู่เทียบเบรลล์เซลล์เดี่ยวทั่วไป

เบรลล์	แอสกี	อักษรปกติ
⠛	g	ก
⠅	k	ข
⠥	u	ค
⠻	]	ง
⠚	j	จ
⠌	/	ฉ
⠬	+	ช
⠮	!	ซ
⠙	d	ด
⠹	\	ต
⠞	t	ถ
⠾	)	ท
⠝	n	น
⠧	v	บ
⠷	&	ป
⠏	p	ผ

เบรลล์	แอสกี	อักษรปกติ
⠭	x	ฝ
⠿	?	พ
⠽	$	ฟ
⠍	m	ม
⠽	y	ย
⠗	r	ร
⠇	l	ล
⠺	w	ว
⠎	s	ส
⠓	h	ห
⠕	o	อ
⠿	=	ฮ
⠒	:	ไ
⠼	1	ๆ
⠴	0	์

(2) คู่เทียบสระทั่วไป

คู่เทียบสระทั่วไป เป็นคู่เทียบที่ใช้จับคู่สระในเบรลล์ไทยกับสระในไทยปกติ และอักษรเบรลล์ในคู่เทียบเหล่านี้ยังนำมาใช้เป็นเซ็ตของตัวอักษรเพื่อตรวจสอบบริบท เมื่อระบบพบเบรลล์ไทยที่กำกวม ซึ่งจะได้แสดงรายละเอียดในหัวข้อที่ 5.2 ต่อไป

ตารางสำหรับคู่เทียบสระทั่วไปแสดงได้ดังนี้

ตารางที่ 5.4 คู่เทียบสระทั่วไป

เบรลล์	แอสกี	อักษรปกติ
⠷	>	◌ั
⠦	*	า
⠃	b	◌ิ
⠆	2	◌ี
⠯	[	◌ึ
⠢	5	◌ื
⠉	c	◌ุ
⠒	3	◌ู
⠄	'	◌็

(3) คู่เทียบสระพิเศษ

คู่เทียบสระพิเศษเป็นคู่เทียบสระ 2 ตัวที่มีการจัดเรียงแตกต่างกันระหว่างไทยและไทยปกติ เมื่อสระเหล่านี้ปรากฏกับวรรณยุกต์ในอักษรเบรลล์ หลังจากถ่ายถอดแล้วต้องทำการสลับที่ระหว่างวรรณยุกต์กับสระเหล่านี้ในอักษรปกติ (ดูหัวข้อ 4.1.1.1 ประกอบ)

ตารางที่ 5.5 คู่เทียบสระพิเศษ

เบรลล์	แอสกี	อักษรปกติ
⠁	a	ะ
⠵	z	◌ำ

(4) คู่เทียบสระหน้า

คู่เทียบสระหน้าเป็นคู่เทียบสระเดี่ยวที่มีการจัดเรียงเหมือนกับอักษรปกติคือ มีรูปแบบการจัดเรียงเป็น สระ + พยัญชนะต้น แต่แตกต่างจากอักษรปกติ โดยที่อักษรปกติเมื่อกลายเป็นสระสั้นต้องวาง ‘⠐’ (=ะ) ไว้หลังสระเหล่านี้ทันที (ดูหัวข้อ 4.1.1.1 ประกอบ)

สระหน้าในเบรลล์ไทยประกอบด้วยอักขระเบรลล์ดังต่อไปนี้

ตารางที่ 5.6 คู่เทียบสระหน้า

เบรลล์	แอสกี	อักษรปกติ
⠮	f	เ
⠷	<	แ
⠪	i	โ

(5) คู่เทียบวรรณยุกต์

แม้ว่าเบรลล์ไทยและอักษรปกติจะมีการปรากฏของวรรณยุกต์ร่วมกับ สระหน้า สระประสม หรือสระพิเศษได้เช่นเดียวกัน แต่ก็มีการจัดเรียงวรรณยุกต์ในระบบการเขียนทั้ง 2 ระบบแตกต่างกัน ดังนั้นจึงจำเป็นต้องมีคู่เทียบวรรณยุกต์เพื่อใช้กับระบบถ่ายถอดอักษรเบรลล์เพราะคู่เทียบเหล่านี้จะใช้จับคู่วรรณยุกต์ในเบรลล์ไทยกับไทยปกติ (ดูหัวข้อ 4.1.1.1 และ 4.1.1.2.2)

นอกจากนี้ อักขระเบรลล์ในคู่เทียบเหล่านี้ยังนำมาใช้ ในการตรวจสอบบริบทต่าง ๆ เพื่อการตัดสินใจของระบบ ดังจะได้อภิปรายต่อไป

ตารางที่ 5.7 คู่เทียบวรรณยุกต์

เบรลล์	แอสกี	อักษรปกติ
⠹	9	่
⠴	4	้
⠽	7	๊
⠾	8	๋

(6) คู่เทียบเบรลล์เซลล์คู่ทั่วไป

คู่เทียบเบรลล์เซลล์คู่ทั่วไป เป็นคู่เทียบที่ต้องใช้เบรลล์สองเซลล์เพื่อแทนอักษรปกติหนึ่งตัวในเบรลล์ไทย เบรลล์เซลล์คู่จำนวนหนึ่งเกิดจากการเติมอักขระหน้าเบรลล์เซลล์เดี่ยว (หรือบางกรณีเติมข้างหลัง) เพื่อสร้างตัวอักษรใหม่ซึ่งมีลักษณะบางประการโดยเฉพาะลักษณะทางเสียงที่สอดคล้องกับเบรลล์เซลเดี่ยวนั้น ๆ และเบรลล์เซลล์คู่อีกส่วนหนึ่งเป็นเครื่องหมายต่างๆ (ดูหัวข้อ 4.1.1.2.1 ประกอบ)

ดังนั้นเพื่อให้ระบบสามารถถ่ายถอดเบรลล์เซลล์คู่ได้ผู้วิจัยจึงได้สร้างตารางคู่เทียบตัวอักษรสำหรับเบรลล์เซลล์คู่ขึ้น ดังนี้

ตารางที่ 5.8 คู่เทียบเบรลล์เซลล์คู่ทั่วไป

เบรลล์	แอสกี	อักษรปกติ
⠐⠥	,u	ฒ
⠐⠬	,+	ฌ
⠐⠽	,y	ญ
⠐⠙	,d	ฎ
⠐⠳	,|	ฏ
⠐⠸	,\	ฏ
⠐⠞	,t	ฐ
⠐⠾	,)	ฑ
⠤⠾	-)	ฒ
⠐⠝	,n	ณ
⠐⠦	,?	ภ
⠐⠎	,s	ศ
⠤⠎	-s	ษ
⠐⠇	,l	ฬ
⠰⠂	_1	,
⠰⠆	_2	;
⠰⠒	_3	:
⠰⠲	_4	.

⠦	_6	!
⠦	_8	?
⠦	_/	/
⠦	;2	ฯ
⠦	;1	ฯลฯ
⠒	:1	ใ
⠠	,7	[
⠠	,,	..
⠸	7’	]

(7) คู่เทียบสระประสม

คู่เทียบสระประสม เป็นคู่เทียบที่อักขระเบรลล์หนึ่งตัวใช้แทนสระซึ่งต้องเขียนด้วยอักษรหลายตัวในอักษรปกติ (ดูหัวข้อ 4.1.1.2 ประกอบ)

คู่เทียบสระประสมแสดงได้ดังนี้

ตารางที่ 5.9 คู่เทียบสระประสม

เบรลล์	**แอสกี**	**อักษรปกติ**
⠮	6	เ◌า
⠷	(	เ◌ีย
⠿	q	เ◌ือ
⠽	e	◌ัว

(8) คู่เทียบเบรลล์เซลล์เดี่ยวที่กำกวม

ในเบรลล์ไทย อักขระเบรลล์หนึ่งตัวสามารถใช้แทนอักษรปกติได้มากกว่าหนึ่งกรณี โดยที่อักขระเบรลล์ที่กำกวมนั้นอาจปรากฏอยู่ส่วนใดของสายอักขระก็ได้ (ดูหัวข้อ 4.1.2.1 ประกอบ)

คู่เทียบเบรลล์เซลล์เดี่ยวที่กำกวม แสดงดังในตารางต่อไปนี้

ตารางที่ 5.10 คู่เทียบเบรลล์เซลล์เดี่ยวที่กำกวมในตำแหน่งทั่วไป

เบรลล์	แอสกี	อักษรปกติ	
		กรณีที่ 1	กรณีที่ 2
⠷	%	เ◌ิ	เ อ
⠲	4	◌้	. (จุด)

(9) คู่เทียบเบรลล์ที่กำกวมท้ายสายอักขระ

นอกจากเบรลล์กำกวมที่ปรากฏในส่วนใดของสายอักขระก็ได้ ยังมีอักขระเบรลล์ที่เมื่อปรากฏท้ายสายอักขระแล้วสามารถตีความได้มากกว่าหนึ่งกรณี โดยที่จะแปรไปตามบริบทที่อักขระเบรลล์นั้นปรากฏอยู่ (ดูหัวข้อ 4.1.2.1.2 ประกอบ) เบรลล์กำกวมท้ายสายอักขระได้แก่อักขระเบรลล์ที่ตรงกับอักษรปกติในตารางคู่เทียบต่อไปนี้

ตารางที่ 5.11 คู่เทียบเบรลล์ที่กำกวมท้ายสายอักขระ

เบรลล์	แอสกี	อักษรปกติ	
		กรณีที่ 1	กรณีที่ 2
⠾	7	◌์	)
⠴	0	◌์	”

(10) คู่เทียบเบรลล์เซลล์คู่ที่กำกวม

เบรลล์เซลล์คู่ที่กำกวม เป็นอักขระเบรลล์สองตัวที่อยู่ชิดกันแล้วสามารถตีความให้เป็นอักษรปกติหนึ่งตัวหรือหลายตัวก็ได้ ซึ่งต้องตีความตามบริบทที่อักขระเบรลล์คู่เหล่านี้ปรากฏอยู่ (ดูหัวข้อ 4.1.3 ประกอบ) ดังนั้นผู้วิจัยจึงได้สร้างตารางคู่เทียบสำหรับอักขระเบรลล์กลุ่มนี้เพื่อใช้ในการถ่ายถอดอักษร ดังนี้

ตารางที่ 5.12 คู่เทียบสำหรับเบรลล์เซลล์คู่ที่กำกวม

เบรลล์	แอสกี	อักษรปกติ	
		กรณีที่ 1	กรณีที่ 2
⠕⠁	oa	อะ	เ าะ
⠵⠿	0)	์ท	ธ
⠗⠼	r1	รๆ	ฤ
⠍⠎	-s	-ส	ษ

5.1.1.2.2 ข้อมูลที่ใช้แก้ปัญหาความกำกวมในการถ่ายถอดอักษร

ข้อมูลประเภทนี้ มีเพื่อใช้แก้ปัญหาความกำกวมในการถ่ายถอดอักษรเบรลล์ที่ตีความได้มากกว่าหนึ่งกรณี ประกอบด้วยเซ็ตของอักษรเบรลล์ และเซ็ตของอักษรปกติ ที่ระบบจะนำมาใช้ในการตรวจสอบเงื่อนไขเมื่อระบบต้องตัดสินใจเลือกถ่ายถอดให้เป็นกรณีใดกรณีหนึ่ง โดยที่เซ็ตของกลุ่มอักษรที่ผู้วิจัยสร้างขึ้นนี้สามารถแบ่งออกเป็นประเภทต่าง ๆ ได้ดังนี้

(1) เซ็ตพยัญชนะตัวสะกดท้ายพยางค์ในเบรลล์ไทย

‘⠿’ มีรูปในอักษรปกติ 2 รูปคือ "เ◌ิ" และ"เ◌อ" การที่จะถ่ายถอด ‘⠿’ ให้อยู่ในรูปใดนั้นต้องพิจารณาว่า ‘⠿’ มีตัวสะกดท้ายพยางค์ปรากฏตามหลังหรือไม่ ผู้วิจัยจึงได้สร้างเซ็ตพยัญชนะตัวสะกดท้ายพยางค์ขึ้น

ในการสร้างเซ็ตพยัญชนะตัวสะกดท้ายพยางค์ ผู้วิจัยได้สกัดคำจากคลังข้อความภาษาไทยแห่งชาติ ที่มีสระเออในรูป “เ◌ิ” ปรากฏอยู่ แล้วสกัดพยัญชนะที่ปรากฏหลังสระ “◌ิ” มาทำการวิเคราะห์ จากนั้นนำตัวอักษรที่ได้ไปแปลงให้เป็นอักษรเบรลล์เพื่อใช้เป็นเซ็ตพยัญชนะตัวสะกดท้ายพยางค์ ซึ่งประกอบด้วยอักษรเบรลล์และคู่เทียบในอักษรปกติดังนี้

ตารางที่ 5.13 เซ็ตพยัญชนะตัวสะกดท้ายพยางค์สำหรับการถ่ายถอด ‘⠿’ให้เป็น “เ◌ิ”

เบรลล์	แอสกี	อักษรไทยปกติ
⠄⠿	,y	ญ
⠄⠾	,t	ฐ
⠄⠽	,n	ณ
⠄⠞	,s	ศ
⠛	g	ก
⠻	]	ง
⠝	n	น
⠍	m	ม
⠙	d	ด
⠧	v	บ
⠇	l	ล

(2) เซ็ตของพยัญชนะควบที่กำกวมและพยัญชนะควบไม่กำกวม ในอักษรปกติ

ดังที่ได้กล่าวมาแล้วว่า พยัญชนะที่ปรากฏหน้าสระประสมในอักษรเบรลล์ที่มีรูปคล้ายพยัญชนะควบกล้ำอาจเป็นหรือไม่เป็นพยัญชนะควบกล้ำก็ได้ ซึ่งเป็นปัญหาในการถ่ายถอดอักษร นอกจากนี้ยังมีพยัญชนะควบอีกส่วนหนึ่งที่เมื่อถ่ายถอดเป็นอักษรปกติแล้วจะต้องปรากฏระหว่างสระประสมเสมอ (เช่น ‘ฉล’ ในคำว่า “เฉลียง”)

จากการศึกษาพยัญชนะ 2 ตัว ที่ปรากฏระหว่างสระประสมในคลังข้อความภาษาไทยแห่งชาติ ผู้วิจัยพบว่า สามารถแบ่งพยัญชนะแต่ละคู่ได้เป็น 2 ประเภทคือ (ก) พยัญชนะ 2 ตัวที่มีรูปคล้ายพยัญชนะควบกล้ำ โดยที่อักษรแต่ละตัวในคู่เหล่านี้ อาจอยู่ชิดกันเป็นพยัญชนะควบกล้ำ (เช่น “คร” ใน “เครื่องมือ”) หรือปรากฏต่างพยางค์กันโดยที่ตัวหน้าเป็นตัวสะกดท้ายพยางค์ และตัวหลังเป็นพยัญชนะต้นของพยางค์ถัดไป (เช่น “คร” ใน“ภาคเรียน”) ซึ่งใน ‘⠯⠗’ (=คร) ทั้งสองกรณีนี้จะเขียนชิดกันเสมอจึงก่อให้เกิดความกำกวม และ (ข) พยัญชนะ 2 ตัวที่ปรากฏในพยางค์เดียวกันเสมอ เช่น “หล” ใน “เหลือ” โดยที่ “ห” และ “ล” ต้องอยู่ชิดกัน ไม่สามารถปรากฏต่างพยางค์กันได้จึงไม่เกิดความกำกวม

เพื่อให้เกิดความสะดวกในการถ่ายถอดอักษร ผู้วิจัยได้สกัดพยัญชนะ 2 ตัวที่ปรากฏระหว่างสระประสมจากคลังข้อความภาษาไทยแห่งชาติ แล้วนำพยัญชนะแต่ละคู่มาตรวจสอบกับ

คลังข้อความ เพื่อวิเคราะห์ว่าพยัญชนะตัวหน้าสามารถปรากฏต่างพยางค์กับพยัญชนะตัวหลังได้หรือไม่ คือพยัญชนะตัวแรกมีความกำกวมว่าจะเป็นตัวสะกดของพยางค์หน้าได้ ทำให้ได้เซ็ตของพยัญชนะควบ 2 เซ็ต คือเซ็ตของพยัญชนะควบกำกวม และเซ็ตของพยัญชนะควบไม่กำกวม ดังแสดงในตารางต่อไปนี้

ตารางที่ 5.14 เซ็ตของพยัญชนะควบกำกวมและพยัญชนะควบไม่กำกวม

เซ็ตของพยัญชนะควบกำกวม				**เซ็ตของพยัญชนะควบไม่กำกวม**	
กร-	ขล-	ถก-	ศร-	ฉพ-	หง-
กล-	ขว-	ถล-	สด-	ฉล-	หน-
กว-	คร-	ทร-	สน-	ผช-	หม-
กษ-	คล-	ปร-	สม-	ผด-	หย-
ขม-	คว-	ปล-	สร-	ผย-	หร-
ขย-	จร-	พร-	ตร-	ผล-	หล-
ขร-	ชอ-	พล-	ตล-	ผอ-	หว-

(3) เซ็ตอักษรเบรลล์ที่นำมาจากคู่เทียบประเภทต่าง ๆ

นอกจากเซ็ตของอักษรเบรลล์ที่สร้างขึ้นนี้ ผู้วิจัยยังนำอักษรเบรลล์ในตารางคู่เทียบต่าง ๆ มาใช้เป็นเซ็ตของกลุ่มตัวอักษรเพื่อช่วยในการตรวจสอบเงื่อนไขของระบบเมื่อต้องถ่ายถอดอักขระเบรลล์ที่กำกวม เซ็ตอักษรในกลุ่มนี้ประกอบด้วย เซ็ตของสระทั่วไป เซ็ตของสระพิเศษ เซ็ตของสระหน้า เซ็ตของวรรณยุกต์ และเซ็ตของสระประสม (คือ เซ็ตอักษรเบรลล์ที่ปรากฏในตารางที่ 5.4 , 5.5 , 5.6, 5.7 และ 5.9 ตามลำดับ) ซึ่งเซ็ตของอักษรเบรลล์เหล่านี้จะนำมาใช้ในการตัดสินใจของระบบ ซึ่งจะได้แสดงรายละเอียดในหัวข้อที่ 5.2 ต่อไป

(4) เซ็ตที่สกัดจากคลังข้อความภาษาไทยแห่งชาติ

ข้อมูลประเภทนี้ ประกอบด้วยเซ็ตของกลุ่มตัวอักษร 3 ตัว ที่ระบบนำมาใช้ในการตรวจสอบเมื่อต้องการรู้ว่าอักษร 3 ตัวที่ถ่ายถอดแล้วนั้นสามารถปรากฏที่ตำแหน่งท้ายพยางค์หรือท้ายคำได้หรือไม่ หากตรงตามเงื่อนไขใดเงื่อนไขหนึ่ง ที่กล่าวมา ระบบจะดำเนินการไปตามเงื่อนไขที่พบ

ตัวอย่างเช่นในการถ่ายถอด ‘⠷’ หากระบบพบว่า “นธ์” อยู่ในเซ็ตไตรแกรมท้ายพยางค์ แสดงว่า “นธ์” เป็นกลุ่มอักษรที่ปรากฏท้ายคำได้ ระบบก็จะถ่ายถอด ‘⠷’ ที่พบให้เป็นการันต์

ในทางตรงกันข้ามหากนำ “คะ◌์” ไปตรวจสอบกับเซ็ตกลุ่มอักษร แล้วพบว่าไม่อยู่ในเซ็ตดังกล่าว แสดงว่า “คะ◌์” ไม่สามารถปรากฏในตำแหน่งท้ายคำได้ จึงต้องถ่ายถอด ‘◌◌’ ให้เป็น “ ” ” (คำพูดปิด) (ดูหัวข้อ 5.2.2.2.1. ข. ประกอบ)

ในการสร้างเซ็ตไตรแกรมท้ายพยางค์ ผู้วิจัยได้สกัดสามตัวอักษรในบริบทต่าง ๆ จากคลังข้อความภาษาไทยแห่งชาติ (TNC) โดยดำเนินการดังนี้

(1) *สกัดสามตัวอักษรหน้าสระประสม* การสกัดไตรแกรมอักษรที่ปรากฏหน้าสระประสม เพื่อตรวจสอบว่าพยัญชนะ 2 ตัวหน้าสระประสมซึ่งมีรูปคล้ายพยัญชนะควบกล้ำเป็นพยัญชนะควบกล้ำหรือไม่โดยอาศัยเซ็ตของพยัญชนะควบที่กำกวม (ตามตารางที่ 5.14) ในการสกัดจะนำพยัญชนะควบที่กำกวมแต่ละคู่ไปค้นหาในคลังข้อความ โดยกำหนดเงื่อนไขว่าพยัญชนะแรกของคู่จะต้องเป็นตัวสะกดของพยางค์หน้า และพยัญชนะตัวที่สองจะต้องปรากฏภายในสระประสมและเป็นอักษรตัวเดียวในสระประสมเท่านั้น (เช่นนำ “คร”ไปค้นหาในคลังข้อความ ซึ่ง “ค” ต้องเป็นตัวสะกดของพยางค์ก่อนหน้าสระประสม และ “ร” ต้องปรากฏระหว่างสระประสม หากระบบพบคำว่า “ภาคเรียน” ซึ่งตรงตามเงื่อนไขนี้ ระบบจะสกัดสามตัวอักษรคือ“ภาค”มาเก็บไว้ในเซ็ตไตรแกรมท้ายพยางค์เป็นต้น โดยสามตัวอักษรที่หน้าสระประสมนั้นพบเป็นจำนวน 1,251 ชุด (ดูรายละเอียดทั้งหมดในภาคผนวก ก.1)

(2) *สกัดข้อความทุกวรรคที่ลงท้ายด้วย ◌์ (การันต์)* เพื่อระบบจะนำมาใช้ในการตัดสินใจเมื่อพบ ‘◌◌’ ปรากฏท้ายสายอักขระว่าควรจะถ่ายถอดให้เป็น “◌์” หรือ “ ” ” (คำพูดปิด) ซึ่งพบเป็นจำนวน 88 ชุด (ดูรายละเอียดทั้งหมดในภาคผนวก ก.2)

(3) *สกัดข้อความทุกวรรคที่มี ◌้(ไม้โท) ปรากฏภายใน 3 สามตัวอักษรท้ายวรรค* เพื่อระบบนำมาใช้ตัดสินใจเมื่อพบ ‘◌◌’ ปรากฏท้ายสายอักขระว่าควรจะถ่ายถอดให้เป็น “◌้” หรือ “.” (จุด) ซึ่งพบเป็นจำนวน 363 ชุด (ดูรายละเอียดทั้งหมดในภาคผนวก ก.3)

(4) *สกัดข้อความทุกวรรคที่มี ◌๊ (ไม้ตรี) ปรากฏ* โดยสกัดในสองลักษณะคือ

(ก) สกัดทุกวรรคที่ลงท้ายด้วยไม้ตรี เพื่อช่วยในการตัดสินใจว่าจะต้องถ่ายถอด ‘◌◌’ ให้เป็น “◌๊ ” หรือ “) ” ในตำแหน่งท้ายสายอักขระ ซึ่งพบทั้งสิ้น 12 ชุด

(ข) สกัด “◌๊ ” ที่ปรากฏร่วมกับ “ะ” และ “◌๊ ” ที่ปรากฏร่วมกับ“◌ำ” ในตำแหน่งท้ายวรรค เนื่องจาก “ะ” และ “◌ำ” เป็นสระพิเศษ ที่เมื่อปรากฏกับวรรณยุกต์จะมีการจัดเรียงที่แตกต่างกันในสองระบบตัวเขียน และเนื่องจาก ‘◌◌’ เมื่อปรากฏท้ายสายอักขระสามารถตีความให้เป็น “◌๊” หรือ “)”

ดังนั้นเมื่อระบบพบสระพิเศษและ ‘ː’ ปรากฏท้ายสายอักขระ ระบบต้องตัดสินใจว่าจะถ่ายถอดเบรลล์ตัวนี้ให้เป็นรูปใดโดยการตรวจสอบกับไตรแกรมท้ายพยางค์ (ดูรายละเอียดทั้งหมดในภาคผนวก ก.4)

5.1.1.3 ข้อมูลที่ใช้ในการถ่ายถอดเบรลล์อังกฤษ

ข้อมูลที่ใช้ในการถ่ายถอดเบรลล์อังกฤษ ประกอบด้วยคู่เทียบที่ใช้ในการถ่ายถอดเบรลล์อังกฤษรูปย่อ และคู่เทียบที่ใช้ในการถ่ายถอดเบรลล์อังกฤษรูปเต็ม ซึ่งมีรายละเอียดดังนี้

5.1.1.3.1 คู่เทียบเบรลล์อังกฤษรูปย่อ

ในการถ่ายถอดเบรลล์อังกฤษรูปย่อ ระบบต้องอาศัยคู่เทียบประเภทต่าง ๆ ดังนี้

(1) คู่เทียบเบรลล์ต้นสายอักขระ

คู่เทียบกลุ่มนี้ระบบจะใช้ในการถ่ายถอดเบรลล์อังกฤษรูปย่อ และใช้ในการถ่ายถอดเบรลล์อังกฤษรูปเต็มด้วยเช่นเดียวกัน เนื่องจากตัวย่อเบรลล์ที่ผู้วิจัยได้จำแนกไว้นี้มีเครื่องหมายที่ปรากฏในตำแหน่งต้นสายอักขระ ซึ่งพบได้ทั้งเบรลล์อังกฤษรูปเต็มและรูปย่อ เช่น ‘ː’ สามารถปรากฏในตำแหน่งต้นสายอักขระทั้งเบรลล์รูปเต็มและเบรลล์รูปย่อได้เหมือนกัน และต้องตีความให้ใช้แทน “(” (วงเล็บเปิด) เช่นเดียวกัน (ดูบทที่ 4 หัวข้อที่ 4.2.1 และตารางที่ 5.15 ประกอบ)

สำหรับอักขระเบรลล์กลุ่มนี้เมื่อปรากฏในเบรลล์อังกฤษรูปย่อจะแบ่งได้เป็น 2 กลุ่ม คือ อักขระเบรลล์ที่ใช้แทนคำโดด คือใช้แทน to, into, by และอักขระเบรลล์ที่ใช้แทนเครื่องหมายที่ปรากฏหน้าคำ เช่น ((วงเล็บเปิด), “ (คำพูดเปิด) เป็นต้น แต่เมื่อนำอักขระกลุ่มนี้ไปใช้ในการถ่ายถอดเบรลล์อังกฤษรูปเต็ม ระบบจะไม่นำอักขระประเภทที่ใช้แทนคำโดด (to, into, by) มาใช้ในการถ่ายถอด เพราะอักขระประเภทนี้ปรากฏเฉพาะในเบรลล์อังกฤษรูปย่อเท่านั้น

อักษรเบรลล์ต้นสายอักขระสามารถแสดงคู่เทียบได้ตามตารางต่อไปนี้

ตารางที่ 5.15 คู่เทียบเบรลล์ต้นสายอักขระ

เบรลล์	แอสกี	อักษรปกติ
⠼	6	to
⠴	0	by
⠼⠼	96	into
⠲	.	_
⠂	,	(ลบออกจากสายอักขระ)
⠶	7	(
⠦	8	“
⠂⠶	,7	[
⠂⠦	,8	‘

(2) คู่เทียบเครื่องหมายหลังคำ

คู่เทียบกลุ่มนี้ เป็นคู่เทียบที่ใช้ในการถ่ายถอดเบรลล์รูปย่อและเบรลล์รูปเต็มด้วยเช่นกัน เนื่องจากเครื่องหมายเหล่านี้สามารถปรากฏในระบบอักษรเบรลล์อังกฤษทั้ง 2 ระบบได้ (ดูหัวข้อ 4.2.1 และตารางที่ 5.16 ประกอบ) อักขระเบรลล์ที่ใช้แทนเครื่องหมายหลังคำ มีคู่เทียบดังนี้

ตารางที่ 5.16 คู่เทียบเครื่องหมายหลังคำ

เบรลล์	แอสกี	อักษรปกติ
⠈⠒⠿	@3p	%
⠈⠔	@*	*
⠈⠴	@0	%
⠈⠲	@4	$
⠸⠌	_/	/
⠸⠂	_1	,
⠸⠆	_2	;
⠸⠒	_3	:
⠸⠲	_4	.
⠸⠖	_6	!
⠰⠄	0'	`
⠶⠄	7'	]
⠔⠔	99	*
⠤	-	-
⠄	.	.
⠨⠅	.k	=
⠴	0	”
⠂	1	,
⠆	2	;
⠒	3	:
⠲	4	.
⠖	6	!
⠶	7	)
⠦	8	?

หมายเหตุ: เครื่องหมายที่ปรากฏในตาราง บางรูปต้องแสดงด้วยอักขระเบรลล์มากกว่า 1 ตัวไม่สามารถตีความอักขระเบรลล์เหล่านี้ให้แยกจากกันได้เพราะจะกลายเป็นอักษรอื่นที่แตกต่างออกไป เช่น ‘⠨⠏ ⠨⠨⠏’ ต้องใช้แทน “ % ” หากตีความแต่ละตัวจะได้เป็น “@:p” ซึ่งไม่พบในภาษาอังกฤษปกติ

(3) คู่เทียบตัวย่อระดับคำ

ตัวย่อระดับคำเป็นตัวย่อเบรลล์ที่ปรากฏระหว่าง เครื่องหมายเว้นวรรคและ/หรือเครื่องหมายแสดงขอบเขตของคำ ซึ่งในงานวิจัยชิ้นนี้ผู้วิจัยได้จัดให้ตัวย่อระดับคำประกอบด้วย ตัวย่อกลุ่มต่าง ๆ 4 ประเภทซึ่งได้นำเสนอมาแล้วในบทที่ 4 หัวข้อ 4.2.2.2.1 และเพื่อให้รองรับกับการถ่ายถอดตัวย่อกลุ่มนี้ ผู้วิจัยได้สร้างคู่เทียบตัวย่อระดับคำ ดังที่แสดงไว้ในภาคผนวก ข.1

(4) คู่เทียบตัวย่อเบรลล์เซลล์คู่

คู่เทียบตัวย่อเบรลล์เซลล์คู่ เป็นคู่เทียบที่ใช้สำหรับการถ่ายถอดเบรลล์สองเซลล์ ที่ปรากฏในตำแหน่งต่าง ๆ ของคำ ซึ่งอักขระเบรลล์เหล่านี้ต้องตีความเป็นคู่ ไม่สามารถตีความอักขระแต่ละตัวให้แยกจากกันได้ (ดูหัวข้อ 4.2.2.2 ประกอบ) คู่เทียบสำหรับตัวย่อกลุ่มนี้ได้แจกแจงไว้ในภาคผนวก ข.2

(5) คู่เทียบตัวย่อต้นคำ

คู่เทียบตัวย่อต้นคำเป็นคู่เทียบอักขระเบรลล์ที่ใช้ย่อหน่วยคำประเภทอุปสรรค (prefix) ได้แก่ “be” “con” “dis” “com” (ดูหัวข้อ 4.2.2.2 ประกอบ)

ในการถ่ายถอดตัวย่อกลุ่มนี้ ผู้วิจัยได้สร้างตารางคู่เทียบตัวย่อต้นคำซึ่งแสดงได้ดังนี้

ตารางที่ 5.17 คู่เทียบตัวย่อต้นคำ

เบรลล์	แอสกี	อักษรปกติ
⠆	2	be
⠒	3	con
⠖	4	dis
⠤	-	com

(6) คู่เทียบตัวย่อกลางคำ

ตัวย่อกลางคำเป็นอักขระเบรลล์ต่ำที่ปรากฏในตำแหน่งกลางคำ ตัวย่อในกลุ่มนี้จะซ้ำรูปกับตัวย่อกลุ่มอื่น เช่น เครื่องหมายหลังคำ ตัวย่อต้นคำ เป็นต้น แต่เนื่องจากตัวย่อกลุ่มนี้ปรากฏในตำแหน่งกลางคำจึงต้องตีความให้แตกต่างจากตัวย่อกลุ่มอื่น (ดูหัวข้อ 4.2.2.2 ประกอบ)

ตัวย่อกลางคำมีอักขระเบรลล์และคู่เทียบ ซึ่งแสดงได้ดังนี้

ตารางที่ 5.18 คู่เทียบตัวย่อกลางคำ

เบรลล์	แอสกี	อักษรปกติ
⠰	1	ea
⠸	2	bb
⠼	3	cc
⠾	4	dd
⠶	6	ff
⠷	7	gg

(7) คู่เทียบเบรลล์หลังตัวย่อระดับคำ

คู่เทียบเบรลล์หลังตัวย่อระดับคำเป็นคู่เทียบที่ผู้วิจัยสร้างขึ้นโดยเฉพาะ เนื่องจากผู้วิจัยพบว่ามีหน่วยคำประเภทปัจจัย (suffix) อยู่กลุ่มหนึ่งที่มักจะปรากฏกับตัวย่อระดับคำ แล้วก่อให้เกิดปัญหาต่อการถ่ายถอดด้วยกฎ เนื่องจากเมื่ออักขระกลุ่มนี้ปรากฏกับตัวย่อระดับคำจะทำให้ระบบตีความตัวย่อระดับคำให้เป็นตัวอักษรทั่วไป เช่น ระบบตีความให้ ‘⠰⠾⠗⠎’(=friend-s) เป็น “frs” แทนที่จะเป็น “friends” ดังนั้นในการถ่ายถอดจึงต้องทำการถ่ายถอดตัวย่อกลุ่มนี้ที่ทำหน้าที่เหมือนปัจจัยต่างหากก่อน แล้วจึงทำการถ่ายถอดตัวย่อระดับคำจึงจะได้ผลลัพธ์ที่ถูกต้อง (ดูหัวข้อ 4.2.2.2.1 ประกอบ)

เบรลล์หลังตัวย่อระดับคำประกอบด้วยอักขระเบรลล์และคู่เทียบต่อไปนี้

ตารางที่ 5.19 คู่เทียบอักษรเบรลล์หลังตัวย่อระดับคำ

เบรลล์	แอสกี	อักษรปกติ
⠻	]	er
⠙	d	d
⠗	r	r
⠎	s	s
⠄⠙	'd	'd
⠄⠎	's	's
⠄⠞	't	't
⠲⠎	.s	less
⠰⠎	;s	ness
⠢⠌	e/	est
⠊⠇	li	li
⠽	ly	ly
⠄⠇⠇	'll	'll
⠄⠗⠢	're	're
⠄⠧⠢	've	've
⠩⠊⠩	%ip	ship
⠝⠄⠞	n't	n't

(8) คู่เทียบเครื่องหมายที่ใช้คั่นระหว่างคำ

ในกระบวนการถ่ายถอดเบรลล์อังกฤษรูปย่อที่ผู้วิจัยสร้างขึ้นนั้น จะสามารถถ่ายถอดตัวย่อระดับคำได้เพียงครั้งเดียวในแต่ละสายอักขระ เพราะระบบจะนำส่วนของสายอักขระทั้งส่วนไปทำการตรวจสอบกับคู่เทียบตัวย่อระดับคำในขั้นเริ่มต้นของการถ่ายถอดส่วนนี้ และเพื่อให้ระบบสามารถถ่ายถอดตัวย่อระดับคำที่ปรากฏมากกว่า 1 ชุดได้ ระบบจึงต้องตัดแบ่งสายอักขระออกเป็นสายอักขระย่อย แล้วนำแต่ละสายอักขระย่อยไปทำการถ่ายถอดอักษร

โดยในการแบ่งออกเป็นสายอักขระย่อยนั้นวิธีการหนึ่งที่ระบบดำเนินการคือ แบ่งตามเครื่องหมายที่ใช้คั่นระหว่างคำในสายอักขระ (ดูหัวข้อ 4.2.2.2.1 ประกอบ)

เครื่องหมายที่ใช้คั่นระหว่างคำประกอบด้วยเครื่องหมายต่าง ๆ และคู่เทียบดังนี้

ตารางที่ 5.20 คู่เทียบเครื่องหมายที่ใช้คั่นระหว่างคำ

เบรลล์	แอสกี	อักษรปกติ
⠤⠤	--	--
⠰⠂	_1	,
⠰⠆	_2	;
⠰⠒	_3	:
⠰⠲	_4	.
⠰⠖	_6	!
⠰⠌	_/	/
⠤	-	-

(9) คู่เทียบตัวย่อรูปเต็มคงรูป

ตัวย่อที่ปรากฏในคู่เทียบกลุ่มนี้ เป็นตัวย่อที่เมื่อปรากฏร่วมกันตั้งแต่สองตัวขึ้นไปจะต้องเขียนชิดกันเสมอและต้องตีความให้อักขระแต่ละตัวใช้แทนคำโดดหนึ่งคำ (ดูหัวข้อ 4.2.2.3.2 ประกอบ) ตัวย่อรูปเต็มคงรูปประกอบด้วยคู่เทียบต่าง ๆ ดังนี้

ตารางที่ 5.21 คู่เทียบตัวย่อรูปเต็มคงรูป

เบรลล์	แอสกี	อักษรปกติ
⠯	&	and
⠿	=	for
⠷	(	of
⠮	!	the
⠾	)	with
⠁	a	a

(10) คู่เทียบตัวย่อเบรลล์เซลล์เดี่ยวทั่วไป

คู่เทียบกลุ่มนี้ เป็นคู่เทียบอักขระเบรลล์ที่ต้องปรากฏภายในคำเสมอ โดยที่จะใช้แทนกลุ่มอักษรใดอักษรหนึ่งในอักษรปกติโดยเฉพาะ (ดูหัวข้อ 4.2.2.2 ประกอบ)

ตัวย่อเบรลล์เซลล์เดี่ยวทั่วไป ประกอบด้วยอักขระเบรลล์และคู่เทียบในอักษรปกติดังนี้

ตารางที่ 5.22 คู่เทียบตัวย่อเบรลล์เซลล์เดี่ยวทั่วไป

เบรลล์	แอสกี	อักษรปกติ
⠮	!	the
⠸	$	ed
⠩	%	sh
⠡	*	ch
⠌	/	st
⠱	:	wh
⠹	?	th
⠣	[	ow
⠳	\	ou
⠻	]	er
⠫	<	gh
⠜	>	ar
⠢	5	en
⠔	9	in
⠼	#	ble
⠬	+	ing

5.1.1.3.2 คู่เทียบเบรลล์อังกฤษรูปเต็ม

ในการถ่ายถอดเบรลล์อังกฤษรูปเต็ม ระบบไม่จำเป็นต้องอาศัยตารางคู่เทียบตัวอักษร เนื่องจากในการประมวลผลด้วยคอมพิวเตอร์ตัวอักษรเบรลล์ทั้ง 26 ตัว ตรงกับตัวอักษรที่ใช้กับอังกฤษปกติอยู่แล้ว อย่างไรก็ตาม ในเบรลล์อังกฤษรูปเต็มมีการปรากฏของตัวเลขเบรลล์และเครื่องหมายต่าง ๆ ดังนั้นระบบจึงต้องใช้*ตารางคู่เทียบตัวเลข*ซึ่งประกอบด้วยตัวเลขและเครื่องหมายต่างๆ เพื่อจับคู่กับอักษรปกติ

ตารางที่ 5.23 แสดงคู่เทียบตัวเลขเบรลล์

เบรลล์	แอสกี	อักษรปกติ
⠁	a	1
⠃	b	2
⠉	c	3
⠙	d	4
⠑	e	5
⠋	f	6
⠛	g	7
⠓	h	8
⠊	i	9
⠚	j	0
⠖	+	+
⠤	-	-
⠦	*	*
⠌	/	/
⠄	'	'
⠂	1	,
⠒	3	:
⠲	4	.
⠨	.	.[10]

[10] ‘⠲’ มีรูปแทนในอักษรปกติคือ จุดที่ปรากฏท้ายวรรค ‘⠨’ มีรูปแทนในอักษรปกติคือ จุดทศนิยม หรือ จุดเวลา

5.1.2 ข้อมูลที่ใช้ในการทดสอบระบบ

ข้อความเบรลล์ที่ผู้วิจัยนำมาใช้ในการทดสอบระบบ ประกอบด้วยข้อความ 5 ประเภทได้แก่ ข้อความเบรลล์ไทยล้วน ข้อความเบรลล์อังกฤษรูปเต็ม ข้อความเบรลล์อังกฤษรูปย่อ ข้อความเบรลล์ไทยปนอังกฤษรูปเต็ม และ ข้อความเบรลล์ไทยปนอังกฤษรูปย่อ

สำหรับข้อความเบรลล์ 3 ประเภท ได้แก่ เบรลล์ไทย เบรลล์อังกฤษรูปย่อ และเบรลล์ไทยปนอังกฤษรูปเต็มนั้น ผู้วิจัยได้มาจากศูนย์เทคโนโลยีการศึกษาเพื่อคนตาบอด โดยในการคัดเลือกข้อความเบรลล์เพื่อให้เกิดความหลากหลายด้านคำศัพท์เท่าที่จะเป็นไปได้นั้น ผู้วิจัยได้สุ่มตัวอย่างข้อความทุกประเภทที่กล่าวมานี้ตามหมวดหมู่หนังสือที่ทางศูนย์ฯได้จัดทำไว้ทั้งสิ้น 6 หมวดหมู่ ได้แก่ ความรู้ทั่วไป สังคมศาสตร์ ภาษาศาสตร์ วิทยาศาสตร์ วรรณคดี และกฎหมาย เนื่องจากเชื่อว่าข้อความต่างหมวดหมู่กันจะใช้คำศัพท์เฉพาะด้านที่แตกต่างกัน ทำให้ได้แฟ้มข้อมูล 18 แฟ้มข้อมูล รวมทั้งสิ้น 930 หน้าเบรลล์

ข้อความเบรลล์ 3 ประเภท ที่นำมาจากศูนย์จัดทำสื่ออักษรเบรลล์ของคนตาบอดนี้ มีลักษณะการจัดหน้าเอกสารในรูปแบบเฉพาะ คือ แต่ละหน้าเบรลล์ มีข้อความ 25 บรรทัด แต่ละบรรทัดประกอบด้วยอักขระเบรลล์ไม่เกิน 40 เซลล์ และมีเครื่องหมายเว้นวรรค 2 หรือ 4 ตัว ปรากฏในตำแหน่งต้นบรรทัด นั่นคือบรรทัดที่มีเครื่องหมายเว้นวรรค 2 ตัว แสดงถึงข้อความทั่วไป และบรรทัดที่มีเครื่องหมายเว้นวรรค 4 ตัว แสดงว่ามีการขึ้นย่อหน้าใหม่ นอกจากนี้ยังมีการเขียนเลขหน้าของต้นฉบับอักษรปกติไว้ที่มุมบนขวาของแต่ละหน้าเบรลล์อีกด้วย โดยในการระบุเลขหน้าของต้นฉบับจะมีรูปแบบคือ อักษรอังกฤษตามด้วยเลขหน้าเช่น a40, b40 , c40 แสดงว่า ข้อความเบรลล์นั้นตรงกับต้นฉบับหน้า 40 โดยที่ในต้นฉบับหน้า 40 นั้นสามารถถ่ายถอดเป็นอักษรเบรลล์ได้ 3 หน้าเบรลล์ คือ หน้า a , b และ c เป็นต้น

ข้อความไทยปนอังกฤษที่นำมาจากแหล่งข้อมูลอักษรเบรลล์ มีลักษณะของการปนภาษาใน 2 รูปแบบ คือ รูปแบบแรกเป็นการปนภาษาที่มีความถี่ในการสลับภาษาสูง โดยที่ข้อความประเภทนี้นำมาจากหนังสือเรียนภาษาอังกฤษที่อธิบายเกี่ยวกับการใช้คำนาม จึงมีการกล่าวถึงคำศัพท์และคำแปลสลับกันเป็นจำนวนมากเช่น

“... โลหะ ของแข็ง ของเหลว เช่น
air อากาศ cloth ผ้า oil น้ำมัน
wood ไม้ silver เงิน ink หมึก luggage

สัมภาระ water น้ำ cream ครีม

furniture เฟอร์นิเจอร์ cake ขนมเค้ก...”

รูปแบบที่ 2 เป็นข้อความที่มีภาษาอังกฤษและภาษาไทยสลับกันย่อหน้าต่อย่อหน้า ข้อความประเภทนี้ได้มาจาก เอกสารปฏิญญาสากลว่าด้วยสิทธิมนุษยชน พ.ศ. 2553 ดังตัวอย่างต่อไปนี้

“...Article 13 (1) Everyone has the right to

freedom of movement and residence

within the borders of each state.

ข้อ 13 (1) ทุกคนมีสิทธิใน

อิสรภาพแห่งการเคลื่อนย้ายและการอยู่

อาศัยภายในพรมแดนของแต่ละรัฐ...”

จากที่กล่าวมา จะเห็นว่ามีข้อมูลเพียง 3 ประเภทเท่านั้นที่ผู้วิจัยได้จากศูนย์เทคโนโลยีการศึกษาเพื่อคนตาบอด เหตุที่เอกสารเบรลล์ที่ได้จากแหล่งข้อมูลนี้เป็นเพียงตัวบทเบรลล์อังกฤษรูปย่อไม่มีตัวบทที่เป็นอังกฤษรูปเต็ม เนื่องจากโดยทั่วไปแล้วเอกสารเบรลล์ที่เป็นภาษาอังกฤษมักจะอยู่ในรูปของอังกฤษรูปย่อเพื่อเป็นการประหยัดทรัพยากรและเวลาในการผลิตเพราะอักษรมีขนาดใหญ่ นอกจากนี้ข้อความปนภาษาที่ผู้วิจัยได้มามีแค่เพียงไทยปนอังกฤษรูปเต็มเท่านั้น เนื่องจากยังไม่มีโปรแกรมคอมพิวเตอร์ที่รองรับกับการถ่ายถอดเบรลล์ไทยปนอังกฤษรูปย่อที่ใช้กันในวงการคนตาบอดไทย จึงทำให้ขาดข้อความอีก 2 ประเภทที่จำเป็นต่องานวิจัยชิ้นนี้ คือตัวบทเบรลล์อังกฤษรูปเต็ม และตัวบทเบรลล์ไทยปนอังกฤษรูปย่อ

ดังนั้นผู้วิจัยจึงได้นำข้อความเบรลล์อังกฤษรูปย่อทั้งหมดที่ได้จากแหล่งข้อมูลข้างต้น มาทำการถ่ายถอดให้เป็นเบรลล์อังกฤษรูปเต็มโดยใช้โปรแกรม Duxbury ส่วนตัวบทเบรลล์ไทยปนอังกฤษรูปย่อได้จากการนำเอกสารเบรลล์ไทยปนอังกฤษรูปเต็มมาทำการแก้ไขในส่วนที่เป็นอังกฤษรูปเต็มให้เป็นรูปย่อด้วยตัวผู้วิจัยเอง จากนั้นทำการตรวจสอบความถูกต้องของข้อมูลทั้ง 2 ประเภท

ข้อความเบรลล์ทั้ง 5 ประเภท ที่ใช้ในการทดสอบระบบนั้นได้แสดงตัวอย่างไว้แล้วในภาคผนวก ซ. โดยที่ข้อความเบรลล์แต่ละประเภทมีปริมาณดังนี้

ตารางที่ 5.24 ปริมาณของข้อความเบรลล์แต่ละประเภทที่นำมาทดสอบระบบ

ข้อความเบรลล์ไทย	ข้อความเบรลล์อังกฤษรูปย่อ	ข้อความเบรลล์ไทยปนอังกฤษรูปเต็ม	ข้อความเบรลล์อังกฤษรูปเต็ม	ข้อความเบรลล์ไทยปนอังกฤษรูปย่อ
300 หน้าเบรลล์	390 หน้าเบรลล์	240 หน้าเบรลล์	390 หน้าเบรลล์	240 หน้าเบรลล์

เป็นที่น่าสังเกตว่า ข้อความเบรลล์แต่ละประเภทมีจำนวนหน้าเอกสารที่แตกต่างกัน โดยเฉพาะอย่างยิ่งข้อความเบรลล์ไทยปนอังกฤษ ซึ่งมีจำนวนหน้าเอกสารน้อยกว่าข้อความอีก 2 ประเภท เหตุที่เป็นเช่นนี้เนื่องจากแฟ้มข้อมูลที่สุ่มมานั้น เป็นแฟ้มข้อมูลชุดสมบูรณ์ที่ผู้วิจัยไม่ได้ทำการตัดทอนจำนวนหน้าลง จึงมีจำนวนหน้าเอกสารที่แตกต่างกัน แต่ตรงตามเกณฑ์ที่ผู้วิจัยได้กำหนดไว้ อนึ่ง ก่อนที่จะอภิปรายถึงผลของการทดสอบระบบ ผู้วิจัยจะขอนำเสนอระบบการถ่ายถอดเบรลล์ไทยปนอังกฤษโดยใช้กฎทางภาษาที่ผู้วิจัยได้พัฒนาขึ้น ซึ่งมีรายละเอียดดังนี้

5.2 ระบบการถ่ายถอดเบรลล์อักษรโดยใช้กฎ

ในการถ่ายถอดเบรลล์ไทยปนอังกฤษโดยใช้กฎนั้น จะมี 2 ระบบที่รองรับกับข้อความเบรลล์ประเภทต่าง ๆ คือ ระบบที่ใช้ถ่ายถอดเบรลล์ไทยปนอังกฤษรูปเต็ม และระบบที่ใช้ถ่ายถอดเบรลล์ไทยปนอังกฤษรูปย่อ ซึ่งแต่ละระบบมีภาพรวมในการทำงานดังนี้

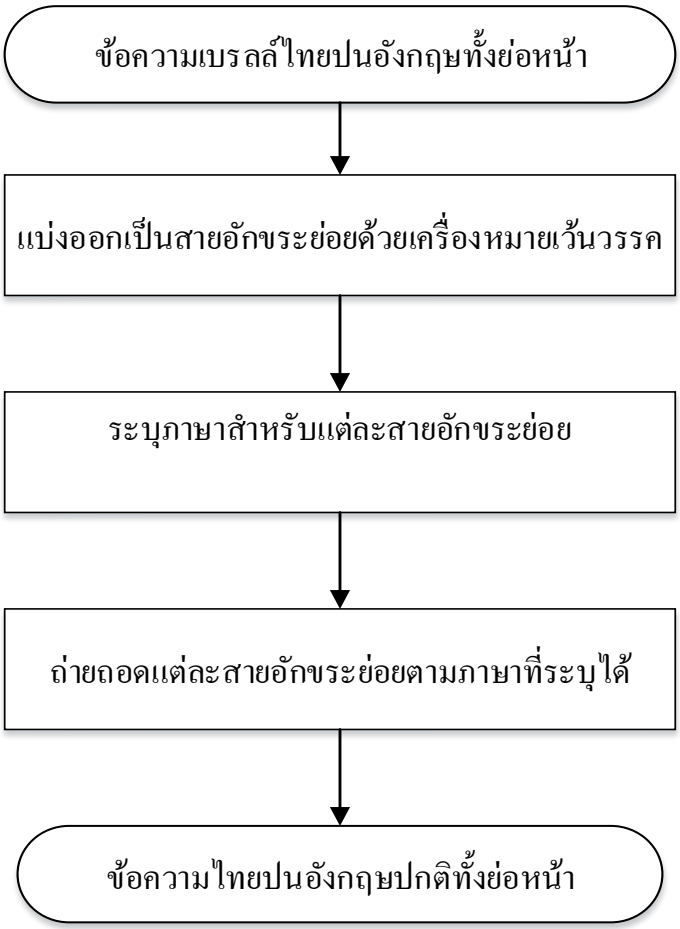
(1) ระบบถ่ายถอดเบรลล์ไทยปนอังกฤษรูปย่อ

ในระบบถ่ายถอดเบรลล์ไทยปนอังกฤษรูปย่อ เริ่มจากการนำเข้าข้อความทีละย่อหน้า แล้วนำแต่ละย่อหน้ามาแยกออกเป็นสายอักขระต่าง ๆ โดยใช้เครื่องหมายเว้นวรรค (space) ที่ปรากฏในย่อหน้าเป็นตัวแบ่งสายอักขระ จากนั้นระบบจะนำแต่ละสายอักขระไปตรวจสอบว่าเป็นภาษาไทยหรือภาษาอังกฤษตามกระบวนการในข้อที่ 5.2.1 โดยอาศัยเซ็ตของสายอักขระที่เป็นเอกลักษณ์ของภาษาไทย (ภาคผนวก ค.3) และเซ็ตของสายอักขระที่เป็นเอกลักษณ์ของภาษาอังกฤษ (ภาคผนวก ค.4) ต่อมาทำการถ่ายถอดแต่ละสายอักขระตามภาษาที่ระบุได้ นั่นคือ หากพบว่าสายอักขระปัจจุบันเป็นภาษาไทยจะ

ดำเนินการตามกระบวนการในข้อที่ 5.2.2 แต่หากพบว่าเป็นภาษาอังกฤษจะดำเนินการตามกระบวนการในข้อที่ 5.2.3

เมื่อดำเนินการระบุภาษาและถ่ายถอดจนครบทุกสายอักขระแล้ว ระบบจะนำสายอักขระที่เป็นอักษรปกติมาจัดเรียงตามลำดับการนำเข้า แล้วใช้เป็นผลลัพธ์ของการถ่ายถอดอักษรทั้งย่อหน้า

ภาพรวมการทำงานของระบบการถ่ายถอดเบรลล์ไทยปนอังกฤษโดยใช้กฎ[11] แสดงได้ดังนี้



[11] ในการถ่ายถอดเบรลล์ไทยปนอังกฤษโดยใช้กฎนั้น ข้อความที่นำเข้าต้องเป็นรูปแบบใดรูปแบบหนึ่งคือ เบรลล์ไทยปนอังกฤษรูปเต็ม หรือเบรลล์ไทยปนอังกฤษรูปย่อ ระบบไม่รองรับกับการปนของเบรลล์อังกฤษรูปเต็มและรูปย่อในข้อความเดียวกัน

ภาพที่ 5.1 ภาพรวมของระบบถ่ายถอดเบรลล์ไทยปนอังกฤษโดยใช้กฎ

(2) ระบบถ่ายถอดเบรลล์ไทยปนอังกฤษรูปเต็ม

ในการถ่ายถอดเบรลล์ไทยปนอังกฤษรูปเต็ม ระบบจะดำเนินการเช่นเดียวกับระบบการถ่ายถอดเบรลล์ไทยปนอังกฤษรูปย่อ แต่ในการระบุภาษานั้นจะใช้อาศัยเซ็ตของสายอักขระที่เป็นเอกลักษณ์ภาษาไทย (ตามภาคผนวก ค.1) และเซ็ตของสายอักขระที่เป็นเอกลักษณ์ภาษาอังกฤษ (ตามภาคผนวก ค. 2) และเพื่อให้ระบบรองรับกับการถ่ายถอดเบรลล์อังกฤษรูปเต็มเท่านั้นจึงจำเป็นต้องแก้ไขกระบวนการถ่ายถอดเบรลล์อังกฤษในส่วนของการระบุประเภทสายอักขระ (ข้อ 5.2.3.2 ง.) ซึ่งแต่เดิมหากระบบพบว่าตรงกับเงื่อนไขในส่วนนี้ระบบจะระบุให้สายอักขระย่อยเป็นเบรลล์รูปย่อแต่จำเป็นต้องแก้ไขให้ระบบระบุสถานะของสายอักขระย่อยเป็นเบรลล์รูปเต็มแทน

ในระบบการถ่ายถอดอักษรเบรลล์ที่กล่าวมาข้างต้นประกอบด้วยกระบวนการต่าง ๆ ซึ่งแต่ละกระบวนการมีรายละเอียดในการทำงานดังนี้

5.2.1 การระบุภาษา

ในการระบุภาษาของข้อความนั้น ระบบจะระบุภาษาทีละสายอักขระ โดยมีการดำเนินการ 2 ขั้น คือ การหาจุดเริ่มต้นของการระบุภาษาในสายอักขระ และการระบุภาษาโดยอาศัยกฎต่าง ๆ

5.2.1.1 การหาจุดเริ่มต้นของการระบุภาษา

ในส่วนต้นของสายอักขระ มักจะมีเครื่องหมายเบรลล์ต้นสายอักขระตามตารางที่ 5.2 ปรากฏอยู่ โดยที่เครื่องหมายเหล่านี้จะพบได้ในสายอักขระเบรลล์ไทย เบรลล์อังกฤษรูปเต็ม และเบรลล์อังกฤษรูปย่อได้เช่นเดียวกัน จึงทำให้ระบบไม่สามารถใช้เซ็ตของสายอักขระที่เป็นเอกลักษณ์ของแต่ละภาษามาใช้ในการระบุภาษาได้ ดังนั้นเพื่อให้สามารถระบุภาษาโดยใช้เซ็ตของสายอักขระที่เป็นเอกลักษณ์ได้ ระบบจึงต้องหาจุดเริ่มต้นของการระบุภาษา โดยตรวจสอบว่ามีอักขระเบรลล์เหล่านี้ปรากฏต้นสายอักขระหรือไม่ หากพบอักขระเหล่านี้ระบบจะเลื่อนตำแหน่งจุดเริ่มต้นออกไปทางขวาเรื่อย ๆ จนกว่าเบรลล์เซลล์ปัจจุบันไม่ตรงกับเครื่องหมายในตารางที่ 5.2 จึงจะเริ่มทำการระบุภาษาด้วยการใช้สายอักขระที่เป็นเอกลักษณ์ของแต่ละภาษา

5.2.1.2 กฎที่ใช้ในการระบุภาษา

หลังจากที่ระบบพบจุดเริ่มต้นของการระบุภาษาแล้วระบบจะนำเบรลล์เซลล์ปัจจุบันและสองเซลล์ถัดไป(รวมเป็น3เซลล์)ไปตรวจสอบกับเซ็ตของสายอักขระที่เป็นเอกลักษณ์ของภาษาไทยและ

ภาษาอังกฤษ โดยที่ในการระบุภาษาสำหรับข้อความเบรลล์ไทยปนอังกฤษรูปเต็ม ระบบจะใช้เซ็ตสายอักขระเบรลล์ไทยที่ใช้กับข้อความเบรลล์ไทยปนอังกฤษรูปเต็มตามภาคผนวก ค.1 ร่วมกับเซ็ตสายอักขระเบรลล์อังกฤษตามภาคผนวก ค.2 ส่วนการระบุภาษาสำหรับข้อความเบรลล์ไทยปนอังกฤษรูปย่อนั้น จะใช้เซ็ตสายอักขระเบรลล์ไทยที่ใช้กับข้อความเบรลล์ไทยปนอังกฤษรูปย่อตามภาคผนวก ค.3 ร่วมกับเซ็ตสายอักขระเบรลล์อังกฤษตามภาคผนวก ค.4

ในการระบุภาษาโดยใช้สายอักขระที่เป็นเอกลักษณ์ ระบบจะดำเนินการตามเงื่อนไขดังนี้

(1) ถ้าอักขระเบรลล์ 1 เซลล์ 2 เซลล์ หรือ 3 เซลล์ปัจจุบัน ตรงกับสายอักขระที่เป็นเอกลักษณ์ของภาษาไทยหรือภาษาอังกฤษระบบจะระบุให้สายอักขระปัจจุบันเป็นภาษานั้นแต่ถ้าอักขระเบรลล์ดังกล่าวไม่ตรงกับสายอักขระที่เป็นเอกลักษณ์ของภาษาใด ระบบจะดำเนินการในขั้นต่อไป

(2) ถ้าสายอักขระก่อนหน้าเป็นภาษาไทย ระบบจะระบุให้สายอักขระปัจจุบันเป็นภาษาไทยเช่นเดียวกัน แต่ถ้าสายอักขระก่อนหน้าไม่ใช่ภาษาไทย ระบบจะดำเนินการในขั้นที่ 3

(3) ถ้าสายอักขระก่อนหน้าเป็นภาษาอังกฤษ ระบบจะตรวจสอบเงื่อนไขคือ หากสายอักขระปัจจุบันมีความยาวมากกว่า 22 ตัว[12] ระบบจะระบุให้เป็นภาษาไทย แต่ถ้ามีความยาว 22 ตัวหรือน้อยกว่านั้นระบบจะระบุให้สายอักขระปัจจุบันเป็นภาษาอังกฤษ โดยที่ค่าความยาวที่กำหนดนี้จะใช้กับทั้งเบรลล์อังกฤษรูปเต็มและเบรลล์อังกฤษรูปย่อ

(4) หากสายอักขระก่อนหน้าไม่สามารถระบุภาษาได้ เนื่องจากสายอักขระปัจจุบันเป็นสายอักขระแรกของแฟ้มข้อมูล ระบบจะตรวจสอบความยาวของสายอักขระปัจจุบันถ้าความยาวมากกว่า 22 ตัวจะระบุให้เป็นภาษาไทย แต่ถ้าความยาว 22 ตัวหรือน้อยกว่านั้นจะระบุให้เป็นภาษาอังกฤษ

หลังจากที่ระบุภาษาของสายอักขระปัจจุบันได้แล้ว ระบบจะนำสายอักขระนี้ไปถ่ายถอดตามภาษาที่ระบุได้ ซึ่งกระบวนการถ่ายถอดอักษรแต่ละภาษามีการทำงานที่แตกต่างกัน ดังจะได้กล่าวต่อไป

5.2.2 การถ่ายถอดเบรลล์ไทย

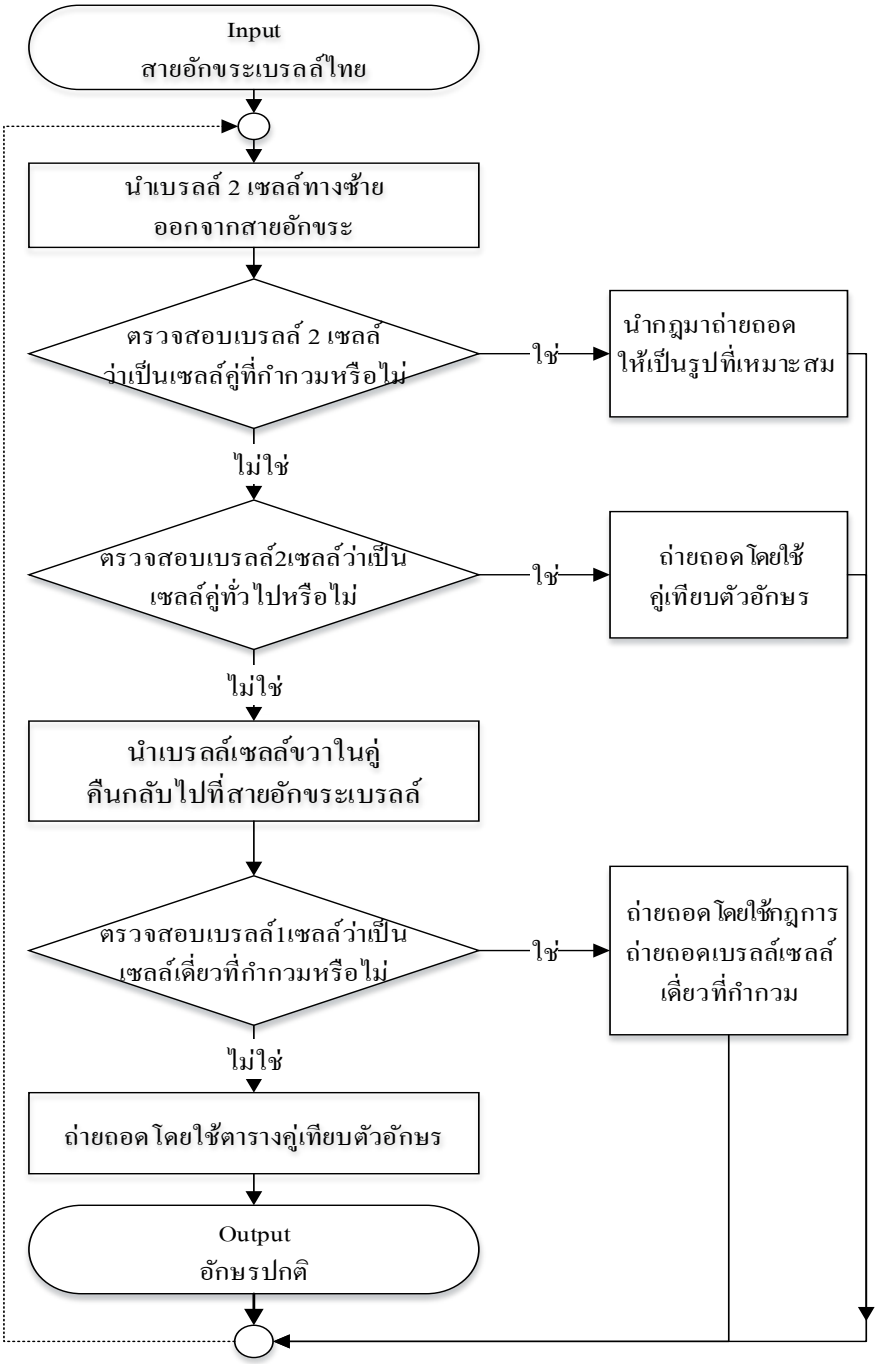
หลังจากระบุได้ว่าสายอักขระปัจจุบันเป็นภาษาไทย ระบบจะนำสายอักขระปัจจุบันมาทำการถ่ายถอดให้เป็นไทยปกติ ซึ่งมีภาพรวมในการทำงานเริ่มจาก การพยายามถ่ายถอดเบรลล์เซลล์คู่ก่อน

---

[12] เหตุที่ผู้วิจัยกำหนดความยาวของสายอักขระในการจำแนกภาษาไว้ที่ 22 เซลล์ เนื่องจากข้อมูลที่นำมาใช้ทดสอบมีวรรค (หรือสายอักขระ) ที่เป็นภาษาอังกฤษที่ยาวที่สุดซึ่งมีขนาด 22 เซลล์คือ ‘⠠⠋⠊⠝⠛⠻⠏⠗⠊⠝⠞⠸⠉⠗⠊⠍⠊⠝⠁⠇⠤⠓⠊⠎⠞⠕⠗⠽’ (=fingerprint/criminal-history) ผู้วิจัยจึงนำค่าความยาวนี้มาช่วยในการระบุภาษา

โดยตรวจสอบว่าเป็นเบรลล์เซลล์คู่ที่กำกวมหรือไม่ หากกำกวม ให้พิจารณาแก้ไขความกำกวมนั้นก่อน หากไม่กำกวมแต่เป็นเบรลล์เซลล์คู่ทั่วไปก็ให้ถ่ายถอดคู่อักษรนั้น แต่หากไม่ใช่เบรลล์เซลล์คู่จะประมวลผลเพียงเซลล์เดียวโดยพิจารณาว่ากำกวมหรือไม่เช่นเดียวกัน

ขั้นตอนการถ่ายถอดเบรลล์ไทย มีภาพรวมดังแผนภาพต่อไปนี้



ภาพที่ 5.2 กระบวนการถ่ายถอดเบรลล์ไทยโดยใช้กฎ

จากแผนภาพข้างต้นจะพบว่าในการถ่ายถอดเบรลล์ไทย ระบบจะนำสายอักขระปัจจุบันไปดำเนินการตามขั้นตอน (โดยที่ในการตรวจสอบอักขระเบรลล์จะอาศัยตารางคู่เทียบอักษรเบรลล์ประเภทต่าง ๆ ซึ่งได้แสดงไว้ในบทที่ 4) ซึ่งการดำเนินการมีรายละเอียดดังนี้

ขั้นตอนที่ 1 ตรวจสอบเบรลล์เซลล์ปัจจุบันและเซลล์ถัดไปว่าเป็นเบรลล์คู่ที่กำกวมหรือไม่

หากพบว่าเป็นเบรลล์คู่ที่กำกวม ระบบจะนำเบรลล์คู่นี้ไปตรวจสอบกับกฎเพื่อถ่ายถอดให้เป็นกรณีใดกรณีหนึ่งตามกฎที่ระบุไว้ แล้วนำเบรลล์ถัดไปอีก 2 เซลล์มาดำเนินการขั้นนี้ซ้ำ แต่หากเบรลล์ 2 เซลล์นี้ไม่กำกวม ระบบจะดำเนินการในขั้นต่อไป

ขั้นตอนที่ 2 ตรวจสอบเบรลล์เซลล์ปัจจุบันและเซลล์ถัดไป ว่าเป็นเบรลล์เซลล์คู่ทั่วไปหรือไม่

หากพบว่าเป็นเซลล์คู่ทั่วไป ระบบจะถ่ายถอดให้เป็นอักษรปกติโดยอาศัยตารางคู่เทียบเบรลล์เซลล์คู่ทั่วไป (ตารางที่ 5.8) แล้วนำเบรลล์ถัดไปอีก 2 เซลล์มาดำเนินการตั้งแต่ขั้นแรก แต่หากไม่ใช่เบรลล์เซลล์คู่ทั่วไป ระบบจะดำเนินการในขั้นต่อไป

ขั้นตอนที่ 3 ตรวจสอบเฉพาะเบรลล์เซลล์ปัจจุบัน ว่าเป็นเบรลล์เซลล์เดี่ยวที่กำกวมหรือไม่

หากพบว่าเป็น ระบบจะทำการถ่ายถอดให้เป็นกรณีใดกรณีหนึ่งโดยอาศัยกฎที่ระบุไว้ แล้วนำเบรลล์ถัดไปอีก 2 เซลล์มาดำเนินการตั้งแต่ขั้นแรก แต่หากไม่ใช่เบรลล์เซลล์เดี่ยวที่กำกวม ระบบจะดำเนินการในขั้นต่อไป

ขั้นตอนที่ 4 ถ่ายถอดเบรลล์เซลล์ปัจจุบันให้เป็นอักษรไทยปกติ

ในการถ่ายถอดอักษรจะอาศัยตารางคู่เทียบเบรลล์เซลล์ เดี่ยวทั่วไป (ตารางที่ 5.3) แล้วนำเบรลล์ถัดไปอีก 2 เซลล์มาดำเนินการตั้งแต่ขั้นตอนที่ 1 - 4 ไปจนสิ้นสุดสายอักขระปัจจุบัน

จากขั้นตอนที่นำเสนอมาข้างต้นจะเห็นได้ว่าการถ่ายถอดเบรลล์ไทยเป็นไทยปกติประกอบด้วย *การถ่ายถอดเบรลล์เซลล์คู่* และ*การถ่ายถอดเบรลล์เซลล์เดี่ยว* โดยระบบจะดำเนินการดังนี้

5.2.2.1 การถ่ายถอดเบรลล์เซลล์คู่

ในอักษรเบรลล์ อักขระแต่ละตัวจะประกอบด้วยจุดไม่เกิน 6 จุด ดังนั้นจึงไม่สามารถแสดงตัวอักษรไทยได้ครบทุกตัวอักษร ทำให้จำเป็นต้องใช้อักขระเบรลล์ 2 ตัว เพื่อแทนอักขระไทย 1 ตัวอักษร

สำหรับอักษรไทยบางตัว เช่น ‘⠫ ⠻’ (เท่ากับ “ณ”) ‘⠫ ⠽’ (เท่ากับ “ญ”) นอกจากนี้ยังมีอักขระเบรลล์บางคู่ที่สามารถตีความได้ 2 กรณี คือตีความให้อักขระเบรลล์ 2 ตัวนั้น ใช้แทนอักษรปกติ 1 ตัว หรือ 2 ตัวได้ โดยจะแปรไปตามบริบทที่อักขระเบรลล์คู่เหล่านั้นปรากฏอยู่ เช่น ‘⠹ ⠽’ สามารถตีความเป็น “์ท” หรือ “ธ” ได้ จึงทำให้เกิดปัญหาในการถ่ายถอดเบรลล์เซลล์คู่เหล่านี้

ในการถ่ายถอดเบรลล์เซลล์คู่ ระบบจะดำเนินการใน 2 ขั้นตอน คือ ตรวจสอบว่าเบรลล์ 2 เซลล์ปัจจุบันเป็นเบรลล์เซลล์คู่ที่กำกวมหรือไม่ หากเป็นเบรลล์เซลล์คู่ที่กำกวมระบบจะถ่ายถอดตามเงื่อนไขที่ระบุไว้สำหรับเบรลล์แต่ละคู่โดยอาศัยตารางที่ 5.12 แต่หากไม่ใช่เบรลล์เซลล์คู่ที่กำกวม ระบบจะตรวจสอบว่าเป็นเบรลล์เซลล์คู่ทั่วไปหรือไม่โดยอาศัยตารางที่ 5.8 ซึ่งการถ่ายถอดเบรลล์เซลล์คู่แต่ละขั้นตอนมีรายละเอียดดังนี้

5.2.2.1.1 การถ่ายถอดเบรลล์เซลล์คู่ที่กำกวม

ดังที่ได้กล่าวมาแล้วในหัวข้อที่ 4.1.1 ว่า ในเบรลล์ไทยมีอักขระเบรลล์อยู่ 4 คู่ ที่สามารถตีความให้เป็นอักษรปกติในลักษณะที่ อักขระเบรลล์ 2 ตัวใช้แทนอักษรไทย 1 ตัวอักษร หรือ อักขระเบรลล์แต่ละตัวใช้แทนอักษรไทยแต่ละตัวได้ ดังนั้น เมื่อระบบพบเบรลล์เซลล์คู่เหล่านี้ระบบจะทำการถ่ายถอดโดยอาศัยบริบทซ้าย และ/หรือ บริบทขวา ที่แวดล้อมเบรลล์เซลล์คู่เหล่านี้โดยจะดำเนินการกับเบรลล์เซลล์คู่แต่ละชุด ดังนี้

ก. การถ่ายถอด ‘⠮ ⠷’ (“อะ” หรือ “เาะ”)

‘⠮ ⠷’ สามารถตีความเป็น “อ-ะ” หรือ “เาะ” ได้ การที่จะตีความว่าเป็นกรณีใด ระบบต้องตรวจสอบเบรลล์คู่นี้ว่าปรากฏในตำแหน่งต้นพยางค์หรือไม่ หากตรวจสอบแล้วพบว่า ‘⠮ ⠷’ ปรากฏต้นพยางค์ระบบจะถ่ายถอดเป็น “อะ” หากไม่ได้ปรากฏต้นพยางค์ระบบจะถ่ายถอดเป็น “เาะ” ซึ่งการที่จะระบุได้ว่า ‘⠮ ⠷’ ปรากฏต้นพยางค์หรือไม่ ต้องทำการตรวจสอบตำแหน่งของอักขระเบรลล์สองตัวนี้ว่าปรากฏในส่วนใดของสายอักขระและมีอักขระใดปรากฏเป็นบริบทซ้ายและ/หรือขวา

ระบบจะระบุให้ ‘⠮ ⠷’ ปรากฏในตำแหน่งต้นพยางค์ได้หรือไม่ โดยดำเนินการดังนี้

(1) ระบบจะทำการตรวจสอบว่า 3 ตัวอักษรที่ปรากฏหน้า ‘⠮ ⠷’ (“อะ” หรือ “เาะ”) ตรงกับเซ็ตไตรแกรมท้ายพยางค์หรือไม่ หากตรงกับเซ็ตดังกล่าวแสดงว่า ‘⠮ ⠷’ อยู่ในตำแหน่งต้นพยางค์ ระบบจึงต้องถ่ายถอด ‘⠮ ⠷’ให้เป็น “อะ” เช่น‘⠭ ⠮ ⠷ ⠭ ⠮ ⠷ ⠇ ⠭ ⠷ ⠭ ⠮’ (=ซ- ื- ้อ-อ-ะ-ไ-ห-ล-่) ที่ใช้เขียนคำว่า “ซื้ออะไหล่”

(2) ระบบจะตรวจสอบว่า ตัวอักษร 1 ตัวข้างหน้า ‘ ’ (“อะ” หรือ “เาะ”) เป็นพยัญชนะไทยหรือไม่ ถ้าอักษรที่ทำการตรวจสอบไม่ใช่พยัญชนะไทย ก็แสดงว่า ‘ ’ปรากฏต้นพยางค์ จึงต้องถ่ายถอด ‘ ’ให้เป็น “อะ” เช่น ‘ ’ (=จ-ะ-ท-ำ-อ-ะ-ไ-ร) “จะทำอะไร”

(3) ระบบจะทำการตรวจสอบว่า 2 ตัวอักษรที่ปรากฏหลัง ‘ ’ (“อะ” หรือ “เาะ”) เป็น ‘ ’ (=ไร) หรือไม่ หากเป็นไปตามเงื่อนไขนี้ระบบจะถ่ายถอด ‘ ’ ให้เป็น “อะ” เช่น ‘ ’ (=ฟ-เ-ร- ์-น-อ-ะ-ไ-ร)

(4) หากระบบตรวจสอบแล้วพบว่า ไม่เป็นไปตามเงื่อนไขข้างต้น ระบบจะระบุให้ ‘ ’ เป็น “เาะ” แต่เนื่องจาก “เาะ” เป็นสระประสม จึงต้องอาศัยกระบวนการถ่ายถอดสระประสมต่อไป

จากข้อ (4) ถ้าพิจารณาจากรูปของ“เาะ” ในอักษรปกติ จะพบว่าสระรูปนี้จะมีองค์ประกอบสองส่วนคือ เ-า และ -ะ โดยที่จะมีพยัญชนะ 1 หรือ 2 ตัว ปรากฏระหว่าง “เ” และ “า” ซึ่งลักษณะเช่นนี้เหมือนกับการจัดเรียงของสระประสมอื่น ๆ เช่น “เ ียะ” ที่ต้องมีพยัญชนะ 1-2 ตัวอยู่ระหว่างสระรูปนี้เช่นกัน เพื่อให้ระบบสามารถถ่ายถอด“เาะ”ได้ระบบต้องแก้ไข ‘ ’ ในตำแหน่งเบรลล์เซลล์ปัจจุบันให้เป็น ‘ ’(=เา) ทำให้เบรลล์ 2 เซลล์ปัจจุบันมีรูปเป็น ‘ ’ แล้วนำอักขระเบรลล์ 2 ตัวนี้ไปถ่ายถอดตามกระบวนการถ่ายถอดสระประสมในหัวข้อ 5.2.2.2.1 1) ต่อไป

ข. การถ่ายถอด ‘ ’ (“ ์ท” หรือ “ธ”)

จากตารางที่ 5.14 ‘ ’ สามารถตีความได้ 2 กรณี คือ “ ์ท” หรือ “ธ” การที่ระบบจะถ่ายถอดให้เป็นรูปใดนั้น จะต้องตรวจสอบบริบทซ้ายและ/หรือขวาของอักขระเบรลล์คู่นี้เช่นเดียวกับการถ่ายถอด ‘ ’ (“อะ” หรือ “เาะ”)

ในการตรวจสอบบริบทที่ปรากฏร่วมกับ ‘ ’ (“ ์ท” หรือ “ธ”) ระบบจะอาศัยเงื่อนไขต่าง ๆ แล้วถ่ายถอดให้ตรงกับรูปอักษรปกติที่ระบุไว้ในเงื่อนไข ดังนี้

(1) หากระบบตรวจสอบบริบททางขวาของ‘ ’ แล้วพบว่าอักขระเบรลล์สองตัวเป็น ‘ ’ (=รร) ระบบจะถ่ายถอด‘ ’ เป็น “ธ”

จากการสำรวจคลังข้อความภาษาไทยแห่งชาติ ผู้วิจัยพบว่า “ธ” มีบริบทขวาเป็น “รร” เป็นจำนวนมาก เช่น ธรรม ธรรมชาติ ธรรมดา ธรรมะ จึงนำผลที่ได้นี้มากำหนดเงื่อนไขในการถ่ายถอด ‘ ’ ให้เป็น “ธ” นอกจากนี้เหตุผลประการสำคัญที่ต้องกำหนดเงื่อนไขข้อนี้ขึ้น เนื่องจากหากไม่มีเงื่อนไขนี้จะทำให้ระบบต้องนำเงื่อนไขข้อที่ 2) มาใช้ซึ่งจะทำให้เกิดข้อผิดในการถ่ายถอดคำบางคำ เช่น ‘ ’ (=จ-ร- ิ-ย-ธ-ร-ร-ม) ซึ่งจะถ่ายถอดผิดเป็น “จริย์ทรรม”

(2) หากระบบตรวจสอบแล้วพบว่า อักษรปกติ 2 ตัว หน้าเบรลล์เซลล์ปัจจุบัน ตามด้วยการันต์[13] แล้วอักษร 3 ตัวนี้ตรงกับอักษรชุดใดในเซ็ตไตรแกรมท้ายพยางค์ระบบจะถ่ายถอด ‘⠸⠸’ ให้เป็น “์ท” เช่น ‘⠸⠸⠸⠸⠸⠸⠸⠸⠸⠸’ (=จ- ั-น-ท-ร- ์-ท- ี- ่) ซึ่งถ่ายถอดเป็น “จันทร์ที่”

(3) ในกรณีอื่น ๆ ระบบจะถ่ายถอด ‘⠸⠸’ เป็น “ธ”

ในบางครั้งระบบพบว่า ‘⠸⠸’ ไม่ตรงกับกรณีที่กล่าวมาข้างต้น เช่น ‘⠸⠸’ ปรากฏในตำแหน่งต้นสายอักขระ ระบบจึงจำเป็นต้องนำเงื่อนไขข้อนี้มาใช้ในการถ่ายถอดอักขระเบรลล์คู่นี้ เช่น ‘⠸⠸⠸⠸⠸⠸⠸⠸⠸’ (=ธ-ง-ช-า-ต- ิ) ซึ่งถ่ายถอดเป็น “ธงชาติ”

ค. การถ่ายถอด ‘⠸⠼’ (“ฤ” หรือ“รๆ”)

‘⠸⠼’ สามารถตีความให้เป็น “ฤ” หรือ“รๆ” ได้ หากเบรลล์ 2 เซลล์ปัจจุบันเป็น ‘⠸⠼’ ระบบจะแทนที่เบรลล์คู่นี้ด้วย‘รๆ’ แล้วนำรูปที่ได้ไปตรวจสอบกับเซ็ตไตรแกรมท้ายพยางค์โดยอาศัยเงื่อนไขต่อไปนี้ หากอักษรปกติ 2 ตัวหน้าเบรลล์คู่นี้ เมื่อนำ “ร” มาต่อท้ายแล้วตรงกับอักษรในเซ็ตไตรแกรมท้ายพยางค์ ระบบจะถ่ายถอดให้เป็น “รๆ” แต่ถ้าไม่ตรงระบบจะถ่ายถอดเป็น “ฤ” ตัวอย่างเช่น ‘⠸⠼⠸⠼⠼’ (=อ-ะ-ไ-ร-ๆ) เมื่อระบบถ่ายถอดมาถึงตำแหน่งเบรลล์คู่นี้ จะมีสองตัวอักษรข้างหน้า ‘⠸⠼’ คือ “ะไ” เมื่อนำ “ร” มาต่อท้ายได้เป็น “ะไร” ซึ่งเมื่อตรวจสอบกับเซ็ตไตรแกรมท้ายพยางค์แล้ว อักษร 3 ตัวนี้ปรากฏในเซ็ต จึงต้องถ่ายถอด ‘⠸⠼’ เป็น รๆ แต่ในกรณี ‘⠸⠸⠸⠸⠸⠸⠸⠸⠸’ (=อ- ั-ง-ก-ฤ-ษ) เมื่อถ่ายถอดมาถึงเบรลล์คู่นี้ จะได้สายอักขระเป็น ‘อังก⠸⠼⠸⠸’ เมื่อรวม “งก” กับ “ร” พบว่า “งกร” ไม่อยู่ในเซ็ตไตรแกรมท้ายพยางค์ จึงต้องถ่ายถอด ‘⠸⠼’ ให้เป็น “ฤ”

ง. การถ่ายถอด ‘⠸⠸’ (“ษ” หรือ“-ส”)

ตามที่ระบุไว้ในตารางที่ 5.12 ‘⠸⠸’ มีรูปแทนในอักษรปกติ 2 รูปคือ “ษ” และ “-ส” แต่เนื่องจาก “ษ” มีโอกาสน้อยมากที่จะปรากฏในตำแหน่งต้นสายอักขระ

ในการถ่ายถอดเบรลล์คู่นี้ ระบบต้องตรวจสอบว่าอักขระเบรลล์คู่นี้ปรากฏต้นสายอักขระหรือไม่ หากระบบพบ ‘⠸⠸’ ปรากฏต้นสายอักขระ จะถ่ายถอด ‘⠸⠸’ให้เป็น “-ส” แต่หากอักขระเบรลล์คู่นี้ปรากฏในตำแหน่งอื่น ๆ ระบบจะถ่ายถอด‘⠸⠸’ให้เป็น “ษ”

---

[13] ตัว “์” เกิดจากการถ่ายถอดเบรลล์เซลล์ปัจจุบันให้เป็นตัวอักษรไทย

5.2.2.1.2 การถ่ายถอดเบรลล์เซลล์คู่ทั่วไป

ในหัวข้อที่ผ่านมา ผู้วิจัยได้นำเสนอวิธีการแก้ไขความกำกวมที่เกิดจากการถ่ายถอดเบรลล์เซลล์คู่ ซึ่งเป็นประเด็นสำคัญในการถ่ายถอดเบรลล์ไทย อย่างไรก็ตามยังมีเบรลล์เซลล์คู่อีกประเภทหนึ่ง ที่อักขระเบรลล์สองตัวใช้แทนอักขระปกติหนึ่งตัว ซึ่งไม่เป็นปัญหาต่อการตีความในอักษรปกติ

ในการถ่ายถอดเบรลล์เซลล์คู่ที่ไม่กำกวม ระบบจะนำเบรลล์สองเซลล์ปัจจุบันไปตรวจสอบกับตารางคู่เทียบเบรลล์เซลล์คู่ทั่วไป (ตามตารางที่ 5.8) หากพบว่าอักขระเบรลล์คู่ที่ทำการตรวจสอบตรงกับอักษรเบรลล์ในตารางนี้ ระบบจะนำตัวอักษรปกติมาแทนที่เบรลล์สองเซลล์ปัจจุบัน แล้วนำเบรลล์สองเซลล์ถัดไปมาดำเนินการตั้งแต่ขั้นแรกตามหัวข้อที่ 5.2.2.1.1 แต่หากไม่ตรงกับอักษรชุดใดในตารางนี้ ระบบจะเข้าสู่กระบวนการถ่ายถอดเบรลล์เซลล์เดี่ยว

5.2.2.2 การถ่ายถอดเบรลล์เซลล์เดี่ยว

ในการถ่ายถอดเบรลล์เซลล์เดี่ยว ระบบจะมีกระบวนการในการทำงานโดยประกอบด้วย การถ่ายถอดเบรลล์เซลล์เดี่ยวกำกวม และ การถ่ายถอดเบรลล์เซลล์เดี่ยวทั่วไป ซึ่งมีรายละเอียดดังนี้

5.2.2.2.1 การถ่ายถอดเบรลล์เซลล์เดี่ยวที่กำกวม

ดังที่ได้กล่าวมาแล้วในหัวข้อที่ 4.1.2 ว่า ความกำกวมที่เกิดกับเบรลล์เซลล์เดี่ยวมี 2 ลักษณะ คือ ลักษณะแรกเป็นความกำกวมที่เกิดจาก อักขระเบรลล์ 1 ตัว สามารถแทนอักขระปกติได้มากกว่า 1 กรณี และอีกลักษณะหนึ่งคือ เบรลล์เซลล์เดี่ยวบางตัว (โดยเฉพาะสระประสม) ส่งผลให้เกิดการจัดเรียงตัวอักษรได้มากกว่า 1 กรณี ในอักษรปกติ

เพื่อแก้ปัญหาความกำกวมที่เกิดขึ้นกับเบรลล์เซลล์เดี่ยวเหล่านี้ ผู้วิจัยจึงขอเสนอให้ใช้กฎในการถ่ายถอดเบรลล์เซลล์เดี่ยวที่กำกวมสำหรับอักษรเบรลล์แต่ละกลุ่ม โดยการนำเบรลล์เซลล์ปัจจุบันมาตรวจสอบกับตารางคู่เทียบเบรลล์เซลล์เดี่ยวที่กำกวมในกลุ่มต่าง ๆ ที่ปรากฏในหัวข้อนี้ หากพบว่าเบรลล์เซลล์ปัจจุบันตรงกับเบรลล์เซลล์เดี่ยวที่กำกวมกลุ่มใด ระบบจะนำเงื่อนไขสำหรับเบรลล์เซลล์เดี่ยวกลุ่มนั้นมาใช้ถ่ายถอดอักษร

อย่างไรก็ตาม สิ่งที่ควรพิจารณาคือ การถ่ายถอดอักขระเบรลล์ที่กำกวมประเภทต่าง ๆ จำเป็นต้องเรียงลำดับตามที่เสนอในงานวิจัยฉบับนี้มิฉะนั้นจะก่อให้เกิดปัญหาต่อการถ่ายถอด เนื่องจากอักขระเบรลล์บางตัวต้องใช้ในกระบวนการถ่ายถอดอักขระเบรลล์มากกว่าหนึ่งกลุ่ม เช่น ‘⠒’ (“.” หรือ

“◌้”) ต้องใช้ในการถ่ายถอดสระประสมและการถ่ายถอด‘◌ะ’ (“.” หรือ “◌้”) เป็นต้น หากไม่ถ่ายถอดตามลำดับที่เสนอนี้ จะทำให้ได้ผลลัพธ์ในอักษรปกติที่คลาดเคลื่อนจากความเป็นจริง

เมื่อระบบดำเนินการมาถึงส่วนของการถ่ายถอดเบรลล์เซลล์เดี่ยวที่กำกวม ระบบจะดำเนินการถ่ายถอดเบรลล์เซลล์ปัจจุบันโดยอาศัยบริบทซ้ายและ/หรือขวา โดยมีรายละเอียดการดำเนินการสำหรับเบรลล์แต่ละกลุ่ม ดังนี้

ก. การถ่ายถอดสระหน้าและสระประสม

ในการถ่ายถอดสระหน้าและสระประสมนั้น ระบบไม่สามารถทำได้ในขั้นตอนเดียวจึงจำเป็นต้องมีกระบวนการถ่ายถอดสำหรับสระสองกลุ่มนี้ ซึ่งใช้กระบวนการในการถ่ายถอดร่วมกัน โดยระบบจะถ่ายถอดเบรลล์เซลล์ปัจจุบันโดยอาศัยตารางคู่เทียบสระหน้า (ตารางที่ 5.6) และตารางคู่เทียบสระประสม (ตารางที่ 5.9) แล้วตรวจสอบบริบทขวาและบริบทซ้ายของสระเหล่านี้ เพื่อจัดเรียงตัวอักษรที่ปรากฏทางขวาและทางซ้ายร่วมกับสระสองกลุ่มนี้[14] ให้ถูกต้องตามหลักอักขรวิธีไทย

ในการถ่ายถอดสระสองกลุ่มนี้ มีขั้นตอนดังนี้

ก. 1 การถ่ายถอดสระประสม

ในการถ่ายถอดสระกลุ่มนี้ ระบบจะตรวจสอบว่าเบรลล์เซลล์ปัจจุบันอยู่ในสระกลุ่มนี้หรือไม่ แต่เนื่องจาก ‘⠪’ (“เ◌ิ” หรือ “เ◌อ”) เป็นสระที่มีคู่เทียบในอักษรปกติ 2 รูป ที่เป็นไปได้ ระบบจึงต้องหารูปที่เหมาะสมที่สุดสำหรับอักขระตัวนี้ก่อนเป็นอันดับแรก แต่หากระบบตรวจสอบแล้วพบว่า เบรลล์เซลล์ปัจจุบันไม่ใช่ ‘⠪’ (“เ◌ิ” หรือ “เ◌อ”) ระบบจึงจะนำเบรลล์เซลล์ปัจจุบันไปตรวจสอบกับสระตัวอื่นในกลุ่มนี้ต่อไป

1) การถ่ายถอด ‘⠪’

‘⠪’ มีรูปแทนในอักษรปกติ 2 รูป คือ “เ◌ิ” หรือ “เ◌อ” (ตามตารางที่ 5.10) การที่ ‘⠪’ จะมีรูปใดนั้นสามารถทำนายได้จากบริบทที่อักขระเบรลล์ตัวนี้ปรากฏอยู่ ดังนั้นระบบจะทำการตรวจสอบบริบททางขวาของอักขระเบรลล์นี้ 3-4 ตัว โดยอาศัยเงื่อนไขต่อไปนี้

---

[14] เพื่อความสะดวกในการอ้างอิง ผู้วิจัยจะใช้คำว่า “สระประสม” ในการอ้างถึงสระทั้งสองกลุ่มนี้

(1) { ◌ะ } ⟶ { เ◌ิ | ____ (Tone) [Final] [Front Vowel] - ◌ะ

(2) { ◌ะ } ⟶ { เ◌ิ | ____ (Tone) [Final] $\begin{Bmatrix} -Tone \\ -NormalVowel \\ -ComplexVowel \\ -SpecialVowel \end{Bmatrix}$

(3) { ◌ะ } ⟶ { เ◌อ | elsewhere

จากกฎที่ระบุไว้ข้างต้น สามารถอธิบายได้ดังนี้

(1) ‘◌ะ’ จะมีรูปเป็น เ◌ิ ก็ต่อเมื่อ

(ก) อักษรถัดจาก‘◌ะ’ เป็นตัวสะกดท้ายพยางค์[15] หรือวรรณยุกต์[16]ตามด้วยตัวสะกดท้ายพยางค์

(ข) อักษรถัดจากตัวสะกดท้ายพยางค์ เป็นสระหน้า และตัวต่อไปไม่ใช่ ‘◌ะ’ (=ะ)

(ค) อักษร 3 ตัวถัดจาก‘◌ะ’ ไม่ใช่ ‘◌ะ ◌ะ ◌ะ’ (=นโย) และไม่ใช่ ‘◌ะ ◌ะ ◌ะ’ (=มโน)

(2) ‘◌ะ’ จะมีรูปเป็น เ◌ิ ก็ต่อเมื่อ

(ก) อักษรถัดจาก ‘◌ะ’ เป็นตัวสะกดท้ายพยางค์ หรือวรรณยุกต์ตามด้วยตัวสะกดท้ายพยางค์

(ข) อักษรถัดจากตัวสะกดท้ายพยางค์ต้องอยู่ในเงื่อนไขใดเงื่อนไขหนึ่งต่อไปนี้

(ข.1) <u>ไม่ใช่</u>อักขระเบรลล์กลุ่มต่อไปนี้คือ สระทั่วไป[17] สระประสม[18] สระพิเศษ[19] และวรรณยุกต์

(ข.2) อักษรถัดจากตัวสะกดท้ายพยางค์เป็น ‘◌ะ’ (= ◌ิ์ หรือ) (วงเล็บปิด)) และอยู่ในตำแหน่งสุดท้ายของสายอักขระหรือตามด้วยเบรลล์กำกวมท้ายสายอักขระ(กรณีนี้ต้องตีความ ‘◌ะ’ ให้เป็น “)”)

(3) ในกรณีอื่น ๆ ‘◌ะ’ ต้องถูกถ่ายถอดให้เป็น “เ◌อ”

---

[15] ตัวสะกดท้ายพยางค์คือเซ็ตของอักขระเบรลล์ที่ปรากฏในตารางที่ 5.13

[16] วรรณยุกต์คือเซ็ตของอักขระเบรลล์ที่ปรากฏในตารางที่ 5.7

[17] สระทั่วไปคือเซ็ตของอักขระเบรลล์ในตารางที่ 5.4

[18] สระประสมคือเซ็ตของอักขระเบรลล์ในตารางที่ 5.9

[19] สระพิเศษคือเซ็ตของอักขระเบรลล์ในตารางที่ 5.5

‘◌ิ’ มีรูปในอักษรปกติ 2 รูป โดยที่จะมีรูปใดนั้นต้องพิจารณาบริบทที่ปรากฏร่วมกับอักขระเบรลล์ตัวนี้ กล่าวคือจะเป็น “เ◌ิ” ต่อเมื่อตามด้วยตัวสะกดท้ายพยางค์และจะมีรูปเป็น “เ◌อ” ในกรณีอื่น ดังนั้นในการถ่ายถอด‘◌ิ’ระบบจำเป็นต้องตรวจสอบว่าอักขระหลัง‘◌ิ’เป็นตัวสะกดท้ายพยางค์หรือไม่

อย่างไรก็ตาม การพิจารณาอักขระเบรลล์ที่ปรากฏหลัง ‘◌ิ’ เพียงตัวเดียว ยังไม่เพียงพอที่จะระบุได้ว่าอักขระนั้นเป็นตัวสะกดท้ายพยางค์หรือไม่ เนื่องจากอักขระเบรลล์หลัง ‘◌ิ’ อาจมีรูปเหมือนพยัญชนะตัวสะกดท้ายพยางค์ แต่อาจเป็นพยัญชนะต้นของพยางค์ถัดไปก็ได้ ดังจะเห็นได้ว่าใน ‘◌ิ◌ิ◌ิ◌ิ◌ิ’ (= ธ-เ อ-<u>ม</u>-) “เธอมี” มีพยัญชนะที่มีรูปคล้ายตัวสะกดท้ายพยางค์ปรากฏหลัง ‘◌ิ’ (“เ◌ิ” หรือ “เ◌อ”) ระบบจึงจำเป็นต้องตรวจสอบตัวอักษรอื่นเพิ่มเติม ซึ่งในกรณีนี้จะพบ‘◌ิ’ (= ี) ปรากฏหลัง ‘◌ิ’ (= ม) แต่เนื่องจาก ‘◌ิ’ เป็นสระทั่วไปที่ต้องปรากฏหลังพยัญชนะต้นพยางค์เสมอ จึงเป็นผลให้ ‘◌ิ’(=ม) เป็นพยัญชนะต้นของพยางค์ถัดจาก‘◌ิ’ (“เ◌ิ” หรือ “เ◌อ”) ระบบจึงต้องเลือกรูปของ ‘◌ิ’ เป็น “เ◌อ” ดังนั้นในการถ่ายถอด‘◌ิ◌ิ◌ิ◌ิ◌ิ’ (= ธ-เ อ-<u>ม</u>-) จึงต้องอาศัยกฎข้อที่ (1) ข้างต้น นอกจากนี้หากพิจารณากฎในข้อ (1) จะพบว่าจำเป็นต้องรวมอักษร 3 ตัว ที่ปรากฏหลัง ‘◌ิ’ ไว้ในกฎด้วย ได้แก่ ‘◌ิ◌ิ◌ิ’ (=นโย) และ ‘◌ิ◌ิ◌ิ’ (=มโน) หากไม่มีเซ็ตของไตรแกรมนี้ก็จะทำให้กฎข้อ (1) ถ่ายถอดคำว่า “เสนอนโยบาย” ผิดเป็น “เสนินโยบาย”

ในกฎข้อ (2) จำเป็นต้องมีเงื่อนไขย่อยต่าง ๆ เพื่อแก้ไขปัญหาการเลือกรูปของ ‘◌ิ’ ที่เหมาะสม ในกฎข้อนี้ระบบจะตรวจสอบบริบทขวา 2 ตัวของ ‘◌ิ’ ว่าเป็นตัวสะกดท้ายพยางค์ตามด้วยวรรณยุกต์หรือไม่ หากอักษรตัวแรกเป็นตัวสะกดท้ายพยางค์แต่ตัวหลังไม่ใช่วรรณยุกต์ ระบบจะเลือกรูปของ ‘◌ิ’ เป็น “เ◌ิ” ซึ่งก่อให้เกิดปัญหาในบางครั้ง เมื่อระบบพบ ‘◌ิ◌ิ’ ซึ่งระบบตีความเป็นวรรณยุกต์เสมอไม่ว่าจะปรากฏในตำแหน่งใดก็ตาม แต่ในความเป็นจริง‘◌ิ◌ิ’ เมื่อปรากฏในตำแหน่งท้ายสายอักขระอาจเป็น “◌ิ๊” หรือ “)” ก็ได้ แต่โดยส่วนใหญ่แล้วจะใช้แทน “)” จึงทำให้ระบบจำต้องตรวจสอบว่า ‘◌ิ◌ิ’ ปรากฏในตำแหน่งท้ายสายอักขระหรือไม่โดยอาศัยกฎข้อ (2) หาก ‘◌ิ◌ิ’ ปรากฏท้ายสายอักขระระบบจะดำเนินการตามกฎข้อ(2) (ข) นั่นคือตีความให้‘◌ิ◌ิ’ เป็นวงเล็บปิด เป็นผลให้พยัญชนะที่ปรากฏข้างหน้าเป็นตัวสะกดท้ายพยางค์ ทำให้ระบบต้องถ่ายถอด ‘◌ิ’ เป็น “เ◌ิ” เช่นสายอักขระที่ลงท้ายด้วย ‘◌ิ◌ิ◌ิ◌ิ’ (=ง-เ◌ิ-น-)) ระบบจะใช้กฎดังกล่าวทำให้ต้องเลือกรูปแทนของ ‘◌ิ’ เป็น “เ◌ิ” เป็นต้น

2) การถ่ายถอดสระประสมอื่น ๆ

หากเบรลล์เซลล์ปัจจุบันเป็นสระหน้าและสระประสมอื่น ๆ ที่ไม่ใช่ ‘⠽’ ระบบจะถ่ายถอดโดยใช้ตารางคู่เทียบสระหน้า[20]และตารางคู่เทียบสระประสม เช่น หากระบบพบ ‘⠬’ (=เีย) ระบบจะนำคู่เทียบที่ตรงกับอักขระเบรลล์ตัวนี้ คือ “เีย” ไปแทนที่เบรลล์เซลล์ปัจจุบันในสายอักขระ

ก. 2 การตรวจสอบบริบทขวาของสระเพื่อทำการถ่ายถอดและจัดเรียง

ในขั้นนี้ ระบบจะทำการตรวจสอบบริบททางขวาของสระ 1-2 ตัว แล้วทำการถ่ายถอดและจัดเรียงอักษร โดยดำเนินการตามเงื่อนไขดังนี้

(1) หากระบบพบว่าได้ถ่ายถอด ‘⠽’ เป็น “เิ” ในขั้นที่ 1 แล้ว เมื่อตรวจสอบเพิ่มเติมพบว่าอักษรทางขวาเป็นวรรณยุกต์ ระบบจะนำวรรณยุกต์ไปถ่ายถอด แล้วนำมาต่อท้ายสระที่ถ่ายถอดแล้วในขั้นที่ผ่านมา เช่น “เิ◌้” จัดเรียงได้เหมือนเดิมเป็น “เิ◌้”

(2) หากอักขระ 2 ตัวทางขวาของสระเป็น ‘⠱’ (=ะ) ตามด้วยวรรณยุกต์ โดยที่วรรณยุกต์ไม่ใช่ ‘⠿’ (◌็ หรือ)) ที่ปรากฏในตำแหน่งท้ายสายอักขระ[21] ระบบจะนำอักขระเบรลล์ 2 ตัวนี้ไปถ่ายถอดแล้วนำผลลัพธ์ที่ได้ไปดำเนินการตามเงื่อนไขคือ

(ก) หากถ่ายถอดในขั้นที่ 1 แล้วได้สระหน้า จะนำอักษรปกติที่ได้มาจัดเรียงใหม่เป็น สระหน้า + วรรณยุกต์ + “ะ” เช่น “โะ◌็” จัดเรียงเป็น “โ◌็ะ”

(ข) หากถ่ายถอดในขั้นที่ 1 ได้สระประสม ระบบจะนำวรรณยุกต์ไปแทรกในตำแหน่งรองสุดท้ายของสระประสมที่ถ่ายถอดแล้ว จากนั้นนำ“ะ”ไปต่อท้ายสระประสม เช่น “เีย◌ะ◌็” จัดเรียงเป็น “เี◌็ย◌ะ”

(3) หากอักขระทางขวาของสระประสมเป็นวรรณยุกต์ และไม่ใช่‘⠿’(“◌็” หรือ “)”) ที่ปรากฏตำแหน่งท้ายสายอักขระและไม่ใช่‘⠿⠷’ (=) +สัญลักษณ์กำกับเครื่องหมาย) ระบบจะถ่ายถอดวรรณยุกต์แล้วนำไปแทรกในตำแหน่งก่อนตัวอักษรสุดท้ายของสระประสมที่ถ่ายถอดแล้วเช่น “เีย◌้” จัดเรียงเป็น "เี◌้ย"

(4) หากอักขระทางขวาของสระประสมเป็น ‘⠱’ (=ะ) ระบบจะนำเบรลล์นี้ไปถ่ายถอดเป็น “ะ” จากนั้นนำ “ะ” ไปต่อท้ายสระประสมที่ถ่ายถอดแล้ว เช่น “เียะ” จัดเรียงเช่นเดิมเป็น “เียะ”

---

20 สระหน้าคือเซ็ตของอักษรเบรลล์ในตารางที่ 5.6

21 ‘⠿’ เมื่อปรากฏท้ายสายอักขระอาจตีความเป็น “◌็” หรือ “)” ก็ได้ จึงต้องมีเงื่อนไขส่วนนี้

(5) หากถ่ายถอดในขั้นที่ 1 ได้เป็น “เ-” “แ-” หรือ “โ-” (ตามตารางที่ 5.6) แล้วไม่พบ ‘◌ะ’ (=ะ) และ/หรือวรรณยุกต์ปรากฏหลังสระประสม ระบบจะสิ้นสุดการทำงานในส่วนการถ่ายถอดสระหน้าและสระประสม แล้วนำเบรลล์ 2 เซลล์ต่อไปมาดำเนินการตั้งแต่ขั้นแรกของการถ่ายถอดเบรลล์ไทย (ตามหัวข้อ 5.2.2.1.1)

ก. 3 การตรวจสอบบริบทซ้ายของสระประสมเพื่อจัดเรียงให้ถูกต้อง

ระบบอาจดำเนินการในขั้นนี้หรือไม่ก็ได้ ขึ้นอยู่กับว่าสระประสมที่ถ่ายถอดมาแล้วในขั้นที่ 1 ขึ้นต้นด้วย "เ-” หรือไม่ เนื่องจากในอักษรปกติสระประสมที่มี“เ-” เป็นส่วนประกอบจะต้องมีพยัญชนะต้น 1-2 ตัวแทรกระหว่างสระประสม ระบบจึงต้องทำการสลับที่ระหว่างพยัญชนะ 1-2 ตัวกับ “เ-”

การดำเนินการในขั้นนี้จะเกิดขึ้นหากระบบพบว่า สระประสมที่ถ่ายถอดมาแล้วขึ้นต้นด้วย “เ-” โดยระบบจะตรวจสอบตัวอักษรที่ปรากฏหน้าสระ“เ-” 2-5 ตัว แล้วดำเนินการในข้อใดข้อหนึ่งต่อไปนี้

(1) สลับที่พยัญชนะ 1 ตัว กับ “เ-” เมื่อพบว่าอักษร 2 ตัว ข้างหน้าสระเป็นพยัญชนะควบที่กำกวมตามตารางที่ 5.14 และตรงกับเงื่อนไขต่อไปนี้

(ก) ถ้าอักษรที่ปรากฏหน้าพยัญชนะควบที่กำกวม เป็น เ- แ- โ- ไ- ใ- เช่น “ไพรเาะ” จัดเรียงเป็น “ไพเราะ”

(ข) ถ้าอักษร 3 ตัว เริ่มจากตัวที่ 4 หน้าเบรลล์เซลล์ปัจจุบันอยู่ในเซ็ตไตรแกรมท้ายพยางค์ เช่น “ทัพรเือ” จัดเรียงเป็น “ทัพเรือ”

(2) สลับที่พยัญชนะ 2 ตัวกับ “เ-” หากพยัญชนะที่ปรากฏหน้าสระประสม 2 ตัวเป็นพยัญชนะควบที่กำกวมหรือเป็นพยัญชนะควบไม่กำกวม เช่น “หลเือ” จัดเรียงเป็น “เหลือ” เนื่องจาก หล” เป็นพยัญชนะควบไม่กำกวม

(3) หากเป็นกรณีอื่น ๆ ระบบจะนำพยัญชนะหนึ่งตัวหน้าสระประสมมาแทรกหลังสระ เ- เช่น “มเื่อ” จัดเรียงเป็น “เมื่อ”

หลังจากที่ระบบถ่ายถอดสระประสมและบริบททางซ้ายที่ปรากฏร่วมกับสระประสมแล้วระบบจะนำเบรลล์สองเซลล์ถัดไปมาดำเนินการตั้งแต่ขั้นแรกของการถ่ายถอดเบรลล์ไทยต่อไป อย่างไรก็ตาม หากระบบพบว่าเบรลล์เซลล์ปัจจุบันไม่ใช่สระหน้าหรือสระประสม ระบบจะดำเนินการในขั้นต่อไป

ข. การถ่ายถอดเบรลล์ที่กำกวมท้ายสายอักขระ

จากตารางที่ 5.11 เบรลล์ที่กำกวมท้ายสายอักขระได้แก่ ‘⠶’ (=) หรือ ◌ั) และ‘⠒’ (=” หรือ ◌์) เบรลล์ที่กำกวมเหล่านี้สามารถตีความได้ 2 กรณี ในการถ่ายถอดอักษรทั้ง 2 ตัวนี้ ระบบจะต้องอาศัยเซ็ตไตรแกรมของตัวอักษรท้ายพยางค์ (ตามภาคผนวก ก) ที่ผู้วิจัยสร้างขึ้น

ฉะนั้น เมื่อระบบพบอักษรเบรลล์ในกลุ่มนี้ ระบบจะตรวจสอบว่าเป็นไปตามเงื่อนไขหนึ่งต่อไปนี้หรือไม่ คือ เบรลล์เซลล์ปัจจุบันเป็นตัวสุดท้ายในสายอักขระหรือไม่ หรืออักขระที่ตามหลังเบรลล์เซลล์ปัจจุบันเป็น‘⠶’หรือเบรลล์กำกวมท้ายสายอักขระหรือไม่ (เนื่องจากเบรลล์กำกวมนี้อาจตีความเป็นเครื่องหมายได้และเครื่องหมายเหล่านี้ก็สามารถปรากฏเรียงต่อกันได้) หากตรงกับเงื่อนไขใดข้างต้น ระบบจะเลือกถ่ายถอดโดยอาศัยเงื่อนไขเพิ่มเติมต่อไปนี้

(1) การถ่ายถอดเบรลล์กำกวมท้ายสายอักขระเป็นตัวอักษรไทย

หากนำไทยปกติที่ตรงกับเบรลล์เซลล์ปัจจุบัน (คือ ◌์ หรือ ◌ั) มาแทนที่เบรลล์เซลล์ปัจจุบัน แล้วนำบริบททางซ้ายอีก 2 ตัวซึ่งถ่ายถอดเป็นอักษรปกติแล้ว (รวมเป็น 3 ตัว) ไปตรวจสอบกับเซ็ตไตรแกรมท้ายพยางค์แล้วพบว่าอยู่ในเซ็ตนี้ ระบบจะถ่ายถอดเบรลล์ที่กำกวมเป็นตัวอักษรไทย ตัวอย่างเช่น ‘⠎⠜⠶’ (=ศ-ก-ร์) เมื่อระบบถ่ายถอดมาถึง ‘⠶’ (◌์ หรือ ◌ั) จะได้สายอักขระเป็น “ศุกร◌ั” เมื่อนำ “◌์” ไปแทนที่ ‘⠶’ แล้วนำ “กร์”ไปตรวจสอบกับเซ็ตไตรแกรมท้ายพยางค์พบว่า ตรงกับอักษรในเซ็ตนี้ จึงต้องถ่ายถอด ‘⠶’ ให้เป็น ◌์

(2) การถ่ายถอดเบรลล์กำกวมท้ายสายอักขระเป็นเครื่องหมาย

หากระบบนำ 3 ตัวอักษรที่ทดลองถ่ายถอดในข้อ (1) ไปตรวจสอบกับเซ็ตไตรแกรมท้ายพยางค์แล้วพบว่าไม่ตรงกับอักษรชุดใดในเซ็ตดังกล่าว ระบบจะถ่ายถอดเบรลล์เซลล์ปัจจุบันเป็นเครื่องหมาย เช่นคำว่า ‘⠝⠐⠡⠐⠒’ (= น-ะ-ค-ะ-”) เมื่อระบบถ่ายถอดมาถึง ‘⠒’(◌์ หรือ ◌ั) จะได้สายอักขระเป็น “นะคะ◌ั” จากนั้นเมื่อนำ “◌์” ไปแทนที่ ‘⠒’ แล้วนำ “คะ◌์” ไปตรวจสอบกับเซ็ตไตรแกรมท้ายพยางค์ พบว่า ไม่อยู่ในเซ็ตไตรแกรมดังกล่าว ระบบจะถ่ายถอด ‘⠒’ ให้เป็น “ ” ”

ค. การถ่ายถอด ‘⠲’

‘⠲’ สามารถตีความให้เป็น “◌ั้” หรือ “.” ได้ โดยจะแปรไปตามบริบทที่เบรลล์ตัวนี้ปรากฏอยู่ คือหากเป็นส่วนหนึ่งของบริบทพยางค์ในภาษาไทยจะถ่ายถอดเป็น"◌ั้" แต่หากไม่ใช่จะถ่ายถอดเป็น "."

ในการถ่ายถอด ‘⠲’ ระบบจะถ่ายถอดให้เป็น “◌ั้” ทุกครั้งที่พบเบรลล์ตัวนี้ โดยในครั้งแรกระบบจะนำ 3 ตัวอักษรนับจากเบรลล์เซลล์ปัจจุบันไปทางซ้าย ไปตรวจสอบกับไตรแกรมท้ายพยางค์

แล้วนับการตรวจสอบเป็นครั้งที่ 1 หากพบว่ายังไม่ตรงกับอักษรชุดใดในเซ็ต ระบบจะถ่ายถอดเบรลล์เซลล์ต่อไป แล้วนำ 3 ตัวอักษรนับจากเซลล์ปัจจุบันไปทางซ้ายไปตรวจสอบกับเซ็ตอีก นับการตรวจสอบเป็นครั้งที่ 2 เมื่อดำเนินการจนครบ 3 ครั้ง หรือจนสิ้นสุดสายอักขระ แล้วยังไม่ตรงกับอักษรชุดใดในเซ็ตไตรแกรมท้ายพยางค์ ระบบจะแก้ไข ‘⠒’ ให้เป็น "." ในทุกตำแหน่งของอักษรครั้งสุดท้ายที่นำไปตรวจสอบ

ตัวอย่าง ในการถ่ายถอด ‘⠃⠒⠁⠒⠒⠁’ (=บ-◌้-า-น-น-า) ระบบต้องนำสายอักขระเหล่านี้ไปตรวจสอบกับเซ็ตไตรแกรมท้ายพยางค์ เมื่อพบว่า “บ้า” เป็นไตรแกรมที่พบในข้อมูลจริงได้ ก็จะถ่ายถอดเป็นให้ “◌้” แต่ในทางตรงกันข้าม กรณี ‘⠏⠒⠁⠒⠒⠁’(=พ-.-ศ-.) เมื่อนำ 3 ตัวอักษร “◌้ ศ ◌้” ไปตรวจสอบแล้วพบว่าไม่อยู่ในเซ็ตไตรแกรมท้ายพยางค์ ก็จะแก้ไข “◌้” ให้เป็น “.” ในทุกตำแหน่ง

5.2.2.2.2 การถ่ายถอดเบรลล์เซลล์เดี่ยวทั่วไป

ในการถ่ายถอดเบรลล์เซลล์เดี่ยวทั่วไป ระบบจะตรวจสอบว่าเบรลล์เซลล์ปัจจุบันนั้นอยู่ในกลุ่มอักษร 5 กลุ่มต่อไปนี้หรือไม่ ได้แก่ (1) เครื่องหมายเบรลล์ต้นสายอักขระ (2) วรรณยุกต์ (3) สระพิเศษ (4) สระทั่วไป และ (5) เบรลล์เซลล์เดี่ยวปกติ[22]

ในการถ่ายถอดเบรลล์เซลล์เดี่ยวทั่วไป ต้องดำเนินการดังนี้

ก. เครื่องหมายเบรลล์ต้นสายอักขระ

หากระบบพบว่าเบรลล์เซลล์ปัจจุบันเป็นอักษรตัวแรกของสายอักขระ และเมื่อนำไปตรวจสอบกับตารางคู่เทียบเบรลล์ต้นสายอักขระ (ตามตารางที่ 5.15) แล้วพบว่าตรงกับอักษรเบรลล์ในตาราง ระบบจะถ่ายถอดเบรลล์เซลล์ปัจจุบันโดยอ้างอิงข้อมูลจากตารางนี้

ข. การถ่ายถอดวรรณยุกต์

ในระบบการถ่ายถอดเบรลล์เซลล์เดี่ยวที่กำกวม ระบบอาจถ่ายถอด ‘⠒’ ให้เป็น "◌้" ได้ สำหรับการถ่ายถอดวรรณยุกต์อีก 3 รูปที่เหลืออยู่นั้น หากระบบพบว่า เบรลล์เซลล์ปัจจุบันตรงกับวรรณยุกต์อื่น ๆ ระบบจะถ่ายถอดเบรลล์เซลล์ปัจจุบันโดยอาศัยคู่เทียบวรรณยุกต์ (ตามตารางที่ 5.7)

---

[22] ในการถ่ายถอดอักขระเบรลล์เหล่านี้ ต้องทำการถ่ายถอดตามลำดับที่เสนอไว้ในงานวิจัยฉบับนี้ เนื่องจากมีอักขระเบรลล์บางตัวต้องตีความตามตำแหน่งที่ปรากฏ หากจัดเรียงไม่ตรงตามลำดับที่นำเสนอมาข้างต้นจะทำให้การถ่ายถอดอักษรไม่ถูกต้อง เช่นหากสลับที่ระหว่างการถ่ายถอดวรรณยุกต์กับเครื่องหมายหน้าข้อความจะทำให้ถ่ายถอด “)” เป็น “◌๊” เป็นต้น

ค. การถ่ายถอดสระพิเศษ

ในการถ่ายถอดสระพิเศษนั้น ระบบจะต้องตรวจสอบว่ามีวรรณยุกต์ปรากฏร่วมด้วยหรือไม่ และวรรณยุกต์นั้นคือ ‘ ◌ ’ (ซึ่งอาจตีความเป็น "◌้" หรือ ")" ได้ในตำแหน่งท้ายสายอักขระ) ดังนั้นในการถ่ายถอดสระกลุ่มนี้ ระบบจะดำเนินการในกรณีใดกรณีหนึ่ง ต่อไปนี้

(1) การถ่ายถอดสระพิเศษที่ปรากฏกับ ‘ ◌ ’ ในตำแหน่งท้ายสายอักขระ

หากระบบพบว่าเบรลล์เซลล์ปัจจุบันเป็นสระพิเศษ (ตามตารางที่ 5.5) ระบบจะถ่ายถอดเบรลล์เซลล์ปัจจุบัน หลังจากนั้นตรวจสอบเบรลล์เซลล์ขวาว่าเป็น ‘ ◌ ’ ที่ปรากฏในตำแหน่งท้ายสายอักขระหรือไม่ หากใช่ ระบบจะทดลองถ่ายถอดอักขระตัวนี้ให้เป็น "◌้" จากนั้นทดลองจัดเรียงตัวอักษรตามอักขรวิธีไทย ทำให้ได้การจัดเรียงเป็น "◌้ะ" หรือ "◌้ำ" จากนั้นนำ 3 ตัวอักษรนับจากตัวก่อนสระพิเศษ (เช่น "จ ◌้ะ") ไปตรวจสอบกับเซ็ตไตรแกรมท้ายพยางค์ ถ้า 3 ตัวอักษรนี้อยู่ในเซ็ตดังกล่าว ระบบจะใช้ผลลัพธ์ที่ทดลองนี้เป็นผลลัพธ์จริง แต่ถ้าไม่ตรงกับเซ็ตนี้ (เช่นนำ "ค◌้ำ" ไปตรวจสอบแล้วพบว่าไม่อยู่ในเซ็ตนี้) ระบบจะถ่ายถอด‘ ◌ ’ ให้เป็น “)” (เช่นได้ผลลัพธ์เป็น “คำ)”)

(2) การถ่ายถอดสระพิเศษที่ปรากฏกับวรรณยุกต์ หากอักขระถัดจากเบรลล์เซลล์ปัจจุบันเป็นวรรณยุกต์ (ตามตารางที่ 5.7) ระบบจะถ่ายถอดเบรลล์เซลล์ปัจจุบันและเซลล์ถัดไป แล้วทำการสลับที่ระหว่างอักขระทั้ง 2 ตัวดังกล่าว เช่น ‘ ◌ ◌ ◌ ’ (= น- ◌ำ- ◌้) ถ่ายถอดเป็น “น้ำ”

(3) การถ่ายถอดสระพิเศษที่ไม่มีวรรณยุกต์ปรากฏร่วมด้วย หากระบบไม่พบวรรณยุกต์ปรากฏหลังสระพิเศษ ระบบจะถ่ายถอดโดยนำอักษรปกติที่ตรงกับคู่เทียบสระพิเศษ (ตารางที่ 5.5) มาแทนที่เบรลล์เซลล์ปัจจุบันโดยตรง

ง. การถ่ายถอดสระทั่วไป

หากเบรลล์เซลล์ปัจจุบันเป็นสระทั่วไป ระบบจะนำอักษรปกติที่ตรงกันในตาราง คู่เทียบสระทั่วไป (ตารางที่ 5.4) ไปแทนที่เบรลล์เซลล์ปัจจุบัน

จ. การถ่ายถอดเบรลล์เซลล์เดี่ยวทั่วไป

เมื่อระบบนำเบรลล์เซลล์ปัจจุบันไปตรวจสอบกับตารางคู่เทียบเบรลล์เซลล์เดี่ยวทั่วไป (ตารางที่ 5.3) แล้วพบว่าตรงกับอักขระเบรลล์ในตาราง ระบบจะนำคู่เทียบไปแทนที่เบรลล์เซลล์ปัจจุบัน

5.2.2.3 การถ่ายถอดตัวเลข

ระบบจะเริ่มถ่ายถอดตัวเลข เมื่อพบว่าเบรลล์เซลล์ปัจจุบันเป็น ‘ ⠼ ’ (number sign) โดยที่ในขั้นแรก จะลบ ‘ ⠼ ’ ออกจากสายอักขระ หลังจากนั้นนำเบรลล์เซลล์ปัจจุบันไปถ่ายถอดโดยอาศัยตารางคู่

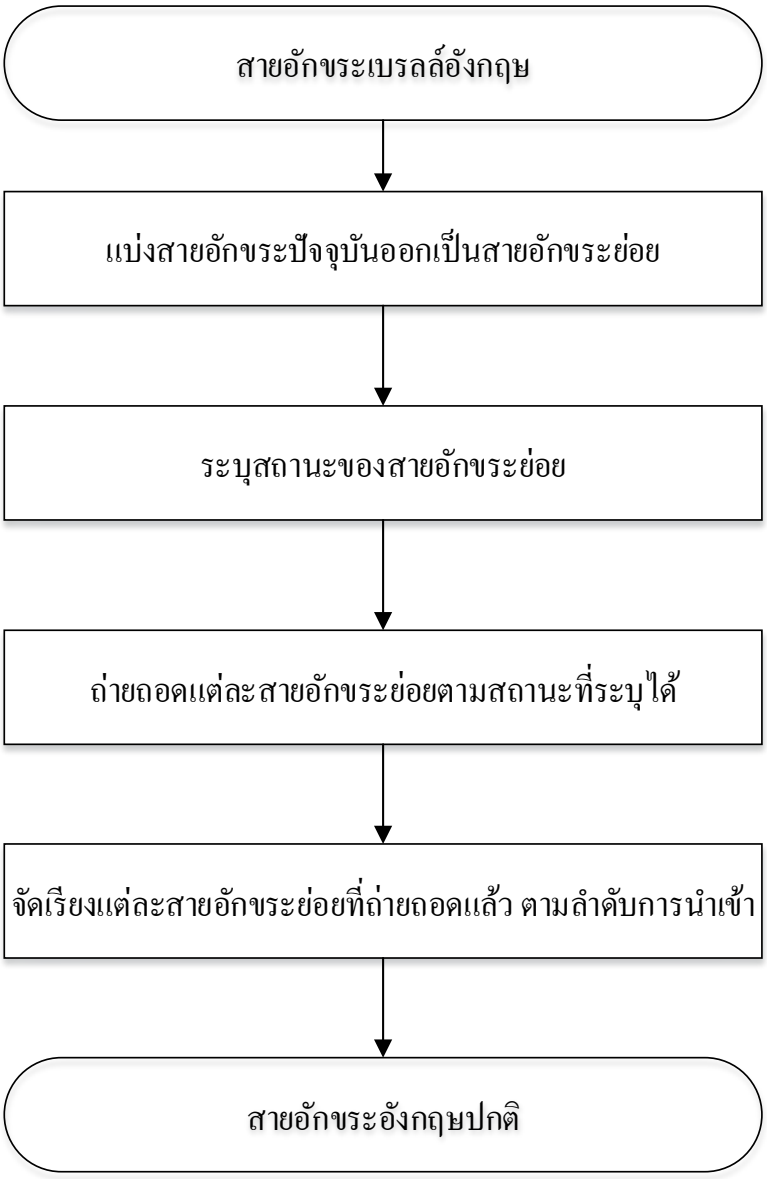
เทียบตัวเลข (ตารางที่ 5.23) จนกระทั่งสิ้นสุดสายอักขระปัจจุบัน หรือจนกระทั่งพบเครื่องหมายประกอบคำที่ปรากฏในหัวข้อที่ 4.2.2.4.1

5.2.3 การถ่ายถอดเบรลล์อังกฤษ

เมื่อกระบวนการระบุภาษาได้ระบุให้สายอักขระปัจจุบันเป็นเบรลล์อังกฤษ ระบบจะส่งสายอักขระนี้มายังกระบวนการถ่ายถอดเบรลล์อังกฤษ ซึ่งในการถ่ายถอดเบรลล์อังกฤษนั้นระบบจะได้รับข้อมูลที่มีการระบุจากผู้ใช้แล้วว่าสายอักขระที่กำลังถ่ายถอดอยู่นั้นเป็นเบรลล์อังกฤษรูปเต็ม หรือเบรลล์อังกฤษรูปย่อ

ขั้นตอนที่จะได้นำเสนอในส่วนนี้ เป็นขั้นตอนสำหรับการถ่ายถอดเบรลล์อังกฤษรูปย่อโดยเฉพาะ สำหรับการถ่ายถอดเบรลล์อังกฤษรูปเต็มนั้น ผู้ใช้ระบบจะต้องแก้ไขสถานะเบรลล์รูปย่อ ในหัวข้อที่ 5.2.3.2 ง. ให้เป็นสถานะเบรลล์รูปเต็ม จึงจะถ่ายถอดเบรลล์อังกฤษรูปเต็มได้ถูกต้อง

ในการถ่ายถอดเบรลล์อังกฤษรูปเต็มและเบรลล์อังกฤษรูปย่อมีกระบวนการโดยภาพรวมคล้ายคลึงกัน ซึ่งสามารถแสดงได้ดังนี้



ภาพที่ 5.3 กระบวนการถ่ายถอดเบรลล์อังกฤษโดยใช้กฎ

จากภาพที่แสดงไว้ข้างต้น ระบบจะดำเนินการตามลำดับคือ ขั้นแรก แบ่งสายอักขระออกเป็นสายอักขระย่อย ต่อมานำแต่ละสายอักขระย่อยไปผ่านการระบุสถานะ จากนั้นถ่ายถอดให้เป็นอักษรปกติตามสถานะที่ระบุได้ และในขั้นสุดท้าย นำสายอักขระอักษรปกติที่ได้มาเรียงต่อกันใหม่ตามลำดับนำเข้า เพื่อใช้เป็นผลลัพธ์

ในการถ่ายถอดเบรลล์อังกฤษแต่ละขั้นตอน มีรายละเอียดในการดำเนินการดังนี้

5.2.3.1 การแบ่งสายอักขระปัจจุบันออกเป็นสายอักขระย่อย

ในการถ่ายถอดเบรลล์อังกฤษ จำเป็นต้องมีการตัดแบ่งสายอักขระปัจจุบันออกเป็นสายอักขระย่อย เพื่อแก้ความกำกวมของอักษรเบรลล์ที่ทำการถ่ายถอด โดยในการตัดแบ่งสายอักขระนั้น ระบบจะเลื่อนตำแหน่งการตรวจสอบไปครั้งละ 1-2 อักขระ เมื่อพบว่าอักขระนั้นตรงกับเงื่อนไขที่กำหนดไว้เพื่อการตัดแบ่งจึงจะทำการตัดแบ่งสายอักขระ โดยดำเนินการเช่นนี้ไปเรื่อย ๆ จนสิ้นสุดสายอักขระปัจจุบัน

การตัดแบ่งสายอักขระจะดำเนินการกับกลุ่มอักษรเบรลล์ต่าง ๆ ดังนี้

(1) <u>การตัดแบ่งสายอักขระในตำแหน่งหลังเครื่องหมายคั่นระหว่างคำ</u>

ในสายอักขระหนึ่งๆอาจมีตัวย่อระดับคำปรากฏมากกว่าหนึ่งแห่งได้ ตัวอย่างเช่น

‘⠆⠿⠤⠿⠆⠿’(=**bef-and-af**) ⟹ “before-and-after”

‘⠆⠿’ (= **bef**ore) และ ‘⠁⠿’ (= **af**ter) เป็นตัวย่อระดับคำ หากนำทั้งสายอักขระไปถ่ายถอด จะเกิดข้อผิดเป็น “bef-and-af” เนื่องจากระบบได้ถ่ายถอดตัวย่อทีละตัวเพราะไม่มีจุดกำกับว่าส่วนใดเป็นคำย่อระดับคำ

อย่างไรก็ตามจะพบว่าในสายอักขระนี้มี ‘⠤’ (= -) ซึ่งเครื่องหมายนี้สามารถใช้เป็นตัวแบ่ง (delimiter) เพื่อแยกแต่ละคำออกจากกันได้ ในการถ่ายถอดจึงจำเป็นต้องแยกสายอักขระนี้ออกเป็น ‘⠆⠿⠤’ (= before-) ‘⠿⠤’ (= and-) และ ‘⠁⠿’ (=after) แล้วนำสายอักขระย่อยเหล่านี้ไปถ่ายถอด จากนั้นนำกลับมาเรียงต่อกันตามลำดับการนำเข้าจึงจะได้ผลลัพธ์ที่ถูกต้อง

ดังนั้นในขั้นนี้ ระบบจะทำการตัดแบ่งสายอักขระปัจจุบันออกเป็นสายอักขระย่อยต่าง ๆ (เช่นเดียวกับตัวอย่างที่แสดงข้างต้น) โดยตรวจสอบเบรลล์เซลล์ปัจจุบันว่าตรงกับเบรลล์ในตารางคู่เทียบเครื่องหมายคั่นระหว่างคำ (ตามตารางที่ 5.20) หรือไม่ ถ้าเบรลล์เซลล์ปัจจุบันตรงกับตารางนี้ และอักขระก่อนเบรลล์เซลล์ปัจจุบัน<u>ไม่ใช่</u>เบรลล์หน้าสายอักขระตามตารางที่ 5.15 (เช่น<u>ไม่ใช่</u>อักขระเบรลล์ที่ใช้แทน “to” “into” “by”) ระบบจะแบ่งสายอักขระเป็น 2 สายอักขระย่อย โดยตัดแบ่งที่ตำแหน่งหลังเครื่องหมายคั่นระหว่างคำ จากนั้นนำสายอักขระย่อยหลังมาดำเนินการขั้นนี้ซ้ำ

(2) <u>การแบ่งสายอักขระสำหรับถ่ายถอดที่อยู่ของเว็บไซต์</u>

ในการเขียนที่อยู่ของเว็บไซต์นั้น จะมีเครื่องหมาย ‘⠒’(= :) ‘⠲’ (= .) และ ‘⠸’ (= /) ปรากฏอยู่ หากนำเครื่องหมายเหล่านี้ไปถ่ายถอดด้วยกระบวนการถ่ายถอดเบรลล์อังกฤษรูปย่อแล้ว จะทำให้เครื่องหมายเหล่านี้ถูกถ่ายถอดเป็นตัวย่อแทนที่จะเป็นเครื่องหมาย เช่น ‘⠒’ (= :) ถูกถ่ายถอดเป็น “cc” หรือ ‘⠲’ (= .) ถูกถ่ายถอดเป็น “dd” เป็นต้น ซึ่งระบบสามารถแก้ไขความกำกวมของเครื่องหมายเหล่านี้ได้โดยการเติม ‘⠐’ (สัญลักษณ์กำกับเครื่องหมาย) หน้าอักขระเบรลล์เหล่านี้

เมื่อระบบพบว่าเบรลล์เซลล์ปัจจุบัน เป็น ‘⠺⠺⠺’ (=www) หรือ ‘⠓⠞⠞⠏’ (=http) หรือ ถ้าสายอักขระปัจจุบันมี ‘⠈’ (=@) ตามด้วยอักษรอังกฤษแล้วตามด้วยอักขระกี่ตัวก็ได้จากนั้นตามด้วย ‘⠲’(จุด) ระบบจะทำการค้นหาเครื่องหมาย ‘⠒’(= :) ‘⠲’ (= .) และ ‘⠸’ (= /) แล้วทำการเติม ‘⠐’ (สัญลักษณ์กำกับเครื่องหมาย) ข้างหน้าเครื่องหมายเหล่านี้ เพื่อขจัดความกำกวมที่เกิดขึ้น

(3) <u>การตัดแบ่งสายอักขระในตำแหน่งหน้าเครื่องหมาย ‘⠶’</u>

‘⠶’ สามารถใช้แทนได้ 3 กรณีคือ หากอยู่ต้นสายอักขระจะใช้แทน “(” หากอยู่กลางสายอักขระจะใช้แทน “gg” และหากอยู่ท้ายสายอักขระจะใช้แทน “)” แต่ในบางครั้งอาจมี ‘⠶’ ปรากฏอยู่ 2 ตำแหน่งในสายอักขระเดียวกัน จึงจำเป็นต้องตีความให้ ‘⠶’ ตัวหน้าเป็น “(” และ ‘⠶’ ตัวหลังเป็น “)” เช่น ‘⠋⠊⠇⠑⠦⠎⠴’ (= f-i-l-e-(-s-)) ซึ่งใช้เขียน “file(s)” ในกรณีนี้ไม่สามารถตีความ‘⠶’ ตัวหน้าให้เป็น “gg” ได้ แม้จะปรากฏในตำแหน่งกลางสายอักขระก็ตาม จึงทำให้ ‘⠶’ ที่ปรากฏกลางสายอักขระเกิดความกำกวมในบางครั้งเพราะอาจตีความให้เป็น “gg” หรือ “(” ได้

หากระบบพบว่ามี ‘⠶’ ปรากฏ 2 ตำแหน่งที่ห่างกันในสายอักขระ ระบบจะทำการตรวจสอบบริบทซ้ายและบริบทขวา‘⠶’ตัวแรก หากระบบพบว่าสามารถตีความ‘⠶’ตัวแรกเป็น“(” ได้ ระบบจะทำการตัดแบ่งสายอักขระปัจจุบันที่ตำแหน่งหน้า‘⠶’(=( )เช่น ‘⠋⠊⠇⠑⠦⠎⠴’ (= f-i-l-e-(-s-)) เมื่อทำการตัดแบ่งตามกระบวนการนี้จะได้ 2 สายอักขระย่อยคือ ‘⠋⠊⠇⠑’ (f-i-l-e) และ‘⠶⠎⠶’ ((-s-))

(4) <u>การตัดแบ่งสายอักขระในตำแหน่งหลังเครื่องหมาย ‘⠸’</u>

‘⠸’ ในเบรลล์รูปย่อสามารถใช้แทนได้ 3 กรณีคือ “/” “st” หรือ“still” (เมื่อปรากฏเป็นคำโดด) จากการวิเคราะห์การปรากฏของ ‘⠸’ ผู้วิจัยพบว่า ‘⠸’ ที่ปรากฏในตำแหน่งต้นคำหรือท้ายคำไม่เป็นปัญหาต่อการถ่ายถอดเพราะต้องตีความให้เป็น “st” เสมอ (ดูหัวข้อ 4.2.2.1.3) แต่เมื่อ ‘⠸’ ปรากฏกลางสายอักขระอาจตีความเป็น “st” หรือ “/” ได้ จึงเกิดความกำกวมขึ้น (เช่น ‘⠔⠎⠞⠁⠇⠇’ (in-st-a-l-l) ซึ่ง

ใช้เขียน“install” และใน ‘⠓⠊⠍⠤⠸⠓⠤⠻’ (him-/-h-er) ซึ่งใช้แทน“him/her”) ด้วยเหตุนี้จึงไม่สามารถถ่ายถอด‘⠸’ให้เป็นกรณีใดกรณีหนึ่งได้เมื่อ‘⠸’ปรากฏกลางสายอักขระ

ในการขจัดความกำกวมที่เกิดขึ้นกับ ‘⠸’ ในตำแหน่งกลางสายอักขระ ระบบจะดำเนินการโดยตรวจสอบว่า‘⠸’สามารถปรากฏเป็นส่วนหนึ่งของพยางค์ได้หรือไม่โดยอาศัยบริบทซ้ายและบริบทขวาที่แวดล้อม ‘⠸’ หาก ‘⠸’ ปรากฏเป็นส่วนหนึ่งของพยางค์ได้ระบบก็จะไม่ตัดแบ่งสายอักขระนั้น แต่หาก‘⠸’ไม่ใช่ส่วนหนึ่งของพยางค์ระบบจึงจะตัดแบ่งสายอักขระออกเป็นสายอักขระย่อย โดยการเติม‘⠐’ (สัญลักษณ์กำกับเครื่องหมาย) ข้างหน้า‘⠸’ เพื่อขจัดความกำกวมของ‘⠸’ก่อนที่จะทำการตัดแบ่งสายอักขระในตำแหน่งหลัง ‘⠸’

ตัวอย่างเช่นเมื่อสายอักขระปัจจุบันเป็น ‘⠔⠌⠞⠁⠇⠇’ (in-st-a-l-l) ซึ่ง “st” เป็นส่วนหนึ่งของพยางค์ “stall”ได้ ระบบจะไม่ตัดแบ่งสายอักขระในตำแหน่งเบรลล์เซลล์ปัจจุบัน แต่หากสายอักขระเป็น ‘⠓⠊⠍⠤⠸⠓⠤⠻’ “him/her” ซึ่งเมื่อถ่ายถอด‘⠸’ เป็น “st” แล้ว “st” ไม่สามารถก่อพยางค์ใน“himst” หรือ “sther” ได้ ในกรณีหลังนี้ระบบจำเป็นต้องตัดแบ่งสายอักขระโดยการเติม‘⠐’ ข้างหน้า ‘⠸’ ก่อนที่จะตัดแบ่งสายอักขระในตำแหน่งหลังเครื่องหมาย ‘⠸’ ทำให้ได้ 2 สายอักขระย่อยเป็น “⠓⠊⠍⠐⠄” (=him/) และ ‘⠓⠻’(=her)

5.2.3.2 การระบุสถานะของสายอักขระย่อย

ดังที่ได้กล่าวมาแล้วว่า ข้อความเบรลล์ไทยปนอังกฤษสามารถเป็นเบรลล์ไทยปนอังกฤษรูปเต็มหรือเบรลล์ไทยปนอังกฤษรูปย่อได้เท่านั้น จึงเป็นที่สงสัยว่าเหตุใดจึงต้องระบุสถานะของสายอักขระย่อยให้เป็นเบรลล์รูปเต็มในการถ่ายถอดข้อความเบรลล์รูปย่อแม้ว่าระบบจะทราบตั้งแต่ขั้นนำเข้าแล้วว่าข้อความเบรลล์นำเข้าเป็นเบรลล์รูปย่อ

เหตุที่เป็นเช่นนี้เนื่องจากในการถ่ายถอดเบรลล์รูปย่อนั้นจำเป็นต้องตีความให้สายอักขระย่อยเป็นเบรลล์รูปเต็มในบางกรณี เนื่องจากหากนำสายอักขระย่อยนั้นไปผ่านกระบวนการถ่ายถอดเบรลล์รูปย่อโดยตรงแล้ว จะทำให้ได้ผลลัพธ์ไม่ตรงกับความเป็นจริง เช่นคำว่า ‘⠏⠲⠍⠲’ (p.m.) หากนำอักขระเหล่านี้ไปทำการถ่ายถอดด้วยกระบวนการถ่ายถอดเบรลล์รูปย่อแล้ว ‘⠲’ (=.) ที่ปรากฏกลางสายอักขระจะถูกถ่ายถอดให้เป็นตัวย่อกลางคำคือ “dd”

เพื่อแก้ปัญหาที่เกิดขึ้นนี้ ผู้วิจัยจึงได้รวมกระบวนการถ่ายถอดเบรลล์รูปเต็มไว้ในระบบ เพื่อให้ระบบสามารถเรียกใช้ได้ตามต้องการและเพื่อให้ระบบสามารถตีความสายอักขระที่ต้องตีความเหมือนเบรลล์รูปเต็มแต่ปรากฏในข้อความเบรลล์รูปย่อได้ถูกต้องดังที่กล่าวมานี้ ระบบจำเป็นต้องระบุสถานะ

ของแต่ละสายอักขระเท่าที่จะทำได้ ฉะนั้น เมื่อระบบตัดแบ่งสายอักขระออกเป็นสายอักขระย่อยต่าง ๆ แล้ว ระบบจะนำแต่ละสายอักขระย่อยไปตรวจสอบสถานะว่าเป็นเบรลล์อังกฤษรูปเต็ม คอมพิวเตอร์เบรลล์ ตัวเลขเบรลล์ หรือเบรลล์อังกฤษรูปย่อ โดยตรวจสอบตามเงื่อนไขสำหรับเบรลล์แต่ละกลุ่ม ดังนี้

ก. เบรลล์อังกฤษรูปเต็ม

สถานะเบรลล์อังกฤษรูปเต็ม เป็นสถานะของสายอักขระย่อยที่กำหนดขึ้นเพื่อระบุว่าระบบจะต้องนำสายอักขระย่อยไปถ่ายถอดเป็นเบรลล์อังกฤษรูปเต็มหรือถ่ายถอดเป็นตัวเลข ซึ่งเบรลล์ทั้ง 2 ประเภทนี้จะใช้กระบวนการร่วมกัน คือ เป็นการจับคู่ระหว่างอักขระเบรลล์กับอักขระปกติแบบหนึ่งต่อหนึ่ง แต่ต่างกันที่เบรลล์ 2 ประเภทนี้จะใช้ตารางคู่เทียบที่ต่างกัน

จากการศึกษาเบรลล์รูปย่อ ผู้วิจัยพบว่าในบางครั้งต้องตีความบางสายอักขระที่ปรากฏในข้อความให้เป็นเบรลล์รูปเต็มจึงจะสื่อความได้ โดยระบบจะต้องตีความสายอักขระย่อยให้เป็นเบรลล์รูปเต็ม เมื่อระบบพบ ‘⠼’ ที่ตีความให้เป็น “เครื่องหมายนำเลข” หรือพบ ‘⠨’ ที่อยู่กลางสายอักขระย่อยแล้วต้องตีความเป็น “.”หรือพบตัวอักษรโดดที่ปรากฏร่วมกับ “(” และ “)” ในสายอักขระย่อยเดียวกัน

(1) ‘⠼’ ที่ปรากฏกลางสายอักขระย่อย

ในเบรลล์อังกฤษรูปย่อ ‘⠼’ สามารถตีความให้เป็น “ble” หรือเครื่องหมายนำเลขได้ เช่น ‘⠭⠊⠼⠃⠇⠑⠙’ (= dis-a-ble-d) ซึ่งใช้เขียนคำว่า “disa**ble**d” ‘⠁⠼⠙’ (= a-เครื่องหมายนำเลข-4) ที่ใช้เขียน “a4” หากระบบตรวจสอบบริบทซ้ายและ/หรือขวาของอักขระเบรลล์ ‘⠼’ แล้วระบุได้ว่า ‘⠼’ ใช้แทนเครื่องหมายนำเลข ระบบจะระบุให้สายอักขระย่อยที่กำลังดำเนินการอยู่มีสถานะเป็นเบรลล์รูปเต็ม

(2) ‘⠨’ ในตำแหน่งกลางสายอักขระย่อย

เมื่อ ‘⠨’ ปรากฏกลางสายอักขระย่อยในเบรลล์อังกฤษรูปย่อ อาจตีความ ‘⠨’ ให้ใช้แทน “dd” หรือ “.” ได้ หากระบบตรวจสอบบริบทซ้ายและ/หรือขวาของอักขระเบรลล์ตัวนี้ แล้วระบุได้ว่า ‘⠨’ ใช้แทน “.” ระบบจะระบุให้สายอักขระย่อยที่ทำการตรวจสอบมีสถานะเป็นเบรลล์รูปเต็ม

(3) อักษรโดด

การเขียนเลขข้อ (numbering) ในอังกฤษเบรลล์รูปย่อ มักจะเขียนโดยไม่มีการกำกับด้วยเครื่องหมาย ‘⠰’ (เครื่องหมายกำกับตัวอักษร) ทำให้เกิดปัญหาในการถ่ายถอด เช่น ระบบจะถ่ายถอดเลขข้อเป็น ข้อ “but.” ข้อ “can.” แทนที่จะเป็นข้อ “b.” หรือข้อ “c.” เป็นต้น ซึ่งประเด็นนี้ยังเป็นปัญหาต่อการถ่ายถอดเบรลล์อังกฤษรูปย่อให้เป็นอักษรปกติ

อย่างไรก็ตาม ในการเขียนเลขข้อนั้นบางครั้งมีการครอบด้วยเครื่องหมายวงเล็บ ดังนั้น ระบบจึงสามารถใช้กฎมาช่วยในการถ่ายถอดเลขข้อที่ปรากฏในวงเล็บได้ ซึ่งต้องตีความอักษรโดดในวงเล็บให้เป็นเบรลล์รูปเต็ม ไม่ใช่เบรลล์รูปย่อ มิฉะนั้น จะทำให้เกิดการถ่ายถอดเลขข้อผิด เช่น ถ่ายถอดผิดเป็น ข้อ “(can)” ข้อ “(do)” เป็นต้น ในการระบุให้สายอักขระย่อยที่มีลักษณะดังกล่าว มีสถานะเป็นเบรลล์อังกฤษรูปเต็มนั้น ระบบต้แงตรวจสอบว่า มี ‘⠣’ (= ( ) และ ‘⠜’ (=)) ปรากฏอยู่ในสายอักขระย่อยเดียวกันหรือไม่ แล้วตรวจสอบเพิ่มเติมว่าระหว่างเครื่องหมายวงเล็บมีอักษรอังกฤษ (ไม่ว่าจะเป็นตัวพิมพ์เล็กหรือตัวพิมพ์ใหญ่ก็ตาม) ปรากฏอยู่เพียงตัวเดียวหรือไม่ หากตรงตามหลักเกณฑ์ทั้ง 2 ข้อนี้ ระบบจะระบุให้สายอักขระย่อยปัจจุบันมีสถานะเป็นเบรลล์รูปเต็ม

ข. คอมพิวเตอร์เบรลล์

ในการเขียนที่อยู่ของเว็บไซต์หรือที่อยู่ของอีเมล์ในอักษรเบรลล์ จะใช้รหัสคอมพิวเตอร์เบรลล์ (ASCII) เพื่อให้เกิดความแตกต่างจากข้อความทั่วไป[23] ดังนั้นเมื่อระบบพบเครื่องหมาย ‘⠸⠼’ (เครื่องหมายกำกับจุดเริ่มต้นคอมพิวเตอร์เบรลล์), ‘⠸⠼’ (เครื่องหมายกำกับจุดสิ้นสุดคอมพิวเตอร์เบรลล์) และ/หรือ ‘⠸⠼’ (เครื่องหมายปรากฏท้ายบรรทัดเพื่อแสดงว่ารหัสเบรลล์ที่ตามมายังคงเป็นคอมพิวเตอร์เบรลล์) ระบบจะระบุสถานะของสายอักขระย่อยปัจจุบันให้เป็นคอมพิวเตอร์เบรลล์

ค. ตัวเลขเบรลล์

ถ้าสถานะของสายอักขระย่อยปัจจุบันเป็นตัวเลขเบรลล์ และสายอักขระย่อย<u>มี</u>เครื่องหมาย ‘⠰’ (= letter sign) หรือ <u>ไม่มี</u>อักขระต่อไปนี้ปรากฏอยู่คือ [‘⠁-⠚’ (=k - z) ‘⠪’ (=ow) ‘⠻’ (=er) ‘⠳’ (=ou) ‘⠫’ (=gh) ‘⠗’ (=ar)] ระบบจะระบุสถานะของสายอักขระย่อยให้เป็นตัวเลขเบรลล์

จากกฎข้อนี้จะเห็นได้ว่า ระบบจำเป็นต้องมีสถานะตัวเลขเบรลล์เพื่อกำกับสายอักขระย่อย ดังจะพบว่าในกระบวนการตัดแบ่งสายอักขระออกเป็นสายอักขระย่อยนั้น หากระบบพบสายอักขระที่มีเครื่องหมาย ‘⠤’ (= -) ปรากฏกลางสายอักขระ สายอักขระนั้นจะถูกแบ่งออกเป็น 2 สายอักขระย่อย โดยไม่คำนึงว่าสายอักขระที่ทำการตัดแบ่งจะเป็นสายอักขระประเภทใด จึงทำให้สายอักขระที่เป็นตัวเลขถูกตัดแบ่งเช่นเดียวกัน หากตัวเลขนั้นมีเครื่องหมาย‘⠤’ปรากฏกลางสายอักขระ

---

[23] รหัสคอมพิวเตอร์เบรลล์หรือรหัสแอสกี เป็นรหัสที่ต้องตีความแบบตัวต่อตัวในอักษรเบรลล์ เช่น ‘⠁⠃⠉’ (ซึ่งปรากฏในรหัสแอสกีคือ a80) หากผู้ใช้อักษรเบรลล์ตระหนักว่าตัวอักษรที่กำลังสัมผัสอยู่นี้เป็นรหัสคอมพิวเตอร์เบรลล์ก็ต้องตีความให้ ‘⠁⠃⠉’ เป็น “a80” ตามที่รหัสแอสกีแสดงบนหน้าจอคอมพิวเตอร์

อย่างไรก็ตาม ในการเขียนตัวเลขที่คั่นด้วย '⠼' นั้น ไม่ต้องเขียนเครื่องหมาย '⠼' หลังเครื่องหมาย '⠼' เช่น '⠼⠁⠂⠉⠄' (= 1-3) หากระบบทำการตัดแบ่งสายอักขระตัวเลขแล้วจะทำให้ได้ 2 สายอักขระย่อยคือ ได้เป็น '⠼⠁⠂⠄' (= 1-) และ'⠉' (= 3) สายอักขระย่อยหลังนี้เองที่เป็นปัญหาต่อการถ่ายถอดเพราะการที่ไม่มี '⠼'(เครื่องหมายนำเลข) กำกับจะทำให้ระบบนำสายอักขระย่อยนี้ไปถ่ายถอดให้เป็นเบรลล์รูปย่อ ทำให้ได้เป็น "1-can" แต่หากกำหนดให้สองสายอักขระย่อยมีสถานะเดียวกัน(คือมีสถานะเป็นเบรลล์รูปเต็มเหมือนกันเนื่องจากเป็นตัวเลข) ก็ยังเกิดปัญหาอยู่เพราะ '⠉' (=3) ตีความในเบรลล์รูปเต็มได้เป็น "c" ทำให้ได้ผลลัพธ์เป็น "1-c" ฉะนั้นจึงต้องกำหนดให้สายอักขระย่อยหลังมีสถานะเป็นตัวเลข และกำหนดให้กระบวนการถ่ายถอดเบรลล์รูปเต็มสามารถถ่ายถอดตัวเลขได้แม้จะไม่พบเครื่องหมาย '⠼' ในสายอักขระย่อยก็ตาม

ง. เบรลล์รูปย่อ

ระบบจะระบุสถานะสายอักขระย่อยให้เป็นเบรลล์รูปย่อ หากเป็นไปตามเงื่อนไขดังนี้

(1) ถ้าสายอักขระย่อยปัจจุบันมี 3 อักขระ

โดยที่อักขระตัวแรกเป็น '⠠' (=single capital sign) ตัวที่ 2 เป็นอักขระใดก็ได้ และตัวที่ 3 เป็น'⠲' (= .) ระบบจะเติม '⠰'(=letter sign) หน้าสายอักขระ แล้วระบุสถานะของสายอักขระย่อย ให้เป็นเบรลล์รูปย่อ สำหรับกฎข้อนี้มีไว้เพื่อใช้กับการถ่ายถอดอักษรตัวพิมพ์ใหญ่ที่ตามด้วยจุด (.) และตัวย่อบางตัว เช่น '⠠⠎⠠⠞' (=ST.) ระบบจำเป็นต้องเติม '⠰' หน้าสายอักขระเพื่อกันไม่ให้ '⠠⠞'เป็นตัวย่อระดับคำ (คือไม่ให้'⠠⠞'ถ่ายถอดเป็น "still")

(2) ในกรณีอื่น สายอักขระย่อยปัจจุบันจะมีสถานะเป็นเบรลล์รูปย่อ

หมายเหตุ: ในการถ่ายถอดเบรลล์รูปเต็ม ผู้ใช้ระบบจะต้องทำการปรับเปลี่ยนเงื่อนไขในข้อ ง. ที่ผ่านมา ให้มีสถานะเป็นเบรลล์รูปเต็มแทนที่จะมีสถานะเป็นเบรลล์รูปย่อ จึงจะถ่ายถอดเบรลล์รูปเต็มได้ถูกต้อง

5.2.3.3 การถ่ายถอดสายอักขระย่อยตามสถานะที่ระบุได้

หลังจากระบบระบุสถานะของสายอักขระย่อยได้แล้ว จะนำแต่ละสายอักขระย่อยไปถ่ายถอดตามสถานะที่ระบุได้ ซึ่งกระบวนการถ่ายถอดในแต่ละสถานะ มีรายละเอียดดังนี้

5.2.3.3.1 การถ่ายถอดเบรลล์อังกฤษรูปเต็ม

ในการถ่ายถอดเบรลล์อังกฤษรูปเต็ม ระบบจะถ่ายถอดสายอักขระย่อยตามสถานะที่ระบุได้คือ สถานะเบรลล์รูปเต็มและสถานะตัวเลข แต่เนื่องจากเบรลล์อังกฤษรูปเต็มมีเครื่องหมายส่วนหนึ่งที่ต้องตีความตามตำแหน่งที่ปรากฏเช่นเดียวกับเบรลล์อังกฤษรูปย่อ ระบบจึงต้องถ่ายถอดเครื่องหมายเหล่านี้ตามตำแหน่งที่ปรากฏด้วยเช่นกัน เช่นเครื่องหมาย ‘⠶’ หากปรากฏต้นสายอักขระต้องตีความเป็น “(” วงเล็บเปิด ถ้าปรากฏท้ายสายอักขระจะต้องตีความให้เป็น “)” วงเล็บปิด (ดูหัวข้อ 4.2.1 ประกอบ)

ดังนั้นในการถ่ายถอดเบรลล์รูปเต็มระบบจึงต้องดำเนินการ 2 ขั้นตอน ได้แก่ *การถ่ายถอดส่วนต้น* และ*การถ่ายถอดส่วนที่เหลือ*

ขั้นที่ 1 การถ่ายถอดส่วนต้น

ระบบจะดำเนินการโดยภาพรวมคือ นำเบรลล์ต้นสายอักขระ 2 เซลล์ไปตรวจสอบกับตารางคู่เทียบเบรลล์ต้นสายอักขระ (ตามตารางที่ 5.15) ถ้าไม่พบ จะลดจำนวนเบรลล์ทางขวาลงเหลือ 1 เซลล์ หากพบว่าเบรลล์ 2 เซลล์ หรือ 1 เซลล์ นั้นตรงกับอักขระเบรลล์ในตารางที่ 5.15 และตรงตามเงื่อนไขที่กำหนดไว้เพิ่มเติม ระบบจะทำการถ่ายถอดเบรลล์ที่ตรงกับตารางที่ 5.15 แล้วนำอักขระต่อไปมาดำเนินการไปเรื่อย ๆ จนพบว่าเบรลล์เซลล์ปัจจุบันไม่ตรงกับตาราง ระบบจึงสิ้นสุดการถ่ายถอดส่วนนี้

อย่างไรก็ตาม สิ่งที่กล่าวมาเป็นเพียงภาพรวมของการถ่ายถอดเบรลล์ส่วนต้นสายอักขระย่อยเท่านั้น แต่ในรายละเอียดระบบจะดำเนินการดังนี้

(1) <u>การลบเครื่องหมายประกอบคำ</u> เนื่องจากเครื่องหมายประกอบคำปรากฏเฉพาะในอักษรเบรลล์เท่านั้น ระบบจึงต้องตรวจสอบเบรลล์เซลล์ปัจจุบันว่าเป็นเครื่องหมายใด แล้วดำเนินการตามเครื่องหมายที่พบ ดังนี้

(ก) หากเบรลล์ 2 เซลล์ปัจจุบันเป็น ‘⠠⠠’ (= double capital sign) ระบบจะลบเบรลล์ 2 เซลล์นี้ออกจากสายอักขระ แล้วกำหนดสถานะของตัวอักษรให้เป็นตัวพิมพ์ใหญ่ทั้งคำ (all capitalized) จากนั้นกำหนดจุดเริ่มต้นตัวพิมพ์ใหญ่ให้มีค่าเท่ากับตำแหน่งเบรลล์เซลล์ปัจจุบัน [24]

(ข) หากเบรลล์เซลล์ปัจจุบันเป็น ‘⠠’ (=single capital sign) ระบบจะลบเบรลล์เซลล์นี้ออกจากสายอักขระ แล้วกำหนดสถานะตัวอักษรให้เป็นตัวพิมพ์ใหญ่ตัวแรก จากนั้นกำหนดจุดเริ่มต้นตัวพิมพ์ใหญ่ให้มีค่าเท่ากับตำแหน่งเบรลล์เซลล์ปัจจุบัน

---

[24] เหตุที่ต้องมีค่าจุดเริ่มต้น เนื่องจากระบบต้องนำค่านี้ไปใช้กับกระบวนการอื่นเช่นกัน

(ค) หากเบรลล์เซลล์ปัจจุบันเป็น ‘⠨’ (=letter sign) ระบบจะลบเครื่องหมายนี้ออกจากสายอักขระ ขณะเดียวกันก็กำหนดสถานะของสายอักขระย่อยให้เป็นเบรลล์รูปเต็ม และกำหนดสถานะของตัวอักษรให้เป็นตัวพิมพ์เล็ก

(2) <u>การถ่ายถอดอักขระเบรลล์ส่วนต้น</u> ในขั้นนี้ ระบบจะนำเบรลล์ 2 เซลล์ไปตรวจสอบกับตารางที่ 5.15 ถ้าไม่พบจะขยับมาตรวจสอบเพียง 1 เซลล์ ถ้าพบว่าเบรลล์ 2 หรือ 1 เซลล์ ตรงกับเบรลล์ในตาราง ระบบจะหยุดตรวจสอบ จากนั้นเปรียบเทียบว่าสายอักขระย่อยมีความยาวมากกว่าจำนวนเซลล์ที่ตรงกับตารางหรือไม่[25] ถ้ามีจำนวนมากกว่า ระบบจะตรวจสอบสถานะตัวอักษร ณ ปัจจุบันแล้วดำเนินการตามเงื่อนไขต่อไปนี้

(ก) หากสถานะปัจจุบันของตัวอักษรเป็น *ตัวพิมพ์ใหญ่ทั้งคำ* ระบบจะถ่ายถอดอักขระเบรลล์ชุดที่ตรงกับตารางที่ 5.15 พร้อมกับแปลงตัวอักษรที่ได้ให้เป็นตัวพิมพ์ใหญ่ทั้งหมด โดยเริ่มจากตำแหน่งจุดเริ่มต้นตัวพิมพ์ใหญ่

(ข) หากสถานะของตัวอักษรเป็น *ตัวพิมพ์ใหญ่ตัวแรก* ระบบจะถ่ายถอดอักขระเบรลล์ชุดที่ตรงกับตารางที่ 5.15 พร้อมกับแปลงตัวอักษรที่ถ่ายถอดแล้วให้เป็นตัวพิมพ์ใหญ่เฉพาะอักขระที่ปรากฏในตำแหน่งจุดเริ่มต้นตัวพิมพ์ใหญ่ แล้วกำหนดสถานะของตัวอักษรให้เป็นตัวพิมพ์เล็ก

(ค) ในกรณีอื่น ระบบจะถ่ายถอดเบรลล์เซลล์ปัจจุบันโดยใช้ตารางที่ 5.15

(3) <u>การตรวจสอบจุดสิ้นสุดส่วนต้นสายอักขระย่อย</u> ในขั้นนี้หากเบรลล์ 1-2 เซลล์ปัจจุบันยังอยู่ในเงื่อนไขข้อ(1) - (2) ระบบจะยังดำเนินการถ่ายถอดเบรลล์ต้นสายอักขระย่อยต่อไปเรื่อย ๆ แต่ทันทีที่ไม่ตรงกับเงื่อนไขใด ระบบจะสิ้นสุดกระบวนการนี้

ขั้นที่ 2 การถ่ายถอดส่วนที่เหลือ

ระบบจะดำเนินการกับอักขระเบรลล์ครั้งละ 2 เซลล์ ถ้าไม่ตรงกับเงื่อนไขจะขยับลงมาดำเนินการเพียง 1 เซลล์ โดยจะทำต่อไปเรื่อย ๆ จนอักขระเบรลล์ที่ดำเนินการไม่ตรงกับตารางจึงจะสิ้นสุดการถ่ายถอดส่วนนี้ ซึ่งมีขั้นตอนดังนี้

(1) <u>การลบเครื่องหมายประกอบคำ</u> ในขั้นนี้ระบบจะดำเนินการกับเบรลล์เซลล์ปัจจุบันเช่นเดียวกับการลบเครื่องหมายประกอบคำหัวข้อการถ่ายถอดส่วนต้นของสายอักขระย่อย

---

[25] เหตุที่ต้องตรวจสอบความยาวของสายอักขระ เพราะมีอักขระเบรลล์บางชุดที่มีความยาว 1 หรือ 2 ตัว โดยที่อักขระเหล่านี้สามารถตีความเป็นคำโดดหรือเครื่องหมายได้ หากไม่ทำการตรวจสอบความยาวสายอักขระจะทำให้อักขระเบรลล์ถูกตีความเป็นเครื่องหมายโดยอัตโนมัติ เป็นผลให้ไม่สามารถถ่ายถอดอักขระเหล่านี้เป็นคำโดดได้ในขั้นต่อไป

(2) <u>การตรวจสอบเครื่องหมายกำกับตัวเลข</u> หากเบรลล์เซลล์ปัจจุบันไม่ใช่เครื่องหมายประกอบคำ ระบบจะตรวจสอบว่าเบรลล์ 1-2 เซลล์ปัจจุบันนี้ เป็นเครื่องหมายกำกับตัวเลข[26] หรือไม่ตามเงื่อนไขดังนี้

(ก) หากเบรลล์ 2 เซลล์ปัจจุบันเป็น ‘⠸⠼’ ระบบจะแทนที่เบรลล์คู่นี้ด้วย “$”

(ข) หากเบรลล์ 2 เซลล์ปัจจุบันเป็น ‘⠨⠼’ ระบบจะแทนที่เบรลล์คู่นี้ด้วย “£”

(ค) หากเบรลล์เซลล์ปัจจุบันเป็น ‘⠼’ ระบบจะลบเครื่องหมายนี้ออกจากสายอักขระ

เมื่อดำเนินการตามเงื่อนไขข้อใดข้อหนึ่งแล้ว ระบบจะทำการกำหนดสถานะของสายอักขระย่อยให้เป็นตัวเลขเบรลล์

(3) <u>การถ่ายถอดเบรลล์เซลล์ปัจจุบัน</u> เมื่อระบุสถานะของสายอักขระย่อยได้แล้ว ระบบจะถ่ายถอดเบรลล์เซลล์ปัจจุบันโดยอาศัยสถานะที่ระบุได้ตามลำดับ ดังนี้

(ก) การถ่ายถอดอักขระเบรลล์พิเศษท้ายตัวเลข

หากเบรลล์เซลล์ปัจจุบันเป็นตัวสุดท้ายของสายอักขระย่อยและมีสถานะตัวเลขเบรลล์ ระบบจะตรวจสอบเบรลล์เซลล์ปัจจุบัน แล้วถ่ายถอดตามเงื่อนไขต่อไปนี้ หากเป็น ‘⠹’ ระบบจะถ่ายถอด ให้เป็น “st” หรือหากเบรลล์เซลล์ปัจจุบันเป็น ‘⠾’ ระบบจะถ่ายถอด ให้เป็น “th”

(ข) การถ่ายถอดตัวเลขเบรลล์

หากสถานะของสายอักขระย่อยเป็นตัวเลขเบรลล์ และเบรลล์เซลล์ปัจจุบันตรงกับเบรลล์ในตารางคู่เทียบตัวเลข (ตารางที่ 5.23) ระบบจะถ่ายถอดให้เป็นตัวเลข แล้วนำเบรลล์เซลล์ต่อไปมาดำเนินการตั้งแต่ขั้น (ก) แต่ถ้าไม่ตรงกับเงื่อนไขเหล่านี้ ระบบจะดำเนินการในขั้น (ค) ต่อไป

(ค) การถ่ายถอดเครื่องหมายต่าง ๆ

หากระบบตรวจสอบเบรลล์ 3 เซลล์ปัจจุบัน, 2 เซลล์ปัจจุบัน และ 1 เซลล์ปัจจุบัน ตามลำดับ แล้วพบว่าอักษรชุดใดตรงกับอักษรเบรลล์ในตารางคู่เทียบเครื่องหมายหลังคำ (ตารางที่ 5.16) เป็นครั้งแรก ระบบจะหยุดตรวจสอบ แล้วถ่ายถอดให้เป็นอักษรปกติ จากนั้นนำเบรลล์เซลล์ต่อไปมาดำเนินการตั้งแต่ขั้นแรก (คือขั้น (ก)) แต่ถ้าไม่มีเบรลล์ชุดใดตรงกับตารางนี้ จะดำเนินการในขั้นต่อไป

---

[26] เครื่องหมายกำกับตัวเลข ประกอบด้วย ‘⠸⠼’ (=$ + ตัวเลข) ‘⠨⠼’’ (=£ + ตัวเลข) และ‘⠼’ (=ตัวเลข)

(ง) การถ่ายถอดตัวอักษรโดด

ระบบจะระบุให้สายอักขระเป็นเบรลล์รูปเต็ม แล้วดำเนินการดังนี้(1) หากพบว่าสถานะของตัวอักษรเป็นตัวพิมพ์ใหญ่ทั้งคำ ระบบจะแปลงเบรลล์เซลล์ปัจจุบันให้เป็นอักษรอังกฤษตัวพิมพ์ใหญ่โดยตรง[27] แล้วนำเบรลล์เซลล์ต่อไปมาดำเนินการ และ (2) หากพบว่าสถานะของตัวอักษรเป็นตัวพิมพ์ใหญ่ตัวแรก ระบบจะถ่ายถอดเบรลล์เซลล์ปัจจุบันให้เป็นตัวพิมพ์ใหญ่ ในขณะเดียวกันกำหนดสถานะของตัวอักษรให้เป็นตัวพิมพ์เล็ก แล้วนำเบรลล์เซลล์ต่อไปมาดำเนินการตั้งแต่ขั้นแรก (ขั้น (ก))

5.2.3.3.2 การถ่ายถอดคอมพิวเตอร์เบรลล์

ในการถ่ายถอดส่วนนี้ ระบบจะดำเนินการกับสายอักขระย่อย 2 กระบวนการ คือ ถ่ายถอดส่วนต้นของสายอักขระย่อย (ตามกระบวนการในหัวข้อ 5.2.3.3.1 ขั้นที่1) หลังจากนั้นถ่ายถอดส่วนท้ายของสายอักขระย่อย (ตามกระบวนการในหัวข้อ 5.2.3.3.4) หลังจากดำเนินการครบทั้ง 2 กระบวนการ ระบบจะนำสายอักขระย่อยปัจจุบันทั้งหมดไปเป็นผลลัพธ์ของการถ่ายถอดอักษร[28]

5.2.3.3.3 การถ่ายถอดตัวเลขเบรลล์

ในการถ่ายถอดตัวเลขเบรลล์ ระบบจะดำเนินการตามกระบวนการถ่ายถอดเบรลล์รูปเต็มในหัวข้อ 5.2.3.3.1 เนื่องจากในบางสายอักขระย่อย มีตัวเลขและตัวอักษรอังกฤษปนกัน เช่น ‘⠼⠃⠁’ (=2a) จึงต้องถ่ายถอดเบรลล์รูปเต็มและตัวเลขโดยอาศัยกระบวนการเดียวกัน

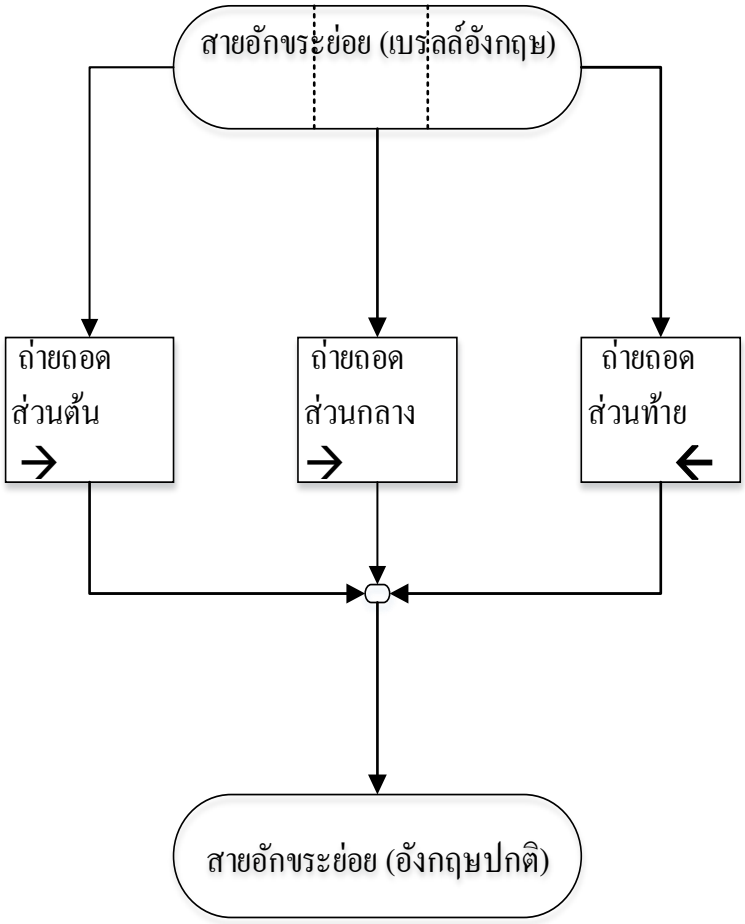
5.2.3.3.4 การถ่ายถอดเบรลล์อังกฤษรูปย่อ

ในการถ่ายถอดเบรลล์อังกฤษรูปย่อจะแบ่งการถ่ายถอดเป็น 3 ส่วน คือ (1) ถ่ายถอดส่วนต้นสายอักขระย่อย (2) ถ่ายถอดส่วนท้ายสายอักขระย่อย และ (3) ถ่ายถอดส่วนกลางสายอักขระย่อยตามลำดับ ดังแสดงในแผนภาพต่อไปนี้

---

[27] คอมพิวเตอร์จะใช้ตัวอักษรภาษาอังกฤษชุดเดียวกัน สำหรับเบรลล์อังกฤษและอังกฤษปกติ

[28] เนื่องจากรหัสของอักษรที่ปรากฏในส่วนกลางของสายอักขระย่อยเป็นสัญลักษณ์แอสกี ซึ่งเป็นสัญลักษณ์คอมพิวเตอร์เบรลล์อยู่แล้วจึงไม่จำเป็นต้องถ่ายถอดตัวอักษรเหล่านี้



ภาพที่ 5.4 ขั้นตอนในการถ่ายถอดเบรลล์อังกฤษรูปย่อ

จากแผนภาพข้างต้น ระบบจะดำเนินการถ่ายถอดเบรลล์อังกฤษรูปย่อตามลำดับขั้นตอน ดังนี้

ขั้นที่ 1 การถ่ายถอดส่วนต้น

ในส่วนนี้ ระบบจะดำเนินการเช่นเดียวกับเบรลล์อังกฤษรูปเต็มในหัวข้อที่ 5.2.3.1.1 แต่เมื่อสิ้นสุดกระบวนการนี้ ระบบจะเก็บตำแหน่งของเบรลล์เซลล์ปัจจุบันไว้เพื่อใช้เป็นจุดเริ่มถ่ายถอด[29]ในกระบวนการต่อไป

---

[29] ***"จุดเริ่มถ่ายถอด"*** เป็นตำแหน่งของอักษรตัวแรกที่ระบบใช้เป็นจุดเริ่มต้นของกระบวนการถ่ายถอดส่วนกลางสายอักขระย่อย ตามหัวข้อที่ 5.2.3.3.4 (ขั้นตอนที่ 3 การถ่ายถอดส่วนกลางสายอักขระย่อย)

ขั้นที่ 2 การถ่ายถอดส่วนท้าย

ในการถ่ายถอดส่วนนี้ ระบบจะดำเนินการในภาพรวมคือ นำ 3 ตัวอักษรสุดท้ายในสายอักขระไปตรวจสอบกับตารางคู่เทียบเครื่องหมายหลังคำและตารางคู่เทียบเบรลล์หลังตัวย่อระดับคำ (ตารางที่ 5.16 และ 5.19) หากไม่ตรงกับอักษรเบรลล์ชุดใดในตารางทั้งสองนี้ระบบจะนำ 2 ตัวอักษรสุดท้ายไปตรวจสอบกับตารางทั้งสอง หากยังไม่ตรงก็จะนำเพียง 1 ตัวอักษรสุดท้ายไปตรวจสอบกับตารางทั้งสอง

หากระบบพบว่า อักขระเบรลล์ 3 ตัว 2 ตัว หรือ 1 ตัว ที่กล่าวมานี้ ตรงกับตารางใดตารางหนึ่ง ระบบจะทำการถ่ายถอดตามตารางนั้น แล้วนำอักขระเบรลล์ชุดทางซ้าย (ของอักขระที่ถ่ายถอดไปแล้ว) มาดำเนินการในลักษณะเดียวกัน จนพบว่า ไม่มีอักขระเบรลล์ชุดใดที่ตรงกับตารางทั้งสองอีกต่อไป จึงสิ้นสุดการถ่ายถอดส่วนท้ายสายอักขระย่อย

สาเหตุที่ระบบต้องถ่ายถอดในส่วนท้ายก่อนถ่ายถอดส่วนกลาง และกระบวนการถ่ายถอดในส่วนนี้ต้องดำเนินการจากขวาไปซ้าย เนื่องจากส่วนท้ายของสายอักขระมักจะมีเครื่องหมายต่าง ๆ ปรากฏ และเครื่องหมายเหล่านี้ก็สามารถจะเรียงต่อกันได้เช่น ‘⠐⠣ ⠒⠕ ⠐⠣ ⠒⠕ ⠐⠣ ⠒⠕’ คือ boy.”) หากทำการถ่ายถอดจากส่วนกลางไปหาส่วนท้ายสายอักขระโดยตรง จะทำให้อักขระบางตัวที่ต้องตีความโดยอาศัยตำแหน่งการปรากฏทำได้ยากยิ่งขึ้น เช่นในกรณีนี้จะตีความ ‘⠒⠒’ (=.) ได้ยาก เนื่องจาก ‘⠒⠒’ ดูเหมือนจะปรากฏในตำแหน่งกลางสายอักขระ ซึ่งต้องตีความให้เป็น “dd” แต่อันที่จริงแล้วต้องตีความให้ ‘⠒⠒’ ปรากฏร่วมกับเครื่องหมายอื่น ๆ ในตำแหน่งท้ายสายอักขระ นอกจากนี้การเริ่มถ่ายถอดจากตำแหน่งขวาสุดของสายอักขระจะช่วยให้การถ่ายถอดส่วนท้ายสายอักขระไม่ซับซ้อนเนื่องจากเครื่องหมายต่าง ๆ ในภาษาอังกฤษมักปรากฏทางขวา และเครื่องหมายต่าง ๆ เหล่านี้สามารถปรากฏเรียงต่อกันได้ หากระบบตรวจสอบแล้วไม่พบเครื่องหมาย ระบบจะใช้ตัวอักษรที่ไม่ใช่เครื่องหมายเป็นจุดสิ้นสุดของกระบวนการนี้ได้

ในการถ่ายถอดส่วนท้ายสายอักขระย่อย มีรายละเอียดการดำเนินการดังนี้

(1) <u>การถ่ายถอดเบรลล์หลังตัวย่อระดับคำ</u> ในการถ่ายถอดเบรลล์ชุดนี้ ระบบจะต้องใช้ตารางที่ 5.19 ซึ่งประกอบด้วยคู่เทียบเบรลล์หลังตัวย่อระดับคำ โดยตรวจสอบเงื่อนไขต่อไปนี้

(ก) *การถ่ายถอดเบรลล์ 3 เซลล์ปัจจุบัน* หากเบรลล์ทั้ง 3 เซลล์นับจากเซลล์ปัจจุบันไปทางซ้ายตรงกับตารางที่ 5.19 และจำนวนอักขระนับตั้งแต่จุดเริ่มถ่ายถอดถึงเซลล์ปัจจุบันมีค่ามากกว่า 3 ระบบจะถ่ายถอด 3 เซลล์ปัจจุบัน แต่หากไม่ตรงกับตารางระบบจะดำเนินการในขั้น (ข) ต่อไป

(ข) *การถ่ายถอดเบรลล์สองเซลล์ปัจจุบัน* ระบบจะถ่ายถอดเบรลล์สองเซลล์ปัจจุบันได้ ก็ต่อเมื่อเบรลล์ 2 เซลล์นี้ตรงกับตารางที่5.18 และเป็นไปตามเงื่อนไขคือ (ข.1) จำนวนอักขระนับตั้งแต่ “จุดเริ่มถ่ายถอดจนถึงเซลล์ปัจจุบัน”มีค่ามากกว่า 3 และเซลล์ที่ 3 (นับจากเซลล์ปัจจุบันไปทางซ้าย) เป็นตัวอักษรอังกฤษ (a-z) (ข.2) จำนวนอักขระตั้งแต่ “จุดเริ่มถ่ายถอดถึงเซลล์ปัจจุบัน”มีค่าเท่ากับ 3 และเซลล์ที่ 3 ไม่ใช่ตัวอักษรอังกฤษ แต่หากไม่ได้ดำเนินการในขั้นนี้ระบบจะไปสู่ขั้นต่อไป

(ค) *การถ่ายถอดเบรลล์เซลล์ปัจจุบัน* ระบบจะดำเนินการขั้นนี้เมื่อพบว่าเบรลล์เซลล์ปัจจุบันตรงกับตารางคู่เทียบหลังตัวย่อระดับคำ (ตารางที่ 5.19) และเป็นไปตามเงื่อนไขหนึ่งต่อไปนี้

(ค.1) จำนวนอักขระนับตั้งแต่จุดเริ่มถ่ายถอดถึงเซลล์ปัจจุบันมีค่ามากกว่า 3 และเบรลล์สามเซลล์หน้าเซลล์ปัจจุบัน ไม่ใช่ ‘⠐⠁⠇’ (= alt)[30]

(ค.2) จำนวนอักขระจาก“จุดเริ่มถ่ายถอดจนถึงเซลล์ปัจจุบัน” มีค่าเท่ากับ 3 และเบรลล์เซลล์ปัจจุบันเป็น ‘⠎’ (= s) และเบรลล์หน้าเซลล์ปัจจุบันเป็นตัวอักษรอังกฤษ

หลังจากถ่ายถอดอักษรหลังตัวย่อระดับคำแล้ว ระบบจะนำ 3 เซลล์ทางซ้ายมาดำเนินการในส่วนของการถ่ายถอดเบรลล์ท้ายสายอักขระตั้งแต่ขั้นแรก (หัวข้อ 5.2.3.3.4 ขั้นตอนที่ 2 (1) ก) แต่ถ้าไม่มีการถ่ายถอดส่วนนี้เกิดขึ้น ระบบจะดำเนินการในขั้นการถ่ายถอดเครื่องหมายหลังคำต่อไป

(2) <u>การถ่ายถอดเครื่องหมายหลังคำ</u> ในการถ่ายถอดเครื่องหมายหลังคำ จะดำเนินการดังนี้

(ก) ตรวจสอบเบรลล์เซลล์ปัจจุบันและเซลล์ซ้าย ว่าตรงกับเบรลล์ในตารางคู่เทียบเครื่องหมายหลังคำ (ตารางที่ 5.16) และสายอักขระปัจจุบันมีความยาวมากกว่า 2 ตัวหรือไม่ หากเป็นไปตามสองเงื่อนไขนี้ระบบจะนำเบรลล์ 2 เซลล์ดังกล่าวไปถ่ายถอดโดยอาศัยตารางที่ 5.16 นี้

(ข) ตรวจสอบเฉพาะเบรลล์เซลล์ปัจจุบันว่าตรงกับตารางที่ 5.16 หรือไม่ และสายอักขระปัจจุบันมีความยาวมากกว่า 1 ตัวหรือไม่ หากเป็นไปตามเงื่อนไขทั้งสองนี้ระบบจะถ่ายถอดเบรลล์เซลล์นี้โดยอาศัยตารางที่ 5.16

---

[30] ระบบไม่สามารถให้ตัวอักษร 3 ตัว หน้าเซลล์ปัจจุบันเป็น “alt” ได้เนื่องจากจะทำให้ถ่ายถอดคำว่า ‘alt)’ (=a-l-t-er) ได้ เนื่องจาก “alt” ใช่ย่อ “altogether” หากยอมให้3อักษรปรากฏได้จะทำให้ระบบถ่ายถอดเป็น “altogetherer”

หลังจากถ่ายถอดเครื่องหมายหลังคำแล้ว ระบบจะ*นำ* 3 เซลล์ทางซ้ายมาดำเนินการตั้งแต่ขั้นแรก (หัวข้อ 5.2.3.3.4 ขั้นตอนที่ 2 (1) ก) แต่ถ้าไม่มีการถ่ายถอดส่วนนี้เกิดขึ้น ระบบจะดำเนินการในขั้นที่ (3) ต่อไป

(3) <u>การสิ้นสุดการถ่ายถอดส่วนท้าย</u> หากไม่ตรงกับเงื่อนไขในทั้งข้อ ก. และข้อ ข. ระบบจะดำเนินการในขั้นสุดท้ายของการถ่ายถอดส่วนท้ายสายอักขระย่อย โดยจะเก็บค่าตำแหน่งเบรลล์เซลล์ปัจจุบันไว้เพื่อดำเนินการในขั้นต่อไป (เพื่อสะดวกต่อการอ้างอิง จะขอเรียกตำแหน่งนี้ว่า "*จุดสิ้นสุดการถ่ายถอด*") จากนั้นตรวจสอบสถานะของตัวอักษร หากมีสถานะเป็นตัวพิมพ์ใหญ่ทั้งคำ และตัวอักษรที่เพิ่งจะถ่ายถอดไม่ใช่ " 's" ระบบจะแปลงตัวอักษรที่ถ่ายถอดมาในส่วนนี้ทั้งหมดให้เป็นตัวพิมพ์ใหญ่แล้วจึงจบการทำงานในส่วนนี้ แล้วดำเนินการในส่วนต่อไป

ขั้นที่ 3 การถ่ายถอดส่วนกลาง

ในกระบวนการนี้ ระบบจะถ่ายถอดอักขระเบรลล์ตั้งแต่จุดเริ่มถ่ายถอด ไปจนถึงจุดสิ้นสุดการถ่ายถอด โดยจะตรวจสอบอักษรทีละขั้น ในขั้นแรก ตรวจสอบอักขระทั้งหมดที่ปรากฏในส่วนกลางสายอักขระย่อย ว่าเป็นตัวย่อระดับคำหรือไม่

หากใช่ตัวย่อระดับคำ จะถ่ายถอดส่วนกลางของสายอักขระย่อยทั้งส่วนแล้วสิ้นสุดกระบวนการถ่ายถอดส่วนกลางของสายอักขระย่อย

แต่หากไม่ใช่ตัวย่อระดับคำ ระบบจะเข้าสู่ขั้นที่ 2 คือ ตรวจสอบเบรลล์ 2 เซลล์ปัจจุบันว่าเป็นตัวย่อเบรลล์เซลล์คู่หรือไม่ <u>หากใช่</u> จะทำการถ่ายถอดเบรลล์เซลล์คู่ดังกล่าว แล้วนำเบรลล์ 2 เซลล์ต่อไปมาดำเนินการในขั้นที่ 2 ซ้ำ แต่หาก<u>ไม่ใช่</u> จะดำเนินการในขั้นที่ 3 คือถ่ายถอดเบรลล์เซลล์ปัจจุบันเพียงเซลล์เดียว แล้วนำเบรลล์อีก 2 เซลล์มาดำเนินการตั้งแต่ขั้นที่ 2 และขั้นที่ 3 ไปเรื่อย ๆ จนถึง*จุดสิ้นสุดการถ่ายถอด*

ขั้นตอนการถ่ายถอดเบรลล์ส่วนกลางสายอักขระย่อย มีรายละเอียดดังนี้

(1) <u>การถ่ายถอดตัวย่อระดับคำ</u> การถ่ายถอดตัวย่อประเภทนี้จะต้องอาศัยตารางคู่เทียบตัวย่อระดับคำ (ภาคผนวก ข.1) หากระบบพบว่า*สถานะของสายอักขระไม่ใช่เบรลล์รูปเต็ม* และอักษรเบรลล์ตั้งแต่*จุดเริ่มถ่ายถอด*จนถึง*จุดสิ้นสุดการถ่ายถอด*ตรงกับอักษรเบรลล์ในตาราง ระบบจะนำอักษรปกติที่ตรงกันมาใช้เป็นผลลัพธ์แล้วดำเนินการในข้อ (4) จากนั้นจบการทำงานในส่วนการถ่ายถอดส่วนกลางสายอักขระย่อย แต่ถ้าไม่เป็นไปตาม 2 เงื่อนไขข้างต้น ระบบจะดำเนินการในขั้นต่อไป

(2) <u>การถ่ายถอดตัวย่อเบรลล์เซลล์คู่</u> ในขั้นนี้ ระบบจะตรวจสอบเบรลล์ 2 เซลล์ปัจจุบันว่าตรงกับอักษรเบรลล์ในตารางคู่เทียบตัวย่อเบรลล์เซลล์คู่ตามภาคผนวก ข.2 หรือไม่ หากตรงกัน ระบบจะถ่ายถอดเบรลล์ 2 เซลล์ปัจจุบันโดยอาศัยตารางดังกล่าว จากนั้นดำเนินการในขั้น (4) อย่างไรก็ตาม หากเบรลล์ 2 เซลล์ปัจจุบันไม่ตรงกับอักษรเบรลล์ในตาราง ระบบก็จะดำเนินการในขั้นต่อไป

(3) <u>การถ่ายถอดตัวย่อเบรลล์เซลล์เดี่ยว</u> การถ่ายถอดเบรลล์เซลล์เดี่ยวแต่ละกลุ่มต้องอาศัยกฎที่แตกต่างกันตามลำดับคือ

(ก) *การถ่ายถอดตัวย่อต้นคำ* ระบบจะตรวจสอบว่าเบรลล์เซลล์ปัจจุบันเป็นไปตามเงื่อนไขทั้ง 2 ข้อนี้หรือไม่ คือ เบรลล์เซลล์ปัจจุบันอยู่ในตำแหน่งจุดเริ่มต้นการถ่ายถอดหรือไม่ และเบรลล์เซลล์ปัจจุบันปรากฏอยู่ในเซ็ตของตัวย่อต้นคำ หรือไม่ หากเป็นไปตาม 2 เงื่อนไขนี้ ระบบจะถ่ายถอดเบรลล์เซลล์ปัจจุบันโดยอาศัยตารางที่ 5.22 จากนั้นดำเนินการในขั้นที่ (4) แต่ถ้าไม่ตรงกับเงื่อนไขเหล่านี้ระบบจะดำเนินการในขั้นการถ่ายถอดตัวย่อกลางคำ

(ข) *การถ่ายถอดตัวย่อกลางคำ* หากตำแหน่งเบรลล์เซลล์ปัจจุบันไม่ใช่ตำแหน่งจุดเริ่มต้นการถ่ายถอด และไม่ใช่จุดสิ้นสุดการถ่ายถอด และเบรลล์เซลล์ปัจจุบันอยู่ในเซ็ตของอักษรในตารางคู่เทียบตัวย่อกลางคำ (ตารางที่ 5.18) ระบบจะแทนที่อักษรเบรลล์ด้วยอักษรปกติที่ตรงกัน จากนั้นดำเนินการในขั้นที่ (4) แต่หากไม่ตรงกับเงื่อนไขที่กล่าวมา ระบบจะดำเนินการในขั้นการถ่ายถอดตัวย่อรูปเต็มคงรูป

(ค) *การถ่ายถอดตัวย่อรูปเต็มคงรูป* ระบบจะนำเบรลล์เซลล์ปัจจุบันไปตรวจสอบกับอักษรเบรลล์ในตารางคู่เทียบตัวย่อรูปเต็มคงรูป (ตารางที่ 5.21) หากพบว่าตรงกัน ระบบจะตรวจสอบบริบทซ้ายและบริบทขวาของเบรลล์เซลล์ปัจจุบัน แล้วดำเนินการดังนี้

หากบริบทซ้ายเป็นเครื่องหมายเว้นวรรค หรือเบรลล์เซลล์ปัจจุบันเป็นตำแหน่งแรกของส่วนกลางสายอักขระ (นั่นคือไม่มีบริบทซ้าย) และบริบทขวาในตำแหน่งแรก<u>มี</u>เครื่องหมายตัวพิมพ์ใหญ่เดี่ยวหรือตัวพิมพ์ใหญ่คู่ หรือไม่มีอักขระใด ๆ ปรากฏในตำแหน่งนี้ และตำแหน่งต่อมามี [‘⠯’(=and) ‘⠿’ (=for) ‘⠷’(=of) ‘⠾’ (=with)] (อักขระเซ็ตที่ 1) หรือ [‘⠁’ (= a) ‘⠮’ (= the)] (อักขระเซ็ตที่ 2) โดยที่อักขระใน 2 ตัวนี้(คืออักขระเซ็ตที่ 2) ต้องอยู่ในตำแหน่งท้ายสุดของส่วนกลางสายอักขระ ระบบจะถ่ายถอดเบรลล์เซลล์ปัจจุบันตามตารางที่ 5.21 แล้วเติมเว้นวรรคหลังอักษรที่ถ่ายถอดแล้ว จากนั้นจะดำเนินการในขั้นที่ (4) อย่างไรก็ตามหากเบรลล์เซลล์ปัจจุบันไม่ใช่ตัวย่อรูปเต็มคงรูป ระบบจะนำเบรลล์เซลล์ปัจจุบันไปดำเนินการในขั้น (ง) ต่อไป

(ง) *การลบเครื่องหมายประกอบคำออกจากสายอักขระย่อย* หากระบบพบเครื่องหมายประกอบคำที่ระบุไว้ในหัวข้อ 5.2.3.3.1(1) ระบบจะดำเนินการเช่นเดียวกับเงื่อนไขในข้อที่กล่าวมานี้ แล้วนำเบรลล์ 2 เซลล์ปัจจุบันไปดำเนินการตั้งแต่การถ่ายถอดเบรลล์เซลล์คู่ในขั้น (2) แต่ถ้าไม่ตรงกับเงื่อนไขที่ระบุไว้ระบบจะดำเนินการในขั้น (จ) ต่อไป

(จ) *การถ่ายถอดตัวย่อเบรลล์เซลล์เดี่ยวทั่วไป* ในการถ่ายถอดตัวย่อกลุ่มนี้ ระบบจะตรวจสอบว่าเบรลล์เซลล์ปัจจุบันอยู่ในเซ็ตของตัวย่อเบรลล์เซลล์เดี่ยวทั่วไปหรือไม่ (ตามตารางที่ 5.3) หากอยู่ในเซ็ตดังกล่าวระบบจะทำการแทนที่เบรลล์ปัจจุบันด้วยอักษรปกติที่ตรงกัน แล้วนำผลลัพธ์ที่ได้ไปดำเนินการในกระบวนการข้อ (4) จากนั้นนำเบรลล์ 2 เซลล์ต่อไปมาดำเนินการในขั้นถ่ายถอดตัวย่อเบรลล์เซลล์คู่ในข้อ (2) ต่อไป

(4) <u>การแปลงอักษรปกติให้เป็นอักษรตัวพิมพ์ใหญ่</u> ระบบจะนำอักษรปกติที่ถ่ายถอดแล้วมาแปลงให้เป็นอักษรตัวพิมพ์ใหญ่โดยอาศัยเงื่อนไขต่อไปนี้

(ก) หากสถานะของตัวอักษรเป็นตัวพิมพ์ใหญ่ทั้งคำ และตำแหน่งจุดเริ่มต้นตัวพิมพ์ใหญ่มีค่าน้อยกว่าหรือเท่ากับค่าของเบรลล์เซลล์ปัจจุบัน ระบบจะตรวจสอบเงื่อนไขเพิ่มเติม คือ

(ก.1) ถ้าตำแหน่งหน้าเบรลล์เซลล์ปัจจุบันเป็นเครื่องหมายเว้นวรรคระบบจะกำหนดให้ค่าจุดเริ่มต้นตัวพิมพ์ใหญ่มีค่าเท่ากับตำแหน่งเบรลล์เซลล์ปัจจุบัน

(ก.2) ในกรณีอื่น ระบบจะแปลงอักขระเบรลล์ตั้งแต่จุดเริ่มต้นตัวพิมพ์ใหญ่จนถึงเบรลล์เซลล์ปัจจุบันให้เป็นตัวพิมพ์ใหญ่ทั้งหมด หลังจากดำเนินการตามเงื่อนไขใดเงื่อนไขหนึ่งแล้ว ระบบจะนำเบรลล์ 2 เซลล์ถัดไปมาดำเนินการถ่ายถอดเบรลล์เซลล์คู่ในขั้นที่ (2) แต่หากสถานะของตัวอักษรไม่ใช่ตัวพิมพ์ใหญ่ ระบบจะดำเนินการในขั้นต่อไป

(ข) หากสถานะของตัวอักษรเป็นตัวพิมพ์ใหญ่ตัวแรก และตำแหน่งจุดเริ่มต้นตัวพิมพ์ใหญ่มีค่าน้อยกว่าหรือเท่ากับค่าของตำแหน่งเบรลล์เซลล์ปัจจุบัน ระบบจะแปลงอักขระเบรลล์ ณ ตำแหน่งจุดเริ่มต้นตัวพิมพ์ใหญ่ให้เป็นอักษรตัวพิมพ์ใหญ่เพียงตัวเดียว จากนั้นกำหนดสถานะของตัวอักษรเป็นตัวพิมพ์เล็ก แล้วนำเบรลล์ 2 เซลล์ถัดไปมาดำเนินการตั้งแต่ขั้นการถ่ายถอดตัวย่อเบรลล์เซลล์คู่ ในขั้นที่ (2) ต่อไป

5.3 การทดสอบระบบการถ่ายถอดอักษรโดยใช้กฎ

ก่อนที่จะทำการทดสอบระบบ ผู้วิจัยได้สร้างต้นฉบับที่เป็นอักษรปกติเพื่อนำมาใช้ตรวจสอบความถูกต้องของการทดสอบระบบ ซึ่งต้นฉบับเหล่านี้ประกอบข้อความ 3 ประเภท คือข้อความไทย ข้อความอังกฤษ และข้อความไทยปนอังกฤษ[31]

โดยในการสร้างต้นฉบับนั้น ได้นำแฟ้มข้อมูลเบรลล์ไทย แฟ้มข้อมูลเบรลล์อังกฤษรูปย่อ และแฟ้มข้อมูลเบรลล์ไทยปนอังกฤษรูปเต็ม ที่ได้จากศูนย์เทคโนโลยีการศึกษาเพื่อคนตาบอด มาถ่ายถอดให้เป็นอักษรปกติด้วยระบบการถ่ายถอดอักษรที่ผู้วิจัยสร้างขึ้น จากนั้นทำการตรวจสอบความถูกต้องในฉบับอักษรปกติแล้วแก้ไขให้สมบูรณ์

ในการทดสอบระบบ ผู้วิจัยได้นำข้อความเบรลล์ 5 ประเภทดังที่ได้ระบุไว้ในหัวข้อที่ 5.1.2 มาทำการทดสอบกับกระบวนการต่าง ๆ ได้แก่ กระบวนการระบุภาษา กระบวนการถ่ายถอดเบรลล์ไทย และกระบวนการถ่ายถอดเบรลล์อังกฤษ โดยเริ่มจากการนำข้อความเบรลล์นำเข้า มาแบ่งออกเป็นย่อหน้า จากนั้นนำแต่ละย่อหน้ามาแบ่งออกเป็นสายอักขระ แล้วนำแต่ละสายอักขระไปทำการทดสอบกับกระบวนการต่าง ๆ หลังจากนั้นทำการประเมินประสิทธิภาพของระบบ โดยประเมินจากความถูกต้องของแต่ละสายอักขระที่ผ่านกระบวนการเหล่านี้

ในการทำการทดลอง ผู้วิจัยเริ่มจากการนำข้อความเบรลล์ที่จัดเตรียมไว้ 5 ประเภท มาทดสอบกับทุกระบบ โดยวิธีการทดสอบจะแยกเป็นสองส่วน ส่วนแรก เป็นการทดสอบประสิทธิภาพของกระบวนการระบุภาษากับกระบวนการถ่ายถอดอักษรโดยใช้ข้อความเบรลล์ที่ไม่ปนภาษา ส่วนที่สองเป็นการทดสอบประสิทธิภาพของระบบกับข้อมูลจริงที่มีการปนภาษา ดังนี้

---

[31] เหตุที่มีข้อความเพียง 3 ประเภท เพราะ ข้อความเบรลล์อังกฤษรูปเต็มกับข้อความเบรลล์รูปย่อเป็นข้อความเดียวกัน และข้อความเบรลล์ไทยปนอังกฤษรูปย่อกับข้อความเบรลล์ไทยปนอังกฤษรูปเต็มก็เป็นข้อความเดียวกันเช่นกัน จึงใช้ต้นฉบับเดียวกันในการตรวจสอบข้อผิดที่เกิดจากการถ่ายถอดอักษร

ตารางที่ 5.25 การทดสอบประสิทธิภาพของกระบวนการต่าง ๆ ในข้อมูลที่ไม่มีการปนภาษา

ประเภทของการทดสอบ	**ข้อมูลนำเข้า**	**โมดูลที่ใช้**	**วิธีการวัดผล**
การระบุภาษา	ไทยล้วน	โมดูลระบุภาษาไทยปนอังกฤษรูปเต็ม	นับจำนวนสายอักขระไทยที่ระบุภาษาผิด
	ไทยล้วน	โมดูลระบุภาษา ไทยปนอังกฤษรูปย่อ	นับจำนวนสายอักขระไทยที่ระบุภาษาผิด
	อังกฤษรูปเต็ม	โมดูลระบุภาษาไทยปนอังกฤษรูปเต็ม	นับจำนวนสายอักขระอังกฤษที่ระบุภาษาผิด
	อังกฤษรูปย่อ	โมดูลระบุภาษา ไทยปนอังกฤษรูปย่อ	นับจำนวนสายอักขระอังกฤษที่ระบุภาษาผิด
การถ่ายถอดอักษร	ไทยล้วน	โมดูลถ่ายถอดเบรลล์ไทย	นับจำนวนคำไทยที่ถ่ายถอดผิด
	อังกฤษรูปเต็ม	โมดูลถ่ายถอดเบรลล์อังกฤษรูปเต็ม	นับจำนวนคำอังกฤษที่ถ่ายถอดผิด
	อังกฤษรูปย่อ	โมดูลถ่ายถอดเบรลล์อังกฤษรูปย่อ	นับจำนวนคำอังกฤษที่ถ่ายถอดผิด

ส่วนที่ 2 เป็นการทดสอบประสิทธิภาพโดยรวมกับข้อความเบรลล์ปนภาษา โดยนำข้อความเบรลล์ไทยปนอังกฤษรูปเต็ม ไปทดสอบกับระบบถ่ายถอดเบรลล์ไทยปนอังกฤษรูปเต็ม และข้อความเบรลล์ไทยปนอังกฤษรูปย่อ ไปทดสอบกับระบบถ่ายถอดเบรลล์ไทยปนอังกฤษรูปย่อ โดยที่จะประเมินผลเป็น 2 ส่วนคือ ประเมินประสิทธิภาพของการระบุภาษา และประเมินประสิทธิภาพของการถ่ายถอดอักษรในแต่ละระบบ

ตารางที่ 5.26 การทดสอบประสิทธิภาพของกระบวนการต่าง ๆ ในข้อมูลปนภาษา

ประเภทของการทดสอบ	**ข้อมูลนำเข้า**	**ระบบที่ทดสอบ**	**วิธีการวัดผล**	
การระบุภาษา	ไทยปนอังกฤษรูปเต็ม	ระบบถ่ายถอดเบรลล์ไทยปนอังกฤษรูปเต็ม	นับจำนวนสายอักขระไทยที่ระบุภาษาผิด	นับจำนวนสายอักขระอังกฤษที่ระบุภาษาผิด
	ไทยปนอังกฤษรูปย่อ	ระบบถ่ายถอดเบรลล์ไทยปนอังกฤษรูปย่อ	นับจำนวนสายอักขระไทยที่ระบุภาษาผิด	นับจำนวนสายอักขระอังกฤษที่ระบุภาษาผิด
การถ่ายถอดอักษร	ไทยปนอังกฤษรูปเต็ม	ระบบถ่ายถอดเบรลล์ไทยปนอังกฤษรูปเต็ม	นับจำนวนคำไทยที่ถ่ายถอดผิด	นับจำนวนคำอังกฤษที่ถ่ายถอดผิด
	ไทยปนอังกฤษรูปย่อ	ระบบถ่ายถอดเบรลล์ไทยปนอังกฤษรูปย่อ	นับจำนวนคำไทยที่ถ่ายถอดผิด	นับจำนวนคำอังกฤษที่ถ่ายถอดผิด

สำหรับการประเมินประสิทธิภาพของกระบวนการระบุภาษานั้น ผู้วิจัยได้ทำการประเมินในระดับสายอักขระ โดยที่นำสายอักขระที่ถ่ายถอดแล้วของข้อความเบรลล์แต่ละประเภท มานับจำนวนสายอักขระที่ถ่ายถอดผิดจากภาษาหนึ่งไปเป็นอีกภาษาหนึ่ง

ส่วนการประเมินประสิทธิภาพของถ่ายถอดอักษร ผู้วิจัยได้ทำการประเมินประสิทธิภาพในระดับคำ โดยได้นับคำที่ถ่ายถอดถูกต้องเฉพาะในสายอักขระที่ระบุภาษาได้ถูกต้องเท่านั้น ในการนับคำผิดนั้น ผู้วิจัยได้ใช้เกณฑ์ต่าง ๆ คือ หากข้อผิดนั้นปรากฏภายในคำจะนับคำผิดเป็น 1 คำ เช่น “กษัตริย์” ถ่ายถอดเป็น *กษัตริย”* แต่หากข้อผิดนั้นปรากฏระหว่างคำ 2 คำ จะนับคำผิดให้เป็น 2 คำ เช่น “กษัตริย์ทรง” ถ่ายถอดเป็น “*กษัตริยธรง*” เป็นต้น

ในการนับข้อผิดที่เกิดขึ้น ผู้วิจัยได้ดำเนินการโดยนำทุกสายอักขระที่ถ่ายถอดแล้วมาตรวจสอบกับต้นฉบับอักษรปกติที่จัดเตรียมไว้ โดยใช้โปรแกรมที่ผู้วิจัยสร้างขึ้น แล้วนำผลลัพธ์ที่ได้จากโปรแกรมมาตรวจสอบความถูกต้องซ้ำอีกครั้งหนึ่ง ในการนับคำผิดนั้นผู้วิจัยได้นับด้วยตัวผู้วิจัยเองโดยใช้หลักเกณฑ์ที่กล่าวมาข้างต้น จากนั้นจึงนำข้อผิดต่าง ๆ มาทำการวิเคราะห์เพื่อจำแนกออกเป็นข้อผิดประเภทต่าง ๆ แล้วรายงานผลการทดสอบต่อไป

5.3.1 การทดสอบระบบกับข้อความเบรลล์ไม่ปนภาษา

(1) โมดูลระบุภาษา

ในการทดสอบโมดูลระบุภาษาไทยปนอังกฤษรูปเต็มและโมดูลระบุภาษาไทยปนอังกฤษรูปย่อนั้น ผู้วิจัยได้ดำเนินการโดยนำแต่ละสายอักขระไปถ่ายถอด แล้วนับจำนวนสายอักขระที่ถ่ายถอดผิดเป็นอีกภาษาหนึ่ง ซึ่งได้ผลลัพธ์ดังนี้

ตารางที่ 5.27 ผลการระบุภาษาในข้อความเบรลล์ไม่ปนภาษา

ระบบที่ใช้	ประเภทข้อความ	สายอักขระที่ใช้ทดสอบ	สายอักขระที่ระบุภาษาผิด	ระบุภาษาได้ถูกต้อง (%)
โมดูลระบุภาษาไทยปนอังกฤษรูปเต็ม	ไทย	12,913	7	99.94
โมดูลระบุภาษาไทยปนอังกฤษรูปย่อ	ไทย	12,913	43	99.66
โมดูลระบุภาษาไทยปนอังกฤษรูปเต็ม	อังกฤษรูปเต็ม	59,072	1	99.99
โมดูลระบุภาษาไทยปนอังกฤษรูปย่อ	อังกฤษรูปย่อ	56,206	1	99.99

(2) โมดูลถ่ายถอดอักษร

ในการทดสอบโมดูลถ่ายถอดอักษร ผู้วิจัยได้นำข้อความเบรลล์ในแต่ละภาษาไปทำการถ่ายถอดแล้วนับจำนวนคำผิด ซึ่งได้ผลการทดสอบดังนี้

ตารางที่ 5.28 ผลการถ่ายถอดอักษรในข้อความเบรลล์ไม่ปนภาษา

โมดูลถ่ายถอดที่ใช้	**ประเภทข้อความ**	**จำนวนคำที่ใช้ทดสอบ**	**จำนวนคำที่ถ่ายถอดผิด**	**ถ่ายถอดได้ถูกต้อง (%)**
โมดูลถ่ายถอดเบรลล์ไทย	ไทยล้วน	19,090	132	99.31
โมดูลถ่ายถอดเบรลล์อังกฤษรูปเต็ม	อังกฤษรูปเต็ม	59,057	0	100
โมดูลถ่ายถอดเบรลล์อังกฤษรูปย่อ	อังกฤษรูปย่อ	59,057	52	99.91

5.3.2 การทดสอบระบบกับข้อความเบรลล์ปนภาษา

ในการทดสอบระบบการถ่ายถอดอักษร ผู้วิจัยได้ทดสอบระบบกับข้อมูลจริง 2 ประเภทคือ เบรลล์ไทยปนอังกฤษรูปเต็ม และเบรลล์ไทยปนอังกฤษรูปย่อ ในการประเมินประสิทธิภาพของระบบจะแบ่งออกเป็น 2 ส่วน ได้แก่ การระบุภาษา และการถ่ายถอดอักษร

โดยในการประเมินประสิทธิภาพของการระบุภาษาจะพิจารณาสายอักขระที่ถ่ายถอดได้ถูกต้องจากจำนวนสายอักขระทั้งหมดของภาษานั้น โดยคิดความถูกต้องเป็นร้อยละ ส่วนในการวัดประสิทธิภาพของการถ่ายถอดอักษรนั้นจะพิจารณาเฉพาะคำที่ถ่ายถอดได้ถูกต้องจากสายอักขระที่ระบุภาษาได้ถูกต้องแล้วเท่านั้น โดยคิดเป็นร้อยละเช่นกัน

ในการระบุภาษาและการถ่ายถอดอักษรด้วยกฎทางภาษา ได้ผลลัพธ์ความถูกต้องซึ่งแสดงได้ดังนี้

(1) ผลการระบุภาษาของระบบ

ตารางที่ 5.29 ผลการระบุภาษาที่ทดสอบกับข้อความเบรลล์ปนภาษา

ระบบ	ประเภทข้อความ	ภาษา	สายอักขระที่ใช้ทดสอบ	สายอักขระที่ระบุภาษาผิด	ระบุถูกต้องแต่ละภาษา(%)	ระบุถูกต้องรวม (%)
ระบบไทยปนอังกฤษรูปเต็ม	ไทยปนอังกฤษรูปเต็ม	ไทย	7,315	148	97.98	98.67
		อังกฤษ	6,814	40	99.41	
ระบบไทยปนอังกฤษรูปย่อ	ไทยปนอังกฤษรูปย่อ	ไทย	7,315	127	98.26	98.17
		อังกฤษ	6,640	129	98.05	

(2) ผลการถ่ายถอดอักษรเฉพาะสายอักขระที่ระบุภาษาได้ถูกต้อง

ในการทดสอบระบบถ่ายถอดอักษรเฉพาะกับคำที่ระบุภาษาได้ถูกต้อง แสดงได้ดังนี้

ตารางที่ 5.30 ผลการถ่ายถอดอักษรเฉพาะสายอักขระที่ระบุภาษาได้ถูกต้อง

ระบบ	ประเภทข้อความเบรลล์	ภาษาที่ต้องการวัดประสิทธิภาพ	จำนวนคำที่ระบุภาษาถูก	จำนวนคำที่ถ่ายถอดผิด	ถ่ายถอดได้ถูกต้อง (%)
ระบบไทยปนอังกฤษรูปเต็ม	ไทยปนอังกฤษรูปเต็ม	ไทย	11,787	97	99.18
		อังกฤษ	6,531	0	100
ระบบไทยปนอังกฤษรูปย่อ	ไทยปนอังกฤษรูปย่อ	ไทย	11,667	97	99.17
		อังกฤษ	6492	110	98.31

(3) ผลการถ่ายถอดอักษรของระบบโดยภาพรวม

อย่างไรก็ตาม จากการทดสอบประสิทธิภาพของระบบโดยการนับจำนวนคำที่ถ่ายถอดไม่ถูกต้อง ผู้วิจัยพบว่าระบบการถ่ายถอดอักษรทั้งสองระบบนี้สามารถถ่ายถอดข้อความที่มีเนื้อหาเดียวกันแต่ใช้อักษรเบรลล์คนละระบบได้มีประสิทธิภาพต่างกัน ซึ่งแสดงได้ดังนี้

ตารางที่ 5.31 ผลการถ่ายถอดอักษรด้วยการใช้กฎโดยภาพรวม

ระบบ	ประเภทข้อความเบรลล์	คำที่ใช้ทดสอบ	คำที่ถ่ายถอดผิด	ถ่ายถอดได้ถูกต้อง (%)
ไทยปนอังกฤษรูปเต็ม	ไทยปนอังกฤษรูปเต็ม	19,412	338	98.26
ไทยปนอังกฤษรูปย่อ	ไทยปนอังกฤษรูปย่อ	19,412	607	96.87

5.4 อภิปรายผลการทดสอบระบบ

จากผลการทดสอบระบบการถ่ายถอดอักษรโดยใช้กฎ ผู้วิจัยจะขออภิปรายในประเด็นต่าง ๆ ดังนี้

5.4.1 การระบุภาษา

จากการทดสอบประสิทธิภาพการระบุภาษาด้วยข้อความที่ไม่มีการปนภาษา (ดังที่แสดงไว้ในตารางที่ 5.27) ผู้วิจัยพบว่า ระบบสามารถระบุภาษาในข้อความเบรลล์อังกฤษรูปเต็มได้ดีที่สุด รองลงมาคือข้อความเบรลล์อังกฤษรูปย่อ และข้อความเบรลล์ไทย ตามลำดับ เหตุที่เป็นเช่นนี้เพราะการใช้เซ็ตของสายอักขระที่เป็นเอกลักษณ์ของแต่ละภาษา จะใช้ได้ผลต่างกันกับข้อความเบรลล์แต่ละประเภท กล่าวคือจะได้ผลดีที่สุดกับการระบุภาษาในข้อความเบรลล์อังกฤษรูปเต็ม

แต่หากเปรียบเทียบข้อความเบรลล์ปนภาษา (ดังที่แสดงไว้ในตารางที่ 5.29) จะพบว่าในข้อความเบรลล์ไทยปนอังกฤษรูปเต็ม ระบบได้ระบุภาษาอังกฤษได้ดีกว่าภาษาไทย ในขณะที่ในข้อความเบรลล์ไทยปนอังกฤษรูปย่อระบบจะระบุภาษาไทยได้ดีกว่าภาษาอังกฤษ

จากการวิเคราะห์ข้อผิดที่เกิดจากการระบุภาษา ผู้วิจัยพบว่า ข้อผิดส่วนหนึ่งเกิดจากการระบุภาษาไทยให้เป็นอังกฤษ ซึ่งเกิดจากสายอักขระเลขหน้าในเอกสารเบรลล์ไทยที่นิยมเขียนเลขหน้าโดยใช้รูปแบบตัวอักษรอังกฤษตามด้วยตัวเลข เช่นหน้า “a2” เป็นผลให้สายอักขระที่ตามหลังเลขหน้า

เหล่านี้ (ที่ไม่สามารถใช้สายอักขระที่เป็นเอกลักษณ์มาระบุภาษาได้) ถูกระบุให้เป็นภาษาอังกฤษโดยอัตโนมัติ ดังจะพบได้ว่าในข้อความเบรลล์ไทยที่ถ่ายถอดด้วยระบบไทยปนอังกฤษรูปย่อ มีการระบุภาษาผิดถึง 43 สายอักขระ

5.4.2 การถ่ายถอดอักษร

จากการนำข้อความแต่ละภาษามาทำการทดสอบระบบ (ตามตารางที่ 5.28) ผู้วิจัยพบว่าข้อความเบรลล์อังกฤษรูปเต็มสามารถถ่ายถอดได้ถูกต้องมากที่สุด (100%) รองลงมาคือข้อความเบรลล์อังกฤษรูปย่อและข้อความเบรลล์ไทย

การทดสอบระบบกับข้อความที่ปนภาษา (ตามตารางที่ 5.29) ผู้วิจัยพบว่า ระบบจะถ่ายถอดข้อความอังกฤษได้ดีกว่าข้อความไทยในข้อความเบรลล์ไทยปนอังกฤษรูปเต็ม แต่จะถ่ายถอดข้อความไทยได้ดีกว่าข้อความอังกฤษในข้อความเบรลล์ไทยปนอังกฤษรูปย่อ เหตุที่เป็นเช่นนี้เพราะข้อความเบรลล์แต่ละประเภทมีความซับซ้อนในการจัดเรียงและความยุ่งยากในการตีความ แตกต่างกันดังจะเห็นได้ว่า ตัวอักษรที่ปรากฏในเบรลล์อังกฤษรูปเต็มสามารถเทียบกับตัวอักษรอังกฤษปกติได้แบบหนึ่งต่อหนึ่งเป็นส่วนใหญ่ อีกทั้งยังมีลำดับการจัดเรียงตัวอักษรที่เหมือนกับอังกฤษปกติ ทำให้ระบบถ่ายถอดข้อความเบรลล์ประเภทนี้ได้ถูกต้องมากที่สุด

ข้อผิดต่าง ๆ ที่เกิดจากการถ่ายถอดในแต่ละภาษามีรายละเอียดดังนี้

5.4.2.1 ข้อผิดที่เกิดจากการถ่ายถอดด้วยโมดูลไทย

เมื่อทำการวิเคราะห์ข้อผิดในการถ่ายถอดด้วยโมดูลไทย ผู้วิจัยพบว่าข้อผิดที่เกิดขึ้นเหล่านี้เกิดจากสาเหตุ 2 ประการ ได้แก่ ความไม่สมบูรณ์ของเซ็ตไตรแกรมท้ายพยางค์ที่ผู้วิจัยได้สกัดมา และ กฎที่ใช้ในถ่ายถอดยังไม่ครอบคลุมทุกกรณี ดังนี้

ก. ความไม่สมบูรณ์ของเซ็ตไตรแกรมท้ายพยางค์

ในการถ่ายถอดด้วยโมดูลไทย ระบบต้องอาศัยเซ็ตไตรแกรมท้ายพยางค์ที่เกิดจากการสกัดตัวอักษรจากคลังข้อความภาษาไทยแห่งชาติเพื่อช่วยตัดสินใจในกรณีที่ระบบพบอักขระเบรลล์ที่ตีความได้มากกว่าหนึ่งกรณี

เมื่อผู้วิจัยทำการวิเคราะห์ผลลัพธ์ที่ได้ พบว่า เซ็ตไตรแกรมท้ายพยางค์ที่สกัดได้นั้น (1) ขาดสายอักขระที่จำเป็น และ (2) มีสายอักขระที่เกินมา จึงทำให้ระบบตัดสินใจผิด (ข้อผิดที่เกิดขึ้นได้แสดงไว้ในภาคผนวก ฌ)

(1) <u>การขาดสายอักขระที่จำเป็นในเซ็ตไตรแกรมท้ายพยางค์</u>

การที่ระบบขาดสายอักขระที่จำเป็นในเซ็ตไตรแกรมท้ายพยางค์ ทำให้เกิดข้อผิดในการถ่ายถอดข้อความเบรลล์ไทยและไทยปนอังกฤษ ขึ้น ซึ่งแสดงได้ดังนี้

(*ก*) *การถ่ายถอด “ ์ ” เป็น “ ” (คำพูดปิด)*

ระบบพบข้อผิดกับการถ่ายถอด ‘⠒’ มากที่สุด เนื่องจากระบบมักถ่ายถอดอักขระนี้เป็น “ ” ซึ่งพบถึง 97 แห่ง เหตุที่เกิดข้อผิดเป็นจำนวนมากเช่นนี้ เนื่องจากเซ็ตไตรแกรมท้ายพยางค์ มีตัวสะกดการันต์ที่สกัดได้อยู่เพียง <u>88</u> ชุด ซึ่งไม่เพียงพอต่อการถ่ายถอด ทำให้ระบบตัดสินใจผิดเมื่อไม่พบตัวสะกดการันต์อื่น ๆ ในเซ็ตดังกล่าว ตัวอย่างเช่น ‘⠓⠝⠎⠒’ (ห-ง-ส-<u>์ *หรือ ”*</u>) ได้ผลลัพธ์เป็น หงส” ซึ่งในการถ่ายถอด ‘⠓⠝⠎⠒’ (=หงส์) เมื่อระบบถ่ายถอดมาถึง ‘⠒’ (=์) เป็นผลให้ได้สายอักขระ ‘หงส⠒’ ระบบจะนำ ‘⠒’ ไปถ่ายถอดเป็น ์ จากนั้นนำบริบททางซ้ายอีกสองตัวนั่นคือ ‘งส⠒’ ไปตรวจสอบกับเซ็ตไตรแกรมท้ายพยางค์ แต่เนื่องจากในเซ็ตนี้ไม่มี “งส์” ระบบจึงจำเป็นต้องเลือกถ่ายถอด ‘⠒’ เป็น ” (คำพูดปิด) ทำให้ได้ผลลัพธ์ที่ไม่ตรงกับความเป็นจริง

(*ข*) *การถ่ายถอด “์ท” ผิดเป็น “ ธ ”*

ระบบพบข้อผิดที่ระบบถ่ายถอด“์ท” ผิดเป็น “ ธ ” 12 แห่ง โดยพบในเบรลล์ไทยล้วน 1 แห่ง และในเบรลล์ไทยปนอังกฤษ 11 แห่ง ข้อผิดนี้เป็นผลมาจากในเซ็ตไตรแกรมท้ายพยางค์ ไม่มีกลุ่มอักษรที่ลงท้ายด้วยการันต์เพียงพอที่จะนำมาตรวจสอบได้เช่นเดียวกันทำให้ระบบตัดสินใจผิด ตัวอย่างเช่น ‘⠎⠧⠊⠔⠞⠺⠊⠙⠞⠒’ (ส-ว- ิ-ต-ช- <u>ธ หรือ ์ท</u> – ี –) ได้ผลลัพธ์เป็น “สวิตชธี่” ในการถ่ายถอด ‘⠎⠧⠊⠔⠞⠺⠊⠙⠒’ (=สวิตช์ที่) เมื่อดำเนินการมาถึง ‘⠞⠺’ ระบบจะถ่ายถอดเบรลล์คู่นี้ให้เป็น “์ท” แล้วนำอักขระตัวหน้าที่ถ่ายถอดแล้ว (คือ “์”) พร้อมบริบททางซ้ายอีกสองตัว (นั่นคือ “ตช์”) ไปตรวจสอบกับเซ็ตไตรแกรมท้ายพยางค์ แต่เนื่องจากในเซ็ตนี้ไม่มี “ตช์” ทำให้ระบบต้องเลือกถ่ายถอด ‘⠞⠺’ เป็น “ธ” แทนที่จะเป็น “ ์ท”

(ค) *การถ่ายถอด “เาะ” ผิดเป็น “อะ”*

‘⠷⠁’ สามารถถ่ายถอดเป็น **“เาะ”** หรือ **“อะ”** ก็ได้ จากการทดสอบระบบ เกิดข้อผิดจากการถ่ายถอด‘⠷⠁’ เป็น “เาะ” 1 แห่ง ซึ่งเกิดขึ้นกับคำว่า “กรดอะมิโน” ที่ถ่ายถอดเป็น “กรเดาะมิโน” เหตุที่เกิดการถ่ายถอดคำนี้ผิด เพราะระบบไม่พบ ‘กรด’ ในเซ็ตไตรแกรมท้ายพยางค์

(ง) *การถ่ายถอดพยัญชนะ 2 ตัว ที่ปรากฏหน้าสระประสมผิด*

ข้อผิดในการถ่ายถอดพยัญชนะ 2 ตัวที่ปรากฏต่างพยางค์กันเป็นพยัญชนะควบ ได้พบ 1 แห่ง คือ “เจียงไคเชกเริ่ม”ถ่ายถอดเป็น“เจียงไคเชเกริ่ม” การเกิดข้อผิดนี้เนื่องจากเมื่อดำเนินการมาถึง

“กร”ซึ่งเป็นพยัญชนะคู่ที่เป็นปัญหา และเมื่อนำ “เชก” ไปตรวจกับเซ็ตไตรแกรมท้ายพยางค์พบว่าไม่ตรงกับอักษรในเซ็ต ระบบจึงตีความให้“กร”เป็นพยัญชนะควบทำให้เกิดข้อผิดดังกล่าวขึ้น

(2) <u>การปรากฏของสายอักขระเกินในเซ็ตไตรแกรมท้ายพยางค์</u>

นอกจากเซ็ตไตรแกรมท้ายพยางค์ จะขาดกลุ่มตัวอักษรที่จำเป็นต่อการถ่ายถอดแล้ว ยังมีกลุ่มตัวอักษรที่ปรากฏในเซ็ตไตรแกรมท้ายพยางค์บางชุดทำให้ถ่ายถอดอักษรได้ไม่ถูกต้อง ดังนี้

(*ก*) *การถ่ายถอดพยัญชนะควบที่ปรากฏกับสระประสมไม่ถูกต้อง*

ผลการทดสอบระบบแสดงให้เห็นว่า เซ็ตไตรแกรมท้ายพยางค์มีบางสายอักขระที่เป็นปัญหาต่อการถ่ายถอดพยัญชนะควบที่ปรากฏกับสระประสม ทำให้เกิดข้อผิดในลักษณะนี้ 59 คำ (ดังแสดงไว้ในภาคผนวก ฌ.2) เช่น ระบบได้ถ่ายถอด “การเคลื่อนย้าย” เป็น “การคเลื่อนย้าย” เนื่องจากระบบพบสายอักขระ “ารค” ในเซ็ตไตรแกรมท้ายพยางค์ จึงทำให้ถ่ายถอดคำนี้ไม่ถูกต้อง

(*ข*)*การถ่ายถอด “ ◌ ” ผิดเป็น “◌์”*

ระบบได้ถ่ายถอด “ ◌ ” ผิดเป็น “◌์” เป็นจำนวน 12 แห่ง เช่นระบบถ่ายถอด “ชุมชน” เป็น “ชุมชน์” เนื่องจากระบบพบ “ชน์” ในเซ็ตไตรแกรมท้ายพยางค์จึงทำให้ต้องตีความ ‘◌◌’ เป็น “◌์”

(ค) *การถ่ายถอด “ ธ ” ผิดเป็น “ ◌์ท ”*

ในเซ็ตไตรแกรมท้ายพยางค์ ยังมีบางสายอักขระที่ทำให้ถ่ายถอด ‘◌◌◌’ เป็น “◌์ท”ในตำแหน่งที่ควรจะตีความเป็น “ธ” จากการทดสอบระบบ พบข้อผิดลักษณะนี้ในข้อความเบรลล์ไทย 2 แห่ง และในข้อความไทยปนอังกฤษ 3 แห่ง ดังจะเห็นได้ว่าระบบได้ถ่ายถอด “แล้วเธอ” เป็น “แล้ว์เทอ” เนื่องจากระบบพบสายอักขระ “◌้ว์” เกินมาในเซ็ตดังกล่าว

(ง) *การถ่ายถอด “เ-าะ” ผิดเป็น “ อะ ”*

ระบบมีการถ่ายถอด “เ-าะ” ผิดเป็น “ อะ ” 2 แห่งในคำที่ซ้ำกัน โดยที่ระบบได้ถ่ายถอด “มาเกาะ” ผิดเป็น “<u>มาก</u>อะ” เนื่องจากระบบพบสายอักขระ “มาก”ในเซ็ตไตรแกรมท้ายพยางค์ จึงทำให้เกิดข้อผิดดังกล่าวขึ้น

(จ) *การถ่ายถอด “ . ” ผิดเป็น " ◌้ "*

ระบบได้ถ่ายถอด ‘◌◌’ เป็น “◌้” แทนที่จะเป็น “.” เนื่องจากระบบพบสายอักขระ“น้ม” ปรากฏในเซ็ตไตรแกรมท้ายพยางค์ ซึ่งพบเพียง 1 แห่ง คือ “น.ม.ส.” ผิดเป็น “น้ม.ส.”

(ฉ) *การถ่ายถอด “) ” เป็น “ ◌้ ”*

นอกจากข้อผิดที่ได้กล่าวมาข้างต้นแล้ว ยังเกิดข้อผิดกับการถ่ายถอด “)” ผิดเป็น “◌้” ในตำแหน่งท้ายสายอักขระ 1 แห่ง คิดเป็น 1 คำ คือถ่ายถอด “สีดำ)” ผิดเป็น “สีด◌้ำ” เนื่องจากระบบพบสายอักขระ “◌้ำ” ในเซ็ตไตรแกรมท้ายพยางค์นั่นเอง

ข. กฎที่ใช้ถ่ายถอดอักษรรองรับไม่ครบทุกกรณี

จากการทดสอบระบบพบว่ากฎที่สร้างขึ้นบางส่วนยังไม่ครบถ้วนสมบูรณ์ ซึ่งจะขออภิปรายดังนี้

(1) <u>กฎในการถ่ายถอด ‘◌ิ’ ไม่ครบถ้วน</u>

(ก) *การถ่ายถอด “เ◌ิ” เป็น “เ-อ”*

ผู้วิจัยพบว่า ระบบได้ถ่ายถอด “เ◌ิ” ผิดเป็น “เ-อ” 8 แห่ง ซึ่งเกิดขึ้นกับคำทับศัพท์ภาษาต่างประเทศ ได้แก่ “โยเกิร์ต” “เฟิร์น” “เมกะเฮิรตซ์” “เสื้อเชิ้ต” ซึ่งแสดงไว้ในภาคผนวก ฌ.3 เนื่องจากระบบไม่พบตัวสะกดบางรูปในเซ็ตพยัญชนะตัวสะกดท้ายพยางค์

(ข)*การถ่ายถอด “เ-อ” เป็น “เ◌ิ”*

พยัญชนะ 2 ตัวในภาษาไทยคือ “ว” และ “อ” สามารถใช้เป็นสระได้ เช่น “ดวง” “ดอง” อย่างไรก็ตาม เนื่องจากในการพัฒนาระบบ ผู้วิจัยได้จัดให้อักษร 2 ตัวนี้เป็นพยัญชนะเสมอ จึงทำให้เกิดข้อผิดในการถ่ายถอด‘◌ิ’ (= สระเ-อ) ผิดเป็น “เ-” ซึ่งในการทดสอบระบบ ผู้วิจัยได้พบจุดที่ผิด 2 คำ คือ “อำเภิดอยสะเก็ด” และ “อำเภิมวกเหล็ก” ตามลำดับ

(2) <u>กฎในการถ่ายถอด ‘◌◌ ◌◌’ ไม่ครบถ้วน</u>

ในเบรลล์ไทย ‘◌◌ ◌◌’ อาจใช้แทน “-ส” หรือ “ษ” ได้ จากการทดสอบระบบ ผู้วิจัยพบว่า กฎที่ใช้ในการถ่ายถอดอักขระเบรลล์คู่นี้ยังไม่ครอบคลุม เนื่องจากยังมีข้อผิดถึง 6 ตำแหน่ง (ดังแสดงในภาคผนวก ฌ.3)

การถ่ายถอด “ษ” ผิดเป็น “-ส” นั้นเกิดจากข้อความในต้นฉบับพิมพ์คำว่า “ภาษา” แยกจากกันคนละบรรทัดจึงทำให้ระบบถ่ายถอด “ษา” เป็น “-สา” เนื่องจากระบบพบว่าตรงกับกฎที่ผู้วิจัยได้กำหนดไว้คือเมื่อพบ ‘◌◌ ◌◌’ ต้นบรรทัดต้องถ่ายถอดเป็น“-ส” ในทำนองกลับกัน ผู้วิจัยยังพบว่า มีการถ่ายถอด “-ส” ผิดเป็น“ษ”กับข้อความ“*เย็นทั้งกายษุขทั้งใจ*” และเกิดกับการแสดงคำอ่านภาษาไทยที่คั่นด้วย “-” จึงได้ผลลัพธ์ของ (พะ-สก-กะ-นิ-กอน) ผิดเป็น “(*พะษก-กะ-นิ-กอน*)” เนื่องจากระบบไม่สามารถถ่ายถอด ‘◌◌ ◌◌’ ในตำแหน่งกลางสายอักขระให้เป็น “-ส” เพราะไม่มีกฎมารองรับนั่นเอง

5.4.2.2 ข้อผิดที่เกิดจากการถ่ายถอดด้วยโมดูลอังกฤษ

ในการทดลองถ่ายถอดข้อความเบรลล์ทั้งหมด ผู้วิจัยพบข้อผิดในการถ่ายถอดข้อความในส่วนที่ระบบระบุให้เป็นภาษาอังกฤษ ดังนี้

(1) <u>การถ่ายถอด ‘⠤’ ในตำแหน่งต้นคำไม่ถูกต้อง</u>

จากการทดสอบระบบ ผู้วิจัยพบว่ามีการถ่ายถอด “-” ในตำแหน่งต้นคำผิดเป็น “com” เป็นจำนวนมาก ซึ่งเกิดเฉพาะกับข้อความอังกฤษรูปย่อเท่านั้น ตามที่แสดงไว้ในภาคผนวก ฌ.4 ข้อผิดเหล่านี้เกิดขึ้นกับหน่วยคำภาษาอังกฤษประเภทปัจจัย (suffix) ทั้งสิ้น เนื่องจากข้อความที่นำมาใช้ทดสอบระบบเป็นตำราไวยากรณ์ภาษาอังกฤษ จึงทำให้มีคำประเภทนี้ปรากฏเป็นจำนวนมาก แม้กระนั้นก็ตามในการใช้อักษรเบรลล์ในสถานการณ์ทั่วไปคงไม่พบคำประเภทนี้บ่อยนัก

(2) <u>การถ่ายถอด “/” กลางสายอักขระเป็น “st”</u>

การถ่ายถอด ‘⠸’ ที่ปรากฏกลางคำในเบรลล์อังกฤษยังเกิดปัญหาต่อระบบที่ผู้วิจัยสร้างขึ้น ผู้วิจัยพบว่า มีข้อผิดในการถ่ายถอด “/” เป็น “st” จำนวนทั้งสิ้น 10 แห่ง กับคำที่เกิดซ้ำ ๆ กัน เช่นระบบได้ถ่ายถอด ‘⠉⠕⠕⠅⠬⠁⠞⠬’ (= c-o-o-k-ing-<u>/</u>-e-a-t-ing) เป็น “cookingsteating”

(3) <u>การตีความตัวอักษรเป็นตัวย่อ</u>

ในเบรลล์อังกฤษรูปย่อเมื่อตัวอักษรภาษาอังกฤษปรากฏระหว่างช่องว่างและ/หรือเครื่องหมาย (คือปรากฏเป็นคำโดด) จะใช้แทนคำโดดเช่น‘⠉’(=c)ใช้แทน“can”หรือ ‘⠑’(=e) ใช้แทน “every” แต่ในการเขียนเลขข้อ (numbering) นั้นสามารถใช้อักษรตามด้วยจุดได้โดยไม่ต้องมีเครื่องหมายกำกับ เป็นผลให้ระบบตีความว่าเลขข้อเป็นตัวย่ออักษร โดด ดังเช่นเลขข้อที่แสดงด้วยตัวอักษร “c.” “d.” “e.” ซึ่งระบบถ่ายถอดเป็น “can.” “do.” “every.” นอกจากนี้ ในข้อความที่นำมาทดสอบนั้นยังมีเลขโรมันปรากฏอยู่ โดยที่ตัวเลขโรมันเหล่านี้ไม่มีการเขียนเครื่องหมายกำกับตัวอักษร(‘⠼’)ไว้ข้างหน้าทำให้ได้ผลลัพธ์ที่ผิดเป็น“Article Very”, “Amendment It” แทนที่จะเป็น“Article V”, “Amendment X” ซึ่งเป็นคำที่ถูกต้อง ซึ่งข้อผิดในลักษณะนี้ได้แสดงไว้ในภาคผนวก ฌ.4

(4) <u>การถ่ายถอดคำย่อที่เป็นส่วนหนึ่งของคำไม่ถูกต้อง</u>

จากการทดสอบระบบ พบว่ามีการถ่ายถอดคำย่อไม่ถูกต้องรวมทั้งสิ้น 6 คำ เช่น ระบบถ่ายถอด ‘⠆⠓⠢’ “beforehand” ผิดเป็น “<u>bef</u>hand” ได้ เนื่องจาก “bef” เป็นส่วนหนึ่งของคำ แต่ยังใช้แทนคำย่อ “before” อยู่เช่นเดิม แต่เนื่องจากมีคำอื่นมาต่อท้าย ระบบจึงไม่สามารถระบุ ⠆⠓ ได้ว่าเป็นคำย่อ ทำให้ระบบถ่ายถอดอักษรนี้แบบตัวต่อตัวจึงได้ผลลัพธ์ผิดตามที่แสดงไว้ข้างต้น

ในบทนี้ ผู้วิจัยได้กล่าวถึงระบบการถ่ายถอดเบรลล์ไทยปนอังกฤษด้วยกฎที่ผู้วิจัยได้พัฒนาขึ้น ซึ่งประกอบด้วยกระบวนการต่าง ๆ ที่ต้องทำงานร่วมกัน อีกทั้งได้นำเสนอผลการทดสอบประสิทธิภาพของระบบที่ผู้วิจัยได้พัฒนาขึ้น ซึ่งพบว่ายังเกิดข้อผิดกับการถ่ายถอดอักษรอยู่จำนวนหนึ่ง ดังนั้นในบทต่อไป ผู้วิจัยจะได้ชี้ให้เห็นการถ่ายถอดอักษรอีกวิธีหนึ่งที่แตกต่างจากระบบที่ได้นำเสนอมาแล้ว คือ การถ่ายถอดเบรลล์ไทยปนอังกฤษด้วยแบบจำลองเอ็นแกรม ซึ่งให้ผลลัพธ์ที่แตกต่างจากการถ่ายถอดด้วยกฎ ดังจะได้อภิปรายในรายละเอียดต่อไป

บทที่ 6

ระบบการถ่ายถอดเบรลล์ไทยปนอังกฤษโดยใช้แบบจำลองเอ็นแกรม

ในบทนี้ผู้วิจัยจะได้กล่าวถึงการถ่ายถอดเบรลล์ไทยปนอังกฤษโดยใช้แบบจำลองเอ็นแกรม ในหัวข้อที่ 6.1 จะอภิปรายเกี่ยวกับการจัดทำข้อมูลที่ใช้กับระบบถ่ายถอดอักษร จากนั้นในหัวข้อที่ 6.2 จะได้กล่าวถึง ระบบถ่ายถอดอักษรด้วยแบบจำลองเอ็นแกรมที่ผู้วิจัยได้สร้างขึ้น หัวข้อที่ 6.3 จะนำเสนอผลการทดสอบระบบ และหัวข้อที่ 6.4 จะได้อภิปรายถึงผลการทดสอบระบบ ซึ่งแต่ละหัวข้อมีรายละเอียดดังนี้

6.1 การจัดทำข้อมูลเพื่อใช้กับระบบถ่ายถอดอักษร

ข้อมูลที่ใช้ในระบบการถ่ายถอดเบรลล์ไทยปนอังกฤษให้เป็นอักษรปกติด้วยแบบจำลองเอ็นแกรม ประกอบด้วยข้อมูล 2 ประเภท ประเภทแรกเป็น คู่เทียบอักษรที่ใช้จับคู่อักษรเบรลล์กับอักษรปกติ เพื่อสร้างผลลัพธ์ของการถ่ายถอดที่เป็นไปได้ในภาษาไทยและภาษาอังกฤษ ข้อมูลประเภทที่สองเป็นข้อมูลแสดงความถี่ที่อักษรปกติแต่ละตัวปรากฏร่วมกัน ซึ่งระบบจะนำมาคำนวณหาผลลัพธ์ที่มีค่าทางสถิติสูงสุดเพื่อนำมาใช้เป็นผลลัพธ์ที่แท้จริงของการถ่ายถอดอักษร โดยที่ข้อมูลแต่ละประเภทมีรายละเอียดในการจัดทำดังนี้

6.1.1 การจัดทำข้อมูลสำหรับการถ่ายถอดเบรลล์ไทย

ในการจัดเตรียมข้อมูลเพื่อใช้ในการถ่ายถอดเบรลล์ไทยด้วยแบบจำลองเอ็นแกรมนั้น ผู้วิจัยได้แบ่งข้อมูลออกเป็น 2 ประเภท คือ *ตารางคู่เทียบอักษรไทย* ที่ระบบจะนำมาใช้สร้างชุดสายอักขระสำหรับผลลัพธ์ที่เป็นไปได้ของการถ่ายถอดอักษร และ *คลังไตรแกรมตัวอักษรไทย* ที่ระบบจะนำมาใช้คำนวณหาสายอักขระเพียงหนึ่งเดียวที่มีค่าทางสถิติสูงสุดจากชุดสายอักขระที่กล่าวมาข้างต้น

6.1.1.1 การสร้างตารางคู่เทียบอักษรไทย

ในการถ่ายถอดอักษรโดยใช้วิธีการทางสถิติ ก่อนอื่นจำเป็นต้องแจกแจงผลลัพธ์ของการถ่ายถอดอักษรเบรลล์ให้ครบทุกกรณีเท่าที่จะเป็นไปได้ เนื่องจากเบรลล์หนึ่งตัวหรือหนึ่งชุดอาจตีความในอักษรปกติได้มากกว่าหนึ่งกรณี ดังนั้นผู้วิจัยจึงได้สร้างตารางคู่เทียบอักษรสำหรับการประมวลผลภาษาไทย ซึ่งแต่ละคู่เทียบมีองค์ประกอบคือ (1) อักษรเบรลล์ (ไม่ว่าจะเป็นเบรลล์เซลล์เดี่ยว เบรลล์ 2 เซลล์ หรือเบรลล์ 3 เซลล์ก็ตาม) และ (2) อักษรปกติในทุกกรณีที่เทียบได้กับอักษรเบรลล์แต่ละชุด

ในการจัดทำตารางคู่เทียบอักษร มีรายละเอียดในการดำเนินการดังนี้

ก. เบรลล์เซลล์เดี่ยว

ในส่วนหนึ่งของกระบวนการถ่ายถอดเบรลล์ไทย ระบบต้องดำเนินการโดยนำเบรลล์แต่ละเซลล์ไปจับคู่กับตัวอักษรปกติแต่ละตัวโดยตรง ดังนั้นเบรลล์เซลล์เดี่ยวที่ปรากฏในตารางคู่เทียบจึงต้องประกอบด้วย เบรลล์เซลล์เดี่ยว และคู่เทียบตัวอักษรแบบตัวต่อตัว เช่น ‘⠅’ มีคู่เทียบตัวอักษรเป็น “ก”

อย่างไรก็ตาม มีอักขระเบรลล์บางตัวที่สามารถจับคู่กับอักษรปกติได้หลายกรณี เช่น '⠒' จับคู่กับ “.” หรือ "◌้" ได้ ดังนั้นเพื่อให้ครอบคลุมการถ่ายถอดตัวอักษรกลุ่มนี้ เบรลล์เซลล์เดี่ยวที่ปรากฏในตารางคู่เทียบจึงต้องประกอบด้วยคู่เทียบที่เป็นตัวอักษรปกติมากกว่าหนึ่งตัว เช่น ‘⠦’ มีคู่เทียบเป็น "◌้" และ "."

นอกจากนี้ การจัดเรียงสระประสมในอักษรเบรลล์และอักษรปกติมีความแตกต่างกัน กล่าวคือ ในเบรลล์ไทยใช้เบรลล์เพียงเซลล์เดียวเพื่อแทนสระประสม ทำให้เบรลล์ไทยมีโครงสร้างการเรียงต่อตัวอักษรแบบ พยัญชนะต้น (ซึ่งอาจเป็น 1 ตัวหรือ 2 ตัวแล้วแต่กรณี) + สระประสม ซึ่งแตกต่างจากอักษรปกติ ที่พยัญชนะต้นจะปรากฏระหว่างสระประสม เป็นผลให้ในการถ่ายถอดต้องนำพยัญชนะ 1 - 2 ตัว ที่ปรากฏหน้าสระประสมในอักษรเบรลล์ มาแทรกระหว่างสระประสมในอักษรปกติ เช่น ‘⠎⠪’ (=ส-เ◌ีย) ถ่ายถอดได้เป็น “เสีย” ‘⠏⠇⠪’ (=พ-ล-เ◌ีย) ถ่ายถอดได้เป็น “เพลีย”

เพื่อให้ระบบสามารถนำตัวอักษรสระประสมไปแทรกระหว่างสระประสมได้ ผู้วิจัยจึงกำหนดให้คู่เทียบสำหรับสระประสมต้องมีเครื่องหมาย “|” (vertical bar) และ “||”เพื่อประโยชน์ในการสลับตำแหน่งตัวอักษรเช่น '⠪' (= เ◌ีย) มีคู่เทียบตัวอักษรเป็น "|เ◌ีย" เพื่อบ่งชี้ว่าต้องนำอักษร 1 ตัวมาแทรกระหว่างสระประสม และ "||เ◌ีย" เพื่อกำกับว่าต้องนำอักษร 2 ตัวมาแทรกระหว่างสระประสม โดยที่ระบุจะทำการลบเครื่องหมายเหล่านี้ออกจากคู่เทียบหลังจากทำการสลับตำแหน่งตัวอักษรแล้ว

ข. เบรลล์เซลล์คู่

ในภาษาไทย มีอักษรโดดที่ต้องแทนด้วยเบรลล์ 2 เซลล์อยู่จำนวนหนึ่ง ดังนั้นผู้วิจัยจึงสร้าง คู่เทียบตัวอักษรสำหรับเบรลล์เซลล์คู่ โดยจับคู่แบบเบรลล์ 2 เซลล์ต่อ 1 ตัวอักษร เช่น ‘⠁⠃’ มีคู่เทียบตัวอักษรเป็น “ศ” อย่างไรก็ตาม มีข้อจำกัดบางประการในอักษรเบรลล์ที่ทำให้ไม่สามารถจับคู่เบรลล์กับอักษรปกติในลักษณะที่กล่าวมาข้างต้นได้เสมอไป โดยแบ่งออกเป็นกรณีต่าง ๆ

จากการศึกษาอักษรเบรลล์ผู้วิจัยพบว่า มีเบรลล์เซลล์คู่อยู่ 4 ชุด ที่สามารถจับคู่กับอักษรปกติได้สองลักษณะ คือ (1) เบรลล์แต่ละคู่จับคู่กับอักษรปกติได้ 1 ตัวอักษร และ (2) เบรลล์แต่ละเซลล์ภายในคู่เหล่านี้ จับคู่กับอักษรปกติได้โดยตรงแบบตัวต่อตัว ซึ่งอาจเกิดกรณีใดกรณีหนึ่งขึ้นได้ โดยจะแปรไปตามบริบท ซึ่งเบรลล์เซลล์คู่ 4 ชุดนี้ ได้แก่ ‘⠁⠁’ ‘⠁⠁’ ‘⠁⠁’ และ ‘⠁⠁’

เพื่อให้ถ่ายถอดอักษรได้ครบทุกกรณีผู้วิจัยจึงได้นำเบรลล์เหล่านี้ พร้อมคู่เทียบตัวอักษรทั้ง 2 กรณีไปบรรจุไว้ในตารางคู่เทียบอักษร เช่น '⠁⠁' มีคู่เทียบตัวอักษรเป็น "ธ" และ "์ท"

กรณีต่อมา ในอักษรเบรลล์สระตามด้วยวรรณยุกต์ บางคู่ เมื่อถ่ายถอดเป็นอักษรปกติแล้วจำเป็นต้องทำการสลับตำแหน่งระหว่างสระและวรรณยุกต์ เพื่อให้ถูกต้องตามอักขรวิธีไทย เช่น ‘⠁⠁⠁’ (=น-ำ- ้) ต้องถ่ายถอดเป็น “น้ำ” (= น- ้-ำ) และ ‘⠁⠁⠁’= (บ-เ◌ีย- ้) ต้องถ่ายถอดเป็น “เบี้ย” (=เ-บ - ี-้-ย) ดังนั้นเพื่อความสะดวกในการถ่ายถอดเป็นอักษรปกติ ตารางคู่เทียบอักษรสำหรับภาษาไทยจึงต้องมีคู่เทียบตัวอักษรรูปแบบนี้ปรากฏอยู่ เช่น '⠁⠁' (= ำ้) มีคู่เทียบเป็น "้ำ", ‘⠁⠁’มีคู่เทียบเป็น “เ◌ี◌้ย” และ “เ◌ี◌้ย” เป็นต้น

ค. เบรลล์ 3 เซลล์

ดังที่ได้อภิปรายในหัวข้อที่ผ่านมาว่า เบรลล์ไทยและไทยปกติมีการจัดเรียงตัวอักษรที่แตกต่างกัน ผู้วิจัยพบว่าความสอดคล้องในการจัดเรียงสระและวรรณยุกต์ของอักษรเบรลล์และอักษรปกติสามารถเกิดขึ้นในลักษณะเดียวกันกับเบรลล์ 3 เซลล์ได้ โดยเฉพาะอย่างยิ่งเมื่ออักษรเบรลล์มีการจัดเรียงแบบ สระ + “ะ” + วรรณยุกต์ เช่น “⠁⠁⠁⠁” (= จ-เ◌ีย-ะ - ่) ซึ่งในอักษรปกติจัดเรียงเป็น “เจี๋ยะ” (= เ-จ- ี- ๋-ย-ะ)

ในความเป็นจริง การถ่ายถอดอักษรที่เรียงต่อในรูปแบบนี้อาจทำได้โดยนำกฎทางภาษามาใช้ปรับเปลี่ยนตัวอักษรเพื่อให้ได้ผลลัพธ์ที่ถูกต้อง แต่หากนำวิธีการดังกล่าวมาใช้ จะทำให้ไม่สอดคล้องกับวิธีการวิจัยโดยใช้แบบจำลองเอ็นแกรมซึ่งเป็นหัวใจของการวิจัยในส่วนนี้ ดังนั้นผู้วิจัยจึงสร้างคู่เทียบอักษรสำหรับเบรลล์ 3 เซลล์ที่มีการจัดเรียงในลักษณะนี้ขึ้นเพื่อบรรจุลงในตารางคู่เทียบอักษร

เป็นผลให้ตารางนี้มีคู่เทียบตัวอักษรเช่น'⁚⁚ ⁚⁚'(=เ◌ีย-ะ - ◌ั)มีคู่เทียบเป็น "|เ◌ีั◌ยะ" และ "||เ◌ีั◌ยะ" เป็นต้น

6.1.1.2 คลังไตรแกรมตัวอักษรไทย

จากหัวข้อที่ผ่านมาจะพบว่า หากนำตารางคู่เทียบอักษรสำหรับภาษาไทยมาใช้ถอดอักษร จะทำให้ได้ผลลัพธ์ของการถ่ายถอดอักษรหลายรูปแบบ ซึ่งไม่สามารถใช้เป็นผลลัพธ์ได้อย่างแท้จริง เพื่อให้ได้ผลลัพธ์เพียงหนึ่งเดี่ยว จำเป็นต้องคัดเลือกเพียงผลลัพธ์ที่ให้ค่าสถิติสูงสุดเท่านั้น

ในการหาค่าสถิติสูงสุดของผลลัพธ์ผู้วิจัยได้ใช้วิธีการประมวลผลด้วยแบบจำลองเอ็นแกรมของอักขระโดยใช้ข้อมูลไตรแกรม (trigram) เนื่องจากงานวิจัยการประมวลผลทางภาษาหลายชิ้นชี้ให้เห็นว่าการใช้ข้อมูลขนาดนี้ให้ผลลัพธ์ที่มีประสิทธิภาพโดยไม่จำเป็นต้องอาศัยจำนวนแกรมที่มากขึ้น

ในการสร้างแบบจำลองไตรแกรมของอักขระในภาษาไทยนั้น ผู้วิจัยได้ทำการสกัดข้อมูลไตรแกรมจากคลังข้อความภาษาไทยแห่งชาติ (TNC) แต่เมื่อทำการสำรวจคลังข้อความดังกล่าว ผู้วิจัยพบว่ามีภาษาอังกฤษปนในคลังข้อความอยู่จำนวนหนึ่ง อันเป็นผลมาจากข้อความที่นำมาสร้างคลังข้อความนี้ส่วนหนึ่งได้มาจากเอกสารทางวิชาการที่มีศัพท์เทคนิคภาษาอังกฤษปะปนอยู่ ดังนั้นเพื่อให้สามารถประมวลผลได้เฉพาะภาษาไทยและตัวเลข ผู้วิจัยจึงได้สกัดตัวอักษรจากคลังข้อความภาษาไทยแห่งชาติโดยดำเนินการดังนี้

ในขั้นแรก ผู้วิจัยได้นำข้อความทีละย่อหน้ามาแบ่งออกเป็นวรรค แล้วนำแต่ละวรรคมาเติมเครื่องหมายวรรค (space) ข้างหน้าและข้างหลังของแต่ละวรรค จากนั้น เลือกเฉพาะวรรคที่ไม่มีตัวอักษรภาษาอังกฤษตนอยู่ ไม่ว่าจะเป็นตัวพิมพ์เล็กหรือตัวพิมพ์ใหญ่ก็ตาม หลังจากที่ได้ข้อความเป็นวรรคตามหลักเกณฑ์ที่กำหนดไว้แล้ว จึงนำแต่ละวรรคมาสร้างข้อมูลไตรแกรม ทำให้ได้ข้อมูลเบื้องต้น คือ ได้จำนวนไตรแกรมที่ไม่ซ้ำรูปกัน 130,757 ชุด (types) และไตรแกรมที่ปรากฏซ้ำ ๆ จำนวนรวมทั้งสิ้น 131,256,346 ครั้ง (tokens)

เพื่อให้สามารถนำข้อมูลไตรแกรมเหล่านี้ไปใช้กับระบบถ่ายถอดอักษรได้ ผู้วิจัยได้จัดเก็บข้อมูลไตรแกรมในรูปของตาราง โดยที่ช่องแรกเป็นตัวอักษรไตรแกรมที่สกัดได้ และอีกช่องหนึ่งเป็นค่าสถิติที่เกิดจากความถี่ที่พบไตรแกรมนั้นในคลังข้อความภาษาไทยแห่งชาติหารด้วยความถี่ทั้งหมดของทุกไตรแกรมที่สกัดได้จากคลังข้อความนี้

6.1.2 การจัดทำข้อมูลสำหรับการถ่ายถอดเบรลล์อังกฤษ

ในการสร้างชุดสายอักขระสำหรับผลลัพธ์ที่เป็นไปได้สำหรับการถ่ายถอดเบรลล์อังกฤษเป็นอังกฤษปกตินั้น ระบบจะอาศัยข้อมูล 2 ประเภท ได้แก่ (1) ตารางคู่เทียบเบรลล์อังกฤษกับอักษรปกติเพื่อใช้ในการสร้างชุดสายอักขระอังกฤษซึ่งเป็นผลลัพธ์ของการถ่ายถอดในรูปแบบต่าง ๆ เท่าที่จะเป็นไปได้ (2) คลังไตรแกรมตัวอักษรอังกฤษ และคลังไบแกรมคำอังกฤษ เพื่อใช้ในการประมวลผลด้วยแบบจำลองเอ็นแกรม ซึ่งข้อมูลแต่ละประเภทมีรายละเอียดในการจัดทำดังนี้

6.1.2.1 ตารางคู่เทียบเบรลล์อังกฤษ

ในการสร้างตารางคู่เทียบอักษรสำหรับการสร้างชุดสายอักขระอังกฤษ ผู้วิจัยได้ออกแบบคู่เทียบโดยอาศัยความรู้ที่ได้จากบทที่ 2 หัวข้อที่ 2.1.1.2 ซึ่งในเบรลล์อังกฤษรูปย่อนั้น อักษรเบรลล์บางตัวหรือบางชุดสามารถใช้แทนอักษรปกติได้หลายกรณี ฉะนั้นจึงดูเหมือนว่า การสร้างคู่เทียบของอักขระเบรลล์แต่ละตัว ทำได้ไม่ยาก เพียงแค่นำอักขระเบรลล์แต่ละตัวมาแจกแจงคู่เทียบที่เป็นไปได้ทั้งหมดในอักษรปกติ (เช่น ‘⠷’ มีคู่เทียบเป็น “were” “(“ “gg” และ “)” เป็นต้น) แล้วนำไปสร้างผลลัพธ์ที่เป็นไปได้รูปแบบต่าง ๆ ก็จะทำให้ ระบบหาสายอักขระที่มีค่าสถิติความเป็นไปได้สูงสุดมาเป็นผลลัพธ์ของการถ่ายถอดได้อย่างถูกต้อง

อย่างไรก็ตาม ผู้วิจัยได้ทำการทดลองถ่ายถอดเบรลล์อังกฤษรูปย่อ โดยนำตัวย่อเบรลล์แต่ละตัว (หรือแต่ละชุด) ที่เรียงต่อกัน มาจับคู่กับอักษรปกติเท่าที่จะเป็นไปได้ แล้วนำผลลัพธ์ที่ได้ไปหาค่าทางสถิติสูงสุดเพื่อนำมาใช้เป็นผลลัพธ์ของการถ่ายถอดอักษร พบว่า เกิดข้อผิดขึ้นเป็นจำนวนมาก เนื่องจากเมื่อระบบนำคู่เทียบเหล่านี้ไปใช้จับคู่อักษรเบรลล์กับอักษรปกติทำให้ระบบสร้างตัวเลือก (candidate) ขึ้นเป็นทวีคูณ จึงมีโอกาสสูงที่ระบบจะเลือกตัวเลือกที่ไม่ถูกต้องมาใช้เป็นผลลัพธ์ ของการถ่ายถอดอักษร

ดังที่ได้แสดงไว้ในส่วนของการทบทวนวรรณกรรมแล้วว่า ‘⠷’ สามารถใช้แทนอักษรปกติได้หลายกรณี แต่หากอักขระตัวนี้ปรากฏเป็นคำโดดจะใช้แทน “were” หากปรากฏหน้าคำจะใช้แทน “(” หากปรากฏกลางคำใช้แทน “gg” หรือหากปรากฏท้ายคำจะใช้แทน “)” เป็นต้น จึงดูเหมือนว่าทุกครั้งที่พบอักขระเบรลล์ตัวนี้ จะสามารถนำอักษรปกติคือ “(” “)” “gg” และ “were” ไปแทนที่อักขระเบรลล์ตัวนี้ เพื่อสร้างตัวเลือกที่จะนำมาใช้ในการถ่ายถอดอักษรได้ จะเห็นได้ว่าทุกครั้งที่พบ‘⠷’ จะมีตัวเลือกเพิ่มขึ้นอีก 4 กรณี เนื่องจากอักขระตัวนี้สามารถเทียบกับอักษรปกติได้ 4 รูปแบบดังที่ได้กล่าวมาแล้ว

ข้างต้น ดังนั้นหากมีอักขระเบรลล์ตัวนี้ปรากฏในสายอักขระ 3 ตำแหน่ง เช่น '⠨⠎⠥⠛⠛⠑⠎⠞⠊⠕⠝' = (s-u-gg-e-s-tion) ในกรณีนี้จะเกิดตัวเลือกขึ้นมากกว่า 64 ตัวเลือก (= 4*4*4)

เพื่อลดปริมาณตัวเลือกที่สามารถนำมาใช้เป็นผลลัพธ์ ผู้วิจัยจึงแก้ปัญหาโดย สร้างคู่เทียบสำหรับอักขระเบรลล์เฉพาะตำแหน่งขึ้นโดยที่ในการสร้างคู่เทียบลักษณะนี้จำเป็นต้องอาศัยเครื่องหมายเว้นวรรค (space) มาช่วยในการระบุตำแหน่งของอักขระเบรลล์ นั่นคือ

- หากอักขระเบรลล์นั้นมีเครื่องหมายเว้นวรรคที่ข้างหน้าและข้างหลัง แสดงว่าอักขระนั้นปรากฏเป็นคำโดด
- หากมีเครื่องหมายเว้นวรรคอยู่ข้างหน้า แสดงว่าอักขระนั้นปรากฏในตำแหน่งต้นสายอักขระ
- หากมีเครื่องหมายเว้นวรรคปรากฏข้างหลัง แสดงว่าอักขระเบรลล์นั้นปรากฏท้ายสายอักขระ
- หากไม่มีเครื่องหมายเว้นวรรคปรากฏร่วมด้วย แสดงว่าอักขระเบรลล์นั้นปรากฏในตำแหน่งใดก็ได้

ในการสร้างคู่เทียบโดยวิธีการนี้ ผู้วิจัยได้นำตัวย่อแต่ละตัวมาวิเคราะห์ตามตำแหน่งที่ปรากฏใน 4 ลักษณะที่กล่าวมา โดยพิจารณาว่าตัวย่อนั้น ๆ เมื่อปรากฏในแต่ละตำแหน่งจะตีความเป็นกรณีใดได้บ้าง เช่น '⠨⠎' มีคู่เทียบตามตำแหน่งอยู่ 4 ชุด คือคู่เทียบสำหรับ ตำแหน่งต้น ตำแหน่งท้าย ปรากฏเป็นคำโดด และตำแหน่งอื่น ๆ ดังนี้

{1} คู่เทียบสำหรับ '⠨⠎'

ตำแหน่งต้น :	|⠨⠎	=	|(
ตำแหน่งท้าย:	⠨⠎|	=	)|
ปรากฏเป็นคำโดด:	|⠨⠎|	=	|were|
	⠠⠨⠎|	=	Were|
ตำแหน่งอื่น ๆ:	⠨⠎	=	gg , (,)

หมายเหตุ : เพื่อให้เห็นเด่นชัด ในส่วนนี้ ผู้วิจัยจะใช้เครื่องหมาย | แทนเครื่องหมายเว้นวรรค

จากตัวอย่าง ผู้วิจัยได้นำเครื่องหมายเว้นวรรค (ซึ่งในที่นี้จะใช้ | แทนเครื่องหมายเว้นวรรคแต่ละตัว) มาช่วยกำกับตำแหน่งของอักขระเบรลล์ นอกจากนี้ยังได้นำ '⠨' (เครื่องหมายแสดงตัวพิมพ์ใหญ่) และ เครื่องหมายเว้นวรรค มากำกับคำโดด ดังที่ปรากฏในคู่เทียบคำโดด "Were" เนื่องจากเครื่องหมาย '⠨' มักจะเป็นตัวบ่งชี้จุดเริ่มต้นของคำในอักษรเบรลล์ และเครื่องหมายเว้นวรรค สามารถบ่งชี้จุดสิ้นสุดของคำได้

อย่างไรก็ตาม เป็นที่น่าสังเกตว่า '⠶' ในตำแหน่งอื่น ๆ มีคู่เทียบที่เป็นไปได้ 3 กรณีคือ "(" "gg" และ ")" ไม่สามารถกำหนดคู่เทียบนี้ให้เป็นกรณีใดกรณีหนึ่งได้โดยเฉพาะ เนื่องจากทั้ง "(" "gg" และ ")" อาจปรากฏในส่วนใดของคำก็ได้ เช่น '⠿⠊⠇⠑⠦⠎⠴' = f-i-l-e-(-s-)-. ดังนั้นระบบจึงต้องสร้างผลลัพธ์ที่เป็นไปได้สำหรับ '⠶' ให้ครบทุกกรณีเท่าที่จะเป็นไปได้ เพื่อว่าจะได้นำไปใช้เป็นตัวเลือกในการหาผลลัพธ์ที่ดีที่สุดของการถ่ายถอดอักษรในภายหลังได้

ทั้งนี้ เพื่อให้ระบบสามารถนำคู่เทียบอักขระเบรลล์เฉพาะตำแหน่งที่ได้กล่าวมาข้างต้นไปใช้กับข้อความเบรลล์ที่กำลังถ่ายถอดได้ ระบบจะเติมเครื่องหมายเว้นวรรคข้างหน้าและข้างหลังสายอักขระทุกครั้ง ก่อนที่จะทำการถ่ายถอด ซึ่งจะได้อภิปรายในส่วนที่เกี่ยวข้องต่อไป

6.1.2.1.1 ประเภทของคู่เทียบเบรลล์อังกฤษรูปย่อ

ในการสร้างคู่เทียบเบรลล์อังกฤษรูปย่อ ผู้วิจัยได้แบ่งคู่เทียบออกเป็นประเภทต่าง ๆ ดังนี้

(1) คู่เทียบสำหรับตัวย่อทั่วไป

มีตัวย่อบางตัวที่ไม่ว่าจะปรากฏในตำแหน่งใดของสายอักขระก็ตาม ยังคงใช้แทนกลุ่มตัวอักษรเช่นเดิมเสมอ เช่น '⠜' จะใช้แทน "ar" เสมอ ได้แก่ ตัวย่อในกลุ่มต่อไปนี้

- ตัวย่อเบรลล์เซลล์เดี่ยว เช่น ตัวย่อที่ใช้แทน "ar" "er" "gh" "in" ตามภาคผนวก ง
- ตัวอักษรอังกฤษที่ปรากฏในอังกฤษรูปย่อ เช่น '⠃' ใช้เขียน "b" หรือ '⠠⠃' ใช้เขียนตัวอักษร "B"
- ตัวย่อเบรลล์เซลล์คู่ที่นำหน้าด้วย จุด 5 จุด 4-5 หรือจุด 4-5-6 เช่น '⠘⠍' ซึ่งใช้ย่อ "many" เนื่องจากอักขระเบรลล์เหล่านี้ไม่จำเป็นต้องตีความจากตำแหน่งที่ปรากฏ จึงมีคู่เทียบที่ประกอบด้วยอักขระเบรลล์และอักษรปกติที่จับคู่กันแบบหนึ่งต่อหนึ่ง ดังตัวอย่างต่อไปนี้

{2} คู่เทียบสำหรับ '⠻'

⠻ = er

คู่เทียบสำหรับตัวย่อทั่วไป ได้แสดงไว้ในภาคผนวก ฉ.1

(2) คู่เทียบสำหรับคำย่อทั่วไป

สำหรับตัวย่อกลุ่มนี้ เมื่อปรากฏในส่วนใดของคำก็ตาม ยังใช้แทนคำเดิมเสมอ เช่น ‘⠎⠗⠧’ สามารถปรากฏเป็นคำโดดคือ “receive” หรือเป็นส่วนของคำ เช่น ‘⠎⠗⠧⠂’ (= receives) หรือ ‘⠎⠗⠧⠫’ (= received) ได้ ตัวย่อกลุ่มนี้แต่ละชุดจะประกอบด้วยคู่เทียบ 2 คู่ ดังตัวอย่างต่อไปนี้

{3} คู่เทียบสำหรับ ‘⠎⠗⠧’

ก. ⠎⠗⠧ = receive

ข. ⠸⠎⠗⠧ = ⠸receive

ตัวอย่างนี้แสดงให้เห็นว่า จำเป็นต้องมีคู่เทียบ 2 คู่เทียบสำหรับ ‘⠎⠗⠧’ เนื่องจากตัวย่อนี้สามารถปรากฏในตำแหน่งต้นคำตามคู่เทียบข้อ {3} ข. หรือตำแหน่งอื่นตามข้อ (3) ก. แม้ว่าตัวย่อประเภทนี้จะเป็นตัวย่อที่สามารถปรากฏในตำแหน่งใดของคำได้ก็ตาม แต่จำเป็นต้องมีคู่เทียบสำหรับตำแหน่งต้นคำ (นั่นคือมีการเติมเครื่องหมายเว้นวรรคหน้าคู่เทียบ) เนื่องจากในการสร้างผลลัพธ์ที่เป็นไปได้ ระบบจะดำเนินการกับคู่เทียบที่มีเครื่องหมายเว้นวรรคนำหน้าก่อนเป็นอันดับแรก หากระบบไม่พบคู่เทียบที่จำเป็นระบบก็จะดำเนินการกับคู่เทียบอื่น ทำให้ไม่สามารถถ่ายถอดคู่เทียบประเภทนี้ได้ ซึ่งในกรณีนี้หากไม่มีคู่เทียบสำหรับ ‘⠸⠎⠗⠧’ (= space + rcv) ระบบจะดำเนินการกับ ‘⠸⠎’ (=space + r) ทำให้เหลือ ‘⠗⠧’ (= cv) ที่ระบบจะนำไปดำเนินการต่อไป จึงทำให้ไม่สามารถถ่ายถอด ‘⠎⠗⠧’ (= rcv) ซึ่งใช้เป็นคำย่อแทน “receive” ได้ คู่เทียบสำหรับตัวย่อกลุ่มนี้ ได้แสดงไว้ในภาคผนวก ฉ.2

(3) คู่เทียบสำหรับคำย่อหรือตัวอักษร

อักขระเบรลล์กลุ่มนี้ต้องตีความให้เป็นคำย่อเมื่อปรากฏเป็นคำโดด แต่เมื่อปรากฏภายในคำสามารถตีความให้ใช้แทนตัวอักษรหรือใช้แทนคำย่อก็ได้แล้วแต่กรณี ดังจะเห็นได้จาก ‘⠁⠃’ (= ab) เมื่อปรากฏเป็นคำโดดใช้แทน “about” แต่หากปรากฏภายในคำอาจตีความเป็นตัวอักษร “ab” เช่น ‘⠌⠁⠃’ (= st-a-b) หรือ ใช้แทน “about” ที่เป็นส่วนหนึ่งของคำเช่น ‘⠗⠪⠁⠃’ (= r-ound-about) ซึ่งใช้แทน“roundabout”ได้ ดังตัวอย่างข้อ{4}

{4} คู่เทียบสำหรับ ‘⠁⠃’

ก. ⠁⠃ = ab , about

ข. ⠸⠁⠃⠸ = ⠸about⠸

ค. ⠠⠁⠃⠸ = About⠸

คู่เทียบสำหรับตัวย่อกลุ่มนี้ได้แสดงไว้ในภาคผนวก ฉ.3

(4) คู่เทียบที่ตีความต่างกันเมื่อปรากฏเป็นคำโดดหรือเป็นส่วนหนึ่งของคำ

คู่เทียบกลุ่มนี้ประกอบด้วยคู่เทียบ 2 ประเภทคือ (1) คู่เทียบสำหรับตัวย่ออักษรโดด เช่น '⠥' (= u) เมื่อเป็นคำโดดจะใช้แทน "us" แต่เมื่อปรากฏภายในคำจะใช้แทน "u" และ (2) คู่เทียบรูปเต็มกลายรูป เช่น '⠉' เมื่อเป็นคำโดดจะใช้แทน "child" แต่เมื่อปรากฏภายในคำจะใช้แทน "ch" คู่เทียบกลุ่มนี้จะประกอบด้วยคู่เทียบต่าง ๆ ดังตัวอย่างต่อไปนี้

{5} คู่เทียบสำหรับ '⠃'

ก. ⠃ = b

ข. ⠃ | = b|, but|

ค. |⠃ = |b, |but

ง. |⠃| = |but|

จ. ⠀⠃ | = But|

สำหรับอักขระในกลุ่มนี้ อักขระเบรลล์ตัวหนึ่ง ๆ จะประกอบด้วย 5 คู่เทียบ คือคู่เทียบแรกจะแสดงอักขระเบรลล์ที่ปรากฏในส่วนใดของคำก็ได้เช่นตัวอย่างข้อ {5}ก.ซึ่งตีความได้พียงกรณีเดียว คู่เทียบที่ 2 เป็นอักขระเบรลล์ที่ปรากฏในส่วนท้ายของคำ (ดังปรากฏในข้อ {5}ข.) ซึ่งสามารถตีความได้ 2 กรณี คือ ตีความให้เป็นตัวอักษรเมื่อมีอักษรตัวอื่น ๆ ข้างหน้าเช่น '⠏⠥⠃'= pub หรือตีความให้เป็นตัวย่ออักษรโดดในกรณีที่ข้างหน้าเป็นเครื่องหมาย เช่น '⠀⠃' = "but ซึ่งการที่จะได้ผลลัพธ์เป็นกรณีใดนั้น จะต้องอาศัยการประเมินผลด้วยแบบจำลองเอ็นแกรมโดยตรง

คู่เทียบที่ 3 ในตัวอย่างข้อ {5}ค. มีรูปแบบสอดคล้องกับคู่เทียบที่ 2 แต่มีไว้สำหรับอักขระเบรลล์ที่ปรากฏในส่วนต้นของคำ คือหากมีตัวอักษรตามหลังต้องตีความให้อักขระนี้ใช้แทนตัวอักษร แต่หากมีเครื่องหมายตามหลังต้องตีความให้อักขระนี้ใช้แทนตัวย่ออักษรโดด[32] และคู่เทียบที่ 4 - 5 ในตัวอย่างข้อ {5}ง. และ {5}จ. เป็นคู่เทียบ สำหรับอักขระเบรลล์ที่ปรากฏเป็นคำโดด

คู่เทียบสำหรับตัวย่อกลุ่มนี้ ได้แสดงไว้ในภาคผนวก ฉ. 4

(5) คู่เทียบที่เป็นคำย่อเมื่อปรากฏในตำแหน่งท้ายคำหรือปรากฏเป็นคำโดด

อักขระเบรลล์ที่จัดอยู่ในกลุ่มนี้จะตีความให้เป็นคำย่อ เฉพาะเมื่อปรากฏท้ายสายอักขระหรือปรากฏเป็นคำโดด หากปรากฏในตำแหน่งอื่นอาจตีความให้เป็นคำย่อหรือตีความให้เป็นตัวอักษรแบบ

---

[32] ในกรณีที่มีตัวอักษรโดดตามด้วยเครื่องหมาย บางครั้งต้องตีความให้เป็นตัวอักษร ดังเช่นที่ปรากฏกับการเขียนเลขข้อ (numbering) เช่น a. b. c. เป็นต้น

ตัวต่อตัวแล้วแต่บริบท จะพบว่า ‘⠿⠗’ (=fr) เมื่อปรากฏเป็นคำโดดจะใช้แทนคำว่า “friend” และเมื่ออยู่ในตำแหน่งท้ายสายอักขระเช่น ‘⠿⠽⠿⠗’ (=boyfr) จะใช้แทน “boyfriend” แต่เมื่อปรากฏในตำแหน่งอื่นอาจใช้เป็นตัวย่อ (เช่น ‘⠿⠗⠇⠿’ (=frly) ใช้แทน “friendly”) หรือใช้เป็นตัวอักษรแบบตัวต่อตัวได้ (เช่น ‘⠁⠋⠗⠁⠊⠙’(=afraid) ใช้แทน “afraid” ไม่สามารถตีความเป็น “a**friend**aid”) ได้ สำหรับอักขระกลุ่มนี้จะมีคู่เทียบดังตัวอย่างต่อไปนี้

{6} คู่เทียบสำหรับ ⠿⠗

ก. ⠿⠗ = fr , friend

ข. ⠸⠿⠗ = ⠸fr, ⠸friend

ค. ⠿⠗⠸ = friend⠸

คู่เทียบสำหรับตัวย่อกลุ่มนี้ได้แสดงไว้ในภาคผนวก ฉ.5

(6) คู่เทียบที่เป็นคำย่อในตำแหน่งต้นคำหรือปรากฏเป็นคำโดด

อักขระเบรลล์กลุ่มนี้จะเป็นคำย่อเมื่อปรากฏในตำแหน่งต้นคำ หรือปรากฏเป็นคำโดด เช่น ‘⠇⠇’ (=ll) ใช้แทน “little” สำหรับคู่เทียบของอักขระกลุ่มนี้แสดงได้ดังนี้

{7} คู่เทียบสำหรับ ‘⠇⠇’

ก. ⠇⠇ = ll, little

ข. ⠸⠇⠇ = ⠸little

ค. ⠠⠇⠇ = Little

คู่เทียบสำหรับตัวย่อกลุ่มนี้ได้แสดงไว้ในภาคผนวก ฉ.6

(7) คู่เทียบที่เป็นคำย่อในตำแหน่งที่ไม่ใช่ตำแหน่งกลางสายอักขระ

ตัวย่อกลุ่มนี้ต้องตีความให้เป็นคำย่อ เมื่อปรากฏในตำแหน่งต้นคำ ท้ายคำ หรือตำแหน่งคำโดดเท่านั้น แต่เมื่อตัวย่อกลุ่มนี้ปรากฏในตำแหน่งกลางคำไม่สามารถตีความให้เป็นคำย่อได้ เช่น ‘⠡⠝’ (=ch-n) ใช้แทน “children” เมื่อปรากฏเป็นคำโดด และ ‘⠛⠗⠯⠡⠝’ (=g-r-and-ch-n) ใช้แทน “grandchildren” หรือ ‘⠡⠝⠤⠎’ (ch-n-’-s) ใช้แทน “children’s”

แต่ในคำว่า ‘⠞⠑⠡⠝⠊⠟⠥⠑’ (=t-e-ch-n-i-q-u-e) แทน “technique” ไม่สามารถตีความเป็น “*techildrenique*” ได้ สำหรับตัวย่อในกลุ่มนี้มีคู่เทียบดังตัวอย่างต่อไปนี้

{8} คู่เทียบสำหรับ ‘⠨⠨’

ก. ⠨⠨ = chn, children

ข. ⠨⠨| = children|

ค. |⠨⠨ = |children

คู่เทียบสำหรับตัวย่อกลุ่มนี้ได้แสดงไว้ในภาคผนวก ฉ.7

(8) คู่เทียบสำหรับตัวย่อรูปเต็มคงรูป

คู่เทียบตัวย่อรูปเต็มคงรูป เป็นคู่เทียบที่ใช้ถ่ายถอดตัวย่อ and , for, of , the ,with, a ตัวย่อกลุ่มนี้เมื่อเขียนติดกันจะใช้แทนคำโดดในอักษรปกติ ในการสร้างคู่เทียบสำหรับตัวย่อกลุ่มนี้ผู้วิจัยสร้างคู่เทียบใน 2 ลักษณะคือ ลักษณะแรกให้อักขระเบรลล์แต่ละตัวจับคู่กับอักษรปกติโดยตรง เช่น ⠯ มีคู่เทียบเป็น and , ⠿ มีคู่เทียบเป็น for เป็นต้น อีกลักษณะหนึ่งผู้วิจัยได้สร้างคู่เทียบที่มีตัวย่อกลุ่มนี้ปรากฏร่วมกันหลายตัว โดยที่มีคู่เทียบในอักษรปกติที่มีเครื่องหมายเว้นวรรคคั่นระหว่างคำ เนื่องจากคู่เทียบในกลุ่มนี้เมื่อเขียนติดกันจะต้องตีความให้เป็นคำโดดในอักษรปกติ คือ มีเว้นวรรคคั่นระหว่างคำเหล่านี้จึงทำให้ได้คู่เทียบ เช่น ‘⠿⠯’ มีคู่เทียบเป็น “for the” ‘⠿⠯⠗’ มีคู่เทียบเป็น “and for a”

อันที่จริง การสร้างคู่เทียบใน<u>ลักษณะหลัง</u>นี้ อาจสร้างโดยให้ตัวย่อในกลุ่มนี้มีคู่เทียบ 2 กรณี เช่น ‘⠯’ มีคู่เทียบเป็น “and” และ “and|” แต่จากการทดลอง ผู้วิจัยพบว่า การสร้างให้ตัวย่อกลุ่มนี้มีคู่เทียบ 2 กรณี จะทำให้เกิดข้อผิดในการถ่ายถอดอักษรมากกว่าวิธีการที่นำเสนอข้างต้น โดยเฉพาะอย่างยิ่งระบบมักจะถ่ายถอด ‘⠿⠯’ ให้เป็น “ofthe” เสมอ แทนที่จะเป็น “of the”

คู่เทียบสำหรับตัวย่อกลุ่มนี้ได้แสดงไว้ในภาคผนวก ฉ.8

(9) คู่เทียบที่ตีความต่างกัน ในตำแหน่งต้นคำหรือตำแหน่งอื่น

ในภาษาอังกฤษ ‘⠠’ และ‘⠰’ เมื่อปรากฏที่ต้นคำจะใช้เพื่อระบุสถานะตัวอักษรที่ตามมา นั่นคือ ‘⠠’ (เครื่องหมายแสดงตัวพิมพ์ใหญ่) ใช้ระบุว่าอักษรที่ตามหลังอักขระเบรลล์นี้เป็นตัวพิมพ์ใหญ่ และ‘⠰’ (เครื่องหมายกำกับตัวอักษร) ใช้แสดงว่าอักษรที่ตามมาไม่ใช่ตัวย่อระดับคำ (เช่น ‘⠰⠁⠃’ ใช้แทน “ab” ไม่ใช่ “about”) แต่เมื่ออักขระเบรลล์ทั้ง 2 ตัวนี้ปรากฏในตำแหน่งอื่นจะต้องนำจะใช้เป็นส่วนต้นของตัวย่อหน่วยเสริม (ดูบทที่ 2 ตารางที่ 2.17 ประกอบ)

ฉะนั้น เพื่อให้สามารถประมวลผลด้วยแบบจำลองเอ็นแกรมได้ผู้วิจัยจึงได้สร้างคู่เทียบสำหรับตัวย่อกลุ่มนี้ ดังแสดงในภาคผนวก ฉ.9

(10) คู่เทียบตัวย่อที่ตีความตามตำแหน่ง

ตัวย่อกลุ่มนี้เป็นตัวย่อที่ต้องตีความตามตำแหน่งที่ปรากฏ เช่น ‘::’ หากปรากฏต้นสายอักขระจะใช้แทน “(”หากปรากฏที่ท้ายสายอักขระจะใช้แทน “)”หากปรากฏกลางคำจะใช้แทน “gg” หากปรากฏเป็นคำโดดจะใช้แทน “were” เป็นต้น ซึ่งตัวย่อกลุ่มนี้ประกอบด้วยคู่เทียบต่าง ๆ ดังแสดงในภาคผนวก ฉ.10

(11) คู่เทียบเครื่องหมายเบรลล์

เพื่อให้สามารถถ่ายถอดข้อความที่มีเครื่องหมายซึ่งปรากฏในข้อความเบรลล์อังกฤษทั่วไปได้ ผู้วิจัยจึงได้รวบรวมอักขระเบรลล์และคู่เทียบสำหรับเครื่องหมายต่าง ๆ ดังที่แสดงไว้ในภาคผนวก ฉ.11

6.1.2.2 คลังไตรแกรมตัวอักษรอังกฤษ

ในการนำคู่เทียบอักษรที่ได้กล่าวมาในหัวข้อ 6.1.2.1 มาใช้ถ่ายถอดอักษรนั้น จะทำให้เกิดผลลัพธ์ขึ้นเป็นจำนวนมาก ซึ่งไม่สามารถนำมาใช้เป็นผลลัพธ์ของการถ่ายถอดได้อย่างแท้จริง เพื่อให้ได้ผลลัพธ์ที่แท้จริง จำเป็นต้องคัดเฉพาะผลลัพธ์ที่มีค่าทางสถิติสูงสุด ซึ่งในครั้งนี้ผู้วิจัยได้เลือกใช้วิธีการประมวลผลด้วยแบบจำลองไตรแกรม โดยสกัดข้อมูลไตรแกรมจากคลังข้อความ British National Corpus (BNC) ซึ่งดำเนินการดังนี้

ในขั้นแรก ผู้วิจัยได้สกัดข้อความทั้งหมดแล้วนำมาแบ่งออกเป็นย่อหน้า และเพื่อให้ข้อมูลที่ได้เป็นตัวแทนของภาษาอังกฤษอย่างแท้จริง ผู้วิจัยได้แบ่งข้อความแต่ละย่อหน้าออกเป็นวรรค แล้วคัดเลือกเฉพาะวรรคที่มีตัวอักษรอังกฤษ a-z ไม่ว่าจะเป็นตัวพิมพ์เล็กหรือตัวพิมพ์ใหญ่ก็ตาม โดยที่วรรคนั้น ๆ จะต้องไม่มีอักขระใดที่ใช้รหัสคอมพิวเตอร์ร่วมกับอักษรไทยปรากฏอยู่เลย

จากนั้นทำการเติมเครื่องหมายเว้นวรรคทั้งข้างหน้าและข้างหลังของทุกวรรคที่สกัดได้ ต่อมานำข้อความแต่ละวรรคมาสร้างข้อมูลไตรแกรมทำให้ได้ข้อมูลเบื้องต้น คือ ได้จำนวนไตรแกรมที่ไม่ซ้ำรูปกัน 112,411 ชุด (types) และไตรแกรมที่ปรากฏซ้ำ ๆ จำนวนรวม 455,209,441 ครั้ง (tokens)

หมายเหตุ: อาจเป็นที่สงสัยว่า เหตุใดผู้วิจัยจึงไม่สร้างคลังไตรแกรมโดยสกัดตัวอักษรที่มีเครื่องหมายเว้นวรรคปรากฏร่วมด้วย เนื่องจากการสกัดข้อความที่มีเครื่องหมายเว้นวรรคปนอยู่จะทำให้ระบบสามารถนำ บริบท ข้างหน้าตัวอักษรที่กำลังถ่ายถอดมาช่วยถ่ายถอดอักษรในขอบเขตที่กว้างขึ้นได้ เช่น ทำให้ระบบสามารถระบุได้ว่า คำที่กำลังถ่ายถอดเป็นภาษาอังกฤษเนื่องจากคำที่ปรากฏก่อนหน้านี้เป็นภาษาอังกฤษเป็นต้น

ผู้วิจัยได้ทดลองการสกัดข้อความที่มีเครื่องหมายเว้นวรรคปนอยู่ พบว่าได้ผลความถูกต้องน้อยกว่าวิธีการที่ได้นำเสนอมาข้างต้น โดยที่เกิดข้อผิดในหลายลักษณะโดยเฉพาะอย่างยิ่ง เกิดข้อผิดกับการถ่ายถอด ‘⁚⁚’ ในตำแหน่งต้นคำเป็น “-” (แทนที่จะถ่ายถอดเป็น “com”)เป็นจำนวนมาก เช่น ‘⁚⁚’ ที่ในตำแหน่งกลางคำหรือท้ายคำซึ่งควรจะเป็น ation แต่กลับถ่ายถอดเป็น “N” เช่นคำว่า “conversation” “combination” “compensation” ถ่ายถอดเป็น “conversN”, “combinN”, “compensN” เป็นต้น นอกจากนี้ผู้วิจัยยังพบอีกว่า ข้อความในสายอักขระ ก่อนหน้าสายอักขระปัจจุบัน ไม่มีผลต่อการระบุสถานะภาษาของสายอักขระปัจจุบัน ดังจะเห็นได้จากการถ่ายถอดอักษรโดยใช้ข้อมูลที่ได้จาก 2 วิธีนี้ให้ผลลัพธ์ความถูกต้องในการระบุภาษาไม่แตกต่างกัน

เพื่อให้นำข้อมูลที่สกัดจากคลังข้อความมาใช้ถ่ายถอดอักษรได้ ผู้วิจัยได้นำข้อมูลไตรแกรมที่ได้มา มาทำการจัดเก็บในรูปของตาราง โดยที่ช่องแรกเป็นตัวอักษรไตรแกรมที่สกัดได้ และอีกช่องหนึ่งเป็นค่าสถิติที่ได้จากความถี่ที่พบไตรแกรมนั้นในคลังข้อความ BNC หารด้วยจำนวนครั้งที่พบของทุกไตรแกรมทั้งหมด(ที่สกัดได้จากคลังข้อความดังกล่าว) โดยที่ระบบจะนำตารางนี้ไปใช้ในการคำนวณค่าสถิติของสายอักขระที่สร้างขึ้น เพื่อหาสายอักขระที่มีค่าความน่าจะเป็นสูงสุดเพียงหนึ่งสาย มาเป็นผลลัพธ์ของการถ่ายถอดอักษร

6.1.2.3 คลังไบแกรมคำอังกฤษ

จากการทดลองหาค่าสถิติสูงสุดโดยใช้แบบจำลองไตรแกรมตัวอักษรอังกฤษ แต่เพียงลำพัง ผู้วิจัยพบว่าวิธีการนี้ก่อให้เกิดปัญหาต่อการถ่ายถอดตัวย่อบางประเภทในเบรลล์อังกฤษ เนื่องจากบ่อยครั้ง ระบบจะเลือกสายอักขระที่สั้นที่สุดมาเป็นผลลัพธ์ของการถ่ายถอด เพราะสายอักขระดังกล่าวมักจะมีค่าสถิติสูงกว่าสายอักขระอื่น ๆ เช่น ในการถ่ายถอด ‘⁝⁝⁝’ (=frs) ซึ่งใช้ย่อ “friends ” ระบบจะดำเนินการดังนี้

{9} การถ่ายถอด ‘⁝⁝⁝’ โดยอาศัยไตรแกรมตัวอักษรอังกฤษ

|⁝⁝ = |fr , |friend

⁝| = s|, so|

⁝⁝⁝ = |fr , |friend

s|, so| s|, so|

= |frs| . |frso| , |friends| , |friendso|

สายอักขระที่มีค่าทางสถิติสูงสุดเมื่อตรวจสอบกับไตรแกรมตัวอักษรอังกฤษ คือ |frs|

จากการวิจัยครั้งนี้ ผู้วิจัยพบว่า การใช้ข้อมูลไตรแกรมตัวอักษรอังกฤษแต่เพียงลำพังก่อให้เกิดปัญหาในลักษณะนี้ขึ้นเป็นจำนวนมาก ซึ่งปัญหาดังกล่าวสามารถแก้ไขได้โดยการใช้คลังไบแกรมคำอังกฤษ จากตัวอย่างข้างต้นจะพบว่า “ frs ” เป็นคำที่ไม่ปรากฏในภาษาอังกฤษ แต่การประมวลผลด้วยไตรแกรมตัวอักษรอังกฤษไม่สามารถบ่งชี้ความผิดปกติที่กล่าวมานี้ได้ เพื่อแก้ไขปัญหาที่เกิดขึ้น จึงจำเป็นต้องนำคลังไบแกรมคำอังกฤษ มาช่วยในการตัดสินใจ

หากระบบตรวจสอบคำ 2 คำที่เรียงต่อกันโดยอาศัยคลังไบแกรมคำอังกฤษ แล้วพบว่ามีความเป็นไปได้ที่สองคำนี้จะเกิดร่วมกันมากที่สุด ระบบจะสามารถนำคำสองคำนี้มาเป็นผลลัพธ์ที่ถูกต้องของการถ่ายถอดได้ เช่น ‘⠮ ⠋⠗⠎ ⠋⠗⠊⠑⠝⠙⠎’ ระบบจะนำคลังไบแกรมคำอังกฤษมาตรวจสอบ ดังนี้

{10} การถ่ายถอด ‘⠮ ⠋⠗⠎ ⠋⠗⠊⠑⠝⠙⠎’ โดยอาศัยคลังไบแกรมคำอังกฤษ

⠮ = the

⠋⠗⠎ = frs . frso , friends , friendso

⠮ ⠋⠗⠎ = the frs , the frso , the friends , the friendso

สายอักขระที่มีค่าทางสถิติสูงสุดเมื่อตรวจสอบกับคลังไบแกรม คือ the friends

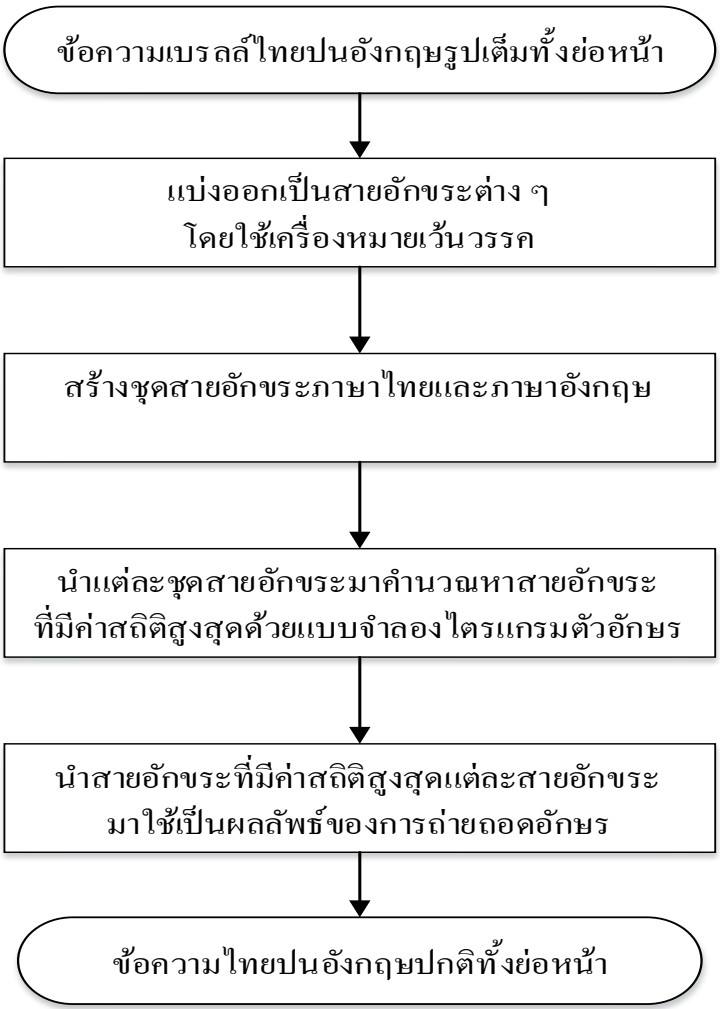
ในการวิจัยครั้งนี้ ผู้วิจัยได้สร้างคลังไบแกรมคำอังกฤษขึ้น โดยการรวบรวมคำที่ปรากฏร่วมกัน 2 คำ จาก British National Corpus โดยในคลังไบแกรมนี้ประกอบด้วย คำแต่ละคู่พร้อมค่าสถิติการปรากฏของคำคู่นั้นในคลังข้อความ BNC ในการสกัดข้อมูลได้จำนวนไบแกรมที่ไม่ซ้ำรูปกัน 4,820,126 ชุด (types) โดยที่ไบแกรมเหล่านี้ปรากฏซ้ำกันจำนวนรวมทั้งสิ้น 78,088,503 ครั้ง (tokens)

6.2 ระบบการถ่ายถอดอักษรโดยใช้แบบจำลองเอ็นแกรม

ในการทำงานของระบบถ่ายถอดอักษรเบรลล์ที่มีการปนกันระหว่างภาษาไทยและภาษาอังกฤษด้วยแบบจำลองเอ็นแกรมนั้น ประกอบด้วยระบบ 2 ระบบ คือ ระบบถ่ายถอดเบรลล์ไทยปนอังกฤษรูปเต็ม และระบบถ่ายถอดเบรลล์ไทยปนอังกฤษรูปย่อ ซึ่งแต่ละระบบมีการทำงานต่างกัน

(1) ระบบถ่ายถอดเบรลล์ไทยปนอังกฤษรูปเต็ม

สำหรับระบบถ่ายถอดเบรลล์ไทยปนอังกฤษรูปเต็มนั้น ระบบจะนำเข้าข้อความเบรลล์ทีละย่อหน้าแล้วนำแต่ละย่อหน้ามาแบ่งออกเป็นสายอักขระต่าง ๆ จากนั้นนำแต่ละสายอักขระไปสร้างชุดสายอักขระ ที่ประกอบด้วย สายอักขระภาษาไทยและสายอักขระภาษาอังกฤษ หลังจากนั้นนำชุดสายอักขระที่สร้างขึ้นไปประมวลผลด้วยแบบจำลองไตรแกรมตัวอักษร แล้วนำค่าสถิติสูงสุดมาใช้เป็นผลลัพธ์ของการถ่ายถอดอักษรสำหรับแต่ละสายอักขระ ดังแสดงในภาพต่อไปนี้



ภาพที่ 6.1 ภาพรวมของระบบถ่ายถอดเบรลล์ไทยปนอังกฤษรูปเต็มโดยใช้แบบจำลองเอ็นแกรม

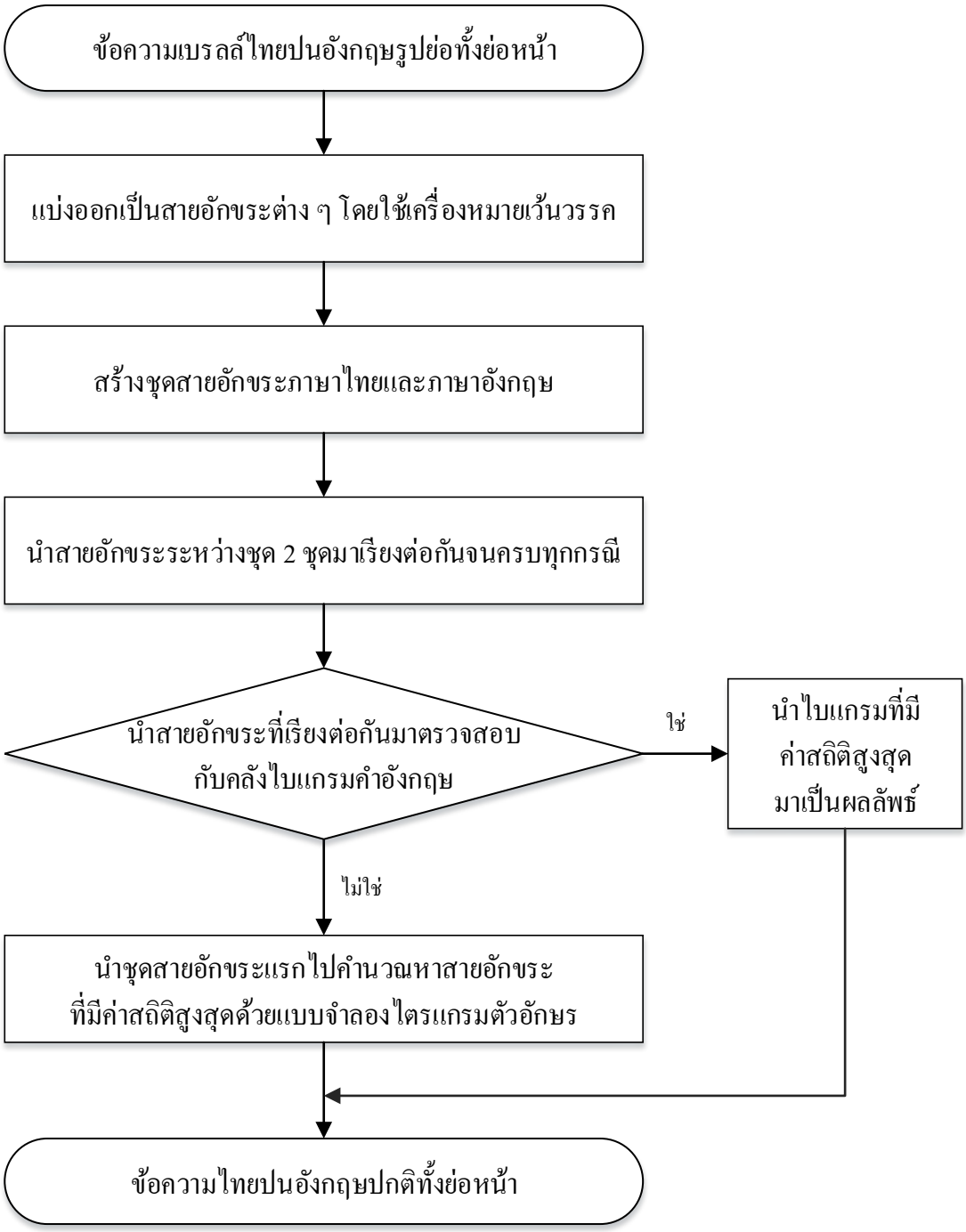
สำหรับการทำงานกับข้อความเบรลล์ไทยปนอังกฤษรูปเต็มนั้น ระบบจะดำเนินการโดยอาศัยคู่เทียบเบรลล์ไทยในภาคผนวก ง และคู่เทียบเบรลล์อังกฤษในภาคผนวก จ.1 โดยระบบจะนำสายอักขระเบรลล์ไปสร้างชุดสายอักขระที่ประกอบด้วยสายอักขระภาษาไทยและภาษาอังกฤษตามกระบวนการในข้อ 6.2.1 แล้วนำชุดสายอักขระที่สร้างขึ้นไปหาสายอักขระที่มีความน่าจะเป็นสูงสุดด้วยแบบจำลองไตรแกรมตัวอักษรตามกระบวนการที่ 6.2.3 เพื่อใช้เป็นผลลัพธ์ของการถ่ายถอดอักษร

(2) ระบบถ่ายถอดเบรลล์ไทยปนอังกฤษรูปย่อ

สำหรับระบบการถ่ายถอดเบรลล์ไทยปนอังกฤษรูปย่อ มีกระบวนการถ่ายถอดที่แตกต่างจากระบบการถ่ายถอดเบรลล์ปนอังกฤษรูปเต็ม โดยในการทำงานของระบบถ่ายถอดเบรลล์ไทยปนอังกฤษรูปย่อ จะเริ่มจากนำเข้าข้อความเบรลล์ทีละย่อหน้า แล้วแบ่งออกเป็นสายอักขระต่าง ๆ จากนั้นนำแต่ละสายอักขระไปสร้างชุดสายอักขระของทั้งภาษาไทยและภาษาอังกฤษจนครบทุกสายอักขระ

หลังจากนั้นระบบจะนำชุดสายอักขระทีละ 2 ชุด มาดำเนินการ โดยนำแต่ละสายอักขระระหว่างชุดมาเรียงต่อกัน แล้วตรวจสอบกับไบแกรมคำอังกฤษว่าสายอักขระที่เรียงต่อกันนั้นสามารถปรากฏเป็นคำอังกฤษ 2 คำที่เรียงต่อกันได้หรือไม่ ถ้าเป็นคำสองคำที่เรียงต่อกันได้และมีความน่าจะเป็นสูงสุด ระบบจะนำคำสองคำที่ได้นี้ไปเป็นผลลัพธ์ แต่หากนำแต่ละสายอักขระระหว่างชุดสายอักขระทั้งสอง มาเรียงต่อกันแล้วไม่สามารถเป็นคำอังกฤษที่เรียงต่อกันได้ ระบบจะนำชุดสายอักขระแรกไปคำนวณหาสายอักขระที่มีความน่าจะเป็นสูงสุดด้วยแบบจำลองไตรแกรมตัวอักษร แล้วนำสายอักขระนี้มาเป็นผลลัพธ์ จากนั้นนำชุดสายอักขระที่สองไปดำเนินการกับชุดสายอักขระที่สาม โดยจะดำเนินการทีละสองชุดสายอักขระเช่นนี้ไปเรื่อย ๆ จนครบทุกชุดสายอักขระ

การถ่ายถอดเบรลล์ไทยปนอังกฤษรูปย่อด้วยแบบจำลองเอ็นแกรม สามารถแสดงได้ด้วยแผนภาพต่อไปนี้



ภาพที่ 6.2 ภาพรวมของระบบถ่ายถอดเบรลล์ไทยปนอังกฤษรูปย่อโดยใช้แบบจำลองเอ็นแกรม

การทำงานของโมดูลแต่ละโมดูลในระบบถ่ายถอดอักษรด้วยแบบจำลองเอ็นแกรมมีรายละเอียดดังนี้

6.2.1 การสร้างชุดสายอักขระสำหรับภาษาไทยและภาษาอังกฤษ

ในขั้นนี้ระบบจะสร้างชุดสายอักขระที่ประกอบด้วยสายอักขระทั้ง 2 ภาษาโดยดำเนินการดังนี้

6.2.1.1 การสร้างชุดสายอักขระสำหรับภาษาไทย

ในการสร้างผลลัพธ์ของการถ่ายถอดอักษรที่เป็นไปได้ในไทยปกติ ระบบจะนำสายอักขระปัจจุบันมาสร้างชุดสายอักขระไทยปกติ ซึ่งมีวิธีการดังนี้

(1) การสร้างชุดสายอักขระตัวอักษรไทย

ในการสร้างชุดสายอักขระตัวอักษรไทย ระบบจะต้องอาศัยตารางคู่เทียบอักษรไทย ตามภาคผนวก ง โดยในขั้นแรกระบบจะนำเบรลล์ 3 เซลล์ 2 เซลล์ หรือ 1 เซลล์ปัจจุบัน[33]จากสายอักขระที่กำลังถ่ายถอด ไปตรวจสอบกับอักษรเบรลล์ในตารางนี้ หากพบว่าอักขระเบรลล์ที่นำมาตรวจสอบตรงกับเบรลล์ชุดใดเป็นอันดับแรก ระบบจะนำอักษรปกติแต่ละกรณีในคู่เทียบที่ตรงกับตารางมาต่อท้ายแต่ละสายอักขระในชุดที่สร้างขึ้นก่อนหน้านี้[34] หลังจากนั้นระบบจะนำเบรลล์เซลล์ต่อไปมาดำเนินการตั้งแต่ขั้นแรก

ตัวอย่างเช่น ถ้าระบบพบเบรลล์ 2 เซลล์ปัจจุบันเป็น ‘⠘⠍’ ซึ่งตรงกับอักขระเบรลล์ในตารางคู่เทียบ โดยที่อักษรชุดนี้มีคู่เทียบในอักษรปกติ 2 กรณี ระบบจะนำกรณีแรกคือ “ธ” ไปต่อท้ายแต่ละสายอักขระในชุดที่สร้างขึ้นก่อนหน้านี้ และนำอีกกรณีหนึ่งคือ“◌์ท”ไปต่อท้ายแต่ละสายอักขระในชุดที่ใช้กับกรณีแรก จากนั้นนำเซลล์ถัดจาก ‘⠘⠍’ มาดำเนินการเช่นเดียวกันตั้งแต่ขั้นแรกไปจนกระทั่งสิ้นสุดสายอักขระ เช่นการถ่ายถอด ‘⠘⠍⠊⠄⠙⠁⠄⠘⠍’ (=อธิบาย) ระบบจะสร้างชุดสายอักขระไทยสำหรับ ‘⠘⠍⠊⠄⠙⠁⠄⠘⠍’ ดังตัวอย่างข้อ {11}

---

[33] ในการดำเนินการครั้งแรก เบรลล์เซลล์ปัจจุบันจะเริ่มตั้งแต่เบรลล์เซลล์แรกของสายอักขระ

[34] ในการดำเนินการครั้งแรก ระบบจะต้องสร้างสายอักขระสำหรับอักษรปกติแต่ละกรณี เนื่องจากไม่มีสายอักขระใดปรากฏอยู่

{11} ‘⠷⠣⠃⠊⠁⠽’ (อธิบาย)

⠷ = อ

⠷⠣⠃ = อธ , อ์ท

⠷⠣⠃⠊ = อธิ , อ์ทิ

⠷⠣⠃⠊⠃ = อธิบ , อ์ทิบ

⠷⠣⠃⠊⠃⠁ = อธิบา , อ์ทิบา

⠷⠣⠃⠊⠃⠁⠽ = อธิบาย , อ์ทิบาย

จากตัวอย่างข้างต้นจะเห็นว่า ระบบจะสร้างชุดสายอักขระไทย ที่ประกอบด้วย 2 สายอักขระคือ “อธิบาย” และ “อ์ทิบาย” ซึ่งระบบจะนำ 2 สายอักขระนี้ไปประมวลผลด้วยแบบจำลองไตรแกรมตัวอักษรตามกระบวนการในหัวข้อที่ 6.2.1.3.3 เพื่อหาว่าสายอักขระใดมีความน่าจะเป็นสูงสุด แล้วนำสายอักขระนั้นมาใช้เป็นผลลัพธ์

อย่างไรก็ตาม หากปรากฏว่า อักขระเบรลล์ที่กำลังถ่ายถอดมีคู่เทียบที่ขึ้นต้นด้วย “|” หรือ “||” ซึ่งแสดงว่าคู่เทียบนั้นเป็นสระประสม ระบบจะดำเนินการเช่นเดียวกับการสร้างชุดสายอักขระที่กล่าวมาข้างต้น แต่จะดำเนินการเพิ่มเติมคือ ตัดเครื่องหมายเหล่านี้ออกจากคู่เทียบที่เป็นอักษรปกติ แล้วสลับที่อักษรตัวแรกของคู่เทียบกับอักษรก่อนหน้า 1 หรือ 2 ตัวขึ้นอยู่กับจำนวนเครื่องหมาย “ | ” ที่ปรากฏในคู่เทียบนั้น ๆ ดังตัวอย่าง

{12} การสร้างชุดสายอักขระสำหรับ ‘⠏⠗⠊⠜’

⠏ = ป

⠏⠗ = ปร

⠏⠗⠊⠜ = ปร|เ◌ีย , ปร||เ◌ีย

= ปเรีย , เปรีย

เพราะฉะนั้น ‘⠏⠗⠊⠜⠃’ = ปเรียบ , เปรียบ

จากตัวอย่าง ‘⠊⠜’ มีคู่เทียบ 2 กรณีคือ |เ◌ีย และ ||เ◌ีย เมื่อระบบพบเครื่องหมาย “ | ” ระบบจะทำการสลับที่ตัวอักษรที่ถ่ายถอดมาแล้วกับคู่เทียบ ‘⠊⠜’ จึงเกิดสายอักขระในไทยปกติขึ้น 2 สายอักขระที่เป็นไปได้ สำหรับการถ่ายถอด ‘⠏⠗⠊⠜⠃’ ดังที่แสดงมาข้างต้น เพื่อให้ได้ผลลัพธ์ของการถ่ายถอด ระบบจะนำ 2 สายอักขระที่ได้ไปผ่านการประมวลผลด้วยแบบจำลองไตรแกรมตัวอักษร เพื่อหาผลลัพธ์ของการถ่ายถอดอักษร

(2) การสร้างชุดสายอักขระสำหรับตัวเลข

ในบทที่ 2 ผู้วิจัยได้นำเสนอมาแล้วว่าในอักษรเบรลล์อังกฤษ ตัวอักษรสิบตัวแรกมีรูปซ้ำกับตัวเลข เช่น “ก” และ “7” ต่างก็แทนด้วย ‘⠛’ ซึ่งโดยทั่วไประบบจะต้องตีความให้รูปที่ซ้ำกันเหล่านี้เป็นตัวอักษร โดยใช้ตารางคู่เทียบอักษรไทยในการสร้างชุดสายอักขระที่เป็นไปได้ในภาษาไทย แต่เมื่อระบบพบว่าเบรลล์เซลล์ปัจจุบันเป็น ‘⠼’ (เครื่องหมายนำเลข) แสดงว่าอักขระที่ตามมาทั้งหมดจนสิ้นสุดสายอักขระปัจจุบัน[35] เป็นตัวเลข

ดังนั้นเมื่อระบบพบเครื่องหมาย ‘⠼’ ระบบจะเปลี่ยนจากการใช้ตารางคู่เทียบอักษรไทยไปใช้ตารางคู่เทียบตัวเลข (ตามภาคผนวก จ. 2) แทน โดยจะดำเนินการในลักษณะเดียวกับการสร้างชุดสายอักขระตัวอักษรไทยในข้อ 6.2.1.1 (1)

6.2.1.2 การสร้างชุดสายอักขระสำหรับภาษาอังกฤษ

เพื่อลดความซับซ้อนในการดำเนินการ ผู้วิจัยได้กำหนดให้การสร้างชุดสายอักขระอังกฤษสำหรับเบรลล์อังกฤษรูปเต็มและเบรลล์อังกฤษรูปย่อใช้กระบวนการเดียวกัน แต่ใช้ตารางคู่เทียบที่แตกต่างกัน คือ เบรลล์อังกฤษรูปเต็มจะดำเนินการโดยใช้ตารางตามภาคผนวก จ.1 สำหรับเบรลล์อังกฤษรูปย่อจะใช้ตารางตามภาคผนวก ฉ โดยในการสร้างชุดสายอักขระอังกฤษมีวิธีการดังนี้

ก. การสร้างสายอักขระที่ประกอบด้วยอักษรอังกฤษ

ในการสร้างสายอักขระที่ประกอบด้วยอักษรอังกฤษ ระบบจะดำเนินการดังนี้

(1) การหาคู่เทียบสำหรับเบรลล์อังกฤษ

ในขั้นแรก ระบบจะนำอักขระ 5 เซลล์ปัจจุบันในสายอักขระที่กำลังดำเนินการ ไปตรวจสอบกับอักษรเบรลล์ในตารางคู่เทียบอังกฤษ ถ้าไม่ตรงกับตาราง ระบบจะทำการลดเซลล์ทางขวาลงทีละเซลล์ จนกว่าอักษรเบรลล์ในสายอักขระจะตรงกับอักษรเบรลล์ในตารางนี้ จึงดำเนินการในขั้นต่อไป

---

35 หลักเกณฑ์ข้อนี้จะใช้เฉพาะการสร้างชุดสายอักขระสำหรับภาษาไทยเท่านั้น เนื่องจากผู้วิจัยยังไม่พบเอกสารเกี่ยวกับเบรลล์ไทยชิ้นใดที่กล่าวถึงการเขียนตัวเลขปนกับตัวอักษรไทยในวรรคเดียวกันว่าควรใช้เครื่องหมายกำกับตัวอักษรหรือไม่ อย่างไร อีกทั้งข้อความเบรลล์ที่นำมาทำการทดสอบนั้น เมื่อพิจารณาจากบริบท ผู้วิจัยพบว่าไม่มีการเขียนตัวเลขปนอักษรไทยจึงไม่เป็นปัญหาต่อการวิจัยครั้งนี้

(2) การนำตัวอักษรในคู่เทียบไปต่อท้ายทุกสายอักขระในชุด

ในขั้นนี้ ระบบจะนำอักษรปกติทีละกรณีที่ตรงกับอักษรเบรลล์ในตาราง ไปต่อท้ายทุกสายอักขระอังกฤษที่สร้างขึ้นก่อนหน้านี้จนครบทุกกรณี จากนั้นระบบจะดำเนินการกับเบรลล์เซลล์ต่อไปในลักษณะเดียวกันต่อไปเรื่อย ๆ จนสิ้นสุดสายอักขระปัจจุบัน ดังตัวอย่างที่ {13}

{13} การสร้างชุดสายอักขระสำหรับ ‘⠎⠉⠓⠕⠕⠇’ [36] (=s-ch-o-o-l)

⠸⠩⠎⠉⠓ = |sch , |such

⠸⠩⠎⠉⠓⠕ = |scho , |sucho

⠸⠩⠎⠉⠓⠕⠕ = |schoo , |suchoo

⠸⠩⠎⠉⠓⠕⠕⠇ = |school| , |suchool| , |schoolike| , |suchoolike|

การดำเนินการของระบบที่กล่าวมาข้างต้น ไม่รองรับกับการถ่ายถอดอักษรตัวพิมพ์ใหญ่และตัวเลขในอักษรเบรลล์ ระบบจึงต้องทำการตรวจสอบเบรลล์เซลล์ปัจจุบันเพื่อช่วยในการกำหนดสถานะของอักขระเบรลล์ โดยจะดำเนินการดังนี้

ข. การตรวจสอบเบรลล์เซลล์ปัจจุบันเพื่อกำหนดสถานะของอักขระเบรลล์

ในส่วนนี้ระบบจะตรวจสอบอักขระเบรลล์ 3 ชุด ในตำแหน่งเซลล์ปัจจุบัน แล้วดำเนินการตามอักขระเบรลล์ที่พบ ดังนี้

(1) ถ้าพบว่าเบรลล์เซลล์ปัจจุบันเป็น ‘⠠’ (เครื่องหมายแสดงตัวพิมพ์ใหญ่เดี่ยว) ระบบจะลบเครื่องหมายนี้ออกจากสายอักขระ แล้วนำอักขระเบรลล์ชุดต่อไปมาดำเนินการตามข้อ ก. หลังจากนั้นแปลงอักษรที่ได้ใหม่ในแต่ละสายอักขระให้เป็นตัวพิมพ์ใหญ่เฉพาะตัวแรก

(2) หากระบบพบว่าเบรลล์ 2 เซลล์ปัจจุบันเป็น ‘⠠⠠’ (เครื่องหมายแสดงตัวพิมพ์ใหญ่คู่) ระบบจะลบเครื่องหมายนี้ออกจากสายอักขระ แล้วถ่ายถอดอักษรตัวต่อไปตามขั้นตอนในข้อ ก. หลังจากนั้นแปลงอักษรที่เกิดขึ้นใหม่ให้เป็นตัวพิมพ์ใหญ่ทั้งหมดจนจบสายอักขระปัจจุบัน

(3) หากระบบพบว่าเบรลล์เซลล์ปัจจุบันเป็น ‘⠼’ (เครื่องหมายนำเลข) ระบบจะทำการคัดลอกอักขระจากตำแหน่งเบรลล์เซลล์ปัจจุบันจนสิ้นสุดสายอักขระไปดำเนินการในส่วนของการถ่ายถอดตัวเลขตามขั้นตอนในข้อ ค. แล้วนำผลลัพธ์ที่ได้มาดำเนินการตามข้อ ก. (2) ในขณะเดียวกันก็นำ

---

[36] ในที่นี้ “|” จะใช้แทนเครื่องหมายเว้นวรรค เหตุที่ต้องมีการเติมเว้นวรรคหน้า-หลังสายอักขระ เนื่องจากข้อมูลที่ใช้ฝึกระบบมีการกำกับคำด้วยเครื่องหมายเว้นวรรคหน้าและหลัง

เบรลล์ตั้งแต่เบรลล์เซลล์ปัจจุบันไปดำเนินการตามข้อ ก. (1) และ ก.(2) จากนั้นนำผลลัพธ์ของสายอักขระทั้งสองชุดมาเรียงตามลำดับสายอักขระที่ได้ ดังตัวอย่างต่อไปนี้

{14} การสร้างชุดสายอักขระสำหรับ '⠁⠇⠫' (=a#4)

|⠁ = |a

⠼⠙| = 4

|⠁⠇ = |able

|⠁⠇⠫| = |abled| , |abledo|[37]

จากตัวอย่างข้อ {14} ระบบจะสร้างเซ็ตชุดสายอักขระโดยในขั้นแรกเมื่อดำเนินการถึง ⠼ จะได้สายอักขระเป็น |⠁ = "|a" ขั้นต่อมาทำการคัดลอกเบรลล์เซลล์ปัจจุบันจนจบสายอักขระแล้วนำเบรลล์ที่คัดลอกไปสร้างชุดสายอักขระตามข้อ ค. จะได้ผลลัพธ์เป็น ⠼⠙| = "4" จากนั้นนำผลที่ได้มาดำเนินการตาม ข้อก.(2) จะได้ผลลัพธ์เป็น |⠁⠼⠙| = "a4" แล้วนำเบรลล์ตั้งแต่เบรลล์เซลล์ปัจจุบันไปดำเนินการตามข้อ ก. (1) และ ก.(2) ทำให้ได้ผลลัพธ์เป็น |⠁⠇ = "|able" และ |⠁⠇⠫| = "|abled|" , "|abledo|"[38] เมื่อนำสายอักขระสองชุดข้างต้นมาเรียงต่อกันได้เป็น |⠁⠼⠙| = "a4", .abled", "abledo"

ค. การสร้างสายอักขระสำหรับตัวเลขในเบรลล์อังกฤษ

ในกระบวนการนี้ระบบจะดำเนินการเช่นเดียวกับข้อ ก. แต่ใช้ตารางคู่เทียบตัวเลข แทน (ตามภาคผนวก จ.2) โดยจะดำเนินการไปเรื่อย ๆ จนพบว่าเบรลล์เซลล์ปัจจุบันไม่ตรงกับอักขระเบรลล์ในตารางนี้ จึงจะกลับไปใช้ตารางคู่เทียบอักษรอังกฤษ (ตามภาคผนวก จ.1 หรือ ฉ. ตามแต่ว่ากำลังถ่ายถอดเบรลล์รูปเต็มหรือรูปย่อ) แต่หากพบเครื่องหมาย '⠼' อีกก็จะกลับมาใช้ตารางคู่เทียบตัวเลข ทั้งนี้ระบบจะทำการสลับตารางทั้งสองกันไปมาโดยอาศัยเบรลล์เซลล์ปัจจุบันเป็นตัวกำหนด จนกว่าจะสิ้นสุดสายอักขระเบรลล์ที่กำลังดำเนินการอยู่

6.2.2 การหาผลลัพธ์สำหรับเบรลล์ไทยปนอังกฤษรูปย่อ

ในการถ่ายถอดเบรลล์ไทยปนอังกฤษรูปย่อ ระบบจะดำเนินการในขั้นนี้ได้ก็ต่อเมื่อระบบได้สร้างชุดสายอักขระจนครบ 2 ชุดสายอักขระแล้ว (จะทำให้ได้ 2 ชุดสายอักขระ โดยที่แต่ละชุด

---

[37] ดูภาคผนวก ฉ.4 ประกอบ

[38] ดูภาคผนวก ฉ.4 ประกอบ

ประกอบด้วยสายอักขระไทยและสายอักขระอังกฤษ) ซึ่งในที่นี้ผู้วิจัยจะอ้างอิงถึงชุดสายอักขระที่เกิดขึ้นนี้ให้เป็นชุดสายอักขระที่ 1 และชุดสายอักขระที่ 2 ตามลำดับ หลังจากนั้นระบบจะดำเนินการ คือ หาคำที่เป็นไปได้ในภาษาอังกฤษจากชุดสายอักขระครั้งละ 2 ชุด

จากการทดลองถ่ายถอดข้อความเบรลล์อังกฤษด้วยแบบจำลองไตรแกรมตัวอักษรแต่เพียงลำพัง ผู้วิจัยพบว่าระบบจะทำการถ่ายถอดคำที่เรียงต่อกันได้ไม่ถูกต้อง เนื่องจากระบบจะเลือกคู่เทียบที่มีขนาดสั้นที่สุดมาเป็นผลลัพธ์ เช่น ถ่ายถอด “best friend.” เป็น “best fr.” เนื่องจาก ‘⠃⠑⠌’ มีคู่เทียบ ได้ 2 กรณี คือ “|fr” และ “|friend” โดยที่คู่เทียบแรกมีขนาดสั้นกว่าอีกคู่เทียบหนึ่ง ระบบจึงมักจะนำคู่เทียบที่สั้นกว่ามาเป็นผลลัพธ์

เพื่อแก้ปัญหาข้างต้น ระบบจะนำแต่ละสายอักขระที่ปรากฏระหว่างชุด 2 ชุดมาเรียงต่อกัน แล้วดูว่าสายอักขระที่เรียงต่อกันนั้นจะปรากฏเป็นคำอังกฤษ2คำที่เรียงต่อกันได้หรือไม่โดยนำไปตรวจสอบกับแบบจำลอง หากพบว่าตรงกับคลังไบแกรมคำอังกฤษ ระบบจะนำ 2 คำที่มีค่าทางสถิติสูงสุดมาเป็นผลลัพธ์ เช่นการหาคำ 2 คำที่ปรากฏร่วมกันใน‘⠛⠕⠕⠙’ (=good) และ‘⠋⠗⠊⠑⠝⠙’ (= friend.) ระบบนำ 2 คำนี้ไปสร้างชุดสายอักขระ 2 ชุด และนำสายอักขระระหว่างชุดมาเรียงต่อกัน ดังนี้

{15} ‘⠛⠕⠕⠙’ ‘⠋⠗⠊⠑⠝⠙’

[⠛⠕⠕⠙ ⠋⠗⠊⠑⠝⠙] = [กด เร้] [กด เร.] [กด friend.] [กด fr.]

[good เร้] [good เร.] [good friend.] [good fr.]

อย่างไรก็ตาม เมื่อตรวจสอบกับคลังไบแกรมคำอังกฤษแล้วพบว่าไม่ตรงกับไบแกรมชุดใด ระบบจะนำชุดสายอักขระแรกไปประมวลผลด้วยแบบจำลองไตรแกรมตัวอักษร (ตามหัวข้อ 6.2.3) แล้วนำสายอักขระที่มีความน่าจะเป็นสูงสุดมาใช้เป็นผลลัพธ์ของการถ่ายถอดอักษร จากนั้นนำชุดสายอักขระต่อไปมาดำเนินการเช่นเดียวกัน

6.2.3 การหาค่าสถิติสูงสุดของสายอักขระโดยใช้แบบจำลองไตรแกรมตัวอักษร

สำหรับกระบวนการนี้ ระบบจะรับชุดสายอักขระที่ได้จากกระบวนการอื่นแล้วนำมาจัดเรียงจากสายอักขระที่มีค่าสถิติสูงสุดไปยังสายอักขระที่มีค่าสถิติต่ำสุด โดยใช้คลังไตรแกรมตัวอักษรไทย (ที่ผู้วิจัยจัดเตรียมไว้ตามหัวข้อ 6.1.1.2) และคลังไตรแกรมตัวอักษรอังกฤษ (ตามหัวข้อ6.1.2.2) แล้วส่งสายอักขระที่จัดเรียงแล้วไปยังกระบวนการอื่นต่อไป

ในการจัดเรียงสายอักขระ ระบบจะคำนวณค่าสถิติที่สามตัวอักษรปรากฏร่วมกัน โดยอาศัยแบบจำลองไตรแกรมตัวอักษรไทยและอังกฤษ สำหรับไตรแกรมที่ระบบไม่พบจากข้อมูลฝึกระบบ ระบบจะนำค่าสถิติต่ำสุดที่ได้จากการฝึกระบบมาหารด้วย 10^{10} มาใช้แทนไตรแกรมที่พบ

6.3 การทดสอบระบบการถ่ายถอดอักษรโดยใช้แบบจำลองเอ็นแกรม

ในการทดสอบระบบ ผู้วิจัยได้นำข้อความเบรลล์ 5 ประเภทซึ่งเป็นข้อความเดียวกับที่นำมาใช้ทดสอบการถ่ายถอดอักษรด้วยกฎที่ได้นำเสนอมาแล้วในหัวข้อ 5.1.2 มาทำการทดสอบกับกระบวนการต่าง ๆ ได้แก่ กระบวนการระบุภาษา กระบวนการถ่ายถอดเบรลล์ไทย และกระบวนการถ่ายถอดเบรลล์อังกฤษ โดยใช้วิธีการประเมินผลเช่นเดียวกับการถ่ายถอดด้วยกฎแต่เปลี่ยนข้อมูลที่ใช้ในการทดสอบระบบเป็นข้อมูลที่มาจากการถ่ายถอดด้วยแบบจำลองเอ็นแกรม

6.3.1 การทดสอบระบบกับข้อความเบรลล์ไม่ปนภาษา

วิธีการทดสอบจะแยกเป็นสองส่วน ส่วนแรกเป็นการทดสอบประสิทธิภาพของการระบุภาษาและถ่ายถอดอักษรโดยใช้ข้อความเบรลล์ที่ไม่ปนภาษา และส่วนที่สองเป็นการทดสอบประสิทธิภาพของระบบกับข้อมูลจริงที่มีการปนภาษา ซึ่งผู้วิจัยจะขอนำเสนอผลในการทดสอบทั้ง 2 ส่วน ดังนี้

(1) การระบุภาษา

ในระบบถ่ายถอดเบรลล์ไทยปนอังกฤษโดยใช้แบบจำลองเอ็นแกรม ผู้วิจัยได้ทำการทดสอบระบบ โดยนำข้อความแต่ละภาษาไปทดลองกับระบบถ่ายถอดไทยปนอังกฤษรูปเต็ม และระบบถ่ายถอดไทยปนอังกฤษรูปย่อ แล้วนำผลลัพธ์ที่ได้จากการถ่ายถอดอักษรในแต่ละการทดลองมาประเมินผลความถูกต้องของการระบุภาษา โดยการนับจำนวนสายอักขระที่ถ่ายถอดผิดจากภาษาหนึ่งไปเป็นอีกภาษาหนึ่ง ทำให้ได้ผลลัพธ์ซึ่งแสดงได้ดังนี้

ตารางที่ 6.1 ผลการทดสอบประสิทธิภาพของการระบุภาษาโดยใช้แบบจำลองเอ็นแกรม

ระบบที่ใช้ทดสอบ	ประเภท ข้อความเบรลล์	สายอักขระ ที่ใช้ทดสอบ	สายอักขระ ที่ระบุภาษาผิด	ระบุภาษา ถูกต้อง (%)
ระบบไทยปนอังกฤษรูปเต็ม	ไทยล้วน	12,913	32	99.75
ระบบไทยปนอังกฤษรูปย่อ	ไทยล้วน	12,913	146	98.86
ระบบไทยปนอังกฤษรูปเต็ม	อังกฤษรูปเต็ม	59,072	30	99.95
ระบบไทยปนอังกฤษรูปย่อ	อังกฤษรูปย่อ	56,206	511	99.09

(2) การถ่ายถอดอักษร

ในการทดสอบประสิทธิภาพของการถ่ายถอดอักษร ได้ผลการทดสอบดังนี้

ตารางที่ 6.2 ผลการทดสอบประสิทธิภาพของการถ่ายถอดอักษรโดยใช้แบบจำลองเอ็นแกรม

ประสิทธิภาพในการถ่ายถอด	ประเภท ข้อความเบรลล์	จำนวนคำที่ ใช้ทดสอบ	จำนวนคำที่ ถ่ายถอดผิด	ถ่ายถอด ได้ถูกต้อง(%)
เบรลล์ไทย	ไทยล้วน	19,090	121	99.36
เบรลล์อังกฤษรูปเต็ม	อังกฤษรูปเต็ม	59,057	6	99.99
เบรลล์อังกฤษรูปย่อ	อังกฤษรูปย่อ	59,057	120	99.80

6.3.2 การทดสอบระบบกับข้อความเบรลล์ปนภาษา

ผู้วิจัยได้ทดสอบระบบถ่ายถอดอักษรกับข้อมูลจริง 2 ประเภทคือ ข้อความเบรลล์ไทยปนอังกฤษรูปเต็ม และข้อความเบรลล์ไทยปนอังกฤษรูปย่อ โดยในการประเมินประสิทธิภาพของระบบจะแบ่งออกเป็น 2 ส่วน ได้แก่การระบุภาษาและการถ่ายถอดอักษร (โดยที่ในการวัดประสิทธิภาพของการถ่ายถอดอักษรจะพิจารณาเฉพาะสายอักขระที่ระบุภาษาได้ถูกต้องเท่านั้น)

(1) ผลการระบุภาษา

การทดสอบประสิทธิภาพของระบบถ่ายถอดอักษรในส่วนของการระบุภาษา แสดงดังนี้

ตารางที่ 6.3 ผลการระบุภาษาของระบบถ่ายถอดอักษรที่ทดสอบกับข้อมูลปนภาษา

ระบบ	ประเภทข้อความเบรลล์	ภาษา	สายอักขระที่ใช้ทดสอบ	ระบุภาษาผิด	ระบุถูกต้องแต่ละภาษา(%)	ระบุถูกต้องรวม (%)
ระบบไทยปนอังกฤษรูปเต็ม	ไทยปนอังกฤษรูปเต็ม	ไทย	7,315	17	99.77	99.37
		อังกฤษ	6,814	69	99.99	
ระบบไทยปนอังกฤษรูปย่อ	ไทยปนอังกฤษรูปย่อ	ไทย	7,315	42	99.43	98.74
		อังกฤษ	6,640	129	98.06	

(2) ผลการถ่ายถอดอักษรเฉพาะข้อความที่ระบุภาษาได้ถูกต้อง

ในการทดสอบระบบถ่ายถอดอักษรเฉพาะกับสายอักขระที่ระบุภาษาได้ถูกต้อง สามารถแสดงได้ดังนี้

ตารางที่ 6.4 ผลการถ่ายถอดอักษรเฉพาะข้อความที่ระบุภาษาถูกต้อง

ระบบ	ประเภทข้อความเบรลล์	ภาษาที่ต้องการวัดประสิทธิภาพ	จำนวนคำที่ระบุภาษาถูก	จำนวนคำที่ถ่ายถอดผิด	ถ่ายถอดถูกต้อง(%)
ระบบไทยปนอังกฤษรูปเต็ม	ไทยปนอังกฤษรูปเต็ม	ไทย	11,921	121	98.98
		อังกฤษ	6,538	0	100
ระบบไทยปนอังกฤษรูปย่อ	ไทยปนอังกฤษรูปย่อ	ไทย	11,896	84	99.29
		อังกฤษ	6,492	119	98.17

(3) ผลการถ่ายถอดอักษรของระบบโดยภาพรวม

ระบบถ่ายถอดอักษรทั้งสองระบบนี้ สามารถถ่ายถอดข้อความที่มีเนื้อหาเดียวกันแต่ใช้อักษรเบรลล์คนละระบบได้มีประสิทธิภาพต่างกัน ซึ่งแสดงได้ดังนี้

ตารางที่ 6.5 ผลการถ่ายถอดอักษรด้วยการใช้แบบจำลองเอ็นแกรมโดยภาพรวม

ระบบ	ประเภทข้อความเบรลล์	จำนวนคำที่ใช้ทดสอบ	จำนวนคำที่ถ่ายถอดผิด	ถ่ายถอดได้ถูกต้อง (%)
ไทยปนอังกฤษรูปเต็ม	ไทยปนอังกฤษรูปเต็ม	19,412	186	99.04
ไทยปนอังกฤษรูปย่อ	ไทยปนอังกฤษรูปย่อ	19,412	260	98.66

6.4 อภิปรายผลการทดสอบระบบ

จากผลการทดสอบระบบในภาพรวม ผู้วิจัยพบข้อผิดประเภทต่าง ๆ จากการถ่ายถอดอักษรด้วยแบบจำลองเอ็นแกรม ซึ่งจะขออภิปรายดังนี้

6.4.1 การระบุภาษา

ในการถ่ายถอดเบรลล์ไทยปนอังกฤษด้วยแบบจำลองเอ็นแกรม ที่ผู้วิจัยได้นำเสนอในหัวข้อที่ 6.2 ไม่มีส่วนของการระบุภาษาอย่างแท้จริง แต่ระบบจะระบุภาษาได้โดยการเลือกสายอักขระที่มีความน่าจะเป็นสูงสุดจากชุดสายอักขระที่ระบบสร้างขึ้น หากสายอักขระที่มีความน่าจะเป็นสูงสุดตรงกับภาษาใด ระบบจะนำสายอักขระนั้นมาเป็นผลลัพธ์ แต่หากสายอักขระที่มีความน่าจะเป็นสูงสุด นั้นไม่ตรงกับภาษาที่ต้องการถ่ายถอด ระบบก็จะเลือกภาษาผิดโดยปริยาย

ในการทดสอบระบบนั้น ผู้วิจัยได้นำข้อความที่ไม่มีการปนภาษาไปทำการทดสอบระบบ ดังแสดงไว้ในตารางที่ 6.1 ผลการทดสอบแสดงให้เห็นว่าระบบสามารถระบุภาษาสำหรับข้อความอังกฤษรูปเต็มที่ถ่ายถอดด้วยระบบไทยปนอังกฤษรูปเต็มได้ดีที่สุด (99.95%) รองลงมาเป็นการระบุภาษาสำหรับข้อความไทยด้วยระบบถ่ายถอดไทยปนอังกฤษรูปเต็ม และอังกฤษด้วยระบบถ่ายถอดไทยปนอังกฤษรูปย่อตามลำดับ

อย่างไรก็ตาม หากเปรียบเทียบการระบุภาษาของระบบไทยปนอังกฤษรูปเต็มและระบบไทยปนอังกฤษรูปย่อโดยทดสอบกับข้อความเบรลล์ไทยล้วน (ตารางที่ 6.1) ผู้วิจัยพบว่า ระบบถ่ายถอดไทยปนอังกฤษรูปเต็มจะสามารถระบุว่าเป็นภาษาไทยได้ดีกว่าระบบถ่ายถอดไทยปนอังกฤษรูปย่อ นั่นคือ ระบบไทยปนอังกฤษรูปเต็มจะระบุเป็นภาษาไทยได้ถูกต้อง 99.75%ในขณะที่ ระบบไทยปนอังกฤษรูปย่อ ระบุถูกต้อง 98.86 % เหตุที่เป็นเช่นนี้เนื่องจากเบรลล์ไทยและเบรลล์อังกฤษรูปเต็ม มีการจัดเรียงอักขระเบรลล์ที่แตกต่างกันเป็นส่วนใหญ่ทำให้ระบบสามารถตีความว่าเป็นภาษาใดภาษาหนึ่ง ได้ง่ายกว่าที่จะเป็นอีกภาษาหนึ่ง

ข้อเท็จจริงที่ว่าระบบถ่ายถอดไทยปนอังกฤษรูปเต็ม ระบุภาษาได้ดีกว่าอีกระบบหนึ่งยังปรากฏให้เห็นจากการถ่ายถอดข้อความปนภาษา จากตารางที่ 6.3 แสดงให้เห็นว่า ระบบไทยปนอังกฤษรูปเต็มสามารถระบุว่าเป็นภาษาอังกฤษได้ถูกต้อง 99.99 % ในขณะที่ระบบปนภาษารูปย่อระบุภาษาให้เป็นภาษาอังกฤษได้ถูกต้อง 98.06 %

อีกประเด็นหนึ่งที่น่าสนใจคือ ระบบถ่ายถอดด้วยแบบจำลองเอ็นแกรมนั้นระบุภาษาใดได้ดีกว่ากันระหว่างเบรลล์ไทยและเบรลล์อังกฤษ ผลการทดสอบแสดงให้เห็นว่า ระบบไทยปนอังกฤษรูปเต็ม จะระบุภาษาอังกฤษได้ดีกว่าภาษาไทย คือระบุให้เป็นภาษาอังกฤษได้ถูกต้อง 99.99% ในขณะที่ระบุให้เป็นภาษาไทยได้ถูกต้อง 99.77% แต่ในทางตรงกันข้ามระบบไทยปนอังกฤษรูปย่อจะระบุภาษาไทยได้ดีกว่าภาษาอังกฤษ นั่นคือ ระบุภาษาไทยได้ถูกต้อง 99.43 ในขณะที่ระบุภาษาอังกฤษได้ถูกต้อง 98.06% โดยที่สายอักขระอังกฤษที่ระบุผิดเป็นไทยนั้นส่วนใหญ่เป็นสายอักขระที่มีความยาว 1 - 5 อักขระ เช่น ‘⠎⠙⠀’ (= sd ซึ่งเป็นคำย่อของ “said”) ถ่ายถอดเป็น “ สด ” เนื่องจากเมื่อระบบนำ ‘⠎⠙⠀’ ไปสร้างชุดสายอักขระที่เป็นไปได้ในอักษรปกติของทั้งภาษาไทยและภาษาอังกฤษสำหรับคำ ๆ นี้แล้ว จะได้ชุดสายอักขระต่าง ๆ เป็นจำนวนมาก แต่สายอักขระที่มีค่าความน่าจะเป็นสูงสุด คือ |สด| ซึ่งตรงกับภาษาไทย ทำให้ระบบเลือกภาษาที่ผิดมาเป็นผลลัพธ์

ใช่ว่าระบบจะระบุภาษาผิดสำหรับสายอักขระที่มีขนาดสั้น ผู้วิจัยยังพบอีกว่า ระบบได้ระบุสายอักขระเบรลล์อังกฤษที่มีขนาดความยาวถึง 14 อักขระ ให้เป็นภาษาไทยคือคำว่า ‘⠎⠊⠍⠥⠇⠞⠁⠝⠑⠕⠥⠎⠇⠽⠲’ (=s-i-m-u-l-t-a-n-e-ou-s-l-y-.) ซึ่งระบบเลือก “*สโมคลถะนัวตสลย.*” เป็นผลลัพธ์ของการถ่ายถอดอักษร แทนที่จะเป็น “simultaneously.” ซึ่งค่าความยาวที่กล่าวมานี้ ผู้วิจัยจะนำไปใช้แก้ปัญหาเกี่ยวกับการระบุภาษาในการถ่ายถอดอักษรโดยวิธีการแบบผสม ที่จะได้นำเสนอในบทต่อไป

6.4.2 การถ่ายถอดอักษร

ตารางที่ 6.2 แสดงให้เห็นว่า การถ่ายถอดข้อความเบรลล์อังกฤษรูปเต็มด้วยแบบจำลองเอ็นแกรม ให้ผลลัพธ์การถ่ายถอดอักษรได้ถูกต้องมากที่สุด (99.99 %) รองลงมาคือการถ่ายถอดข้อความเบรลล์อังกฤษรูปย่อ และการถ่ายถอดข้อความเบรลล์ไทย ตามลำดับ แต่จากการทดสอบการถ่ายถอดข้อความปนภาษาดังที่แสดงไว้ในตารางที่ 6.4 พบว่า ระบบสามารถถ่ายถอดเบรลล์ไทยในข้อความไทยปนอังกฤษรูปย่อได้ถูกต้องมากกว่าเบรลล์อังกฤษในข้อความเดียวกัน ซึ่งข้อผิดต่าง ๆ ที่เกิดจากการถ่ายถอดในแต่ละภาษา มีดังนี้

6.4.2.1 ข้อผิดในการถ่ายถอดเบรลล์ไทย

การถ่ายถอดเบรลล์ไทยด้วยแบบจำลองเอ็นแกรมเกิดข้อผิดอยู่หลายประเภท โดยที่แต่ละประเภทมีจำนวนข้อผิดแตกต่างกัน โดยข้อผิดที่พบมากที่สุดเป็นข้อผิดที่เกิดจากการถ่ายถอด ‘⠮⠪’ (เาะ หรือ อะ) มากที่สุด รองลงมาคือการถ่ายถอด ‘⠪’ (เิ หรือ เ◌อ) ซึ่งแต่ละประเภทมีรายละเอียดดังนี้

ก. ข้อผิดจากการถ่ายถอด ‘⠮⠪’

‘⠮⠪’ ในเบรลล์ไทยสามารถใช้แทน "อะ" หรือ "เาะ" ได้ ฉะนั้นเมื่อระบบพบอักขระเบรลล์คู่นี้ ระบบจะอาศัยคู่เทียบของ ‘⠮⠪’ ซึ่งเทียบได้กับอักษรปกติ 3 กรณีคือ "อะ", "|เาะ" และ "||เาะ" นั่นคือกรณีแรกใช้แทน “อะ” และอีก 2 กรณีใช้แทน "เาะ" โดยที่เมื่อใช้แทน "เาะ" จะต้องมีการสลับที่พยัญชนะ1 ตัวหรือ 2 ตัวหน้า ‘⠮⠪’

ในการทดสอบระบบ พบว่าระบบมักจะถ่ายถอด ‘⠮⠪’ เป็น “อะ” เสียเป็นส่วนใหญ่ โดยที่ข้อผิดนี้ปรากฏในข้อความเบรลล์ไทยล้วน 24 แห่ง และในข้อความเบรลล์ไทยปนอังกฤษ 38 แห่ง ตัวอย่างเช่น ในการถ่ายถอด ‘⠠⠏⠗⠮⠪’ (= ไ-พ-ร-เาะ) ระบบจะสร้างชุดสายอักขระที่เป็นไปได้คือ "ไพรอะ", "ไพเราะ" และ "ไเพราะ" จากนั้นจึงนำสายอักขระที่เป็นไปได้ทั้งหมดไปหาค่าทางสถิติสูงสุดด้วยแบบจำลองไตรแกรมตัวอักษร ซึ่งระบบได้เลือกถ่ายถอด ‘⠠⠏⠗⠮⠪’ เป็น “ไพรอะ” เนื่องจากสายอักขระที่มีคำนี้ปรากฏอยู่ มีค่าสถิติสูงสุด

เป็นที่น่าสังเกตว่า ไม่พบข้อผิดนี้ในการถ่ายถอด‘⠮⠪’ ที่ผิดจาก“อะ” เป็น “เาะ”เลย เนื่องจากโดยมากแล้วระบบจะถ่ายถอดเบรลล์คู่นี้ให้เป็น “อะ”

ข. ข้อผิดจากการถ่ายถอด ‘ ◌ุ ’

‘ ◌ุ ’ สามารถตีความให้เป็น "เ◌ิ" หรือ "เ◌อ" ได้ นอกจากนี้การเรียงต่อของพยัญชนะต้นกับสระตัวนี้ยังมีความแตกต่างกันในเบรลล์ไทยและไทยปกติทำให้บางครั้งเกิดความกำกวมในการตีความ เช่น ‘ ⠏⠑⠥⠷⠗⠉⠁ ’ (=พ-เ◌ือ-◌่-น-ก-ล-เ อ) และ ‘ ⠽⠅⠇⠷⠗ ’ (=ย-ก-ล-เ◌ิ-ก) มีการเรียงต่อของ ‘ ⠅⠇ ’ (=กล) เหมือนกันในอักษรเบรลล์ แต่เมื่อถ่ายถอดเป็นไทยปกติจะต้องจัดเรียง “ก” และ “ล” ใน 2 คำนี้ให้แตกต่างกัน จากการทดสอบระบบ พบว่า มีการถ่ายถอด ‘ ◌ุ ’ เป็น “เ◌อ” ในตำแหน่งที่ควรจะเป็น “เ◌ิ” โดยที่ข้อผิดนี้ปรากฏในข้อความเบรลล์ไทยล้วน 34 แห่ง และในข้อความเบรลล์ไทยปนอังกฤษ 29 แห่ง เช่นระบบถ่ายถอด “กำเริบ” เป็น “กำเรอบ”

เป็นที่น่าสังเกตว่า คำที่ผิดล้วนเป็นคำที่มี “เ◌ิ” ตามด้วยตัวสะกดท้ายพยางค์ทั้งสิ้น ซึ่งระบบควรจะเลือกรูปของ ‘ ◌ุ ’ ได้ถูกต้องเนื่องจากมีบริบทที่ชัดเจนแต่ระบบได้เลือกรูปคำที่ผิดเนื่องจากสายอักขระที่ถูกนำมาเป็นผลลัพธ์มีค่าทางสถิติสูงกว่าชุดสายอักขระอื่น ๆ ที่ระบบสร้างขึ้น

ในทางตรงกันข้าม มีการถ่ายถอด "เ อ” ผิดเป็น "เ◌ิ" ซึ่งเกิดขึ้นเป็นจำนวน 5 ครั้ง กับคำ 2 คำซ้ำ ๆ กัน[39] คือ “พีเดอริน” ถ่ายถอดเป็น “พีเดิริน” (2 ครั้ง) และ “เสนอ” ถ่ายถอดเป็น “เสนิ” (3 ครั้ง)

ค. ข้อผิดจากการถ่ายถอด ‘ ◌ฺ ’

‘ ◌ฺ ’ สามารถตีความให้เป็น “.” หรือ “◌้” ได้ ในการถ่ายถอดอักขระเบรลล์ตัวนี้เป็นปัญหาเพียงเล็กน้อยกับสายอักขระที่มีความยาวมากกว่า 2 อักขระ โดยที่พบว่ามีการถ่ายถอดผิดเพียง 3 คำ คือ “มล.” “"กสท."” และ “น.ม.ส.” ถ่ายถอดเป็น “มล้” "กสท้" และ “น้ม.ส.” ตามลำดับ แต่จะเกิดปัญหาอย่างมากกับสายอักขระที่มีความยาว 2 อักขระ เช่น “ข.” “ต.” ที่ระบบถ่ายถอดเป็น “ข้” “ต้” เป็นต้น ซึ่งพบเป็นจำนวนทั้งสิ้น 30 คำ เหตุที่เป็นเช่นนี้เนื่องจากในคลังข้อความมีตัวอักษรตามด้วยไม้โท (Consonant + Tone) มากกว่าตัวอักษรตามด้วยจุด (Consoant + .)

ง. ข้อผิดจากการถ่ายถอดพยัญชนะ 2 ตัวร่วมกับสระประสม

ในการสร้างผลลัพธ์ที่เป็นไปได้สำหรับสระประสมนั้น ระบบจะอาศัยคู่เทียบซึ่งกำกับด้วยเครื่องหมาย “|” เพื่อช่วยในการสลับตำแหน่งระหว่างตัวอักษรและสระหน้าซึ่งเป็นส่วนหนึ่งของสระประสม อย่างไรก็ตาม จะพบว่าในการสลับตำแหน่งตัวอักษรนั้น ระบบไม่สนใจว่าอักษรที่ปรากฏหน้า

---

[39] ในการนับคำผิดจะถือว่าผิดเป็นจำนวน 5 คำ

สระประสม เป็นพยัญชนะหรือไม่ หรืออักษรที่ทำการสลับที่นั้นสามารถปรากฏระหว่างสระประสมได้หรือไม่ จึงเป็นผลให้ระบบถ่ายถอดพยัญชนะหน้าสระประสมไม่ถูกต้อง

จากการทดสอบระบบ มีการถ่ายถอดพยัญชนะ 2 ตัวที่ปรากฏร่วมกับสระประสมผิด 5 แห่ง โดยที่แบ่งข้อผิดได้เป็น 2 ประเภทคือ (1) ข้อผิดที่เกิดจากการแทรกพยัญชนะที่เป็นไปไม่ได้ระหว่างสระเช่น “ดื่มนมเยอะ” ผิดเป็น “ดื่มนเมยอะ” ในกรณีนี้ “นม”ไม่สามารถปรากฏระหว่างสระได้ และ (2) ข้อผิดที่เกิดจากการนำพยัญชนะที่มีรูปคล้ายพยัญชนะควบกล้ำแต่ไม่ใช่พยัญชนะควบกล้ำมาแทรกระหว่างสระประสม เช่น “ผู้ป่วยด้วยโรคเรื้อรัง” ถ่ายถอดเป็น “ผู้ป่วยด้วยโรเคเรื้อรัง” เป็นต้น

จ. ข้อผิดอื่น ๆ ที่เกิดจากการถ่ายถอดเบรลล์ไทย

นอกจากนี้ยังเกิดข้อผิดประเภทอื่น ๆ นอกเหนือจากที่กล่าวมาข้างต้น ได้แก่การถ่ายถอด ‘⠽⠎’ (ษ หรือ -ส) ‘⠽⠎⠓’ (ศ ห หรือ ธ) ‘⠽⠎’ (ศ หรือ ") ‘⠽⠎’ (ิ หรือ) (วงเล็บปิด)) ‘⠽⠎’ (ิ หรือ “) ซึ่งเกิดจากการที่ระบบถ่ายถอดสายอักขระเบรลล์ที่มีอักษรเหล่านี้ปรากฏอยู่แล้วนำสายอักขระที่มีค่าทางสถิติสูงสุดมาใช้เป็นผลลัพธ์ของการถ่ายถอดอักษร ซึ่งสายอักขระดังกล่าวไม่ตรงกับผลลัพธ์ที่แท้จริง

6.4.2.2 ข้อผิดในการถ่ายถอดเบรลล์อังกฤษ

จากการทดสอบระบบถ่ายถอดเบรลล์ไทยปนอังกฤษ ผู้วิจัยพบข้อผิดจากการถ่ายถอดเบรลล์ไทยปนอังกฤษรูปเต็มและรูปย่อ ซึ่งจะขออภิปรายดังนี้

ก. การถ่ายถอด ‘⠉’ (-) ในตำแหน่งต้นคำไม่ถูกต้อง

ในเบรลล์อังกฤษรูปย่อ หาก ‘⠉’ ปรากฏต้นคำ อาจใช้ย่อ “com” ได้ในกรณีทั่วไป หรืออาจใช้แทน “ - ”ในบางกรณี ดังที่ปรากฏในการเขียนคำประเภทวิภัตและปัจจัย เช่น “-ation” และ “-s”

ปัญหาที่ยังคงไม่สามารถแก้ไขได้ในการถ่ายถอดเบรลล์อังกฤษรูปย่อด้วยแบบจำลองเอ็นแกรม คือ ระบบมักจะถ่ายถอดคำที่ขึ้นต้นด้วย ‘⠉’ เป็น “com” เสมอ เนื่องจากการตีความให้ ‘⠉’ เป็น “com” นั้นมีโอกาสที่จะเกิดขึ้นได้มากกว่าตีความให้เป็น “-” เช่นถ่ายถอด ‘⠉⠕⠍⠬’ (= -ing) เป็น “coming” ไม่ใช่ “-ing” ซึ่งเป็นคำที่ถูกต้อง ข้อผิดประเภทนี้เกิดกับข้อความทั้งที่เป็นอังกฤษล้วนและข้อความไทยปนอังกฤษ เนื่องจากข้อความบางส่วนที่ผู้วิจัยนำมาทดสอบระบบมีเนื้อหาเกี่ยวกับไวยากรณ์อังกฤษ จึงทำให้พบข้อผิดเช่นนี้ขึ้น

ข. การถ่ายถอดตัวอักษรเป็นคำย่อ

ดังที่ได้อภิปรายมาแล้วว่า ในอังกฤษเบรลล์รูปย่อ มีการนำอักขระเบรลล์ 1 ตัวหรือมากกว่านั้นมาใช้แทนคำโดดในภาษาอังกฤษ ซึ่งจะพบได้จากการนำอักขระเบรลล์ที่ใช้แทนตัวอักษรมาใช้แทนคำ 1 คำ เช่น '⠸⠃⠸' ใช้เขียนคำว่า "|but|" นอกจากนี้ ยังมีการนำอักขระเบรลล์มากกว่า 1 ตัวมาใช้ย่อคำโดด เช่น '⠸⠁⠃⠸' ใช้แทน "|about|" เป็นต้น อย่างไรก็ตาม เมื่ออักขระหรือกลุ่มอักขระเบรลล์เหล่านี้ปรากฏภายในคำส่วนใหญ่ต้องตีความให้เป็นตัวอักษรโดดเช่น'⠃⠗⠊⠍' (= b-r-i-m) ไม่สามารถตีความให้เป็น "butrim" ได้ จากการทดลองถ่ายถอดข้อความอังกฤษ พบว่าระบบได้ถ่ายถอดอักษรเป็นตัวย่ออักษรโดด เช่น ระบบถ่ายถอด "⠙⠺⠑⠇⠇" (dwell) เป็น "dowell" เนื่องจากเมื่อ 'd' ปรากฏต้นสายอักขระ มีคู่เทียบในอักษรใน 2 กรณีคือ "d" และ "do" แต่สายอักขระที่มีค่าทางสถิติสูงสุด = "|dowell|" นอกจากนี้ยังพบการถ่ายถอดผิดในลักษณะนี้อีก เช่น ถ่ายถอด "dg40" เป็น "dog40" เป็นต้น

ค. การถ่ายถอดคำย่อเป็นตัวอักษร

ในข้อความเบรลล์อังกฤษรูปย่อ เมื่อตัวอักษรอังกฤษปรากฏระหว่างเครื่องหมายเว้นวรรคและ/หรือเครื่องหมายที่สามารถแสดงขอบเขตของคำจะใช้แทนคำโดด เช่น '⠸⠃⠸' = |but| , '⠰⠃⠸' = "but| ในการถ่ายถอดตัวย่อเหล่านี้ไม่เป็นปัญหา เมื่อตัวย่อปรากฏระหว่างเครื่องหมายเว้นวรรค แต่จะเกิดปัญหาในกรณีที่ตัวย่อเหล่านี้ปรากฏร่วมกับเครื่องหมายต่าง ๆ เช่น ระบบถ่ายถอด |⠰⠃⠸| (เท่ากับ "but) ถ่ายถอดเป็น "b เนื่องจาก |⠰⠃⠸| = |"b|, |"but| ซึ่งระบบมักจะเลือกสายอักขระที่สั้นกว่าเป็นผลลัพธ์ของการถ่ายถอด ทำให้เกิดข้อผิดลักษณะนี้ขึ้นเป็นจำนวนมาก นอกจากนี้ยังเกิดข้อผิดลักษณะนี้กับคำย่อเบรลล์ เช่น '⠋⠗⠝' (friend.) ระบบเลือก "fr." มาเป็นผลลัพธ์ เป็นต้น

ในการออกแบบการถ่ายถอดเบรลล์อังกฤษด้วยแบบจำลองเอ็นแกรม ผู้วิจัยได้พยายามแก้ไขปัญหาที่เกิดขึ้นนี้โดยสร้างคลังไบแกรมคำอังกฤษขึ้นเพื่อหาผลลัพธ์ที่เป็นไปได้มากที่สุด แต่ดูเหมือนว่าจะแก้ไขปัญหาที่เกิดขึ้นได้เพียงบางส่วนเท่านั้น เนื่องจากยังเกิดข้อผิดในลักษณะดังกล่าวขึ้นอีกเป็นจำนวนมาก

ง.การถ่ายถอดตัวเลขเป็นตัวอักษร

ในการถ่ายถอดอักษร ระบบจะสร้างผลลัพธ์ที่เป็นไปได้จากสายอักขระเบรลล์นำเข้า โดยอาศัยตารางคู่เทียบตัวย่อสำหรับการถ่ายถอดอังกฤษรูปย่อ (ตามภาคผนวก ค.1) แต่เมื่อระบบพบ '⠼' (เครื่องหมายนำเลข) ระบบจะนำตารางคู่เทียบตัวเลขมาใช้ร่วมกับตารางคู่เทียบตัวย่อ เนื่องจาก '⠼' สามารถใช้เป็นตัวย่อ "ble" ได้ (เช่น'⠁⠼' ใช้แทน "able") และในขณะเดียวกันก็ใช้เป็นเครื่องหมาย

นำหน้าตัวอักษร “a - j” เพื่อบ่งชี้ว่าตัวอักษรนั้นใช้แทนตัวเลข (เช่น ‘⠁⠼’ ใช้เขียน a1) จึงทำให้ระบบได้ถ่ายถอดตัวเลข ผิดเป็นตัวอักษร เมื่อระบบตีความ ⠼ ผิดเป็น “ble” เป็นผลให้อักษรที่ตามมาถูกถ่ายถอดเป็นตัวอักษรด้วย ตัวอย่างเช่นในการถ่ายถอด ‘⠁⠼⠙’ (=a4) ระบบจะถ่ายถอดเป็น “abled ” แทนที่จะเป็น “a4”(ดูหัวข้อ 6.2.1.2 ค ประกอบ)

จ. การถ่ายถอดตัวอักษรเป็นตัวเลข

นอกจากระบบจะถ่ายถอดตัวเลขเป็นตัวอักษรดังที่แสดงไว้ข้างต้นแล้ว บางครั้งระบบก็ถ่ายถอดตัวอักษรอังกฤษเป็นตัวเลข เหตุที่เกิดข้อผิดในลักษณะนี้ขึ้นเนื่องจาก ในการถ่ายถอดตัวเลขนั้น ระบบจะใช้ตารางคู่เทียบตัวเลขในการจับคู่ ตัวเลขระบบหนึ่งกับอีกระบบหนึ่งจนกว่าพบว่าอักขระเบรลล์ปัจจุบันไม่ตรงกับอักษรเบรลล์ในตารางนี้ จึงจะหันไปใช้ตารางคู่เทียบตัวอักษรแทน

อย่างไรก็ตาม ในข้อความทดสอบมีบางสายอักขระมีโครงสร้าง *ตัวเลข +hyphen + คำ* เช่น “4-acre”, “2-ft” เป็นต้น แต่เนื่องจากในการเขียนตัวเลขที่คั่นกลางด้วย ‘⠤’ (hypen) ในอักษรเบรลล์นั้นไม่ต้องใช้ ‘⠼‘ (เครื่องหมายนำเลข) ข้างหลัง hyphen จึงทำให้ยากที่จะระบุได้ว่าอักขระที่ตามหลัง ‘⠤’ นั้น เป็นตัวอักษรหรือตัวเลข ถ้าเบรลล์นั้นเป็นอักขระในกลุ่ม a - j ซึ่งอาจตีความให้เป็นตัวเลขหรือตัวอักษรก็ได้ ด้วยเหตุนี้ระบบจึงยังคงใช้ตารางคู่เทียบตัวเลขกับอักขระเหล่านี้ แทนที่จะใช้ตารางคู่เทียบตัวอักษร จึงทำให้ถ่ายถอด “4-acre” เป็น “4-13re” หรือถ่ายถอด “2-ft” เป็น “2-6t” เป็นต้น

จากผลการทดลองถ่ายถอดอักษรด้วยวิธีการใช้กฎทางภาษาและการใช้แบบจำลองเอ็นแกรมที่อภิปรายมาแล้ว พอจะสรุปได้ว่า วิธีการทั้งสองนี้ให้ผลลัพธ์ในการระบุภาษาและการถ่ายถอดอักษรต่างกัน โดยในการระบุภาษานั้น การใช้กฎจะมีประสิทธิภาพมากที่สุดสำหรับข้อความที่ไม่ปนภาษา ส่วนการใช้แบบจำลองเอ็นแกรมจะให้ผลดีที่สุดกับข้อความที่มีการปนภาษา

สำหรับการถ่ายถอดอักษรนั้น การใช้แบบจำลองเอ็นแกรมจะถ่ายถอดเบรลล์ไทยได้ถูกต้องมากกว่าการใช้กฎ แต่ยังเกิดข้อผิดกับการถ่ายถอดอักขระเบรลล์ไทยบางตัวเป็นจำนวนมากคือ ‘⠾’ (เ◌ิ หรือ เ◌อ) ‘⠮’ (เ◌าะ หรือ อะ) และ ‘⠲’ (. หรือ ◌๊) ส่วนการถ่ายถอดเบรลล์อังกฤษนั้น พบว่าการใช้กฎจะมีประสิทธิภาพมากกว่าการใช้แบบจำลองเอ็นแกรม ทั้งในข้อความเบรลล์อังกฤษรูปเต็มและข้อความเบรลล์อังกฤษรูปย่อ

ดังนั้นในบทต่อไป ผู้วิจัยจะขอเสนอระบบการถ่ายถอดข้อความเบรลล์ไทยปนอังกฤษรูปเต็มและระบบการถ่ายถอดข้อความเบรลล์ไทยปนอังกฤษรูปย่อโดยวิธีการแบบผสม ซึ่งเกิดจากการผสมผสานกันระหว่างการใช้กฎและการใช้แบบจำลองเอ็นแกรม

บทที่ 7

ระบบการถ่ายถอดเบรลล์ไทยปนอังกฤษโดยใช้วิธีการแบบผสม

จากที่ได้อภิปรายมาแล้วว่า วิธีการถ่ายถอดอักษรโดยใช้กฎและแบบจำลองเอ็นแกรมก่อให้เกิดปัญหาแตกต่างกัน ดังนั้น ในบทนี้ผู้วิจัยจะนำเสนอระบบการถ่ายถอดเบรลล์ไทยปนอังกฤษโดยใช้วิธีการแบบผสม ซึ่งผู้วิจัยได้พัฒนาโดยนำข้อดีของการถ่ายถอดอักษรแต่ละวิธีมาใช้ในระบบ โดยในหัวข้อที่ 7.1 จะกล่าวถึงข้อมูลที่ใช้ในงานวิจัย หัวข้อที่ 7.2 จะกล่าวถึงระบบการถ่ายถอดอักษรที่ผู้วิจัยได้พัฒนาขึ้นซึ่ง ต่อมาในหัวข้อที่ 7.3 จะได้นำเสนอผลการทดสอบระบบ ในหัวข้อที่7.4 จะได้อภิปรายผลการทดสอบระบบ และในหัวข้อที่ 7.5 ผู้วิจัยจะเปรียบเทียบการถ่ายถอดเบรลล์ไทยปนอังกฤษทั้งสามวิธีที่ได้นำเสนอมาในงานวิจัยฉบับนี้ต่อไป

7.1 ข้อมูลที่ใช้ในงานวิจัย

ในการทำการทดลองการถ่ายถอดอักษรเบรลล์ให้เป็นอักษรปกติโดยใช้วิธีการแบบผสมนั้น ผู้วิจัยได้ใช้ข้อมูล 2 ประเภท ได้แก่ ข้อมูลที่ใช้กับระบบถ่ายถอดอักษร และข้อมูลที่นำมาใช้ทดสอบระบบ ซึ่งแต่ละประเภทประกอบด้วยข้อมูลต่าง ๆ ดังนี้

7.1.1 ข้อมูลที่ใช้กับระบบถ่ายถอดอักษร

ข้อมูลที่ใช้ในการถ่ายถอดเบรลล์ไทยปนอังกฤษโดยวิธีแบบผสมแบ่งออกได้เป็น 4 ประเภท ได้แก่ ข้อมูลที่ใช้ในการระบุภาษา ข้อมูลที่ใช้ในการสร้างชุดสายอักขระภาษาไทย ข้อมูลที่ใช้ในการสร้างชุดสายอักขระภาษาอังกฤษ และข้อมูลที่ใช้ประมวลผลเพื่อหาผลลัพธ์ของการถ่ายถอดอักษร

7.1.1.1 ข้อมูลที่ใช้ในการระบุภาษา

ในการระบุภาษาโดยใช้วิธีการแบบผสมนั้น ระบบจะใช้กฎเพื่อระบุภาษาของสายอักขระเบรลล์เป็นอันดับแรก โดยมีการจัดเตรียมข้อมูลตามที่ได้กล่าวมาแล้วในหัวข้อที่ 5.1.1.1 ในการถ่ายถอดอักษร ระบบจะใช้สายอักขระที่เป็นเอกลักษณ์ของสองภาษาคู่กันสำหรับข้อความเบรลล์แต่ละประเภท ดังนี้

(1) การถ่ายถอดเบรลล์ไทยปนอังกฤษรูปเต็ม โดยใช้เซ็ตสายอักขระที่เป็นเอกลักษณ์เบรลล์ไทย และสายอักขระที่เป็นเอกลักษณ์เบรลล์อังกฤษรูปเต็ม ตามภาคผนวก ค.1 และ ค.2

(2) การถ่ายถอดเบรลล์ไทยปนอังกฤษรูปย่อ โดยใช้เซ็ตสายอักขระที่เป็นเอกลักษณ์เบรลล์ไทย และเซ็ตสายอักขระที่เป็นเอกลักษณ์เบรลล์อังกฤษรูปย่อ ตามภาคผนวกที่ ค.3 และ ค.4

7.1.1.2 ข้อมูลที่ใช้ในการสร้างชุดสายอักขระภาษาไทย

ในการสร้างชุดสายอักขระที่เป็นไปได้ในอักษรปกติ เพื่อนำมาใช้เป็นตัวเลือกสำหรับการถ่ายถอดเบรลล์ไทยนั้น ต้องอาศัยตารางคู่เทียบเบรลล์ไทยสำหรับการถ่ายถอดด้วยแบบจำลองเอ็นแกรมตามภาคผนวก ง ซึ่งมีรายละเอียดในการจัดทำตามที่ได้อภิปรายมาแล้วในหัวข้อที่ 6.1.1.1

อย่างไรก็ตาม การใช้คู่เทียบที่กล่าวมานี้กับการถ่ายถอดด้วยแบบจำลองเอ็นแกรมให้ผลความถูกต้องได้ในระดับหนึ่ง ตามที่ได้นำเสนอมาในหัวข้อที่ 6.3 แต่เนื่องจากอักขระเบรลล์บางชุดมีอักษรปกติที่ตรงกันหลายชุด จึงทำให้บ่อยครั้ง ระบบเลือกสายอักขระที่ไม่ถูกต้องมาเป็นผลลัพธ์ของการถ่ายถอดอักษร

จากการทดสอบระบบถ่ายถอดอักษรด้วยแบบจำลองเอ็นแกรม ผู้วิจัยพบว่าระบบมักจะถ่ายถอด ‘⠡⠁’ เป็น “อะ” ในตำแหน่งที่ควรจะเป็น “เาะ” และถ่ายถอด ‘⠡’ (ที่ใช้แทน เ◌ิ หรือ เ◌อ) ไม่ถูกต้อง เนื่องจากอักขระเบรลล์ทั้ง 2 คู่นี้มีคู่เทียบที่เป็นอักษรปกติหลายชุด ดังนั้นเมื่อระบบพบอักขระทั้ง 2 ชุดนี้ ระบบจะพยายามลดคู่เทียบในอักษรปกติให้น้อยลง โดยให้เหลือเฉพาะคู่เทียบที่เหมาะสมเท่านั้นซึ่งจะทำได้โดยการใช้กฎ (ซึ่งกฎเหล่านี้ต้องใช้ข้อมูลในข้อที่ 5.1.1.2.2 ประกอบ) ก่อนที่จะดำเนินการสร้างชุดสายอักขระภาษาไทยต่อไป สำหรับการสร้างชุดสายอักขระภาษาไทย ผู้วิจัยจะได้นำเสนอในหัวข้อที่ 7.2.3 ต่อไป

7.1.1.3 ข้อมูลที่ใช้ในการสร้างชุดสายอักขระภาษาอังกฤษ

ในการสร้างชุดสายอักขระอังกฤษนั้น ระบบจะ*ใช้กฎ*กับการถ่ายถอดส่วนต้นและส่วนท้ายของสายอักขระ จึงต้องใช้ตารางคู่เทียบเบรลล์อังกฤษในหัวข้อ 5.1.1.3 มาช่วยในการถ่ายถอด เมื่อถ่ายถอดมาถึงส่วนกลาง ระบบจึงจะนำส่วนกลางมาสร้างชุดสายอักขระโดยใช้แบบจำลองเอ็นแกรม โดยอาศัยตารางคู่เทียบตามภาคผนวก ฉ. เนื่องจากคู่เทียบที่มีอยู่มีปริมาณมากเกินความจำเป็นซึ่งจะทำให้ระบบถ่ายถอดอักษรผิดได้ จึงต้องปรับเปลี่ยนตารางคู่เทียบให้เหมาะสมกับการถ่ายถอดด้วยวิธีการแบบผสม ซึ่งคู่เทียบที่แก้ไขเพิ่มเติมนั้นได้ระบุไว้ในภาคผนวก ช.

7.1.1.4 ข้อมูลไตรแกรมตัวอักษรภาษาไทยและภาษาอังกฤษ

ในการทำงานของระบบถ่ายถอดอักษรแบบผสมนั้น ระบบยังต้องใช้แบบจำลองไตรแกรมตัวอักษร เพื่อหาค่าทางสถิติสูงสุดในชุดสายอักขระที่นำมาประมวลผลแล้วนำไปใช้เป็นผลลัพธ์ของการถ่ายถอดอักษร โดยใช้คลังไตรแกรมตัวอักษรไทยในข้อที่ 6.1.1.2 และคลังไตรแกรมตัวอักษรอังกฤษในข้อที่ 6.1.2.2 มาใช้กับระบบด้วย

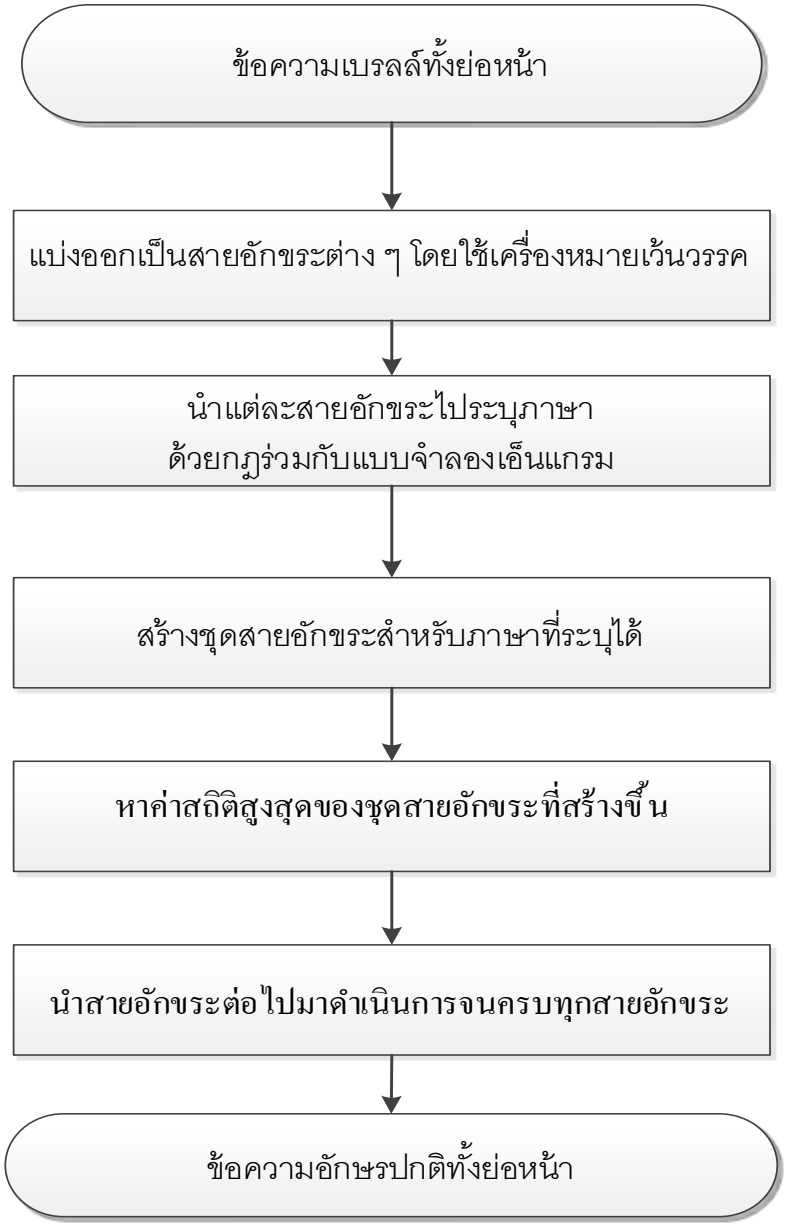
7.1.2 ข้อมูลที่ใช้ในการทดสอบระบบ

ในการทดสอบระบบ ผู้วิจัยได้นำข้อมูลเบรลล์ 5 ประเภทที่ใช้ในการทดสอบกับวิธีการถ่ายถอดด้วยกฎและการถ่ายถอดด้วยแบบจำลองเอ็นแกรม มาใช้กับการถ่ายถอดด้วยวิธีการแบบผสม การใช้ข้อมูลชุดเดียวกันนี้ จะทำให้สะดวกต่อการเปรียบเทียบประสิทธิภาพของการถ่ายถอดอักษรด้วย 3 วิธีนี้ได้อย่างเป็นรูปธรรมและเกิดความคลาดเคลื่อนน้อย

7.2 ระบบการถ่ายถอดอักษรโดยใช้วิธีการแบบผสม

ในการถ่ายถอดเบรลล์ไทยปนอังกฤษด้วยวิธีการแบบผสมนั้น จะแยกระบบการถ่ายถอดออกเป็น 2 ระบบ คือ ระบบถ่ายถอดเบรลล์ไทยปนอังกฤษรูปเต็ม และระบบถ่ายถอดเบรลล์ไทยปนอังกฤษรูปย่อ โดยที่แต่ละระบบมีภาพรวมในการทำงานคล้ายคลึงกัน แต่จะแตกต่างกันในรายละเอียดการดำเนินการ ซึ่งทั้ง 2 ระบบนี้มีภาพรวมในการดำเนินการดังนี้

ในการทำงานของระบบถ่ายถอดเบรลล์ไทยปนอังกฤษด้วยวิธีการแบบผสมนั้น จะเริ่มจากนำเข้าข้อความเบรลล์ทีละย่อหน้า แล้วนำแต่ละย่อหน้าไปแบ่งออกเป็นสายอักขระต่าง ๆ จากนั้นนำแต่ละสายอักขระไปผ่านการระบุภาษาด้วยกฎร่วมกับแบบจำลองเอ็นแกรม ตามหัวข้อที่ 7.2.1 แล้วสร้างชุดสายอักขระสำหรับผลลัพธ์ที่เป็นไปได้ตามภาษาที่ระบุได้ คือ ชุดสายอักขระภาษาไทยตามหัวข้อที่ 7.2.2 และชุดสายอักขระภาษาอังกฤษตามหัวข้อที่ 7.2.3 จากนั้นนำชุดสายอักขระที่สร้างขึ้นไปหาสายอักขระที่มีค่าทางสถิติสูงสุดด้วยแบบจำลองไตรแกรมตัวอักษรเพื่อใช้เป็นผลลัพธ์ของการถ่ายถอดของสายอักขระปัจจุบัน



ภาพที่ 7.1 ภาพรวมของระบบถ่ายถอดเบรลล์ไทยปนอังกฤษโดยใช้วิธีการแบบผสม

7.2.1 การระบุภาษา

ผลการทดลองในบทที่ 5 และบทที่ 6 ชี้ให้เห็นว่า การระบุภาษาด้วยกฎและแบบจำลองเอ็นแกรมให้ผลต่างกัน คือ การระบุภาษาด้วยกฎจะให้ผลดีกับข้อความที่มีการสลับภาษาไม่บ่อยนัก แต่จะเกิดข้อเสียเมื่อข้อความนั้นมีการสลับภาษาถี่ ๆ เช่น ตำราสอนภาษาอังกฤษที่มีคำศัพท์อังกฤษ คำแปลไทย คำอธิบายวิธีการใช้ภาษา เป็นต้น

อย่างไรก็ตาม การระบุภาษาด้วยแบบจำลองเอ็นแกรมนั้น ระบบถ่ายถอดเบรลล์ไทยปนอังกฤษจะระบุภาษาได้ดีกับสายอักขระที่มีขนาดยาวมากกว่า 14 อักขระ แต่จะระบุภาษาผิดในสายอักขระที่มีขนาดสั้น (คือระหว่าง 1-5 อักขระ) โดยมักจะระบุภาษาผิดจากอังกฤษเป็นไทยในข้อความเบรลล์อังกฤษ ดังนั้นผู้วิจัยจึงได้นำข้อดีของการระบุภาษาทั้ง 2 ระบบนี้มาใช้ร่วมกันในวิธีการถ่ายถอดแบบผสม

ในระบบถ่ายถอดเบรลล์ไทยปนอังกฤษรูปเต็ม และระบบถ่ายถอดเบรลล์ไทยปนอังกฤษรูปย่อโดยใช้วิธีการแบบผสม ระบบจะระบุภาษาของสายอักขระเบรลล์ที่กำลังดำเนินการอยู่ โดยเริ่มจากการหาจุดเริ่มต้นของการระบุภาษา ดังที่ได้แสดงไว้ในหัวข้อ 5.2.1.1 หลังจากที่ระบบพบจุดเริ่มต้นของการระบุภาษาแล้ว ระบบจะนำเบรลล์เซลล์ปัจจุบันและเบรลล์ 2 เซลล์ถัดไป (รวม 3 เซลล์) ไปตรวจสอบกับสายอักขระที่เป็นเอกลักษณ์ของภาษาไทยและสายอักขระที่เป็นเอกลักษณ์ของภาษาอังกฤษ

สำหรับการระบุภาษาโดยใช้สายอักขระที่เป็นเอกลักษณ์ของไทยปนอังกฤษรูปเต็ม จะใช้ตารางในภาคผนวก ค.1 และ ค.2 และสำหรับสายอักขระที่เป็นเอกลักษณ์ของไทยปนอังกฤษรูปย่อนั้น จะอาศัยตารางในภาคผนวก ค.3 และ ค.4 ดังที่ได้กล่าวไว้แล้วในการถ่ายถอดเบรลล์ไทยปนอังกฤษด้วยกฎ หากระบบระบุภาษาด้วยสายอักขระที่เป็นเอกลักษณ์ได้ ระบบจะดำเนินการสร้างชุดสายอักขระสำหรับภาษานั้น แต่หากระบบระบุภาษาด้วยวิธีการนี้ไม่ได้ ระบบจะสร้างชุดสายอักขระทั้งไทยและอังกฤษ

ในการระบุภาษาของระบบถ่ายถอดอักษรโดยใช้วิธีการแบบผสม ระบบจะดำเนินการในรายละเอียด โดยอาศัยเงื่อนไขต่อไปนี้

(1) ถ้าเบรลล์ 1 เซลล์ 2 เซลล์ หรือ 3 เซลล์ปัจจุบัน ตรงกับสายอักขระที่เป็นเอกลักษณ์ของภาษาไทยหรือตรงกับสายอักขระที่เป็นเอกลักษณ์ของภาษาอังกฤษ ระบบจะนำสายอักขระเบรลล์ปัจจุบันไปสร้างชุดสายอักขระในภาษาที่พบว่าตรงกัน แต่ถ้าตรวจสอบแล้วไม่ตรงกับสายอักขระในภาษาใดเลย ระบบจะดำเนินการในขั้นต่อไป

(2) ถ้าสายอักขระเบรลล์ปัจจุบันมีความยาวไม่เกิน 14 อักขระ ระบบจะนำสายอักขระเบรลล์ปัจจุบันไปสร้างชุดสายอักขระตามภาษาที่ระบุได้ก่อนหน้านี้ แต่ถ้าไม่ตรงตามเงื่อนไขที่กล่าวมา ระบบจะดำเนินการในขั้นต่อไป

(3) ในกรณีอื่น ๆ ระบบจะสร้างชุดสายอักขระทั้งภาษาไทยและภาษาอังกฤษ

จากเงื่อนไขข้างต้น เหตุที่ผู้วิจัยกำหนดให้ใช้ความยาว 14 ตัวอักษรในการระบุภาษาเมื่อระบบไม่พบสายอักขระที่เป็นเอกลักษณ์ของแต่ละภาษา เนื่องจากผู้วิจัยพบว่ามีคำอังกฤษที่ระบบระบุผิดให้เป็นภาษาไทย ซึ่งมีความยาวมากที่สุดคือ 14 ตัวอักษร ผู้วิจัยจึงใช้ความยาวดังกล่าวนี้มากำหนดเงื่อนไขสำหรับการระบุภาษา

หลังจากที่ระบบระบุภาษาสำหรับแต่ละสายอักขระได้แล้ว หากระบุได้ว่าสายอักขระนั้นเป็นภาษาไทยระบบจะดำเนินการตามกระบวนการในหัวข้อ 7.2.2 แต่หากระบุได้ว่าเป็นภาษาอังกฤษจะดำเนินการตามกระบวนการในหัวข้อ 7.2.3 โดยแต่ละกระบวนการมีการทำงานดังนี้

7.2.2 การสร้างชุดสายอักขระสำหรับภาษาไทย

ในการสร้างชุดสายอักขระสำหรับภาษาไทยนั้น ผู้วิจัยได้*นำกฎมาใช้ร่วมกับการสร้างชุดสายอักขระภาษาไทย* เนื่องจากผู้วิจัยพบว่า ในการถ่ายถอดเบรลล์ไทยด้วยแบบจำลองเอ็นแกรมนั้น เกิดข้อผิดกับการถ่ายถอดอักขระ 3 ชุดต่อไปนี้เป็นจำนวนมากคือ ‘⠽⠽⠀’ (ซึ่งมักจะถ่ายถอดให้เป็น “อะ” ในตำแหน่งที่ควรเป็น “เ◌าะ”) ‘⠀⠽’ (ซึ่งมักถ่ายถอดเป็น “เ◌อ” แทนที่จะเป็น “เ◌ิ”) และ ‘⠽⠽’ (มักถ่ายถอดผิดเป็น “◌ั้” ซึ่งส่วนใหญ่เกิดกับสายอักขระที่มีความยาว 2 ตัวอักษร เช่น “ก.” “ข.” เป็นต้น)

ผู้วิจัยจึงจำเป็นต้องนำกฎมาช่วยในการสร้างคู่เทียบที่เหมาะสมสำหรับอักขระ 3 ชุดนี้ เนื่องจากผลการทดสอบระบบโดยใช้กฎแสดงให้เห็นว่า การนำกฎมาใช้ถ่ายถอดอักษร 3 ชุดนี้จะให้ผลความถูกต้องมากกว่าการถ่ายถอดด้วยแบบจำลองเอ็นแกรม ผู้วิจัยจึงได้นำกฎมาเลือกคู่เทียบที่เหมาะสม เมื่อระบบพบอักขระ 3 ชุดนี้

ในการสร้างชุดสายอักขระภาษาไทย ระบบจะนำสายอักขระเบรลล์ที่กำลังดำเนินการมาสร้างชุดสายอักขระไทยปกติ โดยในขั้นแรกระบบจะตรวจสอบ 3 อักขระนับตั้งแต่เบรลล์เซลล์ปัจจุบันว่าตรงกับเงื่อนไขใดเงื่อนไขหนึ่งที่จะกล่าวต่อไปหรือไม่ หากไม่เป็นไปตามเงื่อนไขใด ระบบจะลดตำแหน่งเบรลล์เซลล์ทางขวาลงครั้งละ 1 เซลล์จนกว่าจะตรงกับเงื่อนไขใดเงื่อนไขหนึ่ง จึงจะดำเนินการตามเงื่อนไขนั้น ๆ จากนั้นนำเบรลล์ชุดต่อไปมาดำเนินการตั้งแต่ขั้นแรกไปเรื่อย ๆ จนสิ้นสุดสายอักขระปัจจุบัน ซึ่งในการสร้างชุดสายอักขระภาษาไทยต้องอาศัยเงื่อนไขดังนี้

(1) เงื่อนไขที่ใช้กับ ‘ ’

ถ้าอักขระเบรลล์ที่นำมาตรวจสอบเป็น ‘ ’ ระบบจะนำชุดสายอักขระที่กำลังดำเนินการ มาหาค่าสถิติสูงสุดด้วยแบบจำลองเอ็นแกรมตัวอักษร จากนั้นตรวจสอบว่า ‘ ’ ควรใช้แทน “อะ” หรือ “เาะ” โดยนำส่วนท้ายของสายอักขระที่มีค่าทางสถิติสูงสุดมาใช้เป็นบริบทซ้าย และอักขระเบรลล์หลัง ‘ ’ มาเป็นบริบทขวา แล้วนำไปตรวจสอบกับเงื่อนไขในการถ่ายถอดเบรลล์เซลล์คู่ที่กำกวมตามหัวข้อ 5.2.2.1.1 ก. ถ้าตรวจสอบแล้วระบุได้ว่า ‘ ’ ใช้แทน “อะ” ระบบจะกำหนดคู่เทียบสำหรับ ‘ ’ ให้เป็น “อะ” แต่ถ้าตรวจสอบแล้วพบว่าใช้แทน “เาะ” ระบบจะกำหนดคู่เทียบ เป็น “|เาะ” และ “||เาะ” แล้วนำคู่เทียบที่ได้ไปดำเนินการตามกระบวนการสร้างชุดสายอักขระสำหรับภาษาไทยตามหัวข้อ 6.2.1.1 จากนั้นนำอักขระเบรลล์ถัดจาก ‘ ’ มาดำเนินการตั้งแต่ขั้นแรก

ตัวอย่างเช่นการถ่ายถอด ‘ ’(=ไ-ด้--อ-ะ-ไ-ร) ระบบจะดำเนินการดังนี้

{1} ‘ ’ (ได้อะไร)

= ไ

= ได

= ได้ , ได. (โดยที่ = . หรือ ้)

เมื่อถึงขั้นนี้จะได้ 2 สายอักขระเป็น “ได้” และ “ได.” อักขระที่ต้องดำเนินการต่อ คือ ‘ ’ เมื่อระบบนำชุดอักขระนี้ไปหาความน่าจะเป็นสูงสุดด้วยแบบจำลองเอ็นแกรมจะได้สายอักขระที่เป็นผลลัพธ์คือ “ได้” ซึ่งสายอักขระนี้จะนำมาใช้เป็นบริบทซ้ายของ ‘ ’ และอักษรที่ถัดจาก ‘ ’ คือ ‘ ’ (=ไร) จะนำมาใช้เป็นบริบทขวา เมื่อได้บริบทซ้าย-ขวาแล้ว ระบบจะนำ ‘ ’ พร้อมบริบทซ้ายและบริบทขวาไปตรวจสอบกับเงื่อนไขในหัวข้อ 5.2.2.1.1 ก. หลังจากตรวจสอบกับเงื่อนไขแล้วจึงระบุได้ว่า ‘ ’ มีคู่เทียบที่เหมาะสมที่สุดคือ “อะ” ต่อมาจึงนำ “อะ” ไปดำเนินการตามกระบวนการสร้างชุดสายอักขระไทยตามหัวข้อ 6.2.1.1

(2) เงื่อนไขที่ใช้กับ ‘ ’

ถ้าอักขระเบรลล์ที่นำมาตรวจสอบเป็น ‘ ’ (เ◌อหรือ เ◌ิ) ระบบจะนำไปตรวจสอบว่าควรจะถ่ายถอดให้เป็น "เ◌อ" หรือ "เ◌ิ" โดยอาศัยเงื่อนไขการถ่ายถอดเบรลล์เซลล์เดี่ยวที่กำกวมตามหัวข้อ 5.2.2.2.1 ก. ซึ่งในการตรวจสอบเงื่อนไขเพื่อหารูปแทนที่เหมาะสมนั้น ต้องอาศัยอักขระเบรลล์ที่ปรากฏหลัง ‘ ’ ในสายอักขระเบรลล์ปัจจุบันเป็นบริบทขวา หากตรวจสอบแล้วพบว่า ‘ ’ ใช้แทน "เ◌ิ" ระบบจะสร้างคู่เทียบสำหรับอักขระนี้เป็น "|เ◌ิ" และ "||เ◌ิ" แต่หากเป็นอีกกรณีหนึ่ง ระบบจะสร้างคู่

เทียบเป็น "|เ◌อ" และ "||เ◌อ" จากนั้น นำคู่เทียบที่ได้ไปดำเนินการตามกระบวนการข้อ 6.2.1 แล้วจึงนำอักขระเบรลล์ถัดจาก ‘⠍’ มาดำเนินการตั้งแต่ขั้นแรกในส่วนนี้ ตัวอย่างเช่นการถ่ายถอด ‘⠍⠊⠩⠎’ (=ม-ี-ธ-เ อ) ระบบจะดำเนินการดังนี้

{2} ‘⠍⠊⠩⠎’ (มีเธอ)

⠍ = ม

⠍⠊ = มี

⠍⠊⠩⠎ = มีธ , มี◌์ท (⠩ = ธ หรือ ◌์ท)

เมื่อดำเนินการถึงขั้นนี้ ระบบจะอยู่ที่เบรลล์เซลล์ปัจจุบันคือ ‘⠍’ ระบบจะนำตัวอักษร 1-3 ตัวถัดจาก ‘⠍’ มาใช้เป็นบริบททางขวา (ในกรณีนี้ไม่มีอักษรใดปรากฏเป็นบริบทขวา) แล้วนำอักขระเบรลล์นี้พร้อมบริบทขวาไปตรวจสอบกับเงื่อนไขการถ่ายถอดเบรลล์เซลล์เดี่ยวที่กำกวม ตามหัวข้อ 5.2.2.2.1 ก. ทำให้ได้คู่เทียบเป็น "|เ◌อ" และ "||เ◌อ" จากนั้นระบบจะนำชุดสายอักขระข้างต้นคือ “มีธ” และ “มี่ท” พร้อมคู่เทียบที่ได้ใหม่ไปดำเนินการในข้อ 6.2.1.1 ต่อไป

(3) เงื่อนไขที่ใช้กับ ‘⠲’

หากเบรลล์เซลล์ปัจจุบันเป็น ‘⠲’ (◌้ หรือ .) ซึ่งปรากฏในตำแหน่งสุดท้ายของสายอักขระและมีอักษรเพียง 1 ตัวปรากฏหน้า ‘⠲’ ระบบจะสร้างคู่เทียบสำหรับ ‘⠲’ เป็น “.” แล้วนำคู่เทียบที่ได้ใหม่ไปดำเนินการตามกระบวนการสร้างชุดสายอักขระภาษาไทยตามหัวข้อ 6.2.1.1 จากนั้นสิ้นสุดกระบวนการสร้างชุดสายอักขระไทยสำหรับสายอักขระเบรลล์ปัจจุบัน

(4) เงื่อนไขที่ใช้กับ ‘⠼’

ถ้าเบรลล์เซลล์ปัจจุบันเป็น ‘⠼’ (เครื่องหมายนำเลข) ระบบจะนำอักขระเบรลล์ปัจจุบันจนถึงอักขระเบรลล์ตัวสุดท้ายในสายอักขระปัจจุบันไปถ่ายถอดตามกระบวนการถ่ายถอดตัวเลขในหัวข้อ 5.2.3.3.3 จากนั้นนำผลลัพธ์ที่ได้ไปดำเนินการตามกระบวนการสร้างชุดสายอักขระภาษาไทยในหัวข้อ 6.2.1.1 แล้วสิ้นสุดกระบวนการสร้างชุดสายอักขระไทยสำหรับสายอักขระเบรลล์ปัจจุบัน

(5) เงื่อนไขที่ใช้กับตารางคู่เทียบเบรลล์ไทย

ถ้าอักขระเบรลล์ที่นำมาตรวจสอบตรงกับตารางคู่เทียบเบรลล์ไทย ที่ใช้กับวิธีการถ่ายถอดด้วยแบบจำลองเอ็นแกรมตามภาคผนวก ง ระบบจะนำคู่เทียบที่ได้จากตารางไปดำเนินการตามกระบวนการ

สร้างชุดสายอักขระตามหัวข้อ 6.2.1.1 หลังจากนั้นจึงดำเนินการกับอักขระเบรลล์ต่อไปตั้งแต่ขั้นแรกในส่วนของการสร้างชุดสายอักขระตามวิธีการแบบผสม

(6) เงื่อนไขในการสลับที่อักษร 2 ตัว หน้าสระประสม

เมื่อระบบพบ "‖" ปรากฏในคู่เทียบอักษรปกติ ที่ตรงกับอักษรเบรลล์ที่กำลังถ่ายถอด ระบบต้องตรวจสอบว่า อักษร 2 ตัวที่อยู่หน้าอักขระเบรลล์นั้น ๆ สามารถปรากฏระหว่างสระประสมได้หรือไม่ โดยนำอักษร 2 ตัวนี้ไปตรวจสอบกับเซ็ตของพยัญชนะควบตามตารางที่ 5.14 ถ้าตรงกับเซ็ตของพยัญชนะควบในตารางแสดงว่าอักษร 2 ตัวนี้สามารถปรากฏระหว่างสระประสมได้ ระบบจึงจะทำการสลับที่ตัวอักษร แต่หากไม่ตรงกับเซ็ตดังกล่าว ระบบจะไม่ดำเนินการในส่วนนี้

ดังที่ได้กล่าวมาแล้วว่า ระบบจะนำกระบวนการสร้างชุดสายอักขระในหัวข้อ 6.2.1.1 มาใช้ในการสร้างชุดสายอักขระสำหรับภาษาไทย แต่ในบางครั้งคู่เทียบที่มีเครื่องหมาย "‖" ก็ก่อให้เกิดปัญหาต่อการถ่ายถอดอักษร ในการสร้างชุดสายอักขระเมื่อระบบพบเครื่องหมายคู่นี้ ระบบจะต้องสลับที่ตัวอักษร 2 ตัวหน้าสระประสม กับอักษรตัวแรกในคู่เทียบที่เป็นสระประสม แต่ใช่ว่าตัวอักษรทุกคู่จะปรากฏระหว่างสระประสมได้เสมอไป หากเกิดการสลับตัวอักษรที่ไม่สามารถปรากฏระหว่าง สระประสมได้ก็จะเกิดข้อผิดขึ้น

ผลจากการทดสอบระบบที่ถ่ายถอดด้วยแบบจำลองเอ็นแกรม แสดงให้เห็นว่ามีการนำตัวอักษร 2 ตัวที่ไม่เหมาะสมกับบริบทมาแทรกในสระประสม เช่น ถ่ายถอด"ดื่มนมเยอะ" เป็น "ดื่มนเมยอะ" เนื่องจากระบบได้ทำการสลับที่ระหว่าง "มย" กับ "เ" ซึ่งเป็นส่วนหนึ่งของ "เ◌อะ"

ฉะนั้น ก่อนที่ระบบจะทำการสลับอักษร 2 ตัวหน้าสระประสม ระบบต้องทำการตรวจสอบว่าสามารถนำอักษร 2 ตัวนี้มาแทรกระหว่างสระประสมได้หรือไม่ โดยอาศัยเซ็ตของพยัญชนะควบตามตารางที่ 5.14 จึงจะไม่เกิดปัญหาในการถ่ายถอดเช่นที่กล่าวมานี้

7.2.3 การสร้างชุดสายอักขระสำหรับภาษาอังกฤษ

ในการสร้างชุดสายอักขระสำหรับภาษาอังกฤษนั้น ระบบถ่ายถอดเบรลล์ไทยปนอังกฤษรูปเต็มและระบบถ่ายถอดเบรลล์ไทยปนอังกฤษรูปย่อจะใช้วิธีการที่ต่างกัน

ในระบบถ่ายถอดเบรลล์ไทยปนอังกฤษรูปเต็ม ระบบจะใช้กระบวนการเดียวกับการถ่ายถอดไทยปนอังกฤษรูปเต็มด้วยวิธีการใช้กฎตามที่นำเสนอมาแล้วในบทที่ 5 ในหัวข้อ 5.2.3 ส่วนระบบถ่ายถอดเบรลล์ไทยปนอังกฤษรูปย่อนั้น ระบบจะดำเนินการตามหัวข้อ 5.2.3 เช่นกัน แต่เมื่อถ่ายถอด

มาถึงเบรลล์ส่วนกลางสายอักขระย่อยตามกระบวนการในหัวข้อ 5.2.3.3.4 ขั้นที่ 3 นั้น ระบบจะนำกระบวนการสร้างชุดสายอักขระสำหรับภาษาอังกฤษตามหัวข้อ 6.2.1.2 มาใช้แทนการถ่ายถอดส่วนกลางสายอักขระ โดยอาศัยตารางคู่เทียบเบรลล์รูปย่อ ตามภาคผนวก ฉ ที่นำมาแก้ไขเพิ่มเติมโดยอาศัยภาคผนวก ช ซึ่งจะทำให้เกิดสายอักขระอังกฤษขึ้นจำนวนหนึ่ง ที่ระบบจะนำไปหาค่าทางสถิติด้วยแบบจำลองเอ็นแกรมในลำดับต่อไป

7.2.4 การหาค่าสถิติสูงสุดด้วยแบบจำลองเอ็นแกรม

ในขั้นนี้ระบบจะนำชุดสายอักขระที่ได้รับจากกระบวนการอื่น มาคำนวณหาสายอักขระที่มีค่าสถิติสูงสุดด้วยแบบจำลองเอ็นแกรม ตามกระบวนการในหัวข้อที่ 6.2.4 แล้วนำสายอักขระที่ได้ส่งให้กระบวนการที่ต้องใช้งานต่อไป

7.3 การทดสอบระบบการถ่ายถอดอักษรโดยใช้วิธีการแบบผสม

ในการทดสอบระบบ ผู้วิจัยได้นำข้อความเบรลล์ 5 ประเภทซึ่งเป็นข้อความเดียวกับที่นำมาใช้ทดสอบการถ่ายถอดอักษรด้วยกฎที่ได้นำเสนอมาแล้วในหัวข้อที่ 5.1.2 โดยทดสอบกับกระบวนการต่าง ๆ ได้แก่ กระบวนการระบุภาษา กระบวนการถ่ายถอดเบรลล์ไทย และกระบวนการถ่ายถอดเบรลล์อังกฤษ โดยใช้วิธีการประเมินผลเช่นเดียวกับการถ่ายถอดด้วยกฎแต่เปลี่ยนข้อมูลที่ใช้ในการทดสอบระบบเป็น ข้อมูลที่มาจากการถ่ายถอดด้วยวิธีการแบบผสม

7.3.1 การทดสอบระบบกับข้อความเบรลล์ไม่ปนภาษา

ในการทดสอบประสิทธิภาพของการระบุภาษาโดยใช้วิธีการแบบผสม วิธีการทดสอบจะแยกเป็น (1) การทดสอบประสิทธิภาพของกระบวนการระบุภาษาและกระบวนการถ่ายถอดอักษรแต่ละประเภท และ (2) การทดสอบประสิทธิภาพของระบบกับข้อมูลจริงที่มีการปนภาษา ซึ่งผู้วิจัยจะขอนำเสนอผลในการทดสอบทั้ง 2 ส่วน ดังนี้

(1) การระบุภาษา

ในการทดสอบระบบโดยใช้ข้อความไทยล้วนและอังกฤษล้วน ผู้วิจัยได้นำข้อความไทยล้วน อังกฤษรูปเต็ม และอังกฤษรูปย่อไปทดสอบการถ่ายถอดของระบบ ซึ่งได้ผลดังนี้

ตารางที่ 7.1 ผลการทดสอบประสิทธิภาพของการระบุภาษาโดยใช้วิธีการแบบผสม

ระบบที่ใช้ทดสอบ	ประเภทข้อความแบรลล์	สายอักขระที่ใช้ทดสอบ	สายอักขระที่ระบุภาษาผิด	ระบุได้ถูกต้อง (%)
ระบบไทยปนอังกฤษรูปเต็ม	ไทย	12,913	0	100
ระบบไทยปนอังกฤษรูปย่อ	ไทย	12,913	31	99.76
ระบบไทยปนอังกฤษรูปเต็ม	อังกฤษรูปเต็ม	59,072	0	100
ระบบไทยปนอังกฤษรูปย่อ	อังกฤษรูปย่อ	56,206	1	99.99

(2) การถ่ายถอดอักษร

ในการทดสอบประสิทธิภาพของการถ่ายถอดอักษร ได้ผลการทดสอบดังนี้

ตารางที่ 7.2 ผลการทดสอบประสิทธิภาพของการถ่ายถอดอักษรโดยใช้วิธีการแบบผสม

ประสิทธิภาพในการถ่ายถอด	ประเภทข้อความแบรลล์	จำนวนคำที่ใช้ทดสอบ	จำนวนคำที่ถ่ายถอดผิด	ถ่ายถอดได้ถูกต้อง (%)
เบรลล์ไทย	ไทยล้วน	19,090	51	99.73
เบรลล์อังกฤษรูปเต็ม	อังกฤษรูปเต็ม	59,057	0	100
เบรลล์อังกฤษรูปย่อ	อังกฤษรูปย่อ	59,057	49	99.92

7.3.2 การทดสอบระบบกับข้อความเบรลล์ปนภาษา

ในการทดลอง ผู้วิจัยได้ทดสอบระบบถ่ายถอดอักษรกับข้อมูลจริง 2 ประเภทคือ ข้อความเบรลล์ไทยปนอังกฤษรูปเต็ม และข้อความเบรลล์ไทยปนอังกฤษรูปย่อ โดยในการประเมินประสิทธิภาพของระบบจะแบ่งออกเป็น 2 ส่วน ได้แก่การระบุภาษาและการถ่ายถอดอักษร

(1) ผลการระบุภาษา

การทดสอบประสิทธิภาพของระบบในส่วนของการระบุภาษาแสดงได้ดังนี้

ตารางที่ 7.3 ผลการระบุภาษาของระบบที่ทดสอบกับข้อมูลปนภาษา

ระบบ	ประเภทข้อความเบรลล์	ภาษา	สายอักขระที่ใช้ทดสอบ	ระบุภาษาผิด	ระบุถูกต้องแต่ละภาษา	ระบุถูกต้องรวม (%)
ระบบไทยปนอังกฤษรูปเต็ม	ไทยปนอังกฤษรูปเต็ม	ไทย	7,315	25	99.66	98.68
		อังกฤษ	6,814	162	97.62	
ระบบไทยปนอังกฤษรูปย่อ	ไทยปนอังกฤษรูปย่อ	ไทย	7,315	16	99.78	99.13
		อังกฤษ	6,640	105	98.42	

(2) ผลการถ่ายถอดอักษรเฉพาะข้อความที่ระบุภาษาได้ถูกต้อง

ในการทดสอบระบบถ่ายถอดอักษรเฉพาะกับสายอักขระที่ระบุภาษาได้ถูกต้องแสดงได้ดังนี้

ตารางที่ 7.4 ผลการถ่ายถอดอักษรเฉพาะข้อความที่ระบุภาษาได้ถูกต้อง

ระบบ	ประเภทข้อความเบรลล์	ภาษาที่ต้องการวัดประสิทธิภาพ	จำนวนคำที่ระบุภาษาถูก	จำนวนคำที่ถ่ายถอดผิด	ถ่ายถอดถูกต้อง (%)
ระบบไทยปนอังกฤษรูปเต็ม	ไทยปนอังกฤษรูปเต็ม	ไทย	11,900	53	99.56
		อังกฤษ	6,459	0	100
ระบบไทยปนอังกฤษรูปย่อ	ไทยปนอังกฤษรูปย่อ	ไทย	11,915	51	99.57
		อังกฤษ	6,516	100	98.47

(3) ผลการถ่ายถอดอักษรของระบบโดยภาพรวม

อย่างไรก็ตาม จากการทดสอบประสิทธิภาพของระบบโดยการนับจำนวนคำที่ถ่ายถอดไม่ถูกต้อง ผู้วิจัยพบว่าระบบถ่ายถอดอักษรทั้งสองระบบนี้สามารถถ่ายถอดข้อความที่มีเนื้อหาเดียวกันแต่ใช้อักษรเบรลล์คนละระบบได้มีประสิทธิภาพต่างกัน ซึ่งแสดงได้ดังนี้

ตารางที่ 7.5 ผลการถ่ายถอดอักษรด้วยการใช้วิธีการแบบผสมโดยภาพรวม

ระบบ	ประเภท ข้อความเบรลล์	จำนวนคำที่ ใช้ทดสอบ	จำนวนคำที่ ถ่ายถอดผิด	ถ่ายถอดได้ ถูกต้อง (%)
ไทยปนอังกฤษรูปเต็ม	ไทยปนอังกฤษรูปเต็ม	19,412	234	98.79
ไทยปนอังกฤษรูปย่อ	ไทยปนอังกฤษรูปย่อ	19,412	253	98.69

7.4 อภิปรายผลการทดสอบระบบ

ดังที่ได้กล่าวมาแล้วว่า วิธีการถ่ายถอดที่ได้นำเสนอข้างต้นนี้ เป็นการนำวิธีการถ่ายถอดอักษรด้วยกฎและแบบจำลองเอ็นแกรมมาใช้ร่วมกัน ดังนั้นข้อผิดต่าง ๆ ที่เกิดขึ้นจากการถ่ายถอดข้อความประเภทต่าง ๆ จึงมีลักษณะคล้ายกับการถ่ายถอดด้วยกฎและด้วยแบบจำลองเอ็นแกรม ซึ่งจะขออภิปรายดังนี้

7.4.1 การระบุภาษา

เมื่อนำข้อความเบรลล์ไทยล้วนและเบรลล์อังกฤษล้วนทั้งรูปเต็มและรูปย่อ มาทำการทดสอบระบบในส่วนของการระบุภาษา (ตามตารางที่ 7.1) จะพบว่าระบบไทยปนอังกฤษรูปเต็มจะระบุภาษาได้ดีกว่าระบบเบรลล์อังกฤษรูปย่อทั้งในการระบุภาษาอังกฤษและภาษาไทย คือระบุได้ถูกต้องอย่างสมบูรณ์

เนื่องจากระบบถ่ายถอดอักษรด้วยวิธีการแบบผสมนี้จะระบุภาษาด้วยกฎ ซึ่งจะทำงานได้ดีกับข้อความที่ไม่มีการปนภาษา เพราะนอกจากจะใช้สายอักขระที่เป็นเอกลักษณ์มาช่วยในการระบุภาษาได้แล้ว ยังสามารถใช้บริบทหน้าเพื่อช่วยในการระบุภาษาสำหรับสายอักขระที่ไม่สามารถระบุภาษาได้โดยตรง

หากพิจารณา การถ่ายถอดข้อความเบรลล์ที่มีการปนภาษาด้วยระบบไทยปนอังกฤษรูปเต็มและระบบไทยปนอังกฤษรูปย่อ (ตามตารางที่ 7.3) จะพบว่า ทั้งสองระบบสามารถระบุให้เป็นภาษาไทยได้ดีกว่าระบุให้เป็นภาษาอังกฤษ แต่เมื่อพิจารณาในแต่ละระบบจะพบว่า ระบบไทยปนอังกฤษรูปเต็มสามารถระบุภาษาไทยได้ดีกว่าภาษาอังกฤษมาก (คือระบุภาษาไทยได้ 99.66% ในขณะที่ระบุภาษาอังกฤษได้ 97.62 %) แต่ในระบบไทยปนอังกฤษรูปย่อนั้น สามารถระบุภาษาทั้งสองได้ถูกต้องไม่ต่างกันมากนัก (คือระบุภาษาไทยได้ 99.78% ในขณะที่ระบุภาษาอังกฤษได้ 98.42 %)

จึงอาจกล่าวได้ว่าระบบทั้งสองนี้จะใช้ได้ดีกับข้อความที่ไม่มีการปนของภาษา แต่หากข้อความนั้นมีการปนภาษา ระบบจะระบุภาษาไทยได้ดีกว่าภาษาอังกฤษ

7.4.2 การถ่ายถอดอักษร

เมื่อทำการทดสอบระบบกับข้อความเบรลล์ที่ไม่มีการปนภาษา เพื่อวัดประสิทธิภาพของการถ่ายถอดอักษร (ตามตารางที่ 7.2) ผู้วิจัยพบว่า ระบบสามารถถ่ายถอดข้อความอังกฤษรูปเต็มได้ดีที่สุด (100%) รองลงมาเป็นข้อความอังกฤษรูปย่อ (99.92%) และข้อความไทยล้วน (99.73%) ตามลำดับ แต่เมื่อทำการเปรียบเทียบข้อความที่มีการปนภาษา (ตามตารางที่ 7.4) จะพบว่า ระบบไทยปนอังกฤษรูปเต็มจะถ่ายถอดภาษาอังกฤษรูปเต็มได้ดีกว่าภาษาไทย ในขณะที่ระบบไทยปนอังกฤษรูปย่อจะถ่ายถอดข้อความเบรลล์ไทยได้ดีกว่าเบรลล์อังกฤษรูปย่อ

เท่าที่ได้กล่าวมา เป็นเพียงการวัดประสิทธิภาพของระบบถ่ายถอดอักษรด้วยวิธีการแบบผสม แต่ประสิทธิภาพของการถ่ายถอดอักษรด้วยวิธีการนี้ แตกต่างจากการถ่ายถอดอักษรอีก 2 วิธี หรือไม่อย่างไร เป็นประเด็นที่จะได้อภิปรายต่อไป

7.5 ผลการเปรียบเทียบวิธีการถ่ายถอดอักษรทั้ง 3 วิธี

วิธีการถ่ายถอดอักษรด้วยกฎ ด้วยแบบจำลองเอ็นแกรม และแบบผสมมีประสิทธิภาพที่แตกต่างกัน ซึ่งจะขออภิปราย ดังนี้

7.5.1 ผลการเปรียบเทียบการระบุภาษา

ผลการเปรียบเทียบการระบุภาษาของทั้งวิธีการถ่ายถอดอักษรทั้ง 3 วิธีสามารถแสดงได้ดังนี้

ตารางที่ 7.6 ประสิทธิภาพของการระบุภาษาด้วยวิธีการทั้ง 3 วิธี

ระบบ	ประเภทข้อความ	กฎ	แบบจำลองเอ็นแกรม	แบบผสม
ระบบไทยปนอังกฤษรูปเต็ม	ไทย	99.94	99.75	100
	อังกฤษรูปเต็ม	99.99	99.95	100
	ไทยปนอังกฤษรูปเต็ม	98.67	99.37	98.68
ระบบไทยปนอังกฤษรูปย่อ	ไทย	99.66	98.86	99.76
	อังกฤษรูปย่อ	99.99	99.09	99.99
	ไทยปนอังกฤษรูปย่อ	98.17	98.74	99.13

หากทำการเปรียบเทียบวิธีการถ่ายถอดอักษรที่นำเสนอมาทั้ง 3 วิธี คือการใช้กฎ แบบจำลองเอ็นแกรมและวิธีการแบบผสมนั้น ผู้วิจัยพบว่าการถ่ายถอดข้อความเบรลล์ที่ไม่มีการปนภาษาจะให้ผลต่างกัน สำหรับข้อความไทยล้วนที่ทดสอบกับระบบไทยปนอังกฤษรูปเต็ม พบว่า วิธีการแบบผสมให้ผลการระบุภาษาดีที่สุดคือระบุภาษาได้ถูกต้องทั้งหมด (100%) รองลงมาคือการใช้แบบจำลองเอ็นแกรม และการใช้กฎ ตามลำดับ ส่วนการนำข้อความไทยล้วนมาทดสอบกับระบบไทยปนอังกฤษรูปย่อ การใช้กฎให้ผลดีที่สุด (99.94%) รองลงมาคือวิธีการผสมและการใช้แบบจำลองเอ็นแกรมตามลำดับ

อย่างไรก็ตาม ในการระบุภาษาสำหรับภาษาอังกฤษนั้น เมื่อนำข้อความเบรลล์อังกฤษรูปเต็มมาทดสอบกับระบบไทยปนอังกฤษรูปเต็ม พบว่า วิธีการแบบผสมให้ผลการระบุภาษาดีที่สุด (100%) วิธีการใช้กฎ และวิธีการใช้แบบจำลองเอ็นแกรม ให้ผลรองลงมาตามลำดับ แต่การถ่ายถอดข้อความ

อังกฤษรูปย่อด้วยระบบปนภาษานั้น พบว่าวิธีการใช้กฎและการใช้แบบจำลองเอ็นแกรมให้ผลการระบุภาษาเท่ากัน (99.99%) และวิธีการแบบจำลองเอ็นแกรมให้ผลรองลงมา

เมื่อนำข้อความปนภาษามาทดสอบกับระบบ พบว่า การระบุภาษาโดยใช้แบบจำลองเอ็นแกรมจะระบุภาษาได้ดีที่สุดสำหรับข้อความไทยปนอังกฤษรูปเต็ม (99.37%) รองลงมาคือวิธีการแบบผสมและวิธีการใช้กฎตามลำดับ ส่วนข้อความไทยปนอังกฤษรูปย่อพบว่าวิธีการแบบผสมจะสามารถระบุภาษาได้ดีที่สุด (99.13%) รองลงมาคือการใช้แบบจำลองเอ็นแกรม และวิธีการใช้กฎตามลำดับ

7.5.2 ผลการเปรียบเทียบการถ่ายถอดอักษร

ผลการเปรียบเทียบการถ่ายถอดอักษรของทั้ง 3 วิธี สามารถแสดงได้ดังนี้

ตารางที่ 7.7 ประสิทธิภาพของการถ่ายถอดอักษรด้วยวิธีการทั้ง 3 วิธี

การถ่ายถอด	ประเภทข้อความ	กฎ	แบบจำลองเอ็นแกรม	แบบผสม
ไทย	ไทย	99.31	99.36	99.73
	ไทยปนอังกฤษรูปเต็ม	99.18	98.98	99.56
	ไทยปนอังกฤษรูปย่อ	99.17	99.29	99.57
อังกฤษรูปเต็ม	อังกฤษรูปเต็ม	100	99.99	100
	ไทยปนอังกฤษรูปเต็ม	99.18	100	100
อังกฤษรูปย่อ	อังกฤษรูปย่อ	99.91	99.80	99.92
	ไทยปนอังกฤษรูปย่อ	98.31	98.17	98.47

จากตารางข้างต้นจะขออภิปรายประสิทธิภาพของการถ่ายถอดเบรลล์ไทย เบรลล์อังกฤษรูปเต็ม และเบรลล์อังกฤษรูปย่อ ซึ่งใช้วิธีการถ่ายถอดอักษรทั้ง 3 วิธี ดังนี้

7.5.2.1 การถ่ายถอดเบรลล์ไทย

ในการนำข้อความเบรลล์ไทยล้วนไปทดสอบกับโมดูลถ่ายถอดเบรลล์ไทยนั้น ผู้วิจัยพบว่าวิธีการแบบผสมจะให้ผลการถ่ายถอดถูกต้องมากที่สุด (99.73%) รองลงมาคือการใช้แบบจำลอง

เอ็นแกรม และการใช้กฎ ตามลำดับ แต่เมื่อทำการเปรียบเทียบข้อความไทยปนอังกฤษในส่วนที่ระบบระบุให้เป็นภาษาไทยเท่านั้นพบว่าวิธีการแบบผสมจะถ่ายถอดเบรลล์ไทยได้ดีที่สุดเช่นกัน จึงกล่าวได้ว่า*วิธีการที่เหมาะสมมากที่สุดที่จะนำไปใช้ในการถ่ายถอดข้อความเบรลล์ไทย คือ การถ่ายถอดด้วยวิธีการแบบผสม* เนื่องจากวิธีการนี้ให้ผลลัพธ์ดีกว่าอีก 2 วิธี ที่กล่าวมา

อย่างไรก็ตามในวิธีการแบบผสมนั้น เกิดจากการพัฒนาระบบโดยอาศัยกฎและแบบจำลองเอ็นแกรม ทำให้ข้อผิดที่พบ ยังคงมีลักษณะคล้ายกับข้อผิดที่เกิดขึ้นจากการถ่ายถอดด้วยกฎและการถ่ายถอดด้วยแบบจำลองเอ็นแกรมดังนี้

(1) <u>ข้อผิดที่คล้ายคลึงกับวิธีการใช้กฎ</u> ในระบบการถ่ายถอดด้วยวิธีการแบบผสม ระบบจะใช้กฎเพื่อหารูปที่เหมาะสม ของ ‘⠪⠢’ (= เาะหรือ อะ) และ ‘⠫’ (=เ◌ิ หรือ เ-อ) จึงทำให้เกิดข้อผิดกับการถ่ายถอด ‘⠪⠢’ (= เาะหรือ อะ) ในลักษณะเดียวกับการถ่ายถอดเบรลล์คู่นี้ด้วยกฎตามที่ได้นำเสนอไว้ในหัวข้อ ที่ 5.4.2.1 1) (ค) นอกจากนี้ยังเกิดข้อผิดกับการถ่ายถอด ‘⠫’ (=เ◌ิหรือ เ-อ) ในลักษณะเดียวกับวิธีการใช้กฎที่ได้อภิปรายไว้ในหัวข้อที่ 5.4.2.1 2) (ง)

(2) <u>ข้อผิดที่คล้ายคลึงกับการใช้แบบจำลองเอ็นแกรม</u> ข้อผิดที่คล้ายคลึงกับการถอดด้วยการใช้แบบจำลองเอ็นแกรม ประกอบด้วย การถ่ายถอด ‘⠹⠹’ (ธ หรือ ◌์ท) ผิด การถ่ายถอด ‘⠹’ (◌์ หรือ ”) ผิด การถ่ายถอด ‘⠹⠱’ (ษ หรือ -ส) ผิด การถ่ายถอด ‘⠷’ ( ) “วงเล็บปิด” หรือ ◌ี) ผิด การถ่ายถอด ‘⠷’ (. หรือ ◌ี้) ผิด การถ่ายถอดพยัญชนะควบที่ปรากฏกับสระประสมไม่ถูกต้อง และ การถ่ายถอดอักษรเป็นเครื่องหมาย (ดูหัวข้อที่ 6.4.2.1 (จ) ประกอบ)

7.5.2.2 การถ่ายถอดเบรลล์อังกฤษรูปเต็ม

จากการทดสอบระบบ ผู้วิจัยพบว่า การถ่ายถอดเบรลล์อังกฤษรูปเต็มไม่เป็นปัญหาในการถ่ายถอดให้เป็นอักษรปกติแต่อย่างใด ตารางที่ 7.7 แสดงให้เห็นว่าประสิทธิภาพการถ่ายถอดเบรลล์อังกฤษรูปเต็มในระบบที่ใช้วิธีการแบบผสมและระบบที่ใช้กฎนั้น สามารถถ่ายถอดเบรลล์อังกฤษรูปเต็มได้ถูกต้องทั้งหมด ทั้งนี้เนื่องจากระบบที่ใช้วิธีการทั้งสองนี้ใช้โมดูลเดียวกันคือโมดูลถ่ายถอดเบรลล์อังกฤษรูปเต็มด้วยกฎ

อย่างไรก็ตาม ในการถ่ายถอดเบรลล์อังกฤษรูปเต็มด้วยแบบจำลองเอ็นแกรมยังเกิดข้อผิดขึ้นจำนวนหนึ่ง เนื่องจากวิธีการนี้ใช้สถิติความน่าจะเป็นมาช่วยในการถ่ายถอด ทำให้ถ่ายถอดอักษรที่กำกวมในบางกรณีไม่ถูกต้อง ซึ่งข้อผิดเหล่านี้อาจแก้ไขได้โดยอาศัยกฎเพิ่มเติมหรือการใช้คลังข้อความในการฝึกระบบ ที่แตกต่างจากงานวิจัยชิ้นนี้

7.5.2.3 การถ่ายถอดเบรลล์อังกฤษรูปย่อ

ตารางที่ 7.7 แสดงให้เห็นว่า การถ่ายถอดข้อความเบรลล์อังกฤษรูปย่อด้วยวิธีการถ่ายถอดทั้ง 3 วิธี ให้ผลลัพธ์ความถูกต้องที่แตกต่างกัน เมื่อผู้วิจัยทำการวิเคราะห์ข้อผิดที่เกิดขึ้นได้พบว่า วิธีการที่เหมาะสมที่สุดที่จะนำไปใช้ในการถ่ายถอดเบรลล์อังกฤษรูปย่อ ยังคงเป็นวิธีการแบบผสม ซึ่งให้ความถูกต้องที่ 99.92% เนื่องจากในการทำงานของโมดูลนี้ระบบได้ใช้กฎร่วมกับแบบจำลองเอ็นแกรมมาช่วยในการถ่ายถอด ทำให้การถ่ายถอดข้อความเบรลล์ประเภทนี้มีความถูกต้องมากกว่าการใช้กฎเพียงลำพัง อย่างไรก็ตาม เมื่อเปรียบเทียบวิธีการแบบผสมกับวิธีการใช้กฎ จะพบว่าปริมาณข้อผิดที่พบในวิธีการแบบผสมไม่ได้ลดลงในปริมาณที่มากนักคือลดลงเพียง .01 %เท่านั้น

จากการวิเคราะห์การถ่ายถอดข้อความเบรลล์ไทยปนอังกฤษรูปย่อโดยพิจารณาเฉพาะข้อความที่ระบบระบุให้เป็นภาษาอังกฤษเท่านั้น ผู้วิจัยพบว่า วิธีการทั้งสามนี้ยังให้ผลสอดคล้องกับรูปแบบที่ได้กล่าวมาข้างต้น กล่าวคือ วิธีการแบบผสมยังคงให้ผลลัพธ์สูงสุดที่ 98.47% ในขณะที่การใช้กฎให้ผลลัพธ์ที่ 98.31% และแบบจำลองเอ็นแกรมให้ผลลัพธ์ที่ 98.17 % ตามลำดับ

ด้วยเหตุที่ระบบถ่ายถอดเบรลล์ไทยปนอังกฤษรูปย่อ ยังคงใช้กระบวนการถ่ายถอดอักษรด้วยกฎทางภาษา ดังนั้นข้อผิดที่พบในวิธีการแบบผสมจึงคล้ายคลึงกับวิธีการใช้กฎ ดังที่ได้แสดงไว้ในหัวข้อที่ 5.4.2.2

อย่างไรก็ตามในการถ่ายถอดส่วนกลางของสายอักขระตามวิธีการแบบผสมนั้น ระบบจะใช้วิธีการถ่ายถอดโดยสร้างชุดสายอักขระเช่นเดียวกับวิธีการใช้แบบจำลองเอ็นแกรม ซึ่งกระบวนการดังกล่าวเป็นผลให้ข้อผิดในการถ่ายถอดคำย่อเบรลล์ มีปริมาณลดลง จากเดิมที่ผิด 6 คำ เหลืออยู่เพียง 3 คำ ได้แก่ “whereabout” “whereabouts” “beforehand” ซึ่งถ่ายถอดผิดเป็น “whereab” “whereabs” และ “befhand” ตามลำดับ

ระบบถ่ายถอดเบรลล์ไทยปนอังกฤษด้วยวิธีการแบบผสม ที่ผู้วิจัยได้นำเสนอมาข้างต้นเป็นระบบที่พัฒนามาจากการใช้กฎร่วมกับแบบจำลองเอ็นแกรม จึงทำให้ระบบที่ใช้วิธีการนี้มีประสิทธิภาพมากกว่า ระบบที่ใช้กฎหรือแบบจำลองเอ็นแกรม แต่เพียงลำพัง แม้ว่าระบบนี้จะให้ผลในการถ่ายถอดเบรลล์ไทยปนอังกฤษได้ดีในระดับหนึ่ง แต่ยังไม่สามารถถ่ายถอดอักขระเบรลล์ที่กำกวมได้ถูกต้องทั้งหมด จึงควรมีการนำระบบนี้ไปพัฒนาต่อเพื่อให้เกิดความสมบูรณ์ยิ่งขึ้นในอนาคต

บทที่ 8

สรุปผล อภิปรายผล และข้อเสนอแนะ

8.1 สรุปผล

จากการศึกษาเอกสารเกี่ยวกับอักษรเบรลล์ไทย ผู้วิจัยพบว่า มีงานวิจัยที่เกี่ยวข้องกับตัวเขียนระบบนี้น้อยมาก จึงทำให้ตัวเขียนภาษาไทยในอักษรเบรลล์เป็นประเด็นที่น่าทำการศึกษา มีเพียงนักวิชาการทางการศึกษาพิเศษได้ตั้งข้อสังเกตเกี่ยวกับตัวอักษรเบรลล์ไทยไว้ว่า พยัญชนะเบรลล์ไทยส่วนหนึ่ง เกิดจากการนำรูปพยัญชนะในเบรลล์อังกฤษที่มีเสียงตรงกับภาษาไทยมาใช้แทนพยัญชนะเบรลล์ไทยนั้น ๆ เช่น ใช้ ‘⠎’ (=s) แทน “ส” เป็นต้น (สมทรง พันธ์สุวรรณ, 2538) นอกจากงานเขียนฉบับนี้แล้ว ยังไม่พบงานวิจัยฉบับใดที่ทำการวิเคราะห์ตัวอักษรเบรลล์ไทยอย่างเป็นระบบ ดังนั้น ผู้วิจัยจึงได้ทำการวิเคราะห์ตัวอักษรไทยในการเขียนระบบดังกล่าว

ในการวิเคราะห์ตัวอักษรเบรลล์ไทยนั้น ผู้วิจัยได้นำตัวอักษรเบรลล์ไทย อังกฤษ และญี่ปุ่น มาทำการเปรียบเทียบรหัสอักษรเบรลล์ที่ปรากฏระหว่างภาษา เหตุที่ต้องนำเบรลล์อังกฤษและญี่ปุ่นมาทำการวิเคราะห์ร่วมด้วย เนื่องจากผู้วิจัยพบว่ามิสเจเนวีฟ คอลฟิลด์ สุภาพสตรีชาวอเมริกันผู้ประดิษฐ์อักษรเบรลล์ไทย เคยพำนักที่ญี่ปุ่นก่อนที่จะเข้ามาบุกเบิกงานด้านการศึกษาคนตาบอดในประเทศไทย ท่านจึงน่าจะนำอักขระเบรลล์บางส่วนจากภาษาอังกฤษและภาษาญี่ปุ่นมาใช้กับระบบอักษรเบรลล์ไทย

นอกจากการเปรียบเทียบอักษรเบรลล์ระหว่างภาษาทั้งสามแล้ว ผู้วิจัยยังได้วิเคราะห์เฉพาะอักษรเบรลล์ไทยเองเพิ่มเติม โดยนำอักษรเบรลล์ไทยแต่ละตัวมาเปรียบเทียบกันเพื่อศึกษาจุดเบรลล์ที่ปรากฏใน เบรลล์ไทยว่ามีรูปแบบการปรากฏของจุดที่สามารถทำนายได้หรือไม่

สำหรับการถ่ายถอดตัวบทเบรลล์ไทยปนอังกฤษซึ่งเป็นประเด็นหลักของงานวิจัยฉบับนี้ มีงานวิจัยที่เกี่ยวข้องจำนวนน้อยมากเช่นกัน งานวิจัยเกี่ยวกับอักษรเบรลล์ในประเทศไทยมักจะเน้นการถ่ายถอดเบรลล์ไทยอย่างเป็นเอกเทศ มีงานวิจัยไม่กี่ชิ้นที่ทำการศึกษาเบรลล์ไทยปนอังกฤษรูปเต็ม ผู้วิจัยพบว่ายังไม่มีงานวิจัยชิ้นใดที่ทำการศึกษาเบรลล์ไทยปนอังกฤษรูปย่อเลย ด้วยเหตุนี้จึงได้ทำการศึกษาการถ่ายถอดตัวบทเบรลล์ไทยปนอังกฤษเป็นอักษรปกติโดยศึกษาทั้งเบรลล์ไทยปนอังกฤษ

รูปเต็มและเบรลล์ไทยปนอังกฤษรูปย่อ นอกจากนี้ ผู้วิจัยยังได้ทำการเปรียบเทียบประสิทธิภาพการถ่ายถอดอักษรระหว่างวิธีการ 3 วิธี คือ การใช้กฎทางภาษา การใช้แบบจำลองเอ็นแกรม และวิธีการแบบผสม

ในการจัดเตรียมข้อมูลเพื่อใช้ในการวิจัย ผู้วิจัยได้นำข้อความเบรลล์จากองค์กรหลักของคนตาบอดมาใช้ในการวิจัย ซึ่งข้อความเหล่านี้ผ่านการตรวจสอบจากผู้เชี่ยวชาญอักษรเบรลล์แล้ว โดยที่ได้แบ่งข้อความออกเป็นสองส่วน ส่วนหนึ่งใช้ในการฝึกระบบ และอีกส่วนหนึ่งใช้ในการทดสอบระบบ

ในการทดสอบระบบถ่ายถอดอักษรนั้น ผู้วิจัยได้ใช้ข้อความเบรลล์ 5 ประเภท ข้อความ 3 ประเภทแรก คือ (1) ข้อความเบรลล์ไทย (2) ข้อความเบรลล์อังกฤษรูปย่อ และ(3) ข้อความเบรลล์ไทยปนอังกฤษรูปเต็ม ได้จากศูนย์เทคโนโลยีการศึกษาเพื่อคนตาบอด ส่วนอีก 2 ประเภท คือ (4) ข้อความเบรลล์อังกฤษรูปเต็ม และ (5) ข้อความเบรลล์ไทยปนอังกฤษรูปย่อ เกิดจากการนำข้อความประเภทที่ (2) และ (3) จากแหล่งข้อมูลที่กล่าวมานี้ มาทำการแก้ไขให้ตรงตามรูปแบบอักษรเบรลล์ที่จำเป็นต้องใช้ในการวิจัย เหตุที่ต้องสร้างข้อความเบรลล์อีก 2 ประเภทขึ้นเพิ่มเติมด้วยตนเองเนื่องจากผู้วิจัยไม่สามารถหาเอกสารสองประเภทหลังนี้ได้เพียงพอต่อการวิจัย

งานวิจัยชิ้นนี้มีหัวข้อวิจัย 2 ประเด็นหลัก ซึ่งผู้วิจัยจะขอเสนอผลการวิจัยที่ได้ค้นพบดังนี้

8.1.1 หน่วยอักขระเบรลล์ไทยและหน่วยอักขระเบรลล์อังกฤษ

ผลการวิจัยแสดงให้เห็นว่า สระในเบรลล์ไทยและเบรลล์ญี่ปุ่นมีรูปเหมือนกันและมีเสียงใกล้เคียงกัน 5 รูปได้แก่ ‘⠡’(ะ) ‘⠩’(ิ) ‘⠹’(ุ) ‘⠱’(เ-) ‘⠕’(โ-) ในขณะที่เบรลล์ไทยมีพยัญชนะ 15 รูป ที่มีรูปและเสียงคล้ายคลึงกับเบรลล์อังกฤษ ได้แก่ ‘⠫’ (=ก) ‘⠣’ (=ข) ‘⠻’ (=จ) ‘⠙’ (=ด) ‘⠮’ (=ถ) ‘⠝’ (=น) ‘⠏’ (=ผ) ‘⠍’ (=ม) ‘⠽’ (=ย) ‘⠗’ (=ร) ‘⠇’ (=ล) ‘⠧’ (=ว) ‘⠎’(=ส) ‘⠓’ (=ห) ‘⠕’ (=อ) อักขระเบรลล์ 20 รูปนี้ จัดเป็นเบรลล์ไทยพื้นฐานซึ่งนำไปขยายเป็นหน่วยอักขระเบรลล์ไทยเพิ่มเติม ใน 2 รูปแบบ

รูปแบบแรกเป็นการปรับเปลี่ยนจุดเบรลล์ภายในเซลล์ ซึ่งใช้ใน 2 วิธีการ ได้แก่ (1) การเติมจุดเบรลล์ภายในเซลล์ ซึ่งพบได้กับอักขระเบรลล์จำนวน 20 รูป ประกอบด้วยสระ 9 รูป และพยัญชนะ 11 รูป และ (2) การเลื่อนตำแหน่งจุดเบรลล์ให้ต่ำลง ซึ่งพบได้กับสระและเครื่องหมาย 10 รูป

รูปแบบที่สองเป็นการเพิ่มอักขระเบรลล์รูปแบบเฉพาะ ซึ่งพบได้ 2 วิธีการ วิธีการแรกเป็นการเพิ่มอักขระเบรลล์หน้าอักขระฐาน ซึ่งพบได้กับพยัญชนะ 16 รูปเท่านั้น และวิธีการที่สองเป็นการเพิ่มอักขระเบรลล์หลังอักขระฐาน ซึ่งพบได้กับสระ จำนวน 11 รูป

นอกจากนี้ยังมีหน่วยอักขระเบรลล์ไทยจำเพาะ ซึ่งประกอบด้วยสระและเครื่องหมาย 7 รูป ได้แก่ ‘⠡’(=ั) ‘⠣’ (=ื) ‘⠩’ (=ิ) ‘⠭’(=ี) ‘⠹’(=ึ) ‘⠽’ (=ุ) ‘⠾’(=ู) และพยัญชนะอีก 2 รูป ได้แก่ ‘⠯’(=ฉ) และ ‘⠿’(=ฮ)

ดังนั้นจึงสอดคล้องกับสมมติฐานข้อที่ 1 ที่ผู้วิจัยตั้งไว้ คือ “หน่วยอักขระเบรลล์ไทยสามารถจำแนกเป็น 3 ประเภท ได้แก่ หน่วยอักขระเบรลล์ไทยพื้นฐาน หน่วยอักขระเบรลล์ไทยดัดแปลง และหน่วยอักขระเบรลล์ไทยจำเพาะ หน่วยอักขระเบรลล์ไทยพื้นฐานแต่ละตัวมีรูปและเสียงสอดคล้องกับหน่วยอักขระเบรลล์ในภาษาอังกฤษและภาษาญี่ปุ่น หน่วยอักขระเบรลล์ไทยดัดแปลงแต่ละตัวได้มาจากการปรับหน่วยอักขระเบรลล์ไทยพื้นฐานที่มีเสียงตรงกัน โดยการกำกับด้วยเครื่องหมาย (glyph) และหน่วยอักขระเบรลล์ไทยจำเพาะเป็นหน่วยอักขระเบรลล์ที่ปรากฏเฉพาะภาษาไทยเท่านั้น” นอกจากรูปแบบการสร้างอักขระเบรลล์ไทยที่ระบุไว้ในสมมติฐานแล้ว ผู้วิจัยยังพบรูปแบบเพิ่มเติมคือการเลื่อนตำแหน่งจุดเบรลล์ให้ต่ำลง ซึ่งแตกต่างจากการสร้างหน่วยอักขระในเบรลล์อังกฤษ ที่มีเพียงรูปแบบเดียว

ในการศึกษาการสร้างหน่วยอักขระเบรลล์อังกฤษนั้น ผู้วิจัยได้ศึกษาจากวรรณกรรมต่าง ๆ ซึ่งระบุว่า หน่วยอักขระเบรลล์อังกฤษประกอบด้วยหน่วยอักขระเบรลล์อังกฤษพื้นฐาน 10 รูป ได้แก่ ‘⠁’ -‘⠊’ (a - j) โดยที่ในการสร้างหน่วยอักขระเบรลล์อังกฤษเพิ่มเติม มีเพียงวิธีการเดียวคือการเติมจุดเบรลล์เข้าไปในหน่วยอักขระฐาน ประกอบด้วย ประเภทแรกเป็นการเติม จุด 3 เข้าไปใน‘⠁’-‘⠊’ (a - j) ทำให้ได้ ‘⠅’ – ‘⠞’ (k - t) ประเภทที่สองเป็นการเติม จุด 3-6 เข้าไปใน ‘⠁’ - ‘⠑’ (a - e) ทำให้ได้ ‘⠥’ – ‘⠵’ (u - z) ยกเว้น ‘⠺’ (w) ในประเภทหลังที่เกิดจากการเติมจุด 6 เข้าไปใน ‘⠚’ (j)

ผลที่ได้ข้างต้นจึงสอดคล้องกับสมมติฐานข้อ 2 ที่ว่า “หน่วยอักขระเบรลล์อังกฤษสามารถจำแนกเป็น 2 ประเภท ได้แก่ หน่วยอักขระเบรลล์อังกฤษพื้นฐาน และหน่วยอักขระเบรลล์อังกฤษดัดแปลง หน่วยอักขระเบรลล์อังกฤษพื้นฐานประกอบด้วยหน่วยอักขระเบรลล์ 10 ตัวแรกในภาษาอังกฤษ (a - j) และหน่วยอักขระเบรลล์อังกฤษดัดแปลงได้มาจากการกำกับหน่วยอักขระเบรลล์อังกฤษพื้นฐานแต่ละตัวด้วยเครื่องหมาย (glyph)”

8.1.2 ระบบการถ่ายถอดเบรลล์ไทยปนอังกฤษ

ในงานวิจัยฉบับนี้ ผู้วิจัยได้เปรียบเทียบวิธีการถ่ายถอดอักษรเบรลล์ไทยปนอังกฤษ 3 วิธี คือ วิธีการใช้กฎ วิธีการใช้แบบจำลองเอ็นแกรม และวิธีการแบบผสม โดยในแต่ละวิธีการ จะแยกการถ่ายถอดอักษรออกเป็น 2 ระบบ คือระบบเบรลล์ไทยปนอังกฤษรูปเต็ม และระบบเบรลล์ไทยปนอังกฤษรูปย่อ ซึ่งการถ่ายถอดอักษรแต่ละวิธีจะดำเนินการดังนี้

วิธีการแรก เป็นการถ่ายถอดอักษรโดยใช้กฎทางภาษา ในวิธีการนี้จะมีภาพรวมในการดำเนินการในภาพรวมเหมือนกัน ทั้งระบบถ่ายถอดเบรลล์ไทยปนอังกฤษรูปเต็มและระบบถ่ายถอดเบรลล์ไทยปนอังกฤษรูปย่อ คือ เริ่มจากนำเข้าข้อความเบรลล์เป็นย่อหน้า แล้วนำแต่ละย่อหน้ามาตัดแบ่งออกเป็นสายอักขระต่าง ๆ หลังจากนั้นแต่ละสายอักขระจะถูกระบุภาษาด้วยการใช้กฎ จากนั้นนำสายอักขระที่ผ่านการระบุภาษาแล้วไปถ่ายถอดอักษรตามภาษาที่ระบุได้

ในส่วนของการถ่ายถอดเบรลล์ไทยด้วยกฎนั้น ระบบจะถ่ายถอดอักขระเบรลล์ครั้งละ 2 หรือ 1 เซลล์ ตามแต่ว่าอักขระจำนวนเท่าใดจะตรงกับคู่เทียบในตารางคู่เทียบเบรลล์ไทยกับไทยปกติ โดยจะเริ่มจากการตรวจสอบ เบรลล์เซลล์คู่ที่กำกวม เบรลล์เซลล์คู่ทั่วไป เบรลล์เซลล์เดี่ยวที่กำกวม หรือเบรลล์เซลล์เดี่ยวทั่วไป ถ้าพบว่าอักขระเบรลล์ที่ทำการตรวจสอบตรงกับอักขระเบรลล์ในตารางคู่เทียบ ระบบจะถ่ายถอดตามกฎที่ระบุไว้สำหรับอักขระเบรลล์แต่ละชุด

ในส่วนการถ่ายถอดเบรลล์อังกฤษด้วยกฎนั้น ระบบจะเริ่มจากการแบ่งสายอักขระออกเป็นสายอักขระย่อย จากนั้น นำแต่ละสายอักขระย่อยไประบุสถานะของอักษรเบรลล์ ได้แก่ เบรลล์อังกฤษรูปเต็ม คอมพิวเตอร์เบรลล์ ตัวเลขเบรลล์ หรือเบรลล์อังกฤษรูปย่อ แล้วถ่ายถอดตามสถานะที่ระบุได้

วิธีการที่สอง เป็นการถ่ายถอดอักษรโดยใช้แบบจำลองเอ็นแกรม วิธีการนี้จะเริ่มจากนำเข้าข้อความเบรลล์เป็นย่อหน้า แล้วนำแต่ละย่อหน้ามาตัดแบ่งออกเป็นสายอักขระต่าง ๆ จากนั้นนำแต่ละสายอักขระไปสร้างชุดสายอักขระไทยและอังกฤษปกติ เมื่อได้ชุดสายอักขระที่มีสองภาษาปนกันแล้ว ระบบจะเริ่มประมวลผล โดยที่การถ่ายถอดเบรลล์ไทยปนอังกฤษรูปเต็มและเบรลล์ไทยปนอังกฤษรูปย่อจะดำเนินการต่างกัน ดังนี้

ในระบบการถ่ายถอดเบรลล์ไทยปนอังกฤษรูปเต็ม ระบบจะนำชุดสายอักขระที่สร้างขึ้นไปคำนวณหาชุดสายอักขระที่มีค่าทางสถิติสูงสุดด้วยแบบจำลองไตรแกรมตัวอักษรไทยและอังกฤษ แล้วนำสายอักขระที่มีค่าทางสถิติสูงสุดมาใช้เป็นผลลัพธ์ของการถ่ายถอดอักษร

สำหรับระบบการถ่ายถอดไทยปนอังกฤษรูปย่อ ระบบจะประมวลผลชุดสายอักขระต่าง ๆ เพื่อหาผลลัพธ์ของการถ่ายถอดอักษร โดยนำแต่ละสายอักขระระหว่างชุดสายอักขระ 2 ชุดมาเรียงต่อกัน เพื่อหาว่าเมื่อเรียงต่อกันแล้วสามารถเป็นคำภาษาอังกฤษ 2 คำที่เรียงต่อกันได้หรือไม่ โดยการตรวจสอบกับแบบจำลองไบแกรมของคำอังกฤษ หากพบว่าเป็นไปได้ ระบบจะนำสองคำที่มีสถิติสูงสุด มาเป็นผลลัพธ์ของการถ่ายถอดอักษร แต่หากพบว่าเมื่อนำสายอักขระมาเรียงต่อกันระหว่างชุดแล้วไม่สามารถเป็นคำอังกฤษสองคำที่เรียงต่อกันได้ ระบบจะนำชุดสายอักขระแรกไปประมวลผลด้วยแบบจำลองไตรแกรมตัวอักษรไทยและอังกฤษ เพื่อหาสายอักขระที่มีความน่าจะเป็นสูงสุดมาใช้เป็นผลลัพธ์ของการถ่ายถอดสำหรับชุดสายอักขระแรก แล้วดำเนินการกับชุดสายอักขระ 2 ชุดต่อไปในลักษณะเดียวกันไปเรื่อย ๆ จนครบทั้งย่อหน้า

วิธีการที่สาม เป็นการถ่ายถอดอักษรโดยใช้วิธีการแบบผสม ในวิธีการนี้จะดำเนินการโดยภาพรวมเหมือนกันทั้ง 2 ระบบ โดยจะเริ่มจาก นำเข้าข้อความเบรลล์เป็นย่อหน้า แล้วนำแต่ละย่อหน้ามาตัดแบ่งออกเป็นสายอักขระต่าง ๆ จากนั้นนำแต่ละสายอักขระไประบุภาษาด้วยการใช้กฎ หากระบุภาษาด้วยกฎได้สำเร็จ จะนำสายอักขระนั้นไปสร้างชุดสายอักขระตามภาษาที่ระบุได้ แต่หากระบุภาษาด้วยกฎไม่ได้ ระบบจะนำสายอักขระนั้นไปสร้างชุดสายอักขระของทั้งสองภาษา หลังจากที่สร้างชุดสายอักขระได้แล้ว ระบบจะนำแบบจำลองไตรแกรมตัวอักษรไทยและอังกฤษ มาหาสายอักขระที่มีความน่าจะเป็นสูงสุด เพื่อนำมาใช้เป็นผลลัพธ์ของการถ่ายถอดอักษร

ผลการศึกษา การถ่ายถอดข้อความเบรลล์ไทยปนอังกฤษโดยใช้ทั้งวิธีการทั้ง 3 วิธี ที่กล่าวมานี้แสดงให้เห็นว่า วิธีการแบบผสม ให้ผลลัพธ์ในการถ่ายถอดอักษรได้ถูกต้องมากที่สุด แต่ในการระบุภาษานั้น จะพบว่า วิธีการแบบผสมระบุภาษาในข้อความเบรลล์ไทยปนอังกฤษรูปย่อได้ดีที่สุด ในขณะที่วิธีการใช้แบบจำลองเอ็นแกรมระบุภาษาในข้อความเบรลล์ไทยปนอังกฤษรูปเต็มได้ดีที่สุด

ดังนั้นในงานวิจัยฉบับนี้ ผลการวิจัยส่วนใหญ่จึงสอดคล้องกับสมมติฐานข้อ 3 ที่ผู้วิจัยตั้งไว้ว่า “ระบบการถ่ายถอดอักษรสำหรับตัวบทเบรลล์ไทยปนอังกฤษเป็นอักษรปกติ ที่อาศัยกฎทางภาษาร่วมกับแบบจำลองเอ็นแกรม จะให้ผลลัพธ์ที่ถูกต้องมากกว่าการใช้วิธีการใดวิธีการหนึ่ง”

8.2 อภิปรายผล

เท่าที่ผ่านมา มีงานวิจัยไม่กี่ชิ้นที่ได้ศึกษาเกี่ยวกับการถ่ายถอดข้อความเบรลล์ไทยปนอังกฤษ หนึ่งในงานวิจัยที่ทำการศึกษาประเด็นนี้คืองานวิจัยของ รวิศม์ วงษ์สมาน (2551) ที่ศึกษาการถ่ายถอดข้อความเบรลล์ไทยปนอังกฤษรูปเต็มแบบไปกลับโดยใช้กฎและแบบจำลองเอ็นแกรม โดยที่ในการฝึกระบบสำหรับการใช้แบบจำลองเอ็นแกรมนั้น รวิศม์ได้ใช้คลังข้อความอักษรเบรลล์ที่สร้างขึ้นเองโดยการแปลงข้อความจากเว็บไซต์ ต่อมาทศวัฒน์ ชุณหวิริยะธีระ (2555) ได้ศึกษาการถ่ายถอดอักขระเบรลล์ไทยที่มีความกำกวมโดยใช้แบบจำลองเอ็นแกรมซึ่งฝึกระบบโดยใช้อักษรเบรลล์ที่มีการตัดแบ่งคำแล้ว

จากการศึกษางานวิจัยที่ผ่านมา ผู้วิจัยพบว่า งานวิจัยเหล่านี้ไม่ได้อธิบายธรรมชาติของอักษรเบรลล์ไทยอย่างลึกซึ้ง เป็นแต่เพียงการพยายามแก้ปัญหาที่เกิดขึ้นกับการถ่ายถอดอักษรเบรลล์เป็นกรณี ๆ ไปเท่านั้น ซึ่งแตกต่างจากงานวิจัยชิ้นนี้ที่เกิดจากการประมวลความรู้เกี่ยวกับอักษรเบรลล์ในแง่มุมต่าง ๆ เท่าที่จะเป็นไปได้ มาสร้างแบบจำลองการถ่ายถอดอักษร ทำให้นอกจากแบบจำลองนี้จะใช้ประโยชน์ได้จริงแล้ว องค์ความรู้อักษรเบรลล์ที่นำมาใช้ในการพัฒนาระบบ ยังเป็นประโยชน์ต่อผู้ที่สนใจอักษรเบรลล์อีกด้วย

สิ่งที่น่าสังเกตอีกประการหนึ่งคือ งานวิจัยในอดีตนำอักษรเบรลล์มาฝึกและทดสอบระบบ โดยที่ไม่มีผู้เชี่ยวชาญตรวจสอบความถูกต้องก่อนที่จะนำมาใช้ในการวิจัย จึงเป็นผลให้เกิดข้อผิดบางส่วนในการถ่ายถอดอักษร

เพื่อแก้ปัญหาที่อาจเกิดขึ้นได้จากการใช้อักษรเบรลล์ในการวิจัย ผู้วิจัยได้นำข้อความเบรลล์จากองค์กรหลักที่ผลิตสื่ออักษรเบรลล์มาใช้ ซึ่งข้อความเหล่านี้ผ่านการตรวจสอบจากผู้เชี่ยวชาญอักษรเบรลล์ และมีการตีพิมพ์เผยแพร่ในวงการคนตาบอดมาแล้ว จึงเป็นการยืนยันในระดับหนึ่งว่า ผลที่ได้จากการวิจัยนี้ มีความคลาดเคลื่อนน้อย

นอกจากนี้ ผู้วิจัยได้ตระหนักถึงปัญหาที่เกิดขึ้นจากการใช้อักษรเบรลล์ในการฝึกระบบ จึงได้พยายามใช้อักษรเบรลล์มาฝึกระบบในปริมาณที่น้อยที่สุดเท่าที่จะเป็นไปได้ และหันไปใช้คลังข้อความอักษรปกติที่มีคุณภาพซึ่งผ่านการตรวจสอบและกลั่นกรองจากผู้เชี่ยวชาญมาแล้วได้แก่ คลังข้อความภาษาไทยแห่งชาติ และคลังข้อความ British national corpus จึงเป็นการยืนยันผลความถูกต้องอีกระดับหนึ่ง

งานวิจัยที่ผ่านมาค้นพบว่า ในการใช้อักษรเบรลล์เพื่อฝึกระบบการถ่ายถอดอักษรนั้น ต้องใช้ข้อมูลขนาดสี่แกรมจึงจะให้ผลดี ผู้วิจัยพบว่า เหตุที่งานวิจัยเหล่านี้ต้องใช้ข้อมูลที่มีขนาดใหญ่ เนื่องจากในอักษรเบรลล์มีอักขระจำนวนจำกัด ทำให้อักษรปกติบางตัวต้องแทนด้วยอักขระเบรลล์ถึง 2 ตัว จึงต้องใช้ข้อมูลที่มีขนาดสี่แกรมในการประมวลผล นอกจากนี้บ่อยครั้งยังมีการนำอักขระเบรลล์ 1 ตัวมาใช้แทนอักษรปกติได้หลายกรณี ทำให้ต้องอาศัยบริบทในอักษรเบรลล์ที่กว้างขึ้นในประมวลผลจึงจะตีความเป็นอักษรปกติได้ถูกต้อง

ในการวิจัย ผู้วิจัยได้ใช้ข้อความอักษรปกติที่มีขนาดเพียงสามแกรมในการฝึกระบบ เนื่องจากพบว่า ในไทยปกติมีจำนวนอักขระมากกว่าเบรลล์ไทยหลายเท่า จึงไม่จำเป็นต้องใช้ข้อมูลขนาดใหญ่เช่นเดียวกับการประมวลผลด้วยอักษรเบรลล์ อีกทั้งผลการทดสอบระบบในงานวิจัยนี้ยังแสดงให้เห็นว่าการใช้ข้อมูลขนาดสามแกรม ให้ผลลัพธ์การถ่ายถอดอักษรที่มีประสิทธิภาพดีอยู่แล้ว และด้วยเหตุที่ข้อมูลสามแกรมมีขนาดเล็กกว่ามากเมื่อเทียบกับข้อมูลสี่แกรม จึงทำให้ไม่สิ้นเปลืองทรัพยากรในการจัดเก็บและหน่วยความจำในการประมวลผล ทำให้เหมาะที่จะนำไปใช้งานกับอุปกรณ์แสดงผลอักษรเบรลล์ที่มีอยู่ในปัจจุบันซึ่งยังมีเนื้อที่ในการจัดเก็บข้อมูลและหน่วยความจำที่จำกัดอยู่ ดังนั้นผู้วิจัยจึงใช้ข้อมูลอักษรปกติขนาดสามแกรมกับระบบการถ่ายถอดอักษรที่เสนอในงานวิจัยฉบับนี้

แม้กระนั้นก็ตาม ผู้วิจัยได้พบปัญหากับข้อความที่นำมาทดสอบระบบ เนื่องจากข้อความที่ได้จากองค์กรคนตาบอด ถูกจัดเตรียมขึ้นเพื่อใช้ในการพิมพ์บนหน้ากระดาษที่รองรับแค่ 40 อักขระเบรลล์ต่อบรรทัด และยังมีการเติมเครื่องหมายเว้นวรรค 2 ตัว ในตำแหน่งต้นบรรทัดเพื่อเพิ่มเนื้อที่ให้สามารถเย็บเล่มทางซ้ายได้ ทำให้ข้อความในแต่ละบรรทัดที่นำมาถ่ายถอดได้จริง มีขนาดเพียง 38 อักขระเท่านั้น ข้อความในแต่ละสายอักขระจึงมีขนาดสั้น (คือไม่เกิน 38 อักขระ) เป็นผลให้วิธีการระบุภาษาโดยใช้ความยาวของสายอักขระที่เสนอในงานวิจัยนี้ไม่สามารถทำงานได้อย่างเต็มประสิทธิภาพ จึงทำให้เกิดข้อผิดในการระบุภาษาส่วนหนึ่งขึ้น

อย่างไรก็ตาม ในการใช้อักษรเบรลล์ในชีวิตจริงนั้น ผู้ใช้มีแนวโน้มที่จะพบสายอักขระเบรลล์ที่มีความยาวมากกว่า 38 อักขระ เนื่องจากในปัจจุบันเครื่องแสดงผลอักษรเบรลล์ (Braille display) หาได้ง่ายขึ้น ซึ่งอุปกรณ์เหล่านี้สามารถรองรับกับสายอักขระที่มีขนาดยาวได้ ผู้วิจัยจึงเชื่อว่า การระบุภาษาด้วยความยาวของสายอักขระที่ผู้วิจัยเสนอมานี้จะได้ผลดีเมื่อนำไปใช้กับอุปกรณ์แสดงผลอักษรเบรลล์ในอนาคต

เมื่อเปรียบเทียบประสิทธิภาพในการถ่ายถอดอักษรด้วยวิธีการทั้ง 3 วิธี คือการใช้กฎทางภาษา การใช้แบบจำลองเอ็นแกรม และการใช้วิธีการแบบผสม จะพบว่า วิธีการเหล่านี้ให้ผลความถูกต้องที่แตกต่างกันไม่มากนัก แต่จากการสังเกตอย่างคร่าว ๆ เกี่ยวกับระยะเวลาและทรัพยากรที่ใช้ในการประมวลผล ผู้วิจัยพบว่า วิธีการใช้กฎทางภาษาสามารถประมวลผลได้รวดเร็วที่สุดและใช้หน่วยความจำน้อยที่สุด รองลงมาคือวิธีการแบบผสม และวิธีการใช้แบบจำลองเอ็นแกรมตามลำดับ

นอกจากนี้ จากการวิเคราะห์ข้อผิดที่เกิดจากการถ่ายถอดอักษรด้วยวิธีการทั้ง 3 วิธี ที่กล่าวมานี้ ผู้วิจัยพบว่า ข้อผิดเหล่านี้เป็นข้อผิดที่สามารถตรวจจับได้ง่าย เนื่องจากเป็นข้อผิดที่เกิดขึ้นในระดับอักขระ ไม่ว่าจะเป็น ข้อผิดจากการสลับที่อักขระ การแทนที่อักขระ หรือการสูญหายของอักขระบางตัว ดังนั้นควรนำแบบจำลองการถ่ายถอดอักษรที่เสนอนี้ไปใช้ร่วมกับโปรแกรมตรวจสอบคำผิด เพื่อให้ได้ผลลัพธ์การถ่ายถอดอักษรที่ดียิ่งขึ้น

8.3 ข้อเสนอแนะ

ในการวิจัยครั้งนี้ ผู้วิจัยได้สัมผัสกับอักษรเบรลล์ผ่านการวิเคราะห์ตัวอักษรเบรลล์ไทยและตัวอักษรเบรลล์อังกฤษ และจากการทดลองถ่ายถอดข้อความเบรลล์ไทยปนอังกฤษไม่ว่าจะเป็นเบรลล์อังกฤษรูปเต็มหรืออังกฤษรูปย่อ ทำให้ผู้วิจัยได้พบปัญหาและอุปสรรคในการใช้อักษรเบรลล์ ไม่ว่าจะเป็นการใช้โดยมนุษย์ หรือการตีความโดยเครื่องคอมพิวเตอร์ ผู้วิจัยจึงขอเสนอแนะประเด็นต่าง ๆ เกี่ยวกับอักษรเบรลล์และการถ่ายถอดอักษร ดังนี้

8.3.1 การแก้ไขตัวอักษรและอักขรวิธีในเบรลล์ไทย

การศึกษาครั้งนี้ชี้ให้เห็นว่า มีรหัสอักษรเบรลล์ไทยหลายตัวที่กำกวม ทำให้เกิดปัญหาต่อการถ่ายถอดอักษรด้วยคอมพิวเตอร์ หากมีการแก้ไขรหัสอักษรเบรลล์เหล่านี้ก็จะทำให้การประมวลผลด้วยคอมพิวเตอร์สะดวกยิ่งขึ้น อีกทั้งเป็นประโยชน์ต่อผู้ใช้อักษรเบรลล์ในอนาคต

อักขระตัวแรกที่ควรทำการแก้ไขรหัสในอักษรเบรลล์คือ ‘⠸’ เนื่องจากรหัสเบรลล์ตัวนี้มีความกำกวมเป็นอย่างมาก เพราะใช้แทนอักษรปกติได้ถึง 4 กรณีคือ ในกรณีทั่วไปใช้แทน “◌์” และเมื่อปรากฏท้ายวรรคสามารถใช้แทน “◌์” หรือ “ ” ” (คำพูดปิด)ได้ นอกจากนี้ อักขระเบรลล์ตัวนี้ยังปรากฏหน้า ‘⠸⠓’ (= ท) ซึ่งสามารถตีความอักขระที่อยู่ชิดกันนี้ (คือ ‘⠸⠓’) ให้เป็น “ธ” หรือ "◌์ท" ได้

การแก้ไขอักขระเบรลล์ที่ใช้เขียน "◌์" จะทำให้ลดปัญหาความกำกวมลงได้มาก ประการแรกจะลดความกำกวมของอักขระเบรลล์ตัวนี้เมื่อปรากฏท้ายวรรคเพราะจะทำให้ "◌์" และ " " มีรหัสอักษรเบรลล์ที่แตกต่างกัน จึงไม่เกิดความกำกวมเมื่อปรากฏท้ายวรรคอีกต่อไป นอกจากนี้ ยังเป็นผลให้ "ธ" และ "◌์ท" ซึ่งในปัจจุบันใช้รหัสอักษรเบรลล์ชุดเดียวกัน (คือ '⠸⠞') มีรหัสอักษรเบรลล์ที่แตกต่างกัน จึงสามารถตีความอักขระเบรลล์คู่นี้ได้ถูกต้องไม่ว่าจะปรากฏในส่วนใดของวรรคก็ตาม

ในเบรลล์ไทย มีการยืมเครื่องหมาย '⠸' (= -) จากภาษาอังกฤษมาใช้ เป็นผลให้อักษรเบรลล์ที่ขึ้นต้นด้วย '⠸' ได้แก่ '⠸⠽' (= ย) และ '⠸⠍' (= ฒ) เกิดความกำกวม เนื่องจากทั้ง "ย" และ "ฒ" เกิดจากการเติม '⠸' เข้าไปในอักขระฐานคือ "ส" และ "ท" ตามลำดับ ทำให้สามารถตีความให้อักษรเบรลล์ 2 คู่นี้เป็น "-ส" และ "-ท" ได้ ดังนั้นจึงควรมีการแก้ไข '⠸' ซึ่งใช้เติมหน้าพยัญชนะฐาน สำหรับพยัญชนะ 2 ตัวนี้ โดยใช้รหัสอักษรเบรลล์อื่น เพื่อลดความกำกวมที่เกิดขึ้น

นอกจากอักขระเบรลล์ 2 ตัวที่กล่าวมาแล้ว ควรมีการแก้ไข '⠒' ในเบรลล์ไทย ที่ใช้แทน "เ◌ิ" หรือ "เ◌อ"ได้ เนื่องจากการให้อักขระตัวนี้สามารถแทนได้ 2 กรณีนั้นก่อให้เกิดปัญหาต่อการถ่ายถอดอักษรสำหรับคำต่าง ๆ โดยเฉพาะคำยืมจากภาษาต่างประเทศ ดังเช่นระบบได้ถ่ายถอด "เฟิร์น" ผิดเป็น "เฟอร์น" และเหตุผลประการสำคัญคือ ในโลกปัจจุบันคนตาบอดต้องสื่อสารกับโลกภายนอกด้วยการพิมพ์บนแป้นพิมพ์ปกติมากขึ้น ทำให้คนตาบอดต้องเรียนรู้การสะกดคำในอักษรปกติ การกำหนดอักขระตัวนี้แทนได้ 2 กรณี จะทำให้ผู้ใช้อักษรเบรลล์ในชีวิตประจำวันจดจำตัวสะกดของคำที่ประกอบด้วยสระเออได้ยากยิ่งขึ้น เพราะในเบรลล์ไทยมีอักขระเบรลล์แทนสระเออเพียงรูปเดียวในขณะที่ในอักษรปกติสามารถแทนได้ 2 รูป ดังนั้น จึงควรทำการแก้ไขรหัสอักษรเบรลล์ให้ '⠒' แทน "เ◌อ" เท่านั้น แล้วเขียนสระเออที่มีรูปเป็น "เ◌ิ" แบบตัวต่อตัวเหมือนอักษรปกติ เพราะจะสามารถแก้ปัญหา 2 ประการที่กล่าวมาข้างต้นได้

ดังที่ได้กล่าวมาแล้วว่า ในบางกรณี การประสมอักษรในเบรลล์ไทยที่เกิดกับสระบางตัวมีความแตกต่างจากอักษรปกติแม้ว่าสระเหล่านี้จะเทียบกับอักษรปกติได้แบบตัวต่อตัวก็ตาม ดังที่ปรากฏกับการประสมอักษร ที่มี "ะ" และ "◌ำ" ปรากฏร่วมกับวรรณยุกต์ ที่แตกต่างจากไทยปกติ ดังนั้นเพื่อให้คนตาบอดสามารถเรียนรู้การสะกดคำในอักษรปกติได้ง่ายยิ่งขึ้น ควรยึดการประสมอักษรในลักษณะนี้เหมือนกับอักษรปกติ เพราะการประสมอักษรแบบอักษรปกติใช้จำนวนเซลล์ปริมาณเท่ากันจึงไม่ก่อให้เกิดปัญหาเพราะยังใช้เนื้อที่ในการแสดงอักษรเบรลล์เท่าเดิม ในทางตรงกันข้าม กลับเป็นผลดีที่ทำให้ผู้ใช้อักษรเบรลล์ปรับตัวได้ง่ายยิ่งขึ้นกับโลกปัจจุบันที่การสะกดคำมีความสำคัญเพิ่มขึ้นเรื่อย ๆ

8.3.2 การแก้ไขตัวอักษรเบรลล์อังกฤษ

ผลการวิจัยแสดงให้เห็นว่า เบรลล์อังกฤษรูปเต็มไม่ก่อให้เกิดปัญหาต่อการถ่ายถอดให้เป็นอักษรปกติมากเท่าใด แต่การถ่ายถอดเบรลล์รูปย่อยังคงเกิดปัญหากับอักขระเบรลล์บางตัวที่มีความกำกวม จึงจำเป็นต้องแก้ไขรหัสเบรลล์เหล่านี้

ดังที่ได้กล่าวมาแล้วว่า ‘⠤’ เมื่อปรากฏในเบรลล์ไทยอาจเป็นปัญหาต่อการถ่ายถอดอักษร จากการทดลองถ่ายถอดเบรลล์อังกฤษรูปย่อ ผู้วิจัยพบว่า อักขระเบรลล์ตัวนี้ยังก่อให้เกิดปัญหาในเบรลล์อังกฤษรูปย่อเช่นเดียวกัน แม้อักขระตัวนี้ไม่ก่อให้เกิดปัญหาหากปรากฏในตำแหน่งกลางและท้ายวรรค แต่จะเป็นปัญหาเมื่อปรากฏต้นวรรค เพราะสามารถตีความให้เป็น “com” (ในกรณีทั่วไป) หรือ “-” (โดยเฉพาะเมื่อเขียนนำหน้าคำประเภทปัจจัยในภาษาอังกฤษ) ได้ ทำให้ในบางกรณีก็ยากที่จะระบุได้ว่า ‘⠤’ ใช้แทน “-” หรือ “com” โดยเฉพาะอย่างยิ่งการระบุด้วยคอมพิวเตอร์ เช่น ‘⠤⠒⠿’ ต้องตีความให้เป็น “-ing” หรือ “coming” ดังนั้น เพื่อขจัดความกำกวมที่เกิดขึ้น จึงไม่ควรนำ ‘⠤’ มาใช้ย่อ “com” อีกต่อไป ควรจะเขียน “com” ให้อยู่ในรูปเต็ม

นอกจากนี้ ในเบรลล์อังกฤษรูปย่อ ‘⠸’ สามารถใช้แทน “/” หรือ “st” ได้ จากการทดลอง ระบบได้ถ่ายถอดผิดจากกรณีหนึ่งเป็นอีกกรณีหนึ่งในหลายแห่ง ดังนั้น ในการพิมพ์ข้อความเบรลล์ที่เป็นอังกฤษรูปย่อ จึงควรใส่ ‘⠸’ (= สัญลักษณ์กำกับเครื่องหมายเบรลล์) ไว้ข้างหน้า ‘⠸’ เพื่อระบุว่า ‘⠸’ ใช้แทน “/” ในข้อความเบรลล์รูปย่อมาตรฐานเสมอ

ในอังกฤษรูปย่อ ‘⠠’ (capital sign) เมื่อปรากฏในตำแหน่งต้นคำ จะใช้กำกับว่าอักษรที่ตามมาเป็นอักษรตัวใหญ่ แต่เมื่อปรากฏในตำแหน่งอื่นของคำสามารถใช้เป็นตัวย่อสำหรับตัวย่อ 2 ชุด คือ ‘⠠⠝’ ใช้ย่อ “ation” และ ‘⠠⠁’ ใช้ย่อ “ally” อย่างไรก็ตาม จากข้อความที่ใช้ทดสอบระบบยังแสดงให้เห็นว่า ‘⠠⠝’ ที่ปรากฏกลางคำใช้แทน “N” ได้ เช่น ‘⠠⠕⠝⠑⠠⠝⠕⠞⠑’ ใช้แทน “OneNote” ไม่ใช่ “Oneationote” ดังนั้น เพื่อลดความกำกวมของตัวย่อทั้ง 2 ชุดที่กล่าวมา จึงควรแก้ไข ‘⠠’ ที่ปรากฏหน้าตัวย่อชุดเหล่านี้ให้เป็นรหัสอักษรเบรลล์อื่นแทน

8.3.3 การกำกับภาษาในอักษรเบรลล์

เนื่องจากเบรลล์ไทยและเบรลล์อังกฤษมีการซ้ำรูปกันเป็นจำนวนมาก และบางครั้งการเรียงต่อของตัวอักษรเหล่านี้ในสองภาษาก็คล้ายคลึงกัน ทำให้บางครั้งผู้ใช้อักษรเบรลล์ต้องคาดเดาจากบริบทที่

อักษรนั้นปรากฏอยู่ว่าเป็นภาษาใดจึงทำให้ต้องใช้เวลาในการตีความมากขึ้น และในขณะเดียวกันก็ไม่สะดวกต่อการนำไปถ่ายถอดอักษรด้วยคอมพิวเตอร์เพราะต้องใช้กระบวนการที่ซับซ้อนจึงจะถ่ายถอดได้ถูกต้อง

เพื่อขจัดปัญหาเหล่านี้ จึงควรมีเครื่องหมายกำกับภาษาในอักษรเบรลล์สำหรับภาษาไทยและภาษาอังกฤษ อันจะทำให้ผู้ใช้ทราบโดยทันทีว่ากำลังใช้อักษรเบรลล์ภาษาใดอยู่ อีกทั้งยังสะดวกต่อการนำอักษรเบรลล์ไปประมวลผลด้วยคอมพิวเตอร์ ทำให้ไม่ต้องใช้กระบวนการที่ซับซ้อนในการระบุภาษาอีกต่อไป

8.3.4 การวิจัยในอนาคต

มีนักวิชาการตาบอดไทยกลุ่มหนึ่งได้สร้างอักษรเบรลล์ไทยรูปย่อขึ้นโดยที่หวังว่าจะทำให้การใช้อักษรเบรลล์ไทยเกิดประสิทธิภาพมากยิ่งขึ้น แต่ระบบตัวย่อที่สร้างขึ้นนี้ยังไม่ได้มีการศึกษาถึงความเหมาะสมและความสะดวกในการใช้งาน ด้วยเหตุนี้ผู้วิจัยจึงไม่ได้นำเบรลล์ไทยรูปย่อมาเป็นส่วนหนึ่งของการวิจัย อย่างไรก็ตามเพื่อให้เบรลล์ไทยรูปย่อเป็นประโยชน์ต่อคนตาบอดอย่างแท้จริง จึงควรมีการศึกษาถึงความเหมาะสมในการใช้อักษรเบรลล์ประเภทนี้ เพื่อปรับเปลี่ยนแก้ไขให้คนตาบอดส่วนใหญ่สามารถเข้าถึงระบบการเขียนนี้ได้

ในวงการคนตาบอดเอเชียตะวันออกเฉียงใต้มีการแลกเปลี่ยนองค์ความรู้เกี่ยวกับอักษรเบรลล์ของแต่ละชาติกันน้อยมาก จึงทำให้อักษรเบรลล์แต่ละภาษาแทบไม่เป็นที่รู้จักกันในประเทศกลุ่มสมาชิกเลย ในความเป็นจริงแล้วภาษาเหล่านี้มีตัวเขียนส่วนหนึ่งคล้ายคลึงกัน จึงน่าสนใจว่าอักษรเบรลล์ที่ปรากฏในภาษาเหล่านี้มีความเหมือนหรือต่างกันอย่างไร

ในการศึกษาครั้งนี้ ผู้วิจัยได้ทำการศึกษาการถ่ายถอดเบรลล์ไทยปนอังกฤษรูปเต็มและการถ่ายถอดเบรลล์ไทยปนอังกฤษรูปย่อแยกออกจากกันเป็น 2 ระบบ โดยผู้ใช้ต้องส่งข้อมูลที่เป็นเบรลล์ไทยปนอังกฤษรูปย่อ หรือเบรลล์ไทยปนอังกฤษรูปเต็มให้ถูกต้องกับระบบเอง แต่ส่วนนี้ควรมีงานวิจัยในอนาคตที่วิเคราะห์ข้อความที่เข้ามา แล้วระบุว่าเป็นเบรลล์ไทยปนอังกฤษรูปย่อหรือเบรลล์ไทยปนอังกฤษรูปเต็มโดยอัตโนมัติและส่งเข้าในระบบที่เหมาะสมได้ ซึ่งจะทำให้ระบบการถ่ายถอดอักษรที่ผู้วิจัยสร้างขึ้นมีสมรรถภาพในระดับที่สามารถแยกเบรลล์อังกฤษรูปเต็มออกจากเบรลล์อังกฤษรูปย่อได้ นอกจากนี้ในอนาคต ควรมีงานวิจัยที่ศึกษาการถ่ายถอดอักษรเบรลล์ที่รองรับการปนภาษาทั้งเบรลล์ไทย อังกฤษรูปเต็มและอังกฤษรูปย่อในระบบเดียวกัน

เท่าที่ผ่านมา ได้มีงานวิจัยที่ศึกษาการถ่ายถอดอักษรเบรลล์ด้วยคอมพิวเตอร์อยู่จำนวนหนึ่ง แต่งานวิจัยเหล่านี้ดำเนินการโดยการสร้างคลังข้อความเบรลล์ขึ้นเอง เพราะผู้ทำการวิจัยไม่สามารถหาข้อความเบรลล์ที่เป็นมาตรฐานได้ เพื่อให้งานวิจัยเกี่ยวกับอักษรเบรลล์ทำได้สะดวกยิ่งขึ้นในอนาคตจึงควรมีการจัดทำคลังข้อความเบรลล์ไทย คลังข้อความเบรลล์อังกฤษรูปเต็ม และคลังข้อความเบรลล์อังกฤษรูปย่อ เพราะนอกจากจะช่วยให้ได้งานวิจัยที่มีประสิทธิภาพมากยิ่งขึ้นแล้ว ยังเป็นประโยชน์ต่อผู้ที่สนใจอักษรเบรลล์ที่สามารถใช้เป็นแหล่งข้อมูลในการศึกษาอักษรเบรลล์ได้อีกรูปแบบหนึ่ง

ปัจจุบันได้มีการนำอักษรเบรลล์ไปใช้กับสาขาวิชาต่าง ๆ เช่น วิศวกรรมศาสตร์ ภาษาศาสตร์ คณิตศาสตร์ ดนตรี ฯลฯ แต่เท่าที่ผู้วิจัยได้ทำการศึกษาจากเอกสารต่าง ๆ พบว่ามีงานวิจัยที่ศึกษาเกี่ยวกับการถ่ายถอดสัญลักษณ์เบรลล์เฉพาะทางเหล่านี้เป็นจำนวนน้อย จึงควรมีงานวิจัยมารองรับกับการถ่ายถอดสัญลักษณ์เบรลล์ในสาขาวิชาต่าง ๆ อันจะทำให้คนตาบอดสามารถแลกเปลี่ยนความรู้ในด้านต่าง ๆ กับโลกภายนอกได้สะดวกยิ่งขึ้นผ่านทางอักษรเบรลล์

รายการอ้างอิง

ภาษาไทย

กานดา รุณนะพงศา และปโยธร อุราธรรมกุล. 2549. *การตัดคำภาษาไทยโดยการปรับปรุงกฎและพจนานุกรมแบบใหม่*. รายงานวิจัย คณะวิศวกรรมศาสตร์ มหาวิทยาลัยขอนแก่น.

ทศวัฒน์ ชุณหวิทยะธีระ. 2555. *การใช้เทคนิคเอ็นแกรมช่วยในการเพิ่มประสิทธิภาพการตัดสินใจแปลงอักษรเบรลล์ในกรณีที่มีคำควบกล้ำ สระประสม และอักษรเบรลล์สองเซลล์*. วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยสงขลานครินทร์.

ทัศนวรรณ ศูนย์กลาง. 2543. *การเข้ารหัสคำทับศัพท์ภาษาไทย/อังกฤษ เพื่อการค้นคืนข้ามภาษาด้วยเทคนิคนิวรอลเน็ตเวิร์ก*. วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาวิทยาศาสตร์คอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

รวิศม์ วงษ์สมาน. 2551. *การออกแบบและพัฒนาตัวแปลงภาษาไปกลับอังกฤษผสมไทยเป็นเบรลล์*. วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

วรพล ทินกรสูติบุตร. 2551. *การแปลงเบรลล์และแอสกีภาษาไทยแบบทันทีทันใด*. วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยสงขลานครินทร์.

วัลย์วรา ไชยฤกษ์. 2547. *การพัฒนาโปรแกรมการถอดอักษรภาษาอังกฤษเป็นภาษาไทยโดยใช้คลังคำทับศัพท์ของราชบัณฑิตยสถาน*. วิทยานิพนธ์ปริญญาอักษรศาสตรมหาบัณฑิต. ภาควิชาภาษาศาสตร์ คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

วิโรจน์ อรุณมานะกุล. 2551. *อักขรวิธีไทยและการถอดอักษรระหว่างภาษาไทยและภาษาอังกฤษ*. กรุงเทพ. คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

_____. 2548. *ระบบการถอดอักษรระหว่างภาษาไทยและภาษาอังกฤษ:แนวทางและการพัฒนา*. รายงานวิจัย สำนักกองทุนสนับสนุนการวิจัย.

วีรชัย อำพรไพบูลย์. 2546. *คู่มือตัวย่อเบรลล์อังกฤษทางลัด*. กรุงเทพ. อินเด็กซ์แอสโซซิเอด.

สมทรง พันธุ์สุวรรณ. 2538. *เอกสารประกอบการสอนวิชาการอ่าน เขียน พิมพ์อักษรเบรลล์.* กรุงเทพ. จงเจริญการพิมพ์.

สุขชาตรี ประสมสุข. 2555. "*Chapter7_DecisionTreeLearning*" [Online]. Available from: www.ict.up.ac.th/skchatri/subject6.pdf

อัครพล เอกวงศ์อนันต์. 2548. *การระบุคำไทยและคำทับศัพท์ด้วยแบบจำลองเอ็นแกรม.* วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาภาษาศาสตร์ คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

อาทิตยา บุญมาก และคณะ. 2555. *คู่มือการอ่าน การเขียน และการผลิตเอกสารอักษรเบรลล์ขั้นพื้นฐาน.* งานการศึกษาพิเศษ กองพัฒนานักศึกษา มหาวิทยาลัยเชียงใหม่

ภาษาอังกฤษ

Bosch, A.Vanden and Dealemans, W. 1993. *Data-Oriented Method for Grapheme-to-Phoneme Conversation.* In Processing of the 6th conference of the European Chapter of the Association for Computational Linguistics. Utrecht. Netherland.

Bussmannm, Hadumod. 1998. *Routledge Dictionary of Language and Linguistics.* Routledge Press.

California Community Colleges. 2010. *Uncontracted Braille Training Manual.* California. [Online] Available: http://www.htctu.net

Carney, Edward. 1994. *A Survey of English Spelling.* London. Routledge Press.

Caulfied, Genevieve. 1960. *The Kingdom Within*. New York. Harper and Brother.

Cavnar, Willian. B. and Trenkle, John. M. 1994. N-Gram-Based Text Categorization. In *Proceedings of Third Annual Conference on Document Analysis and information Retrieval (SDAIR).* Las Vegas.

Deep, Kamal and Goyal, Vishal. 2011. Development of a Punjabi to English Transliteration System. *International Journal of Computer Science and Communication*. (2)2 July-December 2011.

Duxburysystems. 2012. *History of Duxbury Systems, Inc.* duxburysystems.com/duxhist.asp. Online Resources. [Accedded: 2015 June 5]

Dunning, Ted. 1994. *Statistical Identification of Language*. Technical Report MCCS. Computing Research Laboratory. New Mexico State University.

Englebretson, Robert. 2009. An Overview of IPA Braille: An Updated Tactile Representation of the International Phonetic Alphabet. *Journal of the International Phonetic Association.* (39/1)

Gotwals, R. Robert. 2008. *Braille Through Remote Learning: Introduction to Braille.* The Shodor Education Foundation. Durham. NC.

Goyal, Vishal and Lehal, Gurpreet Singh. 2009. Hindi-Punjabi Machine Transliteration System. *George Ronchi Foundation Journal*, Italy. 64 (1)

Grefenstette, Gregory. 1995. *Comparing Two Language Identification Schemes.* In *Proceedings of the 3rd International Conference on Statistical Analysis of Texual Data.* Rome.

Grothe, L., De Luca, E. W. & Nürnberger, A. 2008. A Comparative Study on Language Identification Methods. In *Proceedings of the 6th International Conference on Language Resources and Evaluation.* Marrakech.

Japanese Braille. 2014. http://www.jfbkk.or.th/pdf/JL/2014. [Accedded: 22-Apr-2015]

Jia, Yuxiang Zhu, Danquing and Yu Shiwen. 2009. *Noisy Channel Model for Grapheme-based Machine Transliteration.* In Proceeding of the 2009 Named Entities Workshop. Suntec. Singapore.

Kang, Byung-Ju and Choi, Key-Sun. 2000. Two Approaches for the Resolution of Word Mismatch Problem Caused by English Words and Foreign Words in Korean Information Retrieval. In *Proceeding of the 2nd International Conference on Language Resources and Evaluation.* KAIST. Korea.

Kaur, Kamaljeet and Singh, Parminder. 2014. Review of Machine Transliteration Techniques. *International Journal of Computer Applications (107) 20.*

Klenk, A. Jessicia and Pufpaff, A. Lisa. 2011. *A Case Study of Tack Tiles Literacy Instruction for a Student with Multiple Disabilities Including Congenital Blindness.* Muncie. Ball State University.

Knight, Kevin and Graehl, Jonathan. 1997. Machine Transliteration. In *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and 8th Conference of the European Chapter of the Association for Computational Linguistics.*

Lauenstein, Christine. 2007. *On the Compatibility of the Braille Code and Universal Grammar.* Ph.D. Dissertation. Institute of English Linguistics. Stuttgart University.

Lee, Chun-Jen and Chang, Jason S. 2003. *A cquistion of English-Chinese Transliterated Word Pairs from Parallel Aligned Texts using a Statistical Machine Transliteration Model.* National Tsing Hua University. Taiwan.

Lee, Jae Sung and Choi, Key-Sun. 1998. English to Korean Statistical Transliteration for Information Retrieval. *Computer Processing of Oriental Languages. (12)1.*

Lee, Jae Sung. 1999. *An English-Korean Transliteration and Retransliteration Models for Cross-Lingual Information Retrieval.* Ph.D. Dissertation. Computer Science Dept. Korean Advanced Institute of Science and Technology (KAIST).

Martino, Michael .John. and Paulsen Robert Charles. 2001. Natural Language Determination using Partial Words. *U.S. Patent Documents No.6216102 B1.*

Oard, O. W. 1998. *A Comparative Study of Query and Document Translation for Cross-Language Information Retrieval.* College of Libraries and Information Services. University of Marryland.

Oh, Jong-Hoon and Choi, Key-Sun. 2002. *An English-Korean Transliteration Model, using Pronunciation and Contextual Rules.* In Proc. of the 19th International Conference on Computational Linguistics.

Oh, Jong-Hoon, Choi, Key-Sun and Isahara, Hitoshi. 2006. A Machine Transliteration Model Based on Correspondence between Graphemes and Phonemes. *Asian Language Information Processing. (5) 3.*

Pager, J. M. 1999. *Linguini: Language Identification for Multilingual Documents.* In Proceedings of the 32nd Hawaii International Conference on System Sciences.

Pethő, Gergely. and Mózes, Eszter. 2014. An-N-Gram-Based Language Identification Algorithm for Variable–length and Variable-language Texts. *Argumentum*
10 (2014). Debreceni Egyetemi Kiadó. University of Debrecen.

Petzold, Charles. 2000. *The Hidden Language of Computer Hardware and Software.* Microsoft Press. Washington.

Risjord, Constance. 2009. *Instruction Manual for Braille Transcribing.* Fifth Edition, National Library Service for the Blind and Physically Handicapped. The Library of Congress.

Surana, Harshit and Singh, Anil Kumar. 2001. A more Discerning and Adaptable Multilingual Transliteration Mechanism for Indian Languages. *In Proceeding of 3rd International Joint Conference on Natural Language Processing in NLP.* Hyderbad. India.

Souter, Clive. Chuecher, J. Hayes, J. and Johnson, S. 1994. Natural Language Identification using corpus-based models. *Hermes Journal of Linguistics.* (2)13

Sullivan, M.J. 1999. *Duxbury Braille translator (DBT) for Windows*. Duxbury Systems, Inc. Massachusetts.

Suvanvisat, Prayut and Prasitjutrakkul, Somchai. 1998. *Thai-English Cross-Language Transliterated Word Retrieval using Soundex Technique*. The National Computer Science and Engineering Conference. Kasetsart University, Bangkok.

Tennessee Council of the Blind (TCB). 2008. *What is Braille?* Knoxville. Tennessee. [Online] Available from: http://www.tcb.org [Accedded: 21-Apr-2015]

The Braille Authority of North America (BANA). 2002. *English Braille American Edition 1994 (Revised 2002).* Pittsburgh. Pennsylvania. [Online] Available from: http://www.brailleauthority.org [Accedded: 21-Apr-2015]

ภาคผนวก

รายการภาคผนวก

ภาคผนวก ก. เซ็ตไตรแกรมท้ายพยางค์เพื่อใช้ในการถ่ายถอดเบรลล์ไทยด้วยกฎ

ภาคผนวก ข. คู่เทียบสำหรับเบรลล์อังกฤษรูปย่อเพื่อใช้ในการถ่ายถอดเบรลล์อังกฤษด้วยกฎ

ข.1 คู่เทียบคำย่อเบรลล์อังกฤษ

ข.2 คู่เทียบตัวย่อเบรลล์เซลล์คู่

ภาคผนวก ค. เซ็ตของสายอักขระที่เป็นเอกลักษณ์ของแต่ละภาษาเพื่อใช้ระบุภาษาด้วยกฎ

ค.1 เซ็ตของสายอักขระที่เป็นเอกลักษณ์เบรลล์ไทยที่ใช้กับข้อความเบรลล์ไทยปนอังกฤษรูปเต็ม

ค. 2 เซ็ตของสายอักขระที่เป็นเอกลักษณ์เบรลล์อังกฤษ ที่ใช้กับข้อความเบรลล์ไทยปนอังกฤษรูปเต็ม

ค.3 เซ็ตของสายอักขระที่เป็นเอกลักษณ์เบรลล์ไทย ที่ใช้กับข้อความเบรลล์ไทยปนอังกฤษรูปย่อ

ค.4 เซ็ตของสายอักขระที่เป็นเอกลักษณ์เบรลล์อังกฤษ ที่ใช้กับข้อความเบรลล์ไทยปนอังกฤษรูปย่อ

ภาคผนวก ง. คู่เทียบเบรลล์ไทยสำหรับระบบการถ่ายถอดด้วยแบบจำลองเอ็นแกรม

ภาคผนวก จ. คู่เทียบเบรลล์อังกฤษสำหรับระบบการถ่ายถอดด้วยแบบจำลองเอ็นแกรม

ภาคผนวก ฉ. คู่เทียบเบรลล์อังกฤษรูปย่อสำหรับระบบการถ่ายถอดด้วยแบบจำลองเอ็นแกรม

ภาคผนวก ช. คู่เทียบเบรลล์อังกฤษรูปย่อสำหรับระบบการถ่ายถอดวิธีการแบบผสม

ภาคผนวก ซ. ตัวอย่างข้อความเบรลล์ทั้ง 5 ประเภทที่ใช้ทดสอบระบบ

ภาคผนวก ฌ. ข้อผิดจากการทดสอบระบบการถ่ายถอดอักษรด้วยกฎ

ภาคผนวก ญ. ข้อผิดจากการทดสอบระบบการถ่ายถอดอักษรด้วยแบบจำลองเอ็นแกรม

หมายเหตุ: | ใช้เพื่อแสดงว่าในการถ่ายถอดอักษรต้องทำการสลับที่ตัวอักษรของคู่เทียบกับอักษรที่นำหน้ามาก่อน

ภาคผนวก ก. เซ็ตไตรแกรมท้ายพยางค์เพื่อใช้กับระบบการถ่ายถอดด้วยกฎ

ภาคผนวก ก.1 กลุ่มไตรแกรมอักษรหน้าสระผสม

ไตรแกรมอักษรหน้าสระผสม											
ตก	"บท	กชก	กีส	ขาก	ฆษก	งโต	จ็ค	ซัก	ณนก	ดัค	ยรก
ทศ	"ยก	กดก	กึก	ขิต	ฆาต	งโป	จ็ท	ซัส	ณยก	ดัพ	ยรถ
นก	"แก	กถก	กุส	ขไป	งจก	งใจ	จ๊ก	ซาก	ณรถ	ดาก	ยรส
บก	"ไป	กนก	กเก	คนต	งดก	งใส	จ๊ค	ซิก	ณิก	ดาจ	ยวก
บท	(ยก	กบก	กเต	คยก	งตก	งไต	ฉาก	ซิป	ณิต	ดาป	ยศพ
ปก	(ไป	กบท	กแก	ครก	งถก	งไท	ฉาพ	ซิส	ณิศ	ดิก	ยอก
ภพ	,ยก	กภพ	กโก	ครต	งน	งไป	ฉีก	ซีก	ณีต	ดิค	ยอส
ยก	-บท	กยก	กโต	ครส	กง	งไพ	ฉไป	ซุป	ณโต	ดิต	ยัก
รก	-ยก	กรก	กโป	คอก	บก	งไส	ชดก	ซโก	ณไป	ดีต	ยาก
รถ	-อก	กรถ	กใจ	คอท	งบท	งๆก	ชนก	ซโค	ณไพ	ดึก	ยาท
รส	-แก	กรส	กใส	คอส	งปก	จบท	ชรถ	ซโจ	ด ป	ดุก	ยิก
ลก	-โต	กฤต	กไค	คัก	งภพ	จยก	ชัก	ซไป	ดตก	ดุจ	ยุค
ศพ	.บท	กวก	กไต	คัต	งยก	จรถ	ชัส	ซ็ต	ดนก	ดูก	ยเก
อก	.ยก	กศก	กไท	คัพ	งรก	จรส	ชาก	ซ็ท	ดบท	ดแค	ยแก
เก	.ใจ	กศพ	กไป	คัส	งรถ	จวก	ชาต	ซ็ป	ดพก	ดโต	ยแท
เจ	.ไป	กหก	กไพ	คิก	งวก	จศพ	ชิก	ซ้ทญาต	ดยก	ดใจ	ยแพ
แก	.ไพ	กอก	กไส	คีต	งศพ	จอก	ชิต	ญิก	ดยศ	ดใส	ยโต
แค	/ยก	กอส	ก็ก	คึก	งสก	จัก	ชิพ	ญเข	ดร	ดไถ	ยใจ
แพ	14ก	กัก	ก็ต	คุก	งหก	จาก	ชีพ	ญแจ	ดรต	ดไป	ยใส
โก	3ไป	กัป	ก่ป	คโก	งอก	จาค	ชีส	ญใจ	ดรถ	ดไพ	ยไป
โค	4 ก	กัส	ก๊ก	คโป	งัก	จิก	ชุก	ญไป	ดรส	ด์ก	ยไพ
โต	[ยก	กาก	ก๊ต	คไต	งิก	จิต	ชเค	ฎยก	ดลก	ด์จ	รกต
โพ	‘บท	กาจ	ก๊ส	คไท	งเก	จุก	ชไป	ฐยก	ดวก	ด้ก	รจก
ใจ	‘ยก	กาศ	ขบถ	คไป	งเป	จใจ	ช็ก	ฐไป	ดศพ	ด๊ป	รดก
ไต	‘ไป	กาส	ขยก	ค็ท	งแก	จไป	ช็ค	ฐไพ	ดสต	ยมก	รตก
ไป	’บท	กิจ	ขลก	ค้ก	งแค	จไพ	ช้ค	ฑิต	ดอก	ยมส	รบก
ไพ	กกก	กิต	ขอก	ค้ต	งแพ	จ็ก	ซอก	ฑูต	ดัก	ยยก	รภพ

ไตรแกรมอักษรหน้าสระผสม (ต่อ)											
รยก	ริค	รไพ	ลีค	วตก	ศยก	สเต	อกก	อเป	ฯยก	ัักก	ารค
รยศ	ริช	ร์ช	ลีส	วบท	ศรถ	สเป	องพ	อแก	ฯไป	ัตถ	ารถ
รรก	ริต	ร์ต	ลึก	วภพ	ศอก	สโก	อชก	อแป	ฯไพ	ันก	ารพ
รรค	ริป	ร์ท	ลุก	วยก	ศาจ	สโค	อตก	อแพ	ะจก	ัรต	ารส
รรต	ริส	ร์ป	ลุค	วรถ	ศึก	สใจ	อนก	อโก	ะดก	ัลส	าฤก
รรถ	รึก	ร์ส	ลูก	วรส	ศึต	สใส	อบท	อโต	ะตก	ัสส	าลก
รรพ	รีต	ลถก	ลูต	วศพ	ศึก	สไค	อพก	อโท	ะถก	ั่ก	าวก
รรส	รีท	ลนก	ลเก	วหก	ษณพ	สไป	อภพ	อโพ	ะนก	ั๊ก	าศพ
รลพ	รีพ	ลยก	ลแก	วอก	ษบก	สไพ	อยก	อใจ	ะบท	ากก	าอก
รวก	รีส	ลวก	ลแข	วัก	ษหก	สี็ค	อรก	อไข	ะยก	าคก	าฮก
รวจ	รึก	ลอก	ลโก	วัต	ษัท	ส่จ	อรถ	อไต	ะรก	าคต	าเก
รศพ	รุก	ลอค	ลโต	วัส	ษิต	หงก	อรส	อไป	ะรถ	าฆพ	าเข
รสพ	รุต	ลอต	ลใจ	วาท	ษไป	หนก	อลก	อไพ	ะวก	าจก	าแก
รหก	รุป	ลอพ	ลไก	วาส	ษไพ	หมก	อวก	อ้ก	ะศพ	าฐก	าแพ
รอก	รูป	ลอส	ลไป	วิก	สบถ	หยก	อศพ	อ๊ค	ะสพ	าดก	าโก
รอส	รโค	ลัก	ลไพ	วิต	สยก	หรก	ออก	อ๊ต	ะอก	าตก	าโต
รัก	รโต	ลัส	ล็ต	วิท	สยป	หลก	อัก	ฮอป	ะเป	าถก	าโท
รัค	รใจ	ลาก	ล้ค	วิป	สัก	หวก	อัจ	ฮัก	ะเส	านก	าโป
รัช	รไป	ลาท	ล้ส	วิส	สาก	หวต	อัช	ฮัพ	ะแก	าบก	าใจ
รัต	ร็ก	ลาศ	ล์ก	วีท	สาท	หสถ	อัพ	ฮิจ	ะแต	าบถ	าใส
รัท	ร็ค	ลาส	ล์ค	วีป	สาป	หอก	อาก	ฮิต	ะแส	าบท	าไก
รัส	ร็จ	ลิก	วเค	วใป	สิก	หัก	อาจ	ฮีต	ะโค	าพก	าไป
راก	ร็ต	ลิค	วเจ	วใส	สิค	หัส	อิก	ฮุก	ะโต	าภพ	าไพ
راช	ร็ป	ลิต	วแพ	วไถ	สิต	หาก	อิต	ฮูก	ะโส	ามก	ำกก
راต	ร็พ	ลิป	วโต	วไท	สึก	หิต	อีก		ะใจ	ายก	ำคก
راพ	ร๊ก	ลิศ	วใจ	วไป	สุก	หุก	อึก		ะไต	ารก	ำตก
راศ	ร์ก	ลิส	ว จ	วไพ	สุข	หุต	อุส		ะไป		ำบท
راส	ร์ค	ลีก	ว พ	ว๊ก	สูท		อเค		ะไพ		ำยก
ริก	ร์จ	ล็ก	วกส				อเท				ำรถ

ไตรแกรมอักษรหน้าสระผสม (ต่อ)											
ำรส	ิไต	ีั๊ป	ูั๊ป	เบก	เสพ	แลต	ใจจ	่รส	้อก	์ถก	ใคร (กลุ่มตัวอักษรหน้า ไม้ยมก(ๆ))
ำศพ	ิไป	ีั้ก	เกจ	เบค	เอก	แวก	ใจต	่วก	้อค	์นก	
ำหก	ิั้ก	ือก	เกต	เบต	เอท	แสก	ไกป	่ศพ	้อต	์บก	
ำอก	ิั้ต	ุตก	เกส	เบส	เอส	แอก	ไรส	่หก	้อป	์บท	
ำโพ	ิั๊ก	ุยก	เขต	เปก	แกต	แฮก	ไหส	่อก	้าก	์มก	
ำใจ	ิั๊ป	ุศพ	เขป	เปค	แขก	โคก	ๆใจ	่อป	้าส	์ยก	
ำใส	ีตก	ุหก	เคป	เพช	แคท	โงก	ๆไป	่าก	้แก	์รก	
ำไป	ีนก	ุโส	เคส	เพศ	แจก	โจก	็กป	่าส	้โค	์รถ	
ำไพ	ีบท	ุไป	เจก	เพส	แฉก	โชก	็ตก	่เก	้โต	์ศพ	
ิฎก	ียก	ุไพ	เชส	เฟค	แชต	โชค	็ยก	่เค	้ใจ	์อก	
ิฏก	ียจ	ุั๊ก	เซต	เฟต	แซก	โตก	็รส	่เจ	้ไข	์เส	
ิตก	ียต	ุั๊ค	เซส	เภท	แดก	โทส	็อก	่เพ	้ไท	์แก	
ิบต	ียส	ุั๊ป	เฌศ	เมค	แตก	โบก	็อค	่แก	้ไป	์โต	
ิภพ	ีรถ	ูชก	เดก	เมส	แทก	โพก	็อต	่โก	้ไพ	์ใจ	
ิยก	ีรส	ูนก	เดต	เรก	แทค	โภค	็อท	่โต	ั๊บส	์ไป	
ิยท	ีศพ	ูบท	เดท	เรศ	แทต	โมค	็อป	่โป	ั๊ปส	์ไพ	
ิลก	ีสป	ูรก	เตต	เรส	แนก	โมต	็อพ	่ใจ	ั๊ยก		
ิลป	ีหก	ูรถ	เตท	เลข	แนต	โมท	็โก	่ไป	ั๊อก		
ิหก	ีแค	ูรส	เตส	เลต	แนท	โยก	็ไป	่ไพ	ั๊อค		
ิเก	ีโพ	ูศพ	เทค	เลส	แบก	โยค	็ไพ	้ตก	ั๊อต		
ิแก	ีใจ	ูเก	เทต	เวช	แบค	โรก	่ชก	้นก	ั๊อป		
ิแส	ีใส	ูแก	เทป	เวต	แฝก	โรค	่ดก	้นค	ั๊าก		
ิโก	ีไข	ูโก	เทพ	เวท	แพท	โรป	่ตก	้บท	ั๊าค		
ิโต	ีไป	ูโต	เทศ	เวศ	แมค	โรส	่นก	้ยก	ั๊ไป		
ิโป	ีไพ	ูใจ	เทส	เศส	แยก	โลก	่บท	้รก	ั๊ไพ		
ิโส	ี่ก	ูไป	เนต	เษก	แรก	โศก	่ยก	้รถ	์ จ		
ิใจ	ี้ต	ูไพ	เนท	เสก	แรต	โสก	่รก	้วก	์ดส		
	ี้พ	ูั๊ก	เนส	เสป	แลก	โสต	่รถ	้หก	์ตก		

ก.2 กลุ่มไตรแกรมที่มีการันต์ท้ายวรรค ก. 3 กลุ่มไตรแกรมที่มี ◌ี้ ปรากฏ

ไตรแกรมที่มีการันต์ท้ายวรรค				ไตรแกรมที่มีไม้โทปรากฏ							
กข์	ดส์	มน์	าร์	กล้	ค้า	ซ้ก	ท้อ	ฝ้า	ริ้	ส้ม	◌ั◌้บ
กซ์	ดิ์	มพ์	าว์	กี้	ค้ำ	ซ้ง	ท้า	พล้	รู้	ส้า	◌ั◌้ม
กต์	ตต์	มส์	าส์	กู้	งี้	ซ้น	นี้	พี้	ร้ง	หม้	◌ั◌้ย
กษ์	ตน์	รค์	าห์	ก้ง	ง้ง	ซ้อ	น้ต	พ้ง	ร้น	หว้	◌ั◌้ว
กส์	ตม์	รณ์	◌ีช์	ก้น	ง้น	ซ้า	น้น	พ้น	ร้ว	หู้	◌ิ◌้ง
คส์	ตย์	รพ์	◌ีณ์	ก้ม	ง้ม	ซ้ำ	น้ม	พ้ย	ร้อ	ห้ง	◌ิ◌้ต
งค์	ตร์	รย์	◌ีย์	ก้ว	ง้อ	ญ้า	น้อ	พ้ว	ร้า	ห้น	◌ิ◌้น
งศ์	ตล์	ลซ์		ก้อ	ง้า	ดี้	น้า	พ้อ	ร้ำ	ห้ว	◌ิ◌้ม
งษ์	ตว์	ลน์		ก้า	ง้ำ	ดี้	น้ำ	ฟี้	ลี้	ห้อ	◌ิ◌้ล
งห์	ทธ์	ลป์		ก้ำ	จี้	ด้ง	บี้	ฟ้ง	ล้ง	ห้า	◌ิ◌้ว
จญ์	ทย์	ลย์		ขว้	จู้	ด้น	ปู้	ฟ้น	ล้น	ห้ำ	◌ิ◌้ด
จน์	ทร์	ลห์		ขี้	จ้ก	ด้า	บ้ท	ฟ้ม	ล้ม	อี้	◌ิ◌้ย
ชญ์	ที์	วน์		ข้ง	จ้ง	ด้ำ	บ้อ	ฟ้อ	ล้ว	อู้	◌ิ◌้ก
ชฎ์	ทุ์	วส์		ข้น	จ้น	ตี้	บ้า	ฟ้า	ล้อ	อ้ก	◌ิ◌้ง
ชน์	ธิ์	ศน์		ข้ม	จ้ว	ตู้	ปล้	ม้ง	ล้า	อ้ง	◌ิ◌้ด
ชย์	ธุ์	ศย์		ข้อ	จ้อ	ต้ง	ปี้	ม้น	ล้ำ	อ้ด	◌ิ◌้น
ชร์	นด์	ศร์		ข้า	จ้ะ	ต้น	ป้ง	ม้ม	ว้	อ้น	◌ิ◌้บ
ชว์	นต์	ศษ์		คร้	จ้า	ต้ม	ป้น	ม้ว	ว้ง	อ้ม	◌ิ◌้ม
ซน์	นท์	ษฎ์		คล้	จ้ำ	ต้ว	ป้ว	ม้อ	ว้ด	อ้อ	◌ิ◌้ด
ญญ์	นธ์	ษฐ์		คู้	ฉ้อ	ต้อ	ป้อ	ม้า	ว้น	อ้า	◌ิ◌้น
ฎฐ์	นม์	ษณ์		ค้ก	ชี้	ต้า	ป้า	ม้ำ	ว้ย	ฮู้	◌ิ◌้ม
ฎร์	นย์	ษย์		ค้ง	ชู้	ต้ำ	ป้ำ	ยี้	ว้อ	ฮ้ย	◌ิ◌้อ
ฎฐ์	นส์	สน์		ค้ช	ช้ก	ถ้ง	ผล้	ย้ง	ว้า	ฮ้ว	◌ุ◌้ง
ฒน์	ปต์	อร์		ค้ด	ช้ค	ถ้อ	ผู้	ย้ม	สร้	ฮ้อ	◌ุ◌้น
ณฑ์	ปร์	อล์		ค้ต	ช้า	ถ้า	ผ้ง	ย้ย	สี้	ฮ้า	◌ุ◌้บ
ณย์	พย์	ะห์		ค้น	ช้ำ	ถ้ำ	ผ้ว	ย้อ	สู้	◌ั◌้ก	◌ุ◌้ม
ดร์	พร์	าข์		ค้อ	ซร้	ท้ง	ผ้า	ย้า	ส้ง	◌ั◌้ง	◌ุ◌้ย
				ค้า	ซี้	ท้น	ผ้ำ	ย้ำ	ส้น	◌ั◌้น	◌ู◌้ด

ก. 3 กลุ่มไตรแกรมที่มี ◌้ (ต่อ)

ไตรแกรมที่มีไม้โทปรากฏ (ต่อ)					
◌ู◌้น	แล้	ไซ้	◌้บก	◌้สด	◌้ใน
◌ู◌้บ	แว้	ได้	◌้บน	◌้สา	◌้ใบ
เก้	แส้	ไต้	◌้ปด	◌้หา	◌้ไข
เข้	แห้	ไถ้	◌้ผง	◌้หู	◌้ไฟ
เง้	แอ้	ไท้	◌้ผม	◌้อก	
เด้	โก้	ไม้	◌้ผี	◌้อง	
เต้	โจ้	ไร้	◌้พล	◌้อต	
เบ้	โซ้	ไล้	◌้ยง	◌้อน	
เป้	โต้	ไว้	◌้ยด	◌้อบ	
เย้	โบ้	ไส้	◌้ยน	◌้อม	
เว้	โป้	ไห้	◌้ยบ	◌้อย	
เอ้	โพ้	ไอ้	◌้ยม	◌้าก	
เฮ้	โม้	ไฮ้	◌้ยร	◌้าง	
แก้	โย้	◌้กง	◌้ยว	◌้าด	
แง้	โล้	◌้กบ	◌้ยศ	◌้าน	
แจ้	โว้	◌้กา	◌้ยา	◌้าบ	
แซ้	โห้	◌้ขม	◌้ยำ	◌้าม	
แด้	โอ้	◌้คน	◌้รบ	◌้าย	
แต้	ใช้	◌้ชม	◌้ลม	◌้าว	
แท้	ใต้	◌้ดำ	◌้ลำ	◌้เต	
แบ้	ใบ้	◌้ดี	◌้วก	◌้แย	
แป้	ใภ้	◌้ตก	◌้วง	◌้โท	
แพ้	ให้	◌้ตด	◌้วน	◌้โผ	
แฟ้	ไก้	◌้ทา	◌้วม	◌้โพ	
แม้	ไข้	◌้ที	◌้วย	◌้โอ	
แย้	ไค้	◌้นก	◌้วา	◌้โฮ	
แร้	ไจ้	◌้นำ	◌้ว์	◌้ใจ	

ก.4 กลุ่มไตรแกรมของ ◌๊

ไตรแกรมที่มีไม้ตรีปรากฏ					
ก๊ะ	กี๊	-	-	-	-
กำ๊	จี๊				
จ๊ะ	ฉู๊				
ด๊ะ	ตู๊				
ดำ๊	ปี๊				
ต๊ะ	ปู๊				
ตำ๊	◌ึ๊				
บ๊ะ	อิ๊				
ป๊ะ	เก๊				
อ๊ะ	เบ๊				
	โป๊				
	ไก๊				

ภาคผนวก ข. คู่เทียบสำหรับเบรลล์อังกฤษรูปย่อเพื่อใช้กับระบบการถ่ายถอดด้วยกฎ

ภาคผนวก ข.1 คู่เทียบคำย่อเบรลล์อังกฤษ

เบรลล์	แอสกี	อักษรปกติ
⠁⠃	ab	about
⠁⠃⠧	abv	above
⠁⠉	ac	according
⠁⠉⠗	acr	across
⠁⠋	af	after
⠁⠋⠝	afn	afternoon
⠁⠋⠺	afw	afterward
⠁⠛	ag	again
⠁⠛⠌	ag/	against
⠁⠇⠍	alm	almost
⠁⠇⠗	alr	already
⠁⠇	al	also
⠁⠇⠦	al?	although
⠁⠇⠞	alt	altogether
⠁⠇⠺	alw	always
⠵	z	as
⠆⠉	2c	because
⠆⠋	2f	before
⠆⠓	2h	behind
⠆⠇	2l	below
⠆⠝	2n	beneath
⠆⠎	2s	beside
⠆⠞	2t	between
⠆⠽	2y	beyond
⠃⠇	bl	blind
⠃⠗⠇	brl	braille
⠃	b	but
⠉	c	can

เบรลล์	แอสกี	อักษรปกติ
⠔	*	child
⠔⠝	*n	children
⠒⠧	3cv	conceive
⠒⠧⠬	3cvg	conceiving
⠙⠙	cd	could
⠙⠁⠙⠃⠇	d1fbl	deafblind
⠙⠧	dcv	deceive
⠙⠧⠬	dcvg	deceiving
⠙⠉⠇	dcl	declare
⠙⠉⠇⠬	dclg	declaring
⠙	d	do
⠑⠊	ei	either
⠢	5	enough
⠑	e	every
⠋⠌	f/	first
⠋⠗	fr	friend
⠋	f	from
⠛	g	go
⠛⠙	gd	good
⠛⠗⠞	grt	great
⠓	h	have
⠓⠋⠻	h]f	herself
⠓⠋⠾	h}f	herself
⠓⠍	hm	him
⠓⠍⠋	hmf	himself
⠦	8	his
⠊⠍⠍	imm	immediate
⠭	x	it

เบรลล์	แอสกี	คู่เทียบ
⠭⠎	xs	its
⠭⠋	xf	itself
⠚	j	just
⠅	k	knowledge
⠇⠁⠽⠁⠃	layab	layabout
⠇⠗	lr	letter
⠇	l	like
⠇⠇	ll	little
⠇⠇⠌	ll/	littlest
⠍⠊⠙⠁⠋⠝	midafn	midafternoon
⠍	m	more
⠍⠳	m*	much
⠍⠌	m/	must
⠍⠽⠋	myf	myself
⠝⠑⠉	nec	necessary
⠝⠑⠊	nei	neither
⠝⠑⠺⠎⠇⠗	newslr	newsletter
⠝	n	not
⠕⠌⠉	o'c	o'clock
⠷⠕⠋	"of	oneself
⠳⠗⠧	\rv	ourselve
⠳	\	out
⠏⠙	pd	paid
⠏	p	people
⠏⠻⠉⠧	p]cv	perceive
⠏⠻⠉⠧⠛	p]cvg	perceiving
⠏⠻⠓	p]h	perhaps
⠟⠅	qk	quick
⠟	q	quite
⠗	r	rather

เบรลล์	แอสกี	คู่เทียบ
⠗⠉⠧	rcv	receive
⠗⠉⠧⠛	rcvg	receiving
⠗⠚⠉	rjc	rejoice
⠗⠚⠉⠛	rjcg	rejoicing
⠗⠏⠙	repd	repaid
⠎⠙	sd	said
⠩	%	shall
⠩⠙	%d	should
⠎	s	so
⠌	/	still
⠎⠳	s*	such
⠞	t	that
⠞⠍⠧	!mv	themselve
⠹	?	this
⠹⠽⠋	?yf	thyself
⠞⠙	td	today
⠞⠛⠗	tgr	together
⠞⠍	tm	tomorrow
⠞⠝	tn	tonight
⠥	u	us
⠧	v	very
⠵	0	was
⠺	7	were
⠒	:	which
⠾	w	will
⠾⠙	wd	would
⠽	y	you
⠽⠗	yr	your
⠽⠗⠋	yrf	yourself
⠽⠗⠧	yrv	yourselve

ภาคผนวก ข.2 คู่เทียบตัวย่อเบรลล์เซลล์คู่

เบรลล์	แอสกี	อักษรปกติ
	"!	there
	"*	character
	":	where
	"?	through
	"\	ought
	"c	christ
	"d	day
	"e	ever
	"f	father
	"g	god
	"h	here
	"k	know
	"l	lord
	"m	mother
	"n	name
	"o	one
	"p	part
	"q	question
	"r	right
	"s	some
	"t	time
	"u	under
	"w	work
	"y	young
	,n	ation
	,y	ally
	.d	ound
	.e	ance
	.n	sion
	.s	less

เบรลล์	แอสกี	อักษรปกติ
	.t	ount
	;e	ence
	;g	ong
	;l	ful
	;n	tion
	;s	ness
	;t	ment
	;y	ity
	@&	&
	^!	these
	^:	whose
	^?	those
	^u	upon
	^w	word
	_!	their
	_c	cannot
	_h	had
	_m	many
	_s	spirit
	_w	world
	.k	=
	@+	+
	@<	<
	@>	>
	_/	/
	_1	,
	_2	;
	_3	:
	_4	.
	_6	!

ภาคผนวก ค.
เซ็ตของสายอักขระที่เป็นเอกลักษณ์ของแต่ละภาษาเพื่อใช้ระบุภาษาด้วยกฎ

ภาคผนวก ค.1 เซ็ตของสายอักขระที่เป็นเอกลักษณ์เบรลล์ไทยที่ใช้กับเบรลล์ไทยปนอังกฤษรูปเต็ม (ตัวอย่าง)

⠆	!	⠒⠂	+:	⠒⠌	/(	⠶⠐	]e	⠷⠑	f$	⠶⠶	g)	⠷⠶	gt
⠰	$	⠒⠄	+<	⠒⠆	/)	⠶⠶	]g	⠷⠷	f&	⠶⠰	g*	⠷⠲	gv
⠠	)	⠒⠔	+>	⠒⠢	/*	⠶⠷	]m	⠷⠠	f)	⠶⠔	g+	⠷⠺	gw
⠂	0	⠒⠖	+?	⠒⠱	/<	⠶⠹	]o	⠷⠂	f+	⠶⠂	g,	⠷⠵	gz
⠒	:	⠒⠶	+]	⠒⠒	/>	⠶⠿	]q	⠷⠒	f,	⠶⠤	g-	⠺⠨	h%
⠢	<	⠒⠁	+a	⠒⠌	/?	⠶⠳	]v	⠷⠢	f-	⠶⠄	g2	⠺⠪	h(
⠶	=	⠒⠃	+b	⠒⠃	/b	⠶⠻	]w	⠷⠖	f/	⠶⠤	g3	⠺⠶	h)
⠦	?	⠒⠉	+c	⠒⠛	/g	⠶⠽	]z	⠷⠊	f0	⠶⠴	g6	⠺⠱	h*
⠳	\	⠒⠙	+d	⠒⠟	/q	⠷⠀	d%	⠷⠿	f=	⠶⠼	g7	⠺⠁	h,
⠨⠨	&%	⠒⠩	+e	⠒⠧	/v	⠷⠈	d(	⠷⠑	f?	⠶⠺	g8	⠺⠂	h2
⠨⠬	&*	⠒⠙	+f	⠒⠯	/z	⠷⠬	d*	⠷⠓	f\	⠶⠹	g9	⠺⠒	h3
⠨⠂	&2	⠒⠛	+g	⠖⠖	]%	⠷⠂	d2	⠷⠒	f]	⠶⠱	g<	⠺⠢	h5
⠨⠔	&9	⠒⠕	+i	⠖⠿	]&	⠷⠒	d3	⠷⠓	fd	⠶⠿	g=	⠺⠆	h6
⠨⠢	&<	⠒⠅	+k	⠖⠿	](	⠷⠢	d5	⠷⠛	fg	⠶⠧	g>	⠺⠔	h9
⠨⠴	&[	⠒⠇	+l	⠖⠬	]*	⠷⠔	d6	⠷⠭	fh	⠶⠯	g?	⠺⠱	h<
⠨⠒	&h	⠒⠍	+m	⠖⠂	]2	⠷⠔	d9	⠷⠝	fj	⠶⠽	g[	⠺⠔	h>
⠨⠖	&j	⠒⠝	+n	⠖⠒	]3	⠷⠖	d:	⠷⠇	fk	⠶⠫	g\	⠺⠢	h[
⠨⠶	&q	⠒⠏	+o	⠖⠶	]4	⠷⠇	d>	⠷⠕	fn	⠶⠶	g]	⠺⠶	h\
⠒⠒	+%	⠒⠿	+q	⠖⠒	]6	⠷⠒	d?	⠷⠕	fp	⠶⠂	gb	⠺⠿	h]
⠒⠶	+&	⠒⠾	+r	⠖⠺	]9	⠷⠺	d[	⠷⠇	fs	⠶⠒	gc	⠺⠆	hb
⠒⠇	+'	⠒⠎	+s	⠖⠎	]<	⠷⠾	d]	⠷⠎	fv	⠶⠒	gd	⠺⠒	hc
⠒⠖	+(	⠒⠥	+u	⠖⠇	]>	⠷⠶	df	⠷⠑	fw	⠶⠖	gf	⠺⠒	hd
⠒⠶	+)	⠒⠾	+v	⠖⠎	][	⠷⠶	dm	⠷⠿	fy	⠶⠶	gg	⠺⠶	hf
⠒⠒	+*	⠒⠿	+w	⠖⠿	]]	⠷⠶	dn	⠷⠑	g$	⠶⠰	gk	⠺⠶	hg
⠒⠔	+3	⠒⠶	+x	⠖⠰	]a	⠷⠾	dq	⠷⠑	g%	⠶⠴	gm	⠺⠔	hj
⠒⠔	+4	⠒⠶	+y	⠖⠰	]b	⠷⠾	dv	⠷⠿	g&	⠶⠴	gn	⠺⠆	hl
⠒⠢	+5	⠒⠶	+z	⠖⠰	]c	⠷⠾	dz	⠷⠳	g'	⠶⠶	gq	⠺⠒	hm
⠒⠔	+9	⠒⠒	/%	⠖⠰	]d	⠷⠆	f!	⠷⠶	g(	⠶⠴	gs	⠺⠔	hn

หมายเหตุ:สัญลักษณ์ที่ปรากฏในตาราง ประกอบด้วยสัญลักษณ์เบรลล์และสัญลักษณ์แอสกี

ภาคผนวก ค.1 (ต่อ)

	hq		j3		k3		l5		m5		n3		o3
	hr		j5		k5		l6		m6		n6		o5
	hw		j6		k6		l9		m9		n9		o6
	hz		j7		k9		l<		m:		n<		o7
	i!		j8		k>		l>		m<		n>		o8
	i$		j9		k[		l?		m>		n?		o9
	i&		j<		k\		l[		m?		n[		o>
	i)		j>		k]		l]		m[		n]		o?
	i+		j[		kb		lc		m\		nb		o[
	i,		j\		kc		ld		m]		nf		o\
	i/		j]		kd		lf		mb		ng		o]
	i0		jb		kf		lg		mc		nh		oe
	i=		jc		kj		lh		md		nk		og
	i?		jd		kk		ll		mf		nl		oh
	i\		jf		kl		lm		mh		nm		oj
	i]		jg		kq		ln		mj		nn		oq
	ih		jj		kr		lq		mk		nq		p%
	ij		jk		ks		lr		ml		nr		p'
	ik		jl		ku		ls		mn		ns		p(
	io		jm		kv		lv		mq		nv		p)
	ip		jn		kw		lw		mr		nw		p*
	iu		jq		ky		lz		ms		ny		p+
	iv		jr		kz		m%		mw		nz		p,
	iw		jv		l%		m'		mz		o%		p2
	iy		jw		l&		m(		n%		o&		p3
	j%		jy		l'		m)		n&		o(		p5
	j&		jz		l(		m*		n'		o)		p6
	j(		k%		l*		m+		n(		o*		p9
	j)		k(		l,		m,		n)		o+		p>
	j*		k)		l1		m0		n*		o,		p[
	j,		k*		l2		m2		n,		o0		p]
	j2		k2		l3		m3		n2		o2		pb

ภาคผนวก ค.1 (ต่อ)

⠏⠋	pf	⠗⠉	rc	⠎⠋	s[	⠥⠄	u%	⠺⠺	uw	⠧⠍	vm	⠾⠇	wl
⠏⠛	pg	⠗⠙	rd	⠎⠳	s\	⠥⠖	u&	⠧⠵	uz	⠧⠝	vn	⠾⠍	wm
⠏⠚	pj	⠗⠋	rf	⠎⠻	s]	⠥⠤	u'	⠧⠦	v%	⠧⠟	vq	⠾⠝	wn
⠏⠟	pq	⠗⠛	rg	⠎⠹	sd	⠥⠶	u(	⠧⠄	v'	⠧⠗	vr	⠾⠎	ws
⠏⠧	pv	⠗⠊	rj	⠎⠋	sf	⠥⠷	u)	⠧⠶	v(	⠧⠧	vv	⠾⠥	wu
⠏⠺	pw	⠗⠍	rm	⠎⠛	sg	⠥⠔	u*	⠧⠷	v)	⠧⠺	vw	⠾⠽	wy
⠏⠭	px	⠗⠝	rn	⠎⠊	sj	⠥⠖	u+	⠧⠔	v*	⠧⠽	vy	⠭⠶	x&
⠏⠽	py	⠗⠟	rq	⠎⠗	sr	⠥⠂	u,	⠧⠖	v+	⠧⠵	vz	⠭⠔	x*
⠏⠵	pz	⠗⠗	rr	⠎⠎	ss	⠥⠆	u2	⠧⠂	v,	⠾⠤	w%	⠭⠂	x,
⠏⠄	r$	⠗⠎	rs	⠎⠧	sv	⠥⠒	u3	⠧⠴	v0	⠾⠶	w&	⠭⠴	x0
⠏⠤	r%	⠗⠞	rt	⠎⠵	sz	⠥⠢	u5	⠧⠃	v2	⠾⠷	w(	⠭⠃	x2
⠏⠦	r&	⠗⠧	rv	⠎⠔	t%	⠥⠒	u6	⠧⠒	v3	⠾⠷	w)	⠭⠒	x3
⠏⠄	r'	⠗⠺	rw	⠎⠶	t(	⠥⠔	u9	⠧⠒	v4	⠾⠔	w*	⠭⠢	x5
⠏⠶	r(	⠗⠭	rz	⠎⠔	t*	⠥⠣	u<	⠧⠒	v6	⠾⠶	w+	⠭⠒	x6
⠏⠷	r)	⠎⠗	s$	⠎⠃	t2	⠥⠃	u>	⠧⠶	v7	⠾⠂	w,	⠭⠔	x9
⠏⠔	r*	⠎⠒	s%	⠎⠒	t3	⠥⠇	u?	⠧⠷	v8	⠾⠶	w-	⠭⠃	x>
⠏⠖	r+	⠎⠶	s&	⠎⠔	t5	⠥⠔	u[	⠧⠔	v9	⠾⠷	w0	⠭⠇	x?
⠏⠂	r,	⠎⠷	s(	⠎⠒	t6	⠥⠣	u\	⠧⠣	v<	⠾⠃	w2	⠭⠣	x[
⠏⠴	r0	⠎⠶	s)	⠎⠔	t9	⠥⠇	u]	⠧⠃	v>	⠾⠒	w3	⠭⠇	x]
⠏⠃	r1	⠎⠔	s*	⠎⠖	t>	⠥⠄	ua	⠧⠇	v?	⠾⠃	w6	⠭⠄	xb
⠏⠆	r2	⠎⠖	s+	⠎⠔	t[	⠥⠆	ub	⠧⠔	v[	⠾⠔	w9	⠭⠤	xc
⠏⠒	r3	⠎⠂	s,	⠎⠇	t]	⠥⠤	uc	⠧⠣	v\	⠾⠣	w<	⠭⠦	xg
⠏⠢	r5	⠎⠦	s0	⠎⠴	tb	⠥⠦	ud	⠧⠇	v]	⠾⠖	w>	⠭⠶	xj
⠏⠒	r6	⠎⠆	s2	⠎⠤	tc	⠥⠴	ue	⠧⠄	vb	⠾⠴	w?	⠭⠷	xn
⠏⠔	r9	⠎⠒	s3	⠎⠶	td	⠥⠆	uf	⠧⠤	vc	⠾⠶	w\	⠭⠔	xo
⠏⠣	r<	⠎⠔	s5	⠎⠤	tf	⠥⠶	ug	⠧⠶	vd	⠾⠷	w]	⠭⠶	xq
⠏⠃	r>	⠎⠢	s6	⠎⠤	tg	⠥⠒	uh	⠧⠆	vf	⠾⠤	wb	⠭⠒	xr
⠏⠇	r?	⠎⠔	s9	⠎⠴	tl	⠥⠒	uj	⠧⠶	vg	⠾⠒	wc	⠭⠴	xs
⠏⠣	r[	⠎⠣	s:	⠎⠤	tm	⠥⠴	uo	⠧⠒	vh	⠾⠒	wd	⠭⠢	xv
⠏⠖	r\	⠎⠔	s<	⠎⠶	tn	⠥⠶	uq	⠧⠶	vj	⠾⠤	wf	⠽⠤	y%
⠏⠶	r]	⠎⠔	s>	⠎⠶	tq	⠥⠷	uu	⠧⠷	vk	⠾⠶	wg	⠽⠶	y(
⠏⠄	rb	⠎⠺	s?	⠎⠴	tv	⠥⠧	uv	⠧⠄	vl	⠾⠔	wj	⠽⠔	y*

ภาคผนวก ค.2 เซ็ตของสายอักขระที่เป็นเอกลักษณ์เบรลล์อังกฤษใช้กับเบรลล์ไทยปนอังกฤษรูปต์ม (ตัวอย่าง)

	'		.#		f4		,da		,t4		dea		dri
	a		.,		fa		,de		,ta		deb		dru
	b		..		fe		,di		,te		dec		dry
	c		.a		fi		,do		,th		ded		dud
	e		.b		hu		,dr		,ti		dee		due
	q		.c		i-		,du		,to		def		dul
	z		.d		ic		,dw		,tr		deg		dum
	,,		.f		if		,n-		,tu		del		dup
	,a		.g		ii		,n4		,tw		dem		dur
	,b		.h		le		,na		,ty		den		dut
	,c		.i		li		,ne		,u4		dep		dwa
	,e		.l		ly		,ni		,un		der		dwe
	,f		.m		ok		,no		,up		des		dye
	,g		.n		ox		,nu		,us		det		dyi
	,h		.o		p#		,s7		,ut		dev		fla
	,i		.p		ph		,sa		,y-		do-		fle
	,j		.r		pt		,sc		,ye		do1		fli
	,k		.s		ru		,se		,yo		do4		flo
	,l		.t		se		,sh		dad		do8		foa
	,m		.w		th		,si		dai		doc		foc
	,o		.y		ti		,sl		dam		doe		foo
	,p		4#		tr		,so		dan		doi		fra
	,q		8a		tu		,sp		dar		doo		fre
	,r		8b		ty		,sq		das		dou		fro
	,v		8c		up		,st		dat		dow		ft4
	,w		8e		ut		,su		dau		doz		fue
	,x		8f		ww		,sy		day		dra		fut
	,z		di		,d4		,t-		dba		dre		gee

ภาคผนวก ค.2 (ต่อ)

	gee		hac		ide		jac		lap		mig		new
	gen		had		ima		jea		las		mil		nex
	geo		hai		imp		jeo		lat		mim		nic
	ger		hal		in1		jet		lay		mir		nig
	ges		ham		in8		jew		loc		mis		nin
	get		han		ina		job		lot		mit		nit
	gho		hap		inc		joi		lou		mix		no4
	gif		har		ind		jou		lug		mob		no6
	gig		has		ine		jud		lum		moc		nob
	gir		hat		inf		jui		lun		mon		nod
	giv		hav		ing		jun		lur		moo		noi
	gla		he'		inh		jur		m4p		mot		nom
	gli		he1		inj		jus		mac		muc		noo
	glu		he8		inl		k,w		maj		muf		nos
	go-		hea		inq		kan		map		mul		not
	go1		hec		ins		kee		mat		mun		nou
	go6		hic		int		kep		me-		mur		now
	goe		hid		iro		ker		me1		mus		nuc
	goi		hig		irr		key		me4		mut		num
	got		hil		is1		kic		me6		mys		nun
	gre		him		is4		kid		me8		nam		nut
	gri		hir		iso		kil		mea		nap		oak
	gru		his		iss		kin		mec		nat		oat
	gua		hit		it'		kis		med		nav		obt
	gue		hob		it-		kit		mee		nea		occ
	gui		hol		it1		kne		mel		nec		odd
	gum		hoo		it4		kni		men		nee		of1
	gun		hop		it6		knu		mes		neg		off
	guy		hor		it8		lab		met		nei		oft
	guz		hot		ita		lac		mi4		ner		old
	ha1		how		itc		lad		mic		net		on-
	hab		hyp		ite		lan		mid		nev		on/

ภาคผนวก ค.2 (ต่อ)

	on1		pay		pou		req		sid		spu		tau
	on4		pdf		pra		res		sil		squ		tax
	on7		pea		pre		ret		sis		sta		tea
	one		ped		pri		reu		sit		ste		tec
	oni		pee		pro		rev		six		sti		tee
	ons		pen		pry		rew		siz		sto		tel
	ont		peo		psy		rib		ske		str		tem
	ooh		pep		pub		ric		ski		stu		ten
	ope		per		pul		rid		sky		sty		ter
	opi		pes		pun		rig		sli		sub		tes
	opp		pet		pur		rip		sma		suc		tex
	opt		pew		pus		ris		sme		sud		to1
	or1		pia		put		roc		sne		suf		to4
	ora		pic		puz		roo		sni		sug		to8
	org		pie		rac		rop		so-		sui		toe
	ori		pig		rak		rot		so1		sul		toi
	oth		pil		rap		row		soc		sum		tol
	oug		pin		rat		sac		sof		sun		tom
	our		pip		re-		sak		sog		sup		too
	out		pit		rea		sam		soi		sus		top
	ova		piz		reb		san		sol		swe		tor
	ove		pli		rec		sat		som		swi		tos
	pac		pne		red		sax		soo		sys		tot
	pag		poc		ree		say		sop		tab		tou
	pai		poi		ref		sca		sor		tac		tow
	paj		pol		reg		sce		sou		tad		twe
	pal		pon		reh		sha		spa		tak		twi
	pan		poo		rei		she		spe		tal		two
	pap		pop		rel		shi		spi		tan		ult
	par		por		rem		sho		spl		tap		una
	pat		pos		ren		shu		spo		tar		unc
	pau		pot		rep		sic		spr		tas		une

ภาคผนวก ค.3 เซ็ตของสายอักขระที่เป็นเอกลักษณ์เบรลล์ไทยใช้กับเบรลล์ไทยปนอังกฤษรูปย่อ (ตัวอย่าง)

⠱⠱	!%	⠳⠱	+%	⠹⠱	:?	⠽⠱	?%	⠿⠁	g*	⠷⠁	r*	⠳⠉⠱	!bn
⠱⠁	!*	⠳⠁	+'	⠹⠉	:d	⠽⠹	?(	⠿⠃	g<	⠷⠉	rq	⠳⠉⠁	!bs
⠱⠃	!2	⠳⠹	+(	⠹⠃	:k	⠽⠁	?*	⠿⠉	gq	⠷⠽	s%	⠳⠽⠱	!c&
⠱⠉	!4	⠳⠁	+*	⠹⠇	:l	⠽⠇	?5	⠿⠱	gz	⠷⠹	u(	⠳⠇⠉	!og
⠱⠇	!5	⠳⠉	+4	⠹⠽	:m	⠽⠉	?b	⠷⠽	h%	⠷⠇	u5	⠳⠇⠉	!oy
⠱⠊	!9	⠳⠃	+5	⠹⠱	:p	⠽⠱	?j	⠷⠹	i)	⠷⠊	u9	⠫⠫⠁	$4*
⠱⠇	!<	⠳⠊	+9	⠹⠇	:v	⠽⠹	?q	⠷⠹	i=	⠷⠇	u>	⠫⠫⠃	$4o
⠱⠹	!>	⠳⠹	+>	⠹⠫	:w	⠽⠫	?v	⠿⠁	j*	⠷⠱	u]	⠫⠇⠱	$54
⠱⠽	![	⠳⠁	+a	⠹⠹	<!	⠽⠱	\%	⠿⠃	jb	⠷⠹	uq	⠫⠇⠁	$bl
⠫⠱	$%	⠳⠃	+b	⠹⠹	<$	⠽⠹	\(	⠿⠇	jv	⠷⠽	v%	⠫⠇⠱	$bw
⠫⠉	$3	⠳⠉	+c	⠹⠹	<&	⠽⠇	\9	⠿⠉	jw	⠷⠹	v(	⠳⠹⠹	&(g
⠫⠇	$>	⠳⠁	+e	⠹⠫	<)	⠽⠽	\[	⠿⠹	jz	⠷⠹	v)	⠳⠹⠹	&(i
⠫⠽	$c	⠳⠃	+l	⠹⠹	<=	⠽⠱	\e	⠿⠱	k*	⠷⠱	v*	⠳⠱⠱	&,\
⠫⠱	$o	⠳⠉	+m	⠹⠽	<?	⠽⠇	\o	⠿⠁	l*	⠷⠉	v6	⠳⠱⠃	&,t
⠱⠱	&%	⠳⠉	+n	⠹⠱	<d	⠽⠹	\w	⠿⠁	l<	⠷⠹	vw	⠳⠉⠱	&3d
⠱⠁	&*	⠳⠃	+q	⠹⠉	<g	⠽⠱	]*	⠿⠃	lq	⠷⠁	vz	⠳⠉⠃	&4o
⠱⠃	&2	⠳⠹	+w	⠹⠁	<h	⠽⠉	]3	⠿⠁	lz	⠷⠱	w*	⠳⠉⠃	&4r
⠱⠁	&9	⠳⠃	+y	⠹⠇	<j	⠽⠁	]>	⠿⠁	m2	⠷⠉	w<	⠳⠉⠱	&64
⠱⠱	)%	⠳⠃	+z	⠹⠃	<k	⠽⠉	]c	⠿⠃	mq	⠷⠃	wb	⠳⠁⠱	&>,
⠱⠃	)(	⠳⠃	/(	⠹⠇	<l	⠽⠉	]d	⠿⠉	n%	⠷⠱	x*	⠳⠁⠉	&>4
⠱⠁	)*	⠳⠁	/*	⠹⠁	<m	⠽⠃	]q	⠿⠱	n*	⠷⠁	x9	⠳⠁⠃	&>]
⠱⠇	)>	⠳⠉	/?	⠹⠁	<n	⠽⠉	d%	⠿⠉	ng	⠷⠇	x[	⠳⠁⠃	&>g
⠱⠃	)b	⠳⠃	/b	⠹⠃	<p	⠽⠱	d*	⠿⠁	nz	⠷⠃	xr	⠳⠁⠃	&>j
⠱⠉	)n	⠳⠁	/v	⠹⠁	<s	⠽⠃	f)	⠿⠱	o5	⠷⠉	y%	⠳⠁⠃	&>s
⠱⠃	)q	⠹⠉	:!	⠹⠃	<v	⠽⠉	f?	⠿⠇	o[	⠷⠃	y(	⠳⠃⠇	&b9
⠱⠇	)r	⠹⠃	:&	⠹⠇	<w	⠽⠁	fk	⠿⠉	oq	⠷⠱	y*	⠳⠃⠉	&bd
⠱⠃	)w	⠹⠃	:)	⠿⠱	=*	⠽⠱	g%	⠿⠱	r%	⠷⠃	yq	⠳⠃⠃	&by
⠱⠃	)z	⠹⠉	:=	⠿⠇	=[	⠽⠃	g)	⠿⠱	r(	⠳⠃⠉	!bg	⠳⠉⠱	&c+
⠱⠇⠇	fh\	⠽⠃⠃	fwl	⠿⠃⠉	gbn	⠽⠇⠇	h4o	⠿⠉⠉	hnc	⠷⠃⠃	j64	⠳⠃⠃	jrb
⠱⠇⠇	fh]	⠽⠃⠱	fy'	⠿⠉⠇	gc4	⠽⠃⠃	h9o	⠿⠉⠇	hno	⠷⠃⠇	j9*	⠳⠃⠃	jrr
⠱⠇⠇	fhl	⠽⠃⠇	fy4	⠿⠃⠱	gl*	⠽⠃⠇	h>n	⠿⠉⠇	hnq	⠷⠃⠱	j>,	⠳⠇⠇	k(4
⠱⠇⠇	fhm	⠽⠃⠱	g'+	⠿⠃⠃	gl4	⠽⠃⠇	h]q	⠿⠇⠱	hr5	⠷⠃⠇	j>\	⠳⠇⠇	k(9
⠱⠃⠱	fj'	⠽⠃⠱	g':	⠿⠃⠃	glb	⠽⠃⠃	hc9	⠿⠇⠃	hw>	⠷⠃⠇	j>]	⠳⠇⠇	k(n

ภาคผนวก ค.3 (ต่อ)

	fj\		g'?		glc		hc]		hwo		j>n		k(w
	fjn		g'\		glm		he:		hy(		j[]		k4*
	fl!		g'o		glq		he?		hyc		j]g		k64
	fl'		g't		glw		hl*		i&r		ja&		k6g
	fl:		g'u		gng		hl2		i,?		ja:		k9*
	flk		g(9		go&		hl6		i\4		ja<		k>4
	fm'		g(n		gr*		hl9		i\g		ja\		k>d
	fm-		g(r		grb		hl>		i\y		jad		k>n
	fm\		g2,		grc		hl]		idn		jaf		k[4
	fmg		g2\		grd		hlb		idy		jc7		k]j
	fmk		g4*		grg		hlc		ihd		jcd		kc9
	fny		g4n		grr		hlo		ihl		jcl		kcn
	fo+		g69		gs)		hlq		ijf		jd)		kdl
	fo]		g6h		gw*		hlw		il9		jdh		ke4
	fpy		g7*		gw4		hm%		ilg		jdj		klc
	fs,		g7g		h*\		hm*		ilh		jm3		kn*
	fs4		g9*		h*d		hm>		ilw		jn)		kn<
	fsd		g>]		h*g		hm[		im\		jn<		knm
	fsm		g>m		h*l		hmd		iog		jn\		ko+
	fsn		g>v		h*m		hmo		ipl		jng		ko:
	fsr		g\b		h*r		hmq		ir]		jnr		ko<
	fss		ga)		h*u		hmw		iu4		jnt		ko]
	kov		lo]		mhb		nwn		ob9		plf		rwm
	koy		m(n		ml*		o%n		oc&		plg		rwv
	ko}		m*&		mnc		o*!		oc)		pll		rz)
	kvw		m*+		mo]		o*)		oc9		pm+		rzl
	kwv		m*,		mr,		o*+		oc\		pmn		s(+
	ky*		m*<		n24		o*,		ocd		pn>		s(,
	ky6		m*?		n29		o*d		ocj		r2s		s(9
	l(4		m*\		n4*		o*g		ocl		r2v		s(:
	l(9		m*d		n4o		o*h		ocv		r3&		s(]
	l(n		m*f		n4z		o*j		od2		r34		s(h
	l2l		m*g		n69		o*n		odc		r39		s(s
	l3g		m*i		n9*		o*u		on[		r3s		s(v
	l4*		m*n		n<a		o*v		os>		r4*		s*f
	l4m		m*r		n>4		o*y		ovr		r4o		s*g
	l4w		m*s		n>9		o,?		ow>		r6\		s*k
	l69		m*u		n>g		o0)		oy*		r6u		s*m
	l>)		m*v		n>v		o2g		oy3		r9*		s*r
	l])		m,n		n[g		o4o		oy9		r9w		s*y
	l]i		m3l		nb)		o4w		oz?		r>,		s,?
	l]u		m4]		nb4		o6:		ozl		r>g		s2h
	lb&		m4r		nb\		o6j		ozn		r>s		s2k
	lb9		m4w		nb]		o6l		p34		r>v		s3,

ภาคผนวก ค.3 (ต่อ)

	lb\		m>0		nbd		o9*		p3g		r?4		s39
	lb]		m>n		nbg		o9o		p4*		rc9		s3]
	lbk		ma$		nbn		o>\		p69		rcn		s4n
	lbm		mb/		nbw		o>]		p>n		rh>		s59
	lc&		mb9		nby		o>n		pbd		ro]		s9]
	lc9		mb:		nc9		o],		pbw		rsn		s9w
	lc?		mbl		nn)		o]c		pl%		rt:		s>&
	lc]		mc9		nq4		o]u		pl&		rtj		s>,
	lcg		mfh		nq9		oa:		pl:		rwb		s>\
	s>j		smg		t*d		ul*		uz?		vcu		y,s
	s>m		smj		t*m		ul2		uz\		vdb		y3!
	s]s		smm		t3g		ul4		uzg		vl3		y34
	sa)		smr		t4*		ul5		uzi		vr'		y3f
	sa?		smv		t4o		ul9		uzk		vr3		y3n
	sb)		smw		t4w		ulq		uzm		vrb		y4z
	sb9		sn%		tb9		umn		uzn		vrm		y5n
	sb]		sn*		tc]		un)		uzo		w(d		y6w
	sbn		sn:		te9		un\		uzs		w(n		y9o
	sbr		so]		tgt		uo&		uzv		w2d		y>]
	sbv		sq9		tnn		uo<		v2)		w2r		y>v
	sc,		sr%		tw*		uom		v2f		w2s		y[d
	scd		srr		u*n		uoy		v4*		w>\		yb9
	scj		st*		u*r		ur2		v9m		w],		ybn
	sck		stb		u2y		ur3		v>\		w]j		yc)
	sgl		suw		u3,		ur>		v>]		wc-		yc?
	sh&		sv*		u39		urq		vah		wl2		yc]
	sl*		swb		u4n		urv		vb]		wno		ycu
	sl>		swd		u4z		uw*		vbd		wrc		ygf
	sm%		swm		ua<		uwr		vbn		x2:		ygm
	sm*		swn		ubd		uwv		vc,		x2d		ygy
	sm+		szf		ubg		uz&		vc8		x3]		yis
	sm9		szh		uc,		uz)		vc?		x>]		yod
	smc		szn		uc4		uz:		vc\		xc9		yom
	smf		szr		ucg		uz<		vcg		xnh		yov

ภาคผนวก ค.4 เซ็ตของสายอักขระที่เป็นเอกลักษณ์เบรลล์อังกฤษใช้กับเบรลล์ไทยปนอังกฤษรูปย่อ (ตัวอย่าง)

	"		)\		,h		.?		0s		?i		l;
	%		,!		,i		/"		0t		@&		l\
	(		,"		,j		/&		0x		\/		li
	*		,$		,k		/1		0y		\1		m"
	2		,%		,m		/e		:5		_!		m$
	3		,&		,o		/i		:9		_+		m/
	4		,(		,p		/r		:;		_/		m1
	5		,*		,q		/u		:>		_c		n"
	6		,,		,r		0!		:a		_h		n1
	7		,-		,v		0"		:e		_m		n\
	8		,/		,w		0#		:i		_s		o!
	9		,0		,x		0&		;,		_w		ox
	>		,2		,z		0,		;e		d#		p.
	[		,3		-&		0/		;f		d1		p1
	^		,4		-5		03		;i		di		p;
	a		,5		-9		09		;o		f.		p\
	b		,6		-;		0:		;s		f1		ph
	c		,7		-=		0?		;y		f3		ru
	e		,8		-]		0\		<e		f9		s"
	q		,9		-a		0^		=!		f>		s.
	z		,:		-b		0_		="		fa		s1
	!s		,=		-c		0a		=&		fe		se
	!y		,>		.!		0b		=1		fi		t1
	$u		,[		."		0d		=7		g5		t\
	&!		,^		.#		0e		=8		h"		ti
	&/		,_		.(		0g		=?		h&		tr
	&1		,a		.)		0h		=e		h1		tu
	&=		,b		.*		0j		=m		hu		ty
	's		,c		.,		0l		=t		ic		u1
	)!		,e		..		0n		?"		if		up
	)/		,f		./		0o		?1		j1		ut

ภาคผนวก ค.4 (ต่อ)

	)1		,g		.9		0r		?;		l+		v#
	w$		,la		,te		?+s		def		foo		gul
	w1		,le		,ti		?40		deg		fr%		guy
	ww		,li		,tm		?91		del		fre		h\r
	x'		,lo		,tn		?9k		dem		frl		h\s
	x1		,n"		,to		?\1		den		fro		h]'
	xy		,n,		,tr		?\8		dep		fut		h]1
	y1		,n9		,tu		?\<		der		g-=		h]4
	!m1		,n<		,tw		?\s		des		g]u		h]d
	!m4		,n[		,ty		?r;		det		gaz		h]f
	!mv		,n\		,uk		?re		dev		gd4		ha?
	)91		,na		,un		?us		doc		gdl		hal
	)94		,ne		,us		\40		doo		gds		ham
	,)\		,ng		,ut		\r'		doz		gee		hap
	,)]		,ni		,y1		\tr		dr9		get		has
	,?+		,no		,y5		\ts		dra		gif		hat
	,?9		,nu		,ye		d5i		dre		gig		hav
	,?\		,s+		,yo		d5y		dru		gir		he8
	,?a		,s1		/,n		d9n		due		giv		he>
	,d"		,s5		/>$		d>k		dug		gl\		hi!
	,d1		,s7		/>+		d[n		dul		gla		hi/
	,d5		,s\		/>s		d]n		dur		gli		hi<
	,d7		,sa		/at		dai		dut		go+		hil
	,d9		,se		/ay		dan		dw>		goe		his
	,d;		,si		/om		dat		f$]		got		hmf
	,d\		,so		:1t		dau		fl$		gr.		hoc
	,da		,sp		:]"		de"		fl[		gr\		hol
	,de		,st		:o1		de$		fl\		gre		hop
	,di		,su		=ce		de/		fla		gri		hor
	,do		,t5		?'r		de]		fle		grt		hyp
	,dr		,t9		?+1		deb		fli		gu>		id1
	,du		,t]		?+7		dec		flo		gue		ide
	,l"		,ta		?+8		dee		foe		gui		im"
	imp		lat		mo/		not		p]s		pr(		rev
	in1		lay		moc		num		pa/		pr9		rew
	iro		lei		mon		nun		pa9		pra		ri*

ภาคผนวก ค.4 (ต่อ)

	ita		lot		mul		oas		pan		pri		rid
	j\r		lu7		mum		obt		pap		pro		ro(
	jeo		m,a		mur		oct		pat		psy		roc
	jnr		m51		mus		old		pay		pub		roo
	jo9		m5;		mut		on1		pd4		pun		rop
	ju/		m9d		my1		oni		pdf		pur		rot
	jud		m>i		n,n		op5		peo		put		s*o
	jun		m>k		n47		op9		pep		pyj		s+]
	jur		m>r		n9?		op]		pi3		r11		s5a
	k9d		m\s		n9e		opp		pic		r1s		s5d
	kan		m]c		n[1		or1		pie		ra%		s5s
	kee		ma*		n[6		ora		pig		rac		s5t
	kep		ma/		nat		org		pik		rat		s9c
	kil		ma9		nav		ori		piz		re-		s9u
	kis		mac		ne$		ov]		pl"		re/		s\?
	kne		maj		ne>		ova		pl$		rea		s\l
	kni		mat		nec		ovu		pl1		reb		s\p
	l&"		me*		neg		p5.		pli		rec		sac
	l1!		me4		nei		p5a		pne		red		sam
	l1/		me;		net		p9c		po/		ree		sax
	l11		mee		new		p9k		poi		ref		sc>
	l1v		met		nex		p9t		pol		reg		sca
	l9k		mi4		ni<		p>5		poo		rel		sd4
	la/		mi<		no?		p>a		pop		rem		si/
	lac		mic		nob		p]=		por		rep		si<
	lad		mig		nom		p]c		pos		req		sic
	lan		mil		noo		p]f		pot		res		sil
	las		mir		nos		p]i		pox		ret		sit
	six		squ		t5u		tom		unu		vis		wi%
	sk9		su3		t[>		too		unw		vot		wid
	sma		su4		t[n		top		unx		w9d		wif
	soc		su6		t]m		tor		uny		w9e		wit
	soi		sub		t]r		tot		us$		w9t		wiv
	sol		sug		tak		tw5		use		w>1		wom
	soo		sui		tal		tw9		usu		w>e		woo
	sop		sum		tan		twe		v47		w>f		wri
	sor		sun		tax		twi		v]m		wa%		ye>

ภาคผนวก ง. ตารางคู่เทียบเบรลล์ไทยสำหรับระบบถ่ายถอดด้วยแบบจำลองเอ็นแกรม

เบรลล์	แอสกี	อักษรปกติ
⠷⠐⠳	fa9	|เ่ะ, ||เ่ะ
⠷⠐⠻	fa4	|เ้ะ, ||เ้ะ
⠷⠐⠿	fa7	|เ๊ะ,||เ๊ะ,|เะ),||เะ)
⠷⠐⠽	fa8	|เ๋ะ, ||เ๋ะ
⠯⠐⠳	<a9	|แ่ะ, ||แ่ะ
⠯⠐⠻	<a4	|แ้ะ, ||แ้ะ
⠯⠐⠿	<a7	|แ๊ะ, ||แ๊ะ, |แะ), ||แะ)
⠯⠐⠽	<a8	|แ๋ะ, ||แ๋ะ
⠵⠐⠳	ia9	|โ่ะ, ||โ่ะ
⠵⠐⠻	ia4	|โ้ะ, ||โ้ะ
⠵⠐⠿	ia7	|โ๊ะ, ||โ๊ะ, |โะ), ||โะ)
⠵⠐⠽	ia8	|โ๋ะ, ||โ๋ะ
⠽⠐⠳	oa9	|อ่ะ, |เ่าะ, ||เ่าะ
⠽⠐⠻	oa4	อ้ะ, |เ้าะ, ||เ้าะ
⠽⠐⠿	oa7	อะ), อ๊ะ, |เ๊าะ, ||เ๊าะ, |เาะ), ||าะ)
⠽⠐⠽	oa8	อ๋ะ, |เ๋าะ, ||เ๋าะ
⠷⠐⠳	(a9	|เ่ียะ, ||เ่ียะ
⠷⠐⠻	(a4	|เ้ียะ, ||เ้ียะ
⠷⠐⠿	(a7	|เ๊ียะ, ||เ๊ีย, |เียะ), ||เียะ)
⠷⠐⠽	(a8	|เ๋ียะ, ||เ๋ียะ
⠷⠐⠳	qa9	|เื่อะ, ||เื่อะ
⠷⠐⠻	qa4	|เื้อะ, ||เื้อะ
⠷⠐⠿	qa7	|เื๊อะ, ||เื๊อะ, |เืะ), ||เืะ)

เบรลล์	แอสกี	อักษรปกติ
⠷⠐⠽	qa8	|เื๋อะ, ||เื๋อะ
⠶⠐⠳	%a9	|เ่อะ, ||เ่อะ
⠶⠐⠻	%a4	|เ้อะ, ||เ้อะ
⠶⠐⠿	%a7	|เ๊อะ, ||เ๊อะ, |เอะ), ||เอะ)
⠶⠐⠽	%a8	|เ๋อะ, ||เ๋อะ
⠵⠐⠳	ea9	ั่วะ
⠵⠐⠻	ea4	ั้วะ
⠵⠐⠿	ea7	ั๊วะ
⠵⠐⠽	ea8	ั๋วะ
⠄⠄⠄	'"	...
⠂⠆	,u	ฆ
⠂⠔	,+	ฌ
⠂⠬	,y	ญ
⠂⠤	,d	ฎ
⠂⠳	,\	ฏ
⠂⠴	,t	ฐ
⠂⠬	,)	ฑ
⠄⠬	-)	ฒ, -ท
⠂⠤	,n	ณ
⠂⠹	,?	ภ
⠂⠎	,s	ศ
⠄⠎	-s	ษ, -ส
⠂⠇	,l	ฬ
⠐⠂	_1	,
⠐⠆	_2	;
⠐⠒	_3	:
⠐⠲	_4	.

เบรลล์	แอสกี	อักษรปกติ
⠳⠒	_6	!
⠳⠢	_8	?
⠳⠌	_/	/
⠳⠖	_+	(ว่าง)
⠈⠖	@+	+
⠳⠱	_:	(ว่าง)
⠳⠿	_&	(ว่าง)
⠨⠅	.k	= , _k , _knowledge
⠳⠆	;2	ฯ
⠳⠮	;1	ฯลฯ
⠱⠂	:1	ใ
⠨⠣	,7	[
⠷⠜	7'	]
⠸⠸	0)	ธ, ์ท
⠪⠨	oa	อะ, เาะ, ||เาะ
⠷⠂	r1	ฤ, ร ๆ
⠱⠱	a9	่ะ
⠱⠿	a4	้ะ
⠱⠷	a7	๊ะ, ะ)
⠱⠸	a8	๋ะ
⠷⠱	z9	่ำ
⠷⠿	z4	้ำ
⠷⠷	z7	๊ำ, ำ)
⠷⠸	z8	๋ำ
⠷⠱	69	|เ่า, ||เ่า
⠷⠿	64	|เ้า, ||เ้า
⠷⠷	67	|เ๊า,||เ๊า, |เา), ||เา)
⠷⠸	68	|เ๋า, ||เ๋า
⠷⠱	(9	|เ็่ย, ||เ็่ย
⠷⠿	(4	|เ็้ย, ||เ็้ย

เบรลล์	แอสกี	อักษรปกติ
⠷⠷	(7	|เ็๊ย, |เีย), ||เ็๊ย , ||เีย)
⠷⠸	(8	|เ็๋ย, ||เ็๋ย
⠿⠱	q9	|เ็่อ, ||เ็่อ
⠿⠿	q4	|เ็้อ, ||เ็้อ
⠿⠷	q7	|เ็๊อ, ||เ็๊อ
⠿⠸	q8	|เ็๋อ, ||เ็๋อ
⠽⠱	%9	|เ่อ, |เิ่, ||เ่,|| เิ่
⠽⠿	%4	|เ้อ, |เิ้, ||เ้,|| เิ้
⠽⠷	%7	|เ๊อ, |เิ๊, |เอ), |เิ), ||เ๊อ,||เิ๊,||เอ),||เิ)
⠽⠸	%8	|เ๋อ, |เิ๋, ||เ๋,|| เิ๋
⠱⠱	e9	ั่ว
⠱⠿	e4	ั้ว
⠱⠷	e7	ั๊ว, ัว)
⠱⠸	e8	ั๋ว
⠿	g	ก
⠅	k	ข
⠥	u	ค
⠷	}	ง
⠷	]	ง
⠯	j	จ
⠌	/	ฉ
⠖	+	ช
⠖	!	ซ
⠙	d	ด
⠸	\	ต

เบรลล์	แอสกี	อักษรปกติ
⠞	t	ถ
⠟	)	ท
⠝	n	น
⠧	v	บ
⠯	&	ป
⠏	p	ผ
⠭	x	ฝ
⠹	?	พ
⠎	$	ฟ
⠍	m	ม
⠽	y	ย
⠗	r	ร
⠇	l	ล
⠺	w	ว
⠌	s	ส
⠓	h	ห
⠕	o	อ
⠿	=	ฮ
⠁	a	ะ
⠵	z	ำ
⠾	>	ั
⠡	*	า
⠃	b	ิ
⠆	2	ี
⠪	{	ึ
⠷	[	ื
⠑	5	ี
⠉	c	ุ
⠒	3	ู
⠒	3	ู

เบรลล์	แอสกี	อักษรปกติ
⠄	'	็
⠱	:	ไ
⠂	1	ๆ
⠲	4	้, .
⠦	8	์, "
⠤	9	่
⠴	0	์, "
⠶	7	ี๊, (,)
⠨⠶	,7	[
⠶⠄	7'	]
⠨⠦	,8	'
⠴⠄	0'	'
⠋	f	เ
⠋⠁	fa	เ|ะ, เ||ะ
⠣	<	แ
⠣⠁	<a	แ|ะ, แ||ะ
⠤⠤	-	-
⠨	.	_
⠐⠐	""	..
⠐	"	\
⠊	i	โ
⠊⠁	ia	โ|ะ, โ||ะ
⠖	6	เ|า, เ||า
⠷	(	เ|ีย, เ||ีย
⠿	q	เ|ือ, เ||ือ
⠼	%	เ|อ, เ|ิ, เ||อ, เ||ิ
⠑	e	ั|ว
⠄⠄⠄	111	...
⠨⠨⠨	,,,	...
⠐⠐⠐	"""	...
⠑⠑⠑	444	...

ภาคผนวก จ. ตารางคู่เทียบเบรลล์อังกฤษสำหรับระบบถ่ายถอดด้วยแบบจำลองเอ็นแกรม

ภาคผนวก จ.1 คู่เทียบตัวเลขเบรลล์

เบรลล์	แอสกี	อักษรปกติ
⠼⠁	#a	1
⠼⠃	#b	2
⠼⠉	#c	3
⠼⠙	#d	4
⠼⠑	#e	5
⠼⠋	#f	6
⠼⠛	#g	7
⠼⠓	#h	8
⠼⠊	#i	9
⠼⠚	#j	0
⠼⠄	#'	'
⠁	a	1
⠃	b	2
⠉	c	3
⠙	d	4
⠑	e	5
⠋	f	6
⠛	g	7
⠓	h	8
⠊	i	9
⠚	j	0
⠂	1	,
⠆	2	;

เบรลล์	แอสกี	อักษรปกติ
⠒	3	:
⠲	4	.
⠖	6	! , to
⠶	7	(,)
⠦	8	"
⠴	0	" , by
⠨	.	.
⠤	-	-
⠨⠌	./	/
⠸	/	/
⠸	/	st
⠠⠶	,7	[
⠲⠹	4#	$
⠶⠸	7'	]
⠈⠼	@*	*
⠈⠖⠯	@3p	%
⠈⠴	@0	%
⠈⠼	@*	*
⠨⠨	99	*
(space)	(space)	(space)

ภาคผนวก จ.2 คู่เทียบเบรลล์อังกฤษรูปเต็มสำหรับการถ่ายถอดด้วยแบบจำลองเอ็นแกรม

เบรลล์	แอสกี	อักษรปกติ
⠁	a	a
⠃	b	b
⠉	c	c
⠙	d	d
⠑	e	e
⠋	f	f
⠛	g	g
⠓	h	h
⠊	i	i
⠚	j	j
⠅	k	k
⠇	l	l
⠍	m	m
⠝	n	n
⠕	o	o
⠏	p	p
⠟	q	q
⠗	r	r
⠎	s	s
⠞	t	t
⠥	u	u
⠧	v	v
⠺	w	w
⠭	x	x
⠽	y	y
⠵	z	z
⠤	-	-
⠲	.	_
⠂	1	,
⠆	2	;
⠒	3	:

เบรลล์	แอสกี	อักษรปกติ
⠲	4	.
⠖	6	!
⠶	7	(,)
⠦	8	?, "
⠴	0	"
⠄	'	'
⠌	/	/
⠴⠄	0'	'
⠂⠦	,8	'
⠂⠶	,7	[
⠶⠄	7'	]
⠒⠂	_1	,
⠒⠆	_2	;
⠒⠒	_3	:
⠒⠲	_4	.
⠒⠰	_:	(ลบออกจากสายอักขระ)
⠒⠖	_+	ลบออกจากสายอักขระ)
⠒⠌	_/	/
⠒⠖	_6	!
⠒⠦	_8	?
⠈	@	@
⠈⠖	@+	+
⠯	&	&
⠹⠲	4#	$
⠇⠲	l#	£
⠼	#	#
⠔⠔	99	*
⠄⠄⠄	'''	...
⠂⠂⠂	111	...
⠐⠐⠐	"""	...
⠠⠠⠠	,,,	...
⠆	;	(ลบออกจากสายอักขระ)

ภาคผนวก ฉ. คู่เทียบเบรลล์อังกฤษรูปย่อสำหรับระบบถ่ายถอดด้วยแบบจำลองเอ็นแกรม

คู่เทียบทั้งหมดในภาคผนวก ฉ เป็นคู่เทียบที่ต้องใช้ร่วมกันในการถ่ายถอดเบรลล์อังกฤษรูปย่อ แต่เพื่อสะดวกต่อการอ้างอิง จึงต้องจำแนกออกเป็นประเภทต่าง ๆ ตามที่แสดงไว้ต่อไปนี้

ภาคผนวก ฉ.1 คู่เทียบสำหรับตัวย่อทั่วไป

เบรลล์	แอสกี	คู่เทียบ
⠁	a	a
⠊	i	i
⠕	o	o
⠜	>	ar
⠫	$	ed
⠻	]	er
⠔	9	in
⠣	<	gh
⠪	[	ow
⠆⠁	;a	a
⠆⠃	;b	b
⠆⠉	;c	c
⠆⠙	;d	d
⠆⠋	;f	f
⠆⠓	;h	h
⠆⠊	;i	i
⠆⠚	;j	j
⠆⠅	;k	k
⠆⠍	;m	m
⠆⠕	;o	= o
⠆⠏	;p	p
⠆⠟	;q	q

เบรลล์	แอสกี	คู่เทียบ
⠆⠗	;r	r
⠆⠥	;u	u
⠆⠧	;v	v
⠆⠺	;w	w
⠆⠭	;x	x
⠆⠵	;z	z
⠠⠁	;,a	A
⠠⠃	;,b	B
⠠⠉	;,c	C
⠠⠙	;,d	D
⠠⠑	;,e	E
⠠⠋	;,f	F
⠠⠛	;,g	G
⠠⠓	;,h	H
⠠⠚	;,j	J
⠠⠅	;,k	K
⠠⠇	;,l	L
⠠⠍	;,m	M
⠠⠝	;,n	N
⠠⠏	;,p	P
⠠⠟	;,q	Q
⠠⠗	;,r	R

ภาคผนวก ฉ.1 (ต่อ)

เบรลล์	แอสกี	คู่เทียบ
⠳⠎	;,s	S
⠳⠞	;,t	T
⠳⠥	;,u	U
⠳⠧	;,v	V
⠳⠺	;,w	W
⠳⠭	;,x	X
⠳⠽	;,y	Y
⠳⠵	;,z	Z
⠐⠉	"c	christ
⠐⠙	"d	day
⠐⠑	"e	ever
⠐⠋	"f	father
⠐⠛	"g	god
⠐⠓	"h	here
⠐⠅	"k	know
⠐⠇	"l	lord
⠐⠍	"m	mother
⠐⠝	"n	name
⠐⠕	"o	one
⠐⠕⠿	"of	oneself
⠐⠏	"p	part
⠐⠟	"q	question
⠐⠗	"r	right
⠐⠎	"s	some
⠐⠞	"t	time
⠻⠮	_!	their
⠐⠥	"u	under

เบรลล์	แอสกี	คู่เทียบ
⠐⠺	"w	work
⠐⠽	"y	young
⠐⠮	"!	there
⠐⠮⠄⠮⠿	"!9af	thereinafter
⠐⠮⠿⠿	"!af	thereafter
⠐⠒	":	where
⠐⠳	"\	ought
⠐⠔	"*	character
⠐⠹	"?	through
⠐⠐	""	…
⠘⠥	^u	upon
⠘⠺	^w	word
⠘⠹	^?	those
⠘⠮	^!	these
⠘⠒	^:	whose
⠸⠉	_c	cannot
⠸⠓	_h	had
⠸⠍	_m	many
⠸⠎	_s	spirit
⠸⠺	_w	world

ภาคผนวก ฉ.2 คู่เทียบสำหรับคำย่อทั่วไป

เบรลล์	แอสกี	คู่เทียบที่ 1	คู่เทียบที่ 2
⠮⠑⠍⠧⠎	!mvs	⠮⠑⠍⠧⠎ = themselves	|⠮⠑⠍⠧⠎ = |themselves
⠹⠮⠑⠋	?yf	⠹⠮⠑⠋ = thyself	|⠹⠮⠑⠋ = |thyself
⠳⠗⠧⠎	\rvs	⠳⠗⠧⠎ = ourselves	|⠳⠗⠧⠎ = |ourselves
⠒⠉⠧	3cv	⠒⠉⠧ = conceive	|⠒⠉⠧ = |conceive
⠒⠉⠧⠛	3cvg	⠒⠉⠧⠛ = conceiving	|⠒⠉⠧⠛ = |conceiving
⠃⠗⠇	brl	⠃⠗⠇ = braille	|⠃⠗⠇ = |braille
⠙⠉⠇	dcl	⠙⠉⠇ = declare	|⠙⠉⠇ = |declare
⠙⠉⠇⠛	dclg	⠙⠉⠇⠛ = declaring	|⠙⠉⠇⠛ = |declaring
⠙⠉⠧	dcv	⠙⠉⠧ = deceive	|⠙⠉⠧ = |deceive
⠙⠉⠧⠛	dcvg	⠙⠉⠧⠛ = deceiving	|⠙⠉⠧⠛ = |deceiving
⠛⠗⠞	grt	⠛⠗⠞ = great	|⠛⠗⠞ = |great
⠓]⠋	h]f	⠓]⠋ = herself	|⠓]⠋ = |herself
⠓⠍⠋	hmf	⠓⠍⠋ = himself	|⠓⠍⠋ = |himself
⠍⠽⠋	myf	⠍⠽⠋ = myself	|⠍⠽⠋ = |myself
⠏]⠉⠧	p]cv	⠏]⠉⠧ = perceive	|⠏]⠉⠧ =|perceive
⠏]⠉⠧⠛	p]cvg	⠏]⠉⠧⠛ = perceiving	|⠏]⠉⠧⠛ =|perceiving
⠏]⠓	p]h	⠏]⠓ = perhaps	|⠏]⠓ = |perhaps
⠟⠅	qk	⠟⠅ = quick	|⠟⠅ = |quick
⠗⠉⠧	rcv	⠗⠉⠧ = receive	|⠗⠉⠧ = |receive
⠗⠉⠧⠛	rcvg	⠗⠉⠧⠛ = receiving	|⠗⠉⠧⠛ =|receiving
⠗⠚⠉	rjc	⠗⠚⠉ = rejoice	|⠗⠚⠉ = |rejoice
⠗⠚⠉⠛	rjcg	⠗⠚⠉⠛ = rejoicing	|⠗⠚⠉⠛ =|rejoicing
⠞⠛⠗	tgr	⠞⠛⠗ = together	|⠞⠛⠗ = |together
⠽⠗⠋	yrf	⠽⠗⠋ = yourself	|⠽⠗⠋ = |yourself
⠽⠗⠧⠎	yrvs	⠽⠗⠧⠎ = yourselves	|⠽⠗⠧⠎ =|yourselves
⠉⠄⠞	c't	⠉⠄⠞ = can't	|⠉⠄⠞ = |can't
⠭⠋	xf	⠭⠋ = itself	|⠭⠋ = |itself
⠭⠎	xs	⠭⠎ = its	|⠭⠎ = |its

ภาคผนวก ฉ.3 คู่เทียบสำหรับคำย่อหรือตัวอักษร

เบรลล์	แอสกี	คู่เทียบที่ 1	คู่เทียบที่ 2	คู่เทียบที่ 3
⠁⠃	ab	⠁⠃ = ab, about	|⠁⠃| =|about|	⠠⠁⠃| =About|
⠁⠉	ac	⠁⠉ = ac, according	|⠁⠉| =|according|	⠠⠁⠉| =According|
⠁⠋	af	⠁⠋ = af, after	|⠁⠋| =|after|	⠠⠁⠋| =After|
⠁⠛	ag	⠁⠛ = ag, again	|⠁⠛| =|again|	⠠⠁⠛| =Again|
⠁⠇⠹	al?	⠁⠇⠹= al?, although	|⠁⠇⠹| =|although|	⠠⠁⠇⠹| =Although|
⠁⠇	al	⠁⠇ = al, also	|⠁⠇| =|also|	⠠⠁⠇| =Also|
⠁⠇⠍	alm	⠁⠇⠍ = alm, almost	|⠁⠇⠍| =|almost|	⠠⠁⠇⠍| =Almost|
⠁⠇⠞	alt	⠁⠇⠞ = alt, altogether	|⠁⠇⠞| =|altogether|	⠠⠁⠇⠞| =Altogether|
⠃⠇	bl	⠃⠇ = bl, blind	|⠃⠇| =|blind|	⠠⠃⠇| =|Blind
⠉⠙	cd	⠉⠙ = cd, could	|⠉⠙| =|could|	⠠⠉⠙| =Could|
⠓⠍	hm	⠓⠍ = hm, him	|⠓⠍| =|him|	⠠⠓⠍| =Him|

ภาคผนวก ฉ. 4 คู่เทียบที่ตีความได้ต่างกัน เมื่อเป็นคำโดดหรือเป็นส่วนหนึ่งของคำ

เบรลล์	แอสกี	คู่เทียบที่ 1	คู่เทียบที่ 2	คู่เทียบที่ 3	คู่เทียบที่ 4	คู่เทียบที่ 5
⠃	b	⠃ = b	|⠃| =|but|	|⠃ =|b, |but	⠃| = b|, but|	⠐⠃|= But|
⠉	c	⠉ = c	|⠉| =|can|	|⠉ =|c, |can	⠉|=c|, can|	⠐⠉|= Can|
⠙	d	⠙ = d	|⠙| =|do|	|⠙ =|d, |do	⠙|=d|, do|	⠐⠙|= Do|
⠑	e	⠑ = e	|⠑| =|every|	|⠑ =|e,|every	⠑|=e|, every|	⠐⠑|= Every|
⠋	f	⠋ = f	|⠋| =|from|	|⠋ =|f, |from	⠋|=f|, from|	⠐⠋|= From|
⠛	g	⠛ = g	|⠛| =|go|	|⠛ =|g, |go	⠛|=g|, go|	⠐⠛|= Go|
⠓	h	⠓ = h	|⠓| =|have|	|⠓ =|h, |have	⠓|=h|, have|	⠐⠓|= Have|
⠚	j	⠚ = j	|⠚| =|just|	|⠚ =|j, |just	⠚|=j|, just|	⠐⠚ |= Just|
⠅	k	⠅ = k	|⠅| = |knowledge|	|⠅ =|k , |knowledge	⠅| = k| , knowledge|	⠐⠅| = Knowledge|
⠇	l	⠇ = l	|⠇| =|like|	|⠇ =|l, |like	⠇|=l|, like|	⠐⠇|= Like|
⠍	m	⠍ = m	|⠍| =|more|	|⠍ =|m, |more	⠍|=m|, more|	⠐⠍|= More|
⠝	n	⠝ = n	|⠝| =|not|	|⠝ =|n, |not	⠝ |=n|, not|;	|⠐⠝ ⠝|= |Not| ⠐⠝ ⠝|=Not|, ation|
⠏	p	⠏ = p	|⠏| =|people|	|⠏ = |p , |people	⠏| =p|, people|	⠐⠏| = People|
⠟	q	⠟ = q	|⠟|=|quite|	|⠟ =|q, |quite	⠟|=q|, quite|	⠐⠟| = Quite|
⠗	r	⠗ = r	|⠗|=|rather|	|⠗ =|r, |rather	⠗| =r|, rather|	⠐⠗| = Rather|
⠎	s	⠎ = s	|⠎| =|s|	|⠎ =|s, |s	⠎|=s|, s|	⠐⠎| = S|
⠞	t	⠞ = t	|⠞| =|that|	|⠞ =|t, |that	⠞|=t|, that|	⠐⠞| = That|
⠥	u	⠥ = u	|⠥| =|u|	|⠥ =|u, |u	⠥|=u|, u|	⠐⠥| = U|
⠧	v	⠧ = v	|⠧| =|very|	|⠧ =|v, |very	⠧|=v|, very|	⠐⠧|= Very|
⠺	w	⠺ = w	|⠺| =|will|	|⠺ =|w, |will	⠺|=w|, will|	⠐⠺|= Will|

ภาคผนวก ฉ. 4 (ต่อ)

เบรลล์	แอสกี	คู่เทียบที่ 1	คู่เทียบที่ 2	คู่เทียบที่ 3	คู่เทียบที่ 4	คู่เทียบที่ 5
⠭	x	⠭ =x	|⠭|=|it|	|⠭ =|x, |it	⠭|=x|, it|	⠠⠭| =|It|
⠽	y	⠽ =y	|⠽|=|you|	|⠽ =|y, |you	⠽|=y|, you|	⠠⠽| =|You|
⠵	z	⠵ =z	|⠵|=|as|	|⠵ =|z, |as	⠵|=z|, as|	⠠⠵| = As|
⠡	*	⠡=ch	|⠡|=|child| , |ch|	|⠡ = |child , |ch	⠡|=ch|, child|	⠠⠡ |=Child|
⠳	\	⠳=ou	|⠳|=|out|	|⠳=|ou, |out	⠳|=ou|, out|	⠠⠳|=Out|
⠩	%	⠩=sh	|⠩|=|shall|	|⠩=|sh, |shall	⠩|=sh|, shall|	⠠⠩|=Shall|
⠌	/	⠌=st, /	|⠌|=|still|	|⠌=|st, |still	⠌|=st|, still|	⠠⠌|=Still|
⠹	?	⠹=th	|⠹|=|this|	|⠹=|th, |this	⠹|=th|, this|	⠠⠹|=This|
⠱	:	⠱=wh	|⠱|=|which|	|⠱=|wh , |which	⠱|=which|	-
⠨⠙	.d	⠨⠙ = ound	|⠨⠙| = |_do|	|⠨⠙ = |_d, |_do	⠨⠙|= _do|, |ound|	|⠠⠨⠙| = |_Do|
⠨⠑	.e	⠨⠑ =ance	|⠨⠑ |=|_every|	|⠨⠑ =|_e|	⠨⠑|=every|, ance|	⠠⠨⠑|= _Every|
⠨⠝	.n	⠨⠝ =sion	|⠨⠝ |=|_not|	|⠨⠝ =|_n	⠨⠝|= _not|, sion|	⠠⠨⠝|=_Not|
⠨⠎	.s	⠨⠎ =less	|⠨⠎|= |_so|	|⠨⠎ = |_s, |_so	⠨⠎|= _so|, less|	⠠⠨⠎|= _So|
⠨⠞	.t	⠨⠞ =ount	|⠨⠞|=|_that|	|⠨⠞ =|_t|, _that	⠨⠞|=_that|, ount|	⠠⠨⠞|=_That|

ภาคผนวก ฉ.5 คู่เทียบที่เป็นคำย่อ เมื่อปรากฏในตำแหน่งท้ายคำหรือปรากฏเป็นคำโดด

เบรลล์	แอสกี	คู่เทียบที่ 1	คู่เทียบที่ 2	คู่เทียบที่ 3
⠆⠒	2c	⠆⠒ = bec, because	|⠆⠒ =|bec, |because	⠆⠒ | =because|
⠆⠖	2f	⠆⠖ = bef, before	|⠆⠖ =|bef, |before	⠆⠖ | =before|
⠆⠢	2h	⠆⠢ = beh, behind	|⠆⠢ =|beh, |behind	⠆⠢ | =behind|
⠆⠶	2l	⠆⠶ = bel, below	|⠆⠶ =|bel, |below	⠆⠶ | =below|
⠆⠔	2n	⠆⠔ = ben, beneath	|⠆⠔ =|ben, |beneath	⠆⠔ | =beneath|
⠆⠎	2s	⠆⠎ = bes, beside	|⠆⠎ =|bes, |beside	⠆⠎ | =beside|
⠆⠞	2t	⠆⠞ = bet, between	|⠆⠞ =|bet, |between	⠆⠞ | =between|
⠆⠽	2y	⠆⠽ = bey, beyond	|⠆⠽ =|bey, |beyond	⠆⠽ | =beyond|
⠈⠆⠧	abv	⠈⠆⠧ =abv, above	|⠈⠆⠧ =|abv, |above	⠈⠆⠧ | =above|
⠈⠆⠉	acr	⠈⠆⠉ =acr, across	|⠈⠆⠉ =|acr, |across	⠈⠆⠉ | =across|
⠈⠆⠇	alr	⠈⠆⠇=alr, already	|⠈⠆⠇ =|alr, |already	⠈⠆⠇ | =already|
⠈⠆⠺	alw	⠈⠆⠺= alw,always	|⠈⠆⠺ =|alw, |always	⠈⠆⠺ | =always|
⠈⠢	ei	⠈⠢ = ei, either	|⠈⠢ =|ei, |either	⠈⠢ | =either|
⠈⠗	fr	⠈⠗ =fr, friend	|⠈⠗ =|fr, |friend	⠈⠗ | =friend|
⠈⠍⠍	imm	⠈⠍⠍ =imm, immediate	|⠈⠍⠍ =|imm, |immediate	⠈⠍⠍| =immediate|
⠏⠙	pd	⠏⠙ =pd, paid	|⠏⠙ =|pd, |paid	⠏⠙ | =paid|
⠝⠑⠉	nec	⠝⠑⠉ =nec, necessary	|⠝⠑⠉ =|nec, |necessary	⠝⠑⠉| =necessary|
⠝⠑⠊	nei	⠝⠑⠊ =nei, neither	|⠝⠑⠊ =|nei, |neither	⠝⠑⠊| =neither|
⠕⠉	o'c	⠕⠉ =o'c, o'clock	|⠕⠉ =|o'c, |o'clock	⠕⠉| =o'clock|
⠎⠙	sd	⠎⠙ =sd, said	|⠎⠙ =|sd, |said	⠎⠙| =said|

ภาคผนวก ฉ. 6 คู่เทียบที่เป็นคำย่อในตำแหน่งต้นคำหรือปรากฏเป็นคำโดด

เบรลล์	แอสกี	คู่เทียบที่1	คู่เทียบที่ 2	คู่เทียบที่ 3
⠹⠎	f/	⠹⠎=fst, first	⠸⠹⠎=|first	⠨⠹⠎= First
⠇⠇	ll	⠇⠇ =ll, little	⠸⠇⠇ =|little	⠨⠇⠇ = Little
⠍⠉	m*	⠍⠉ =mch, much	⠸⠍⠉=|much	⠨⠍⠉= Much
⠍⠌	m/	⠍⠌ =mst, must	⠸⠍⠌=|must	⠨⠍⠌= Must
⠙⠞	td	⠙⠞ =td, today	⠸⠙⠞ =|today	⠨⠙⠞ = Today
⠞⠍	tm	⠞⠍ =tm, tomorrow	⠸⠞⠍ =|tomorrow	⠨⠞⠍ = Tomorrow
⠺⠙	wd	⠺⠙ =wd, would	⠸⠺⠙ =|would	⠨⠺⠙ = Would
⠽⠂⠗⠑	y,re	⠽⠂⠗⠑ =y, re you're	⠸⠽⠂⠗⠑ =|you're	⠨⠽⠂⠗⠑ =You're
⠽⠂⠧⠑	y,ve	⠽⠂⠧⠑ =y, ve, you've	⠸⠽⠂⠧⠑ =|you've	⠨⠽⠂⠧⠑ =You've
⠽⠗	yr	⠽⠗ =yr, your	⠸⠽⠗ =|your	⠨⠽⠗ = Your

ภาคผนวก ฉ. 7 คู่เทียบที่เป็นคำย่อ ในตำแหน่งที่ไม่ใช่ตำแหน่งกลางสายอักขระ

เบรลล์	แอสกี	คู่เทียบที่ 1	คู่เทียบที่2	คู่เทียบที่ 3
⠛⠙	gd	⠛⠙ = gd, good	⠛⠙|= good|	|⠛⠙ = |good
⠇⠗	lr	⠇⠗ = lr, letter	⠇⠗|= letter|	|⠇⠗ = |letter
⠞⠝	tn	⠞⠝ = tn, tonight	⠞⠝|= tonight|	|⠞⠝ = |tonight
⠩⠙	%d	⠩⠓ = sh, should	⠩⠓|= should|	|⠩⠓ = |should
⠡⠝	*n	⠡⠉ = chn, children	⠡⠉|= children|	|⠡⠉ = |children

ภาคผนวก ฉ.8 คู่เทียบสำหรับตัวย่อรูปเต็มคงรูป

เบรลล์	**แอสกี**	**คู่เทียบ**
⠮	!	the
⠯	&	and
⠷	(	of
⠾	)	with
⠿	=	for
⠮⠮	&!	and|the|
⠯⠷	&(	and|of|, and|of
⠯⠾	&)	and|with|
⠯⠿⠮	&=!	and|for|the
⠯⠿	&=	and|for|, and|for
⠯⠁	&a	and|a, anda
⠷⠮	(!	of|the
⠷⠯	(&	of|and
⠷⠁	(a	of|a, ofa
⠾⠮	)!	with|the
⠾⠁	)a	with|a
⠿⠮	=!	for|the
⠿⠯	=&	for|and
⠿⠁	=a	for|a

ภาคผนวก ฉ. 9 คู่เทียบที่ตีความต่างกัน ในตำแหน่งต้นคำหรือตำแหน่งอื่น

เบรลล์	แอสกี	คู่เทียบที่ 1	คู่เทียบที่ 2
⠰⠑	;e	⠰⠑ = ence	|⠰⠑ = |e
⠰⠛	;g	⠰⠛ = ong	|⠰⠛ = |g
⠰⠝	;n	⠰⠝ = tion	|⠰⠝ = |n
⠰⠎	;s	⠰⠎ = ness	|⠰⠎ = |s
⠰⠞	;t	⠰⠞ = ment	|⠰⠞ = |t
⠰⠽	;y	⠰⠽ = ity	|⠰⠽ = |y
⠨⠝	,n	⠨⠝ = ation, N	|⠨⠝ = |N Not
⠨⠽	,y	⠨⠽ =ally , You, Y	|⠨⠽ = |You |Y
⠤	-	⠤ = - , com	|⠤ = |com
⠖	+	⠖= ing , +	|⠖ = |+
⠿	#	⠿ = ble (ว่าง)	|⠿ = |#

ภาคผนวก ฉ. 10 คู่เทียบตัวย่อที่ตีความตามตำแหน่ง

เบรลล์	แอสกี	คู่เทียบที่ 1	คู่เทียบที่ 2	คู่เทียบที่ 3	คู่เทียบที่ 4	คู่เทียบที่ 5
⠂	1	⠂=ea, ,			⠂| = ,|	
⠃	2	⠃=bb, be	|⠃| = |be|	|⠃ = |be	⠃| = ;|	⠠⠃ = Be
⠒	3	⠒=con, cc	|⠒| = |:|	|⠒ = |con	⠒| = :|	⠠⠒ = Con
⠲	4	⠲=dd, .	|⠲| = |.|	|⠲ = |dis	⠲| = .|	⠠⠲ = Dis
⠄	5	⠄=en	|⠄|=|enough|			|⠠⠄| = |Enough|
⠶	6	⠶=ff, to, !	|⠶| = |!|	|⠶ = |to|	⠶|= !|	⠠⠶ = To ⠠⠠⠶= TO
⠿	7	⠿= (, gg,)	|⠿| = |were|	|⠿ = |(	⠿| =)|	|⠠⠿|= |Were| ⠠⠿ = [
⠦	8	⠦= " , ?	|⠦| = |his|	|⠦ = |"	⠦| = ?|	|⠠⠦| = |His| ⠠⠦ = '
⠤	0	⠤= " , by|	|⠤| = |was|	|⠤ = |by|	⠤| = "|	|⠠⠤| = |Was| |⠠⠤ = By| ⠠⠠⠤= BY
⠶⠂	62	⠶⠂=to|be !;		|⠶⠂= |to|be	⠶⠂| = !;|	⠠⠶⠂= To|be
⠶⠒	63	⠶⠒=to|con, !:		|⠶⠒= |to|con	⠶⠒| = !:|	⠠⠶⠒ =To| con
⠶⠲	64	⠶⠲=to|dis, !.		|⠶⠲= |to|dis	⠶⠲| = !.|	⠠⠶⠲= To|dis
⠿⠂	72	⠿⠂= (be,);		|⠿⠂ = |(be	⠿⠂| =);|	
⠿⠄	78	⠿⠄= (",)?		|⠿⠄ = |("	⠿⠄| =)?|	
⠰⠶	96	⠰⠶=into|, in!		|⠰⠶ = |into|	⠰⠶| = in!|	

เบรลล์	แอสกี	คู่เทียบที่ 1	คู่เทียบที่ 2	คู่เทียบที่ 3	คู่เทียบที่ 4	คู่เทียบที่ 5
⠳⠒	02	⠳⠒=by|be, ";		| ⠳⠒= |by|be	⠳⠒| = ";|	
⠳⠤	03	⠳⠤=by|con, ":		| ⠳⠤= |by|con	⠳⠤| = ":|	
⠳⠲	04	⠳⠲=by|dis, ".		| ⠳⠲=|by|dis	⠳⠲| = ".|	
⠳⠶	07	⠳⠶=by|(, ")		| ⠳⠶= |by|(	⠳⠶| = ")|	

ภาคผนวก ฉ. 11 คู่เทียบเครื่องหมายเบรลล์

เบรลล์	แอสกี	อักษรปกติ
⠲	"	.
⠲⠲⠲	"""	...
⠄	'	'
⠄⠄⠄	'''	...
⠂	,	,
⠂⠂⠂	,,,	...
⠨	.	. _
⠨⠅	.k	=
⠸⠨	.	_
⠸⠨⠨	..	__
⠨⠨	..	__
⠄⠄⠄	111	...
⠼⠼	4#	$
⠸⠼⠼	4#	|$
⠸⠼⠽⠼	04#	|by|$
⠸⠼⠾⠼	64#	|to|$
⠔⠔	99	*
⠸	|	|
⠆	;	(ว่าง)
⠈	@	@
⠈⠯	@&	&
⠈⠖	@+	+
⠈⠁	@a	@
⠘	^	^
⠨⠯	_&	|

เบรลล์	แอสกี	อักษรปกติ
⠸⠀	_+	(ว่าง)
⠸⠂	_/	/
⠸⠁	_1	,
⠸⠃	_2	;
⠸⠉	_3	:
⠸⠙	_4	.
⠸⠋	_6	!
⠸⠒	_:	(ว่าง)

ภาคผนวก ช

ตารางคู่เทียบเบรลล์อังกฤษรูปย่อ สำหรับระบบการถ่ายถอดด้วยวิธีการแบบผสม

การถ่ายถอดเบรลล์อังกฤษรูปย่อสำหรับระบบที่ใช้วิธีการแบบผสมนั้น ระบบจะใช้คู่เทียบใน ภาคผนวก ฉ. เพื่อถ่ายถอดส่วนกลางของสายอักขระเช่นเดียวกับการถ่ายถอดด้วยแบบจำลองเอ็นแกรม แต่จะใช้ตารางคู่เทียบต่อไปนี้แทนภาคผนวก ฉ.4

เบรลล์	แอสกี	คู่เทียบที่ 1	คู่เทียบที่ 2
⠃	b	⠃ =b	|⠃|=|but|
⠉	c	⠉ =c	|⠉|=|can|
⠙	d	⠙ =d	|⠙|=|do|
⠑	e	⠑ =e	|⠑|=|every|
⠋	f	⠋ =f	|⠋|=|from|
⠛	g	⠛ =g	|⠛|=|go|
⠓	h	⠓ =h	|⠓|=|have|
⠚	j	⠚ =j	|⠚|=|just|
⠅	k	⠅ =k	|⠅|=|knowledge|
⠇	l	⠇ =l	|⠇|=|like|
⠍	m	⠍ =m	|⠍|=|more|
⠝	n	⠝ =n	|⠝|=|not|
⠏	p	⠏ =p	|⠏|=|people|
⠗	r	⠗ =r	|⠗|=|rather|
⠎	s	⠎ =s	|⠎|=|so|
⠞	t	⠞ =t	|⠞|=|that|
⠥	u	⠥ =u	|⠥|=|us|
⠧	v	⠧ =v	|⠧|=|very|

ภาคผนวก ช (ต่อ)

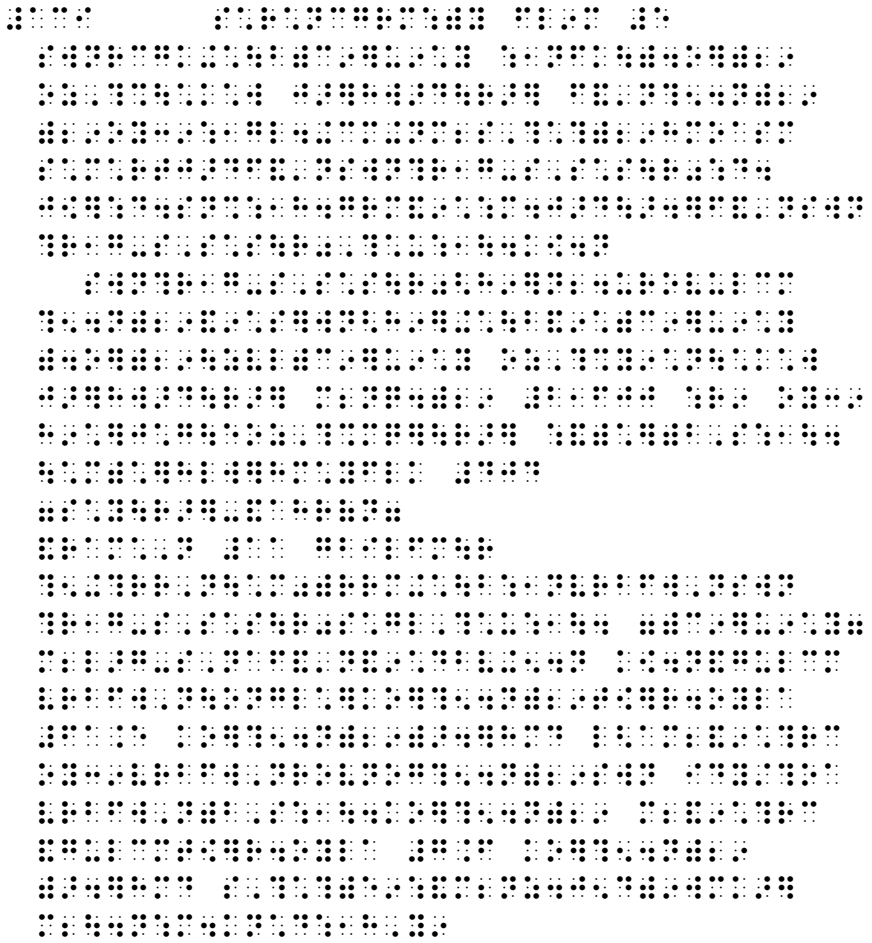
เบรลล์	แอสกี	คู่เทียบที่ 1	คู่เทียบที่ 2
⠺	w	⠺ =w	|⠺|=|will|
⠭	x	⠭ =x	⠭⠄ =it'; |⠭|=|it|
⠵	z	⠵ =z	|⠵|=|as|
⠡	*	⠡=ch	|⠡|=|child|
⠳	\	⠳=ou	|⠳|=|out|
⠩	%	⠩=sh	|⠩|=|shall|
⠹	?	⠹=th	|⠹|=|this|
⠱	:	⠱=wh	|⠱|=|which|

ภาคผนวก ซ ตัวอย่างข้อความเบรลล์ทั้ง 5 ประเภทที่ใช้ทดสอบระบบ

ในที่นี้จะขอแสดงข้อความเบรลล์ที่นำมาทดสอบระบบได้แก่ เบรลล์ไทย เบรลล์อังกฤษรูปเต็ม เบรลล์อังกฤษรูปย่อ เบรลล์ไทยปนอังกฤษรูปเต็ม และเบรลล์ไทยปนอังกฤษรูปย่อ โดยจะแสดงเพียงประเภทละ 1 หน้าเบรลล์เท่านั้น ซึ่งแต่ละข้อความเบรลล์จะแสดงข้อความอักษรปกติที่ตรงกับข้อความเบรลล์นั้น ๆ

<u>ภาคผนวก ซ. 1</u> ข้อความเบรลล์ไทย

(1.1) ข้อความเบรลล์นำเข้า

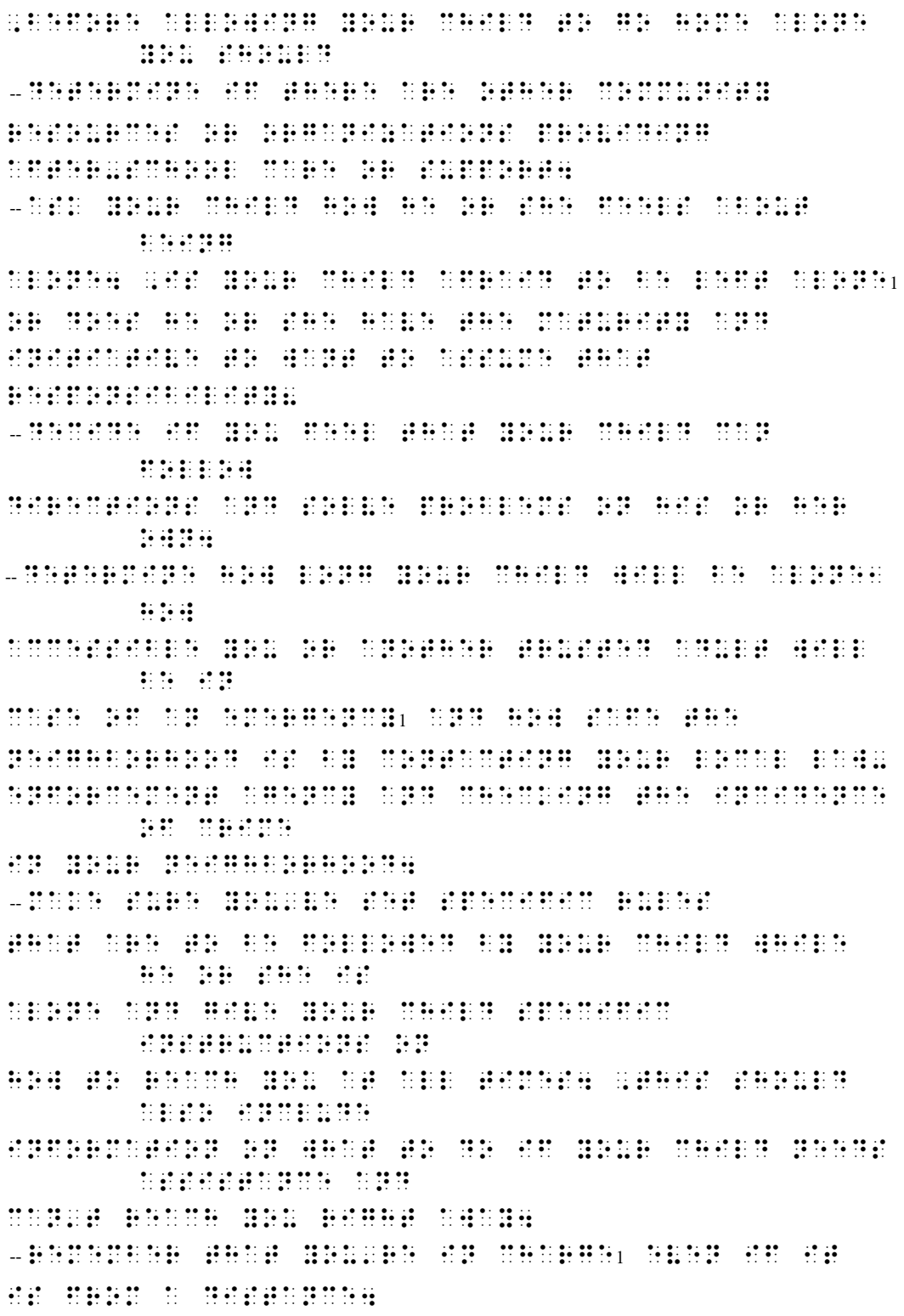


(1.2) ข้อความอักษรปกติที่ตรงกับเบรลล์ไทย

139 สารานุกรมไทย เล่ม 5
สวนรุกขชาติทุ่งค่าย ในเขตท้องที่
อำเภอตาขาว จังหวัดตรัง เป็นพื้นที่
ที่อยู่ใกล้ชุมชนมีสภาพที่เหมาะสม
สามารถจัดเป็นสวนพฤกษศาสตร์ได้
จึงได้เสนอให้กรมป่าไม้จัดตั้งเป็นสวน
พฤกษศาสตร์ภาคใต้ขึ้น
สวนพฤกษศาสตร์แห่งนี้ครอบคลุม
พื้นที่ป่าสงวนแห่งชาติป่าทุ่งค่าย
ท้องที่ตำบลทุ่งค่าย อำเภอย่านตาขาว
จังหวัดตรัง มีเนื้อที่ 2,600 ไร่ อยู่
ห่างจากตัวอำเภอเมืองตรัง ไปทางทิศใต้
ตามทางหลวงหมายเลข 404
(สายตรัง-ปะเหรียน)
ประมาณ 11 กิโลเมตร
พืชพรรณตามธรรมชาติในบริเวณสวน
พฤกษศาสตร์สากลภาคใต้ (ทุ่งค่าย)
มีลักษณะเป็นป่าดิบชื้น ขึ้นปกคลุม
บริเวณตอนกลางของพื้นที่ถึงร้อยละ
61.5 ของพื้นที่ทั้งหมด และมีป่าพรุ
อยู่บริเวณรอบนอกพื้นที่สวน โดยเฉพาะ
บริเวณทิศใต้ของพื้นที่ มีป่าพรุ
ปกคลุมถึงร้อยละ 7.6 ของพื้นที่
ทั้งหมด สภาพทั่วไปมีน้ำจืดท่วมขัง
มีต้นไม้ขนาดใหญ่

<u>ภาคผนวก ซ.2</u> ข้อความเบรลล์อังกฤษรูปเต็ม

(2.1) ข้อความเบรลล์นำเข้า

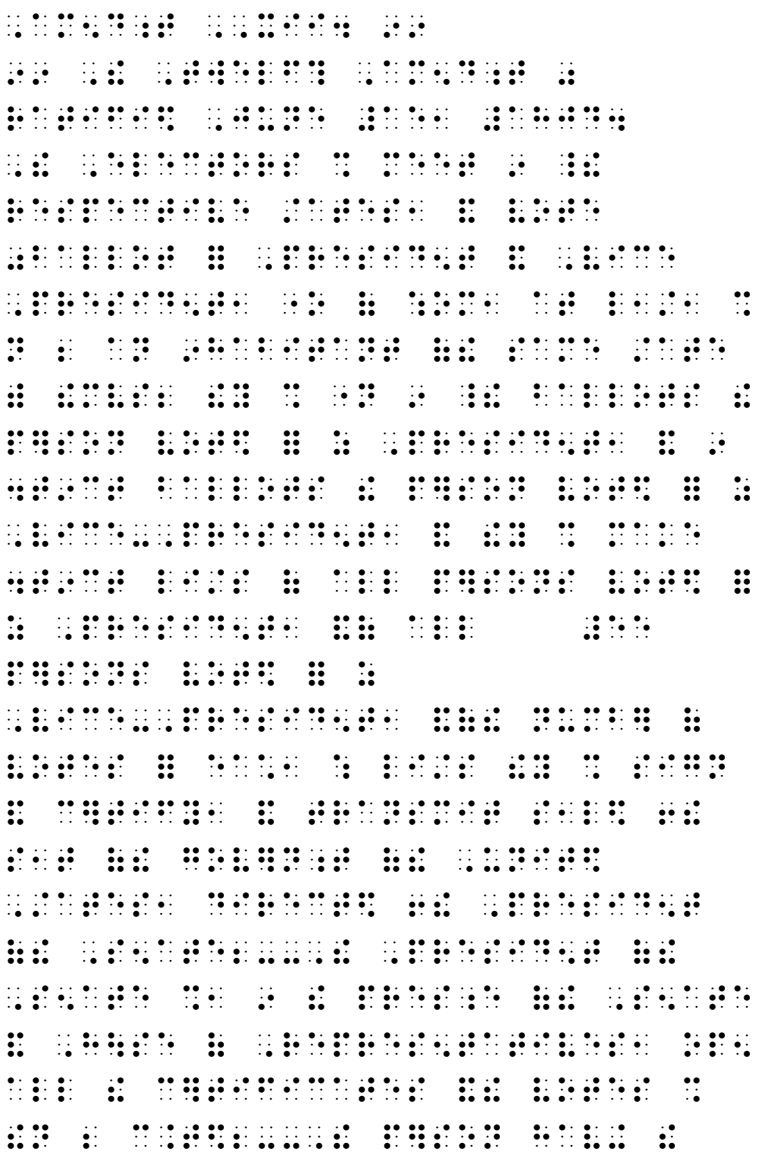


(2.2) ข้อความอักษรปกติที่ตรงกับเบรลล์อังกฤษรูปเต็ม

Before allowing your child to go home alone, you should
-- determine if there are other community
resources or organizations providing
after-school care or support.
-- ask your child how he or she feels about being
alone. Is your child afraid to be left alone,
or does he or she have the maturity and
initiative to want to assume that
responsibility?
-- decide if you feel that your child can follow
directions and solve problems on his or her own.
-- determine how long your child will be alone, how
accessible you or another trusted adult will be in
case of an emergency, and how safe the
neighborhood is by contacting your local law-
enforcement agency and checking the incidence of crime
in your neighborhood.
-- make sure you've set specific rules
that are to be followed by your child while he or she is
alone and give your child specific instructions on
how to reach you at all times. This should also include
information on what to do if your child needs assistance and
can't reach you right away.
-- remember that you're in charge, even if it
is from a distance.

ภาคผนวก ซ.3 ข้อความเบรลล์อังกฤษรูปย่อ

(3.1) ข้อความเบรลล์นำเข้า



(3.2) ข้อความอักษรปกติที่ตรงกับเบรลล์อังกฤษรูปย่อ

Amendment XII. *

* The Twelfth Amendment was

ratified June 15, 1804.

The Electors shall meet in their

respective states, and vote

by ballot for President and Vice

President, one of whom, at least, shall

not be an inhabitant of the same state

with themselves; they shall name in their ballots the

person voted for as President, and in

distinct ballots the person voted for as

Vice-President, and they shall make

distinct lists of all persons voted for

as President, and of all 55

persons voted for as

Vice-President, and of the number of

votes for each, which lists they shall sign

and certify, and transmit sealed to the

seat of the government of the United

States, directed to the President

of the Senate;--The President of the

Senate shall, in the presence of the Senate

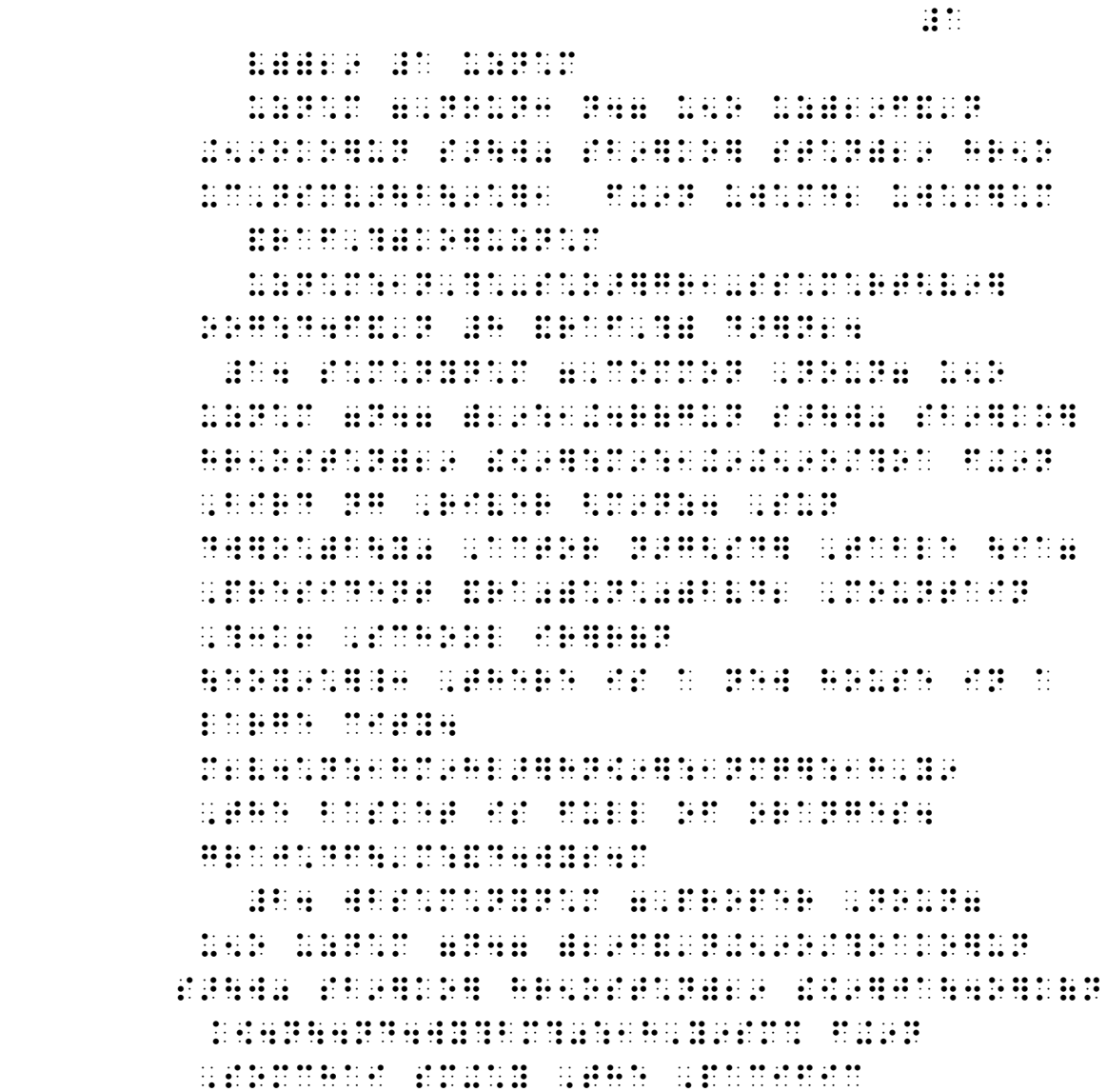
and House of Representatives, open

all the certificates and the votes shall

then be counted;--The person having the

ภาคผนวก ซ.4 ข้อความเบรลล์ไทยปนอังกฤษรูปเต็ม

(4.1) ข้อความเบรลล์นำเข้า



(4.2) ข้อความอักษรปกติที่ตรงกับเบรลล์ไทยปนอังกฤษรูปเต็ม

1

บทที่ 1 คำนาม

คำนาม (Noun: n.) คือ คำที่เป็น ชื่อของคน สัตว์ สิ่งของ สถานที่ หรือ คุณสมบัติต่าง ๆ เช่น ความดี ความงาม

ประเภทของคำนาม

คำนามในภาษาอังกฤษสามารถแบ่ง ออกได้เป็น 8 ประเภท ดังนี้

1. สามานยนาม (Common Noun) คือ คำนาม (n.) ที่ใช้เรียกคน สัตว์ สิ่งของ หรือสถานที่ ซึ่งไม่ใช่ชื่อเฉพาะ เช่น Bird นก River แม่น้ำ Sun ดวงอาทิตย์ Actor นักแสดง Table โต๊ะ President ประธานาธิบดี Mountain ภูเขา School โรงเรียน

ตัวอย่าง: There is a new house in a large city.

มีบ้านใหม่หลังหนึ่งในเมืองใหญ่

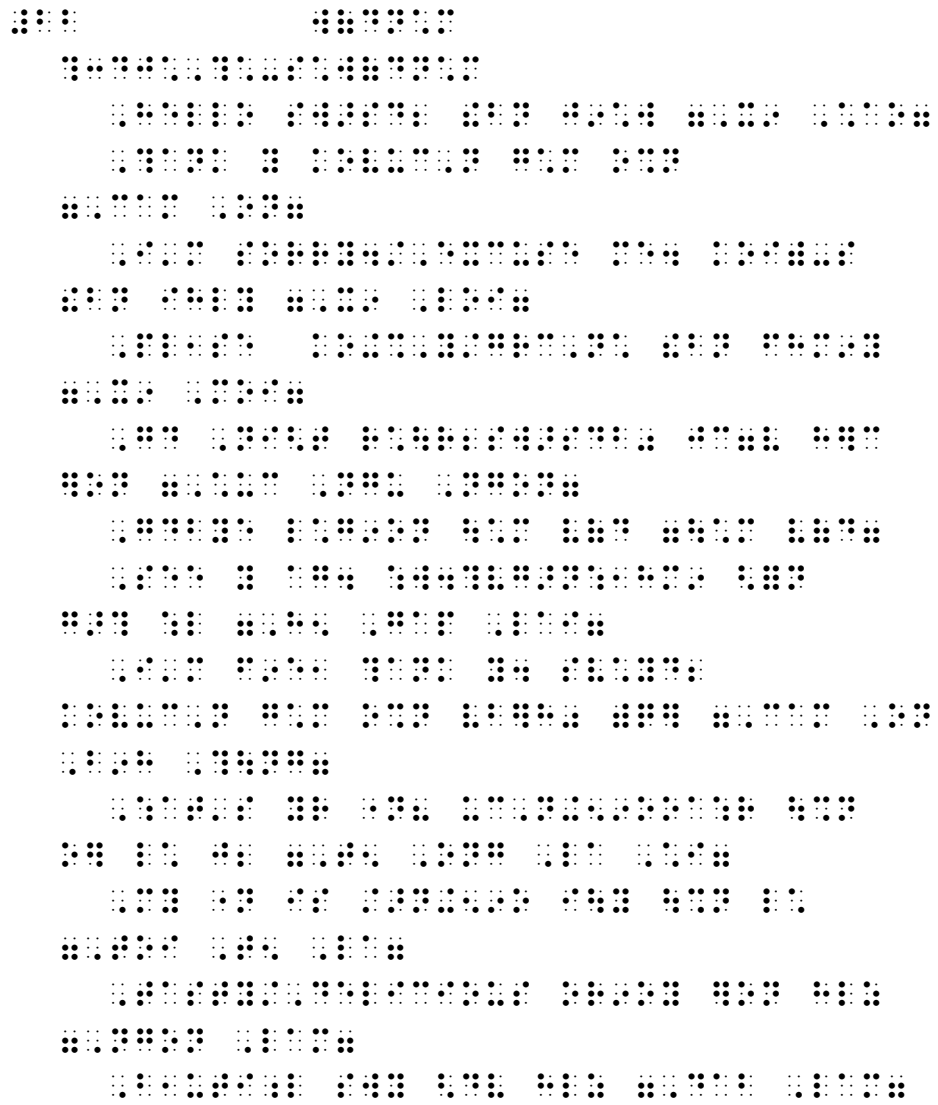
The basket is full of oranges.

กระจาดเต็มไปด้วยส้ม

2. วิสามานยนาม (Proper Noun) คือ คำนาม (n.) ที่เป็นชื่อเฉพาะของคน สัตว์ สิ่งของ หรือสถานที่ ซึ่งจะต้องเขียน ขึ้นต้นด้วยพิมพ์ใหญ่เสมอ เช่น Somchai สมชาย The Pacific

<u>ภาคผนวก ซ.5</u> ข้อความเบรลล์ไทยปนอังกฤษรูปย่อ

(5.1) ข้อความเบรลล์นำเข้า



(5.2) ข้อความอักษรปกติที่ตรงกับเบรลล์ไทยปนอังกฤษรูปย่อ

22 เวียดนาม

พูดจาภาษาเวียดนาม

Hello สวัสดี ซิน จ่าว (Xin Chao)

Thank you ขอบคุณ กาม เอิน (Cam On)

I'm sorry. /Excuse me. ขอโทษ ซิน โหลย (Xin Loi)

Please ขอเชิญ/กรุณา ซิน เหม่ย (Xin Moi)

Good Night ราตรีสวัสดิ์ จุ๊บ หงุ งอน (Chuc Ngu Ngon)

Goodbye ลาก่อน ตาม เบียด (ตาม เบียด)

See you again. ไว้พบกันใหม่ แฮน กาบ ไล (Hen Gap Lai)

I'm fine, thank you. สบายดี ขอบคุณ กาม เอิน บิงห์ เทือง (Cam On Binh Thoung)

What's your name? คุณชื่ออะไร เติน อง ลา จี (Ten Ong La Chi)

My name is ฉันชื่อ โตย เติน ลา (Toi Ten La)

Tasty/Delicious อร่อย งอน หลำ (Ngon Lam)

Beautiful สวย แดบ หลำ (Dab Lam)

ภาคผนวก ฌ. ตัวอย่างข้อผิดจากการถ่ายถอดด้วยกฎ

ภาคผนวก ฌ.1 ข้อผิดในการถ่ายถอดข้อความเบรลล์ไทย

การขาดสายอักขระที่จำเป็นในเซ็ตไตรแกรมท้ายพยางค์

ประเภท	เบรลล์	แอสกี	คำที่ผิด	สาเหตุ	คำที่ถูก
1.การไม่พบตัวสะกดการันต์ในเซ็ตไตรแกรมท้ายพยางค์	⠀	,s>?)0	ศัพท”	พท์	ศัพท์
	⠀	swb\+0	สวิตช”	ตช์	สวิตช์
	⠀	h]s0	หงส”	งส์	หงส์
	⠀	?r*hm,n0	พราหมณ”	มณ์	พราหมณ์
	⠀	w>divst0	วัดโบสถ”	สถ์	วัดโบสถ์
	⠀	o*rm,n0	อารมณ”	มณ์	อารมณ์
	⠀	:v\0	ไบต”	บต์	ไบต์
	⠀	sbrbgb\b0	สิริกิติ”	ติ์	สิริกิติ์
	⠀	urr,?0	ครรภ”	รภ์	ครรภ์
	⠀	u>\v6\0	คัตเอาต”	าต์	คัตเอาต์
	⠀	oc&t>m,?0	อุปถัมภ”	มภ์	อุปถัมภ์
	⠀	,s*n\b0	ศานติ”	ติ์	ศานติ์
	⠀	hr1hrr-s0	หฤหรรษ”	รษ์	หฤหรรษ์
	⠀	s\r(royd0	สเตรียรอยด”	ยด์	สเตรียรอยด์
	⠀	f!ll0	เซลล”	ลล์	เซลล์
2. ข้อผิดที่เกิดจากถ่ายถอด “ ์ท” เป็น “ ธ ”	⠀	pll>?0)0)29	ผลลัพธธี่	ธ์	ผลลัพธ์ที่
	⠀	,s>?)0)29	ศัพทธี่	ท์	ศัพท์ที่
	⠀	g-s>\rby0)r]	กษัตริยธรง	ิย์	กษัตริย์ทรง
	⠀	u>m,?2r0)*]	คัมภีรธาง	ีร์	คัมภีร์ทาง
	⠀	nbnh0)wn	นินหธวน	นห์	นินห์ทวน
	⠀	ivst0)cg	โบสถธุก	สถ์	โบสถ์ทุก
	⠀	swb\+0)29	สวิตชธี่	ตช์	สวิตช์ที่
3. “เาะ” ผิด	⠀	grdoambin	กรเดาะมิโน	กรด	กรดอะมิโน
4.หน้าสระผสม ผิด	⠀	j(]:uf+gr%9m	เจียงไคเชก เริ่ม	เชก	เจียงไคเชก เริ่ม

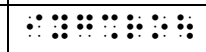
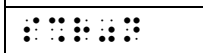
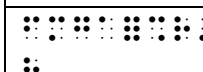
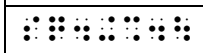
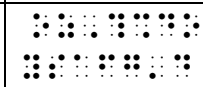
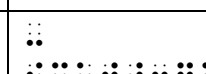
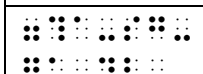
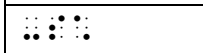
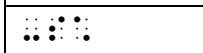
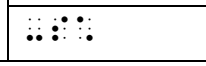
<u>ภาคผนวก ฌ.2</u> ข้อผิดในการถ่ายถอดข้อความเบรลล์ไทย

การปรากฏของสายอักขระเกินในเซ็ตไตรแกรมท้ายพยางค์

ประเภท	เบรลล์	แอสกี	คำที่ผิด	สาเหตุ	คำที่ถูก
1.ถ่ายถอดพยัญชนะหน้าสระผสมผิด	⠠⠛⠗⠎⠝⠄	g*rsn%	การสเนอ	ารส	การเสนอ
	⠠⠒⠁⠋⠞⠉⠇⠄	:1+4urq9]	ใช้คเรื่อง	ช้ค	ใช้เครื่อง
	⠠⠊⠠⠗⠙⠛⠇⠋⠄	I&rdgl64	โปรดกเล้า	รดก	โปรดเกล้า
	⠠⠛⠗⠎⠊⠍⠄	g*rsr%m	การสเริม	ารส	การเสริม
	⠠⠎⠗⠗⠎⠗⠄⠽⠄	srrsr%,y	สรรสเริญ	รรส	สรรเสริญ
	⠠⠛⠗⠇⠟⠝⠽⠋⠽⠄	g*rulq9ny4*y	การคเลื่อนย้าย	ารค	การเคลื่อนย้าย
	⠠⠝⠵⠓⠝⠋⠎⠍⠄	nzhn4*sm%	นำหน้าสเมอ	้าส	นำหน้าเสมอ
	⠠⠕⠛⠗⠢⠇⠐⠣	o*g*r?l(	อาการพเลีย	าร	อาการเพลีย
	⠠⠐⠜⠞⠙⠉⠇⠋⠄	)>n)2urq9]	ทันทีคเรื่อง	ทีค	ทันทีเครื่อง
	⠠⠒⠁⠋⠞⠉⠇⠅⠽⠄	:1+4urqk9*y	ใช้คเรือข่าย	ช้ค	ใช้เครือข่าย
	⠠⠗⠖⠗⠍⠑	r6\r(m	เราตเรียมตัว	ราต	เราเตรียม
	⠠⠧⠭⠄⠠⠛⠎⠢⠎⠝⠄	wc-)bs,?*sn%	วุฒิสภาสเนอ	ภาส	วุฒิสภาเสนอ
2. ถ่ายถอด “ ” ผิดเป็น “ ์”	⠠⠛⠗⠍⠛⠗⠔⠄	grrmg*r0	กรรมการ์	าร์	กรรมการ"
	⠠⠉⠍⠝⠚⠔⠄	+cm+n0	ชุมชน์	ชน์	ชุมชน"
	⠠⠇⠓⠺⠚⠄	<l4w0	หอยทากแล้ว์	้ว ์	แล้ว"
	⠠⠗⠁⠙⠓⠝⠵⠺⠄	r1d3hn*w0	หรือ "ฤดูหนาว์	าว์	ฤดูหนาว์"
	⠠⠙⠑⠥⠍⠝⠍⠄	d59mnm0	หันมาดื่มนม์	นม์	ดื่มนม"
	⠠⠓⠺⠝⠚⠄	<hwn0	ขยาย "แหวน์	วน์	แหวน"
	⠠⠜⠃⠝⠊⠵⠺⠄	!bnj9*w0	ว่า "ซินจ่าว์ ซึ่ง	าว์	ซินจ่าว"
	⠠⠚⠷⠄	j(0	เรียกว่า "เจีย์	ีย์	เจีย"
	⠠⠗⠟⠍⠗⠵⠺⠄	rq9]r*w0	หรือ "เรื่องราว์	าว์	เรื่องราว"

ประเภท	เบรลล์	แอสกี	คำที่ผิด	สาเหตุ	คำที่ถูก
3.ถ่ายถอด “ ◌์ท” ผิดเป็น “ธ”	⠵⠃⠵⠵⠃⠵⠵	<l4w0)%	แล้ว์เทอ	◌้ว์	แล้วเธอ
	⠿⠵⠃⠵⠵⠵⠃⠿	g*r0)zr]	เพื่อการ์ทำรงไว้	าร์	การธำรง
	⠿⠃⠿⠿⠵⠃⠵⠵⠿⠃⠵	j>dg*r0)cra	จัดการ์ทุระ	าร์	จัดการธุระ
	⠵⠿⠃⠿⠵⠵⠿⠵⠵⠃⠿⠵	u6r?0)]+*\b	เคารพ์ทงชาติ	รพ์	เคารพธงชาติ
4. “เาะ”ผิดเป็น “อะ”	⠿⠵⠿⠃⠵	m*goa	มากอะ	มาก	มาเกาะ
5. “.” (จุด) ผิดเป็น “◌้”	⠿⠿⠿⠿⠃⠿	n4m4s4	น้ม.ส.	น้ม	น.ม.ส.
6.“)” ผิดเป็น “◌๋”	s2dz7		สีด๋ำ	ด๋ำ	สีดำ)

ภาคผนวก ฌ.3 ข้อผิดในการถ่ายถอดข้อความเบรลล์ไทย
กฎที่ใช้รองรับไม่ครบทุกกรณี

ประเภท	เบรลล์	แอสกี	คำที่ถ่ายถอดผิด	สาเหตุที่ผิด	คำที่ถูกต้อง
1. “เ◌ิ” ผิดเป็น “เ-อ” จากการไม่พบอักขระบางตัวในเซ็ตพยัญชนะสะกดท้ายพยางค์		iyg%ro\	โยเกอร์ต	ไม่พบ ⠠ (ร)	โยเกิร์ต
		s%r0n	เฟอร์น	ไม่พบ ⠠ (ร)	เฟิร์น
		fmga=%r0\	เมกะเฮอรตซ์	ไม่พบ ⠠ (ร)	เมกะเฮิรตซ์
		sq4+%4\	เสื้อเช้อต	ไม่พบ ⠁ (ต)	เสื้อเชิ้ต
2. “เ-อ” ผิดเป็น“เ◌ิ” จากการไม่พบอักขระในเซ็ตสระทั่วไป		oz,?%doysafg’d	อำเภิดอยสะเก็ด	ไม่พบ ⠠ (อ)	อำเภอดอยสะเก็ด
3. ถ่ายถอด “ษ” ผิดเป็น “-ส” จากการไม่มีกฎการถ่ายถอด ⠠⠎ มารองรับ		-sck)>4]:1j	เย็นทั้งกายษุขทั้งใจ		สุขทั้งใจ
		7?a-sg-ga-nb-gon7	พสกนิกร (พะษก-กะ-นิ-กอน)		(พะ-สก-กะ-นิ-กอน)
		-s*	-สา เมื่อได้ฟังคำ		ษา
		-s*	-สาของบุคคล		ษา
		-s*	-สา มีตำนาน		ษา

ภาคผนวก ฌ .4 ข้อผิดที่พบในการถ่ายถอดข้อความเบรลล์อังกฤษรูปย่อ (ตัวอย่าง)

ประเภท	เบรลล์	แอสกี	คำที่ถ่ายถอดผิด	คำที่ถูกต้อง
1.ถ่ายถอด “-” ผิดเป็น “com”	⠤⠎⠌⠑⠎	-s/-es	coms/comes	-s/-es
	⠤⠽⠔⠽	-ity	comity	-ity
	⠤⠫	-$	comed	-ed
	⠤⠛⠬	-9g	coming	-ing
	⠤⠕	-o	como	-o
	⠶⠤⠶	7-7	(com)	(-)
	⠤⠢⠔	-en	comen	-en
	⠤⠞⠊⠕⠝	-tion	comtion	-tion
	⠤⠁⠊⠁⠝	-ian	comian	-ian
	⠤⠝⠑⠎⠎	-ness	comness	-ness
	⠤⠛⠑	-ge	comge	-ge
	⠤⠑⠎	-es	comes	-es
	⠤⠑⠎⠑	-ese	comese	-ese
	⠤⠁⠞⠊⠕⠝	-ation	comation	-ation
	⠤⠓⠕⠕⠙⠼	-hood1	comhood,	-hood,
	⠤⠊	-i	comi	-i
	⠤⠡⠆	-*1	comch,	-ch,
2.ถ่ายถอด “/” ผิดเป็น “st”	⠠⠕⠉⠞⠊⠍⠎⠺⠊⠞⠝⠑⠎⠎	victim/wit;s	victimstwitness	victim/witness
	⠠⠠⠉⠕⠕⠅⠖⠁⠞⠖⠑⠁⠞⠖	cook+/eat+	cookingsteating	cooking/eating
	⠠⠠⠋⠁⠍⠊⠇⠽⠸⠍⠻⠛⠵⠉⠽	family/ em]g5cy	familystemergency	family/ emergency
	⠠⠚⠕⠃⠸⠁	job/a	jobsta	job/a
	⠠⠕⠗⠸⠝⠑⠊	or/,nei	orstationei	or/Neither
	⠠⠎⠬⠇⠑⠸⠍⠗⠊⠫	s+le/m>ri$	singlestmarried	single/married

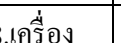
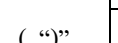
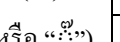
ประเภท	เบรลล์	แอสกี	คำที่ถ่ายถอดผิด	คำที่ถูกต้อง
3. ถ่ายถอดอักษรผิดเป็นตัวย่อ		n4	not.	n.
		v	very	v
		b4---	but. ---	b. ---
		c4---	can. ---	c. ---
		d4---	do. ---	d. ---
		7„cp/,m7	(CP/More)	(CP/M)
		f/fe	from/fe	f/fe
		f4---	from. ---	f. ---
		,>ticle ,v	Article Very	Article V
		e4---	every. ---	e. ---
		,“o,note	Oneationote	OneNote
		,type ,s	Type So	Type S
		,am5d;t ,x	Amendment It	Amendment X
		,r-,p-,e-,w-,e4	RATHER-PEOPLE-EVERY-WILL-EVERY	R-P-E-W-E.
		,c-,h-,e-,s-,t-,e	CAN-HAVE-EVERY-SO-THAT-EVERY	C-H-E-S-T-E

ภาคผนวก ญ. ตัวอย่างข้อผิดที่พบจากการถ่ายถอดด้วยแบบจำลองเอ็นแกรม

ญ.1 ข้อผิดจากการถ่ายถอดข้อความเบรลล์ไทย (ตัวอย่าง)

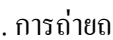
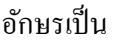
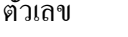
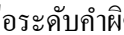
ประเภท	เบรลล์	แอสกี	คำที่ถ่ายถอดผิด	คำที่ถูก
1. ‘⠪⠊’ (“เาะ”หรือ “อะ”)	⠪⠊⠐⠕⠪⠊	joaj]	จอะจง	เจาะจง
	⠪⠊⠐⠕	goa	กอะ	เกาะ
	⠎⠐⠕⠪⠊⠐⠕⠓⠐⠓⠚	s>]uroah0	สังครอะห์	สังเคราะห์
	⠒⠪⠊⠐⠕⠓⠚	:?roa	ไพรอะ	ไพเราะ
	⠎⠪⠊⠐⠕⠓⠚	snoa	สนอะ	เสนาะ
	⠎⠪⠊⠐⠕	soa	สอะ	เสาะ
	⠓⠪⠊⠐⠕⠓⠚	hmoa	หมอะ	เหมาะ
	⠪⠊⠐⠕⠪⠊	ooa	ออะ	เอาะ
	⠎⠐⠕⠪⠊⠐⠕⠇⠇⠕⠁⠐⠶⠪⠊⠐⠕⠇⠇⠕⠁⠐⠶	v>y\clloa=0	บัยตุลเลาะฮ์	บัยตุลเลาะฮ์
	⠞⠐⠕⠇⠐⠕⠓⠚	)aloa	ทะลอะ	ทะเลาะ
	⠞⠊⠐⠕⠞⠊⠐⠕⠪⠊	&oa&(a7	ปอะเปี๊ยะ	เปาะเปี๊ยะ
	⠙⠒⠐⠕⠺⠐⠕⠗⠐⠕⠓⠚	d*wuroah0	ดาวครอะห์	ดาวเคราะห์
	⠎⠙⠙⠐⠕⠚⠐⠕⠚	kcdjoa	ขุดจอะ	ขุดเจาะ
2. ‘⠒⠊’ (“เ◌ิ” หรือ “เ◌อ”)	⠞⠊⠐⠕⠞⠊	gzr%v	กำเรอบ	กำเริบ
	⠞⠒⠐⠕⠞⠊	\%vi\	เตอบโต	เติบโต
	⠞⠒⠐⠕⠇⠒⠐⠕⠇⠝	?l%d?l%n	เพลอดเพลิน	เพลิดเพลิน
	⠞⠊⠐⠕⠞⠊⠐⠕⠝	v%gv*n	เบอกบาน	เบิกบาน
	⠞⠊⠐⠕⠋⠐⠕⠝	$%r0n	เฟอร์น	เฟิร์น
	⠞⠍⠐⠕⠍⠛⠐⠕⠗⠐⠕⠞	fmga=%r0\	เมกะเฮอร์ต	เมกะเฮิร์ต
	⠞⠽⠐⠕⠽⠝⠕	y%nyo	เยอนยอ	เยินยอ
	⠽⠛⠐⠕⠗⠐⠕⠞	Iyg%r0\	โยเกอร์ต	โยเกิร์ต
	⠎⠝⠐⠕	sn%	เสนิ	เสนอ
	⠙⠒⠐⠕⠙⠗⠐⠕⠃⠝	?2d%rbn	พีเดิริน	พีเดอริน

ประเภท	เบรลล์	แอสกี	คำที่ถ่ายถอดผิด	คำที่ถูกต้อง
3. ‘⠻’ (“⠽” หรือ “⠲”)	⠻⠻⠻⠻	m4l4	ม้ล้	ม.ล.
	⠻⠻⠻⠻	gs)4	กสท้	กสท.
	⠻⠻⠻⠻⠻	n4m4s4	น้ม.ส.	น.ม.ส.
	⠻⠻	k4	ข้	ข.
	⠻⠻	\4	ต้	ต.
	⠻⠻	!4	ซ้	ซ.
	⠻⠻⠻	fy4	เย.	เย้
4.การถ่ายถอด พยัญชนะ 2 ตัว ร่วมกับสระผสม	⠻⠻⠻⠻⠻⠻⠻⠻	d59mnmy%a	ดื่มนเมยอะ	ดื่มนมเยอะ
	⠻⠻⠻⠻⠻⠻⠻⠻	nz4*ly%a	น้ำตาเยลอะ	น้ำตาลเยอะ
	⠻⠻⠻⠻⠻⠻⠻⠻	irurq4r>]	โรเครื้อรัง	โรคเรื้อรัง
	⠻⠻⠻⠻⠻⠻⠻⠻	s\rov%r29	สตรเอบอรี่	สตรอเบอรี่
5. ‘⠻⠻’ (“ษ” หรือ “-ส”)	⠻⠻⠻ ⠻⠻⠻⠻⠻	?]-svrb?>\	พง-สบริพัต	พงษบริพัต
	⠻⠻⠻	-s*r	ษาร	-สาร
	⠻⠻⠻⠻⠻⠻⠻	-s*y:$$4*	ษายไฟฟ้า	-สายไฟฟ้า
	⠻⠻⠻	-s2	ษี	-สี
6 ‘⠻⠻’ ("์ท" หรือ "ธ")	⠻⠻⠻⠻ ⠻⠻⠻⠻⠻⠻ ⠻⠻⠻	?r1g-s,s*s\r0)w2+1	พฤกษศาสตรธวีชล	พฤกษศาสตร์ทวีชล
	⠻⠻⠻⠻⠻⠻⠻ ⠻⠻⠻	?h3?jn0)>n)2	พหูพจนธันที	พหูพจน์ทันที
7. ‘⠻’ (“ " ” หรือ “์”)	⠻⠻⠻⠻⠻⠻	rq9]r*w0	เรื่องราว์	เรื่องราว"
	⠻⠻⠻⠻⠻⠻	&raiyu0	ประโยค์	ประโยค"
	⠻⠻⠻⠻⠻⠻⠻	<lg&l(9n0	แลกเปลี่ยน์	แลกเปลี่ยน"

ประเภท	เบรลล์	แอสกี	คำที่ถ่ายถอดผิด	คำที่ถูกต้อง
		d4wy0	ด้วย"	ด้วย"
		i,?+n*g*r0	โภชนาการ์	โภชนาการ"
		)4o]s(0	ท้องเสีย์	ท้องเสีย"
8.เครื่องหมายผิด		r3iv8	รูโบ?	รูโบ๋
9. ‘’ (“)” หรือ “”)		i+ul*,?7	โชคลาภ๊	โชคลาภ)
		grai+g7	กระโชก๊	กระโชก)
		&27	ปี๊	ปี)
		)w2&7	ทวีป๊	ทวีป)
		<olfo7	แอลเอ๊	แอลเอ)
		!*n$r*n!bsig7	ซานฟรานซิสโก๊	ซานฟรานซิสโก)

ญ.2 ข้อผิดจากการถ่ายถอดข้อความเบรลล์อังกฤษรูปย่อ (ตัวอย่าง)

ข้อผิด	เบรลล์	แอสกี	คำที่ถ่ายถอดผิด	คำที่ถูกต้อง
1. การถ่ายถอด ‘⠉’ (“-” หรือ “com”)		-9g	coming	-ing
		-$	comed	-ed
		-$4	comed.	-ed.
		-ev5	comeven	-even
		-acy1	comacy,	-acy,
		-ion	comion,	-ion
2.ถ่ายถอดอักษรเป็นคำย่อ		dwell	dowell	dwell
		dg#dj	dog40	dg40
5. การถ่ายถอดคำย่อเป็นตัวอักษร		b4	b.	but.
		fr4	fr.	friend.
		c4	c.	can.
		d4	d.	do.
		mode/--f	modest--f	modest—from
		,color--t	Color--t	Color--that
5.การถ่ายถอดตัวเลขเป็นตัวอักษร		a#d	able	a4
		,docu;t#a-,micros(t	Documentblea-Microsoft	Document1-Microsoft
		,book#a-,micros(t	Bookblea-Microsoft	Book1-Microsoft

ข้อผิด	เบรลล์	แอสกี	คำที่ถ่ายถอดผิด	คำที่ถูกต้อง
6. การถ่ายถอดอักษรเป็นตัวเลข		#d-acre	4-13re	4-acre
		#c-column	3-3olumn	3-column
		#b-ft	2-6t	2-ft
		Co#b-equivalent	CO2-5quivalent	CO2-equivalent
7. ถ่ายถอดตัวย่อระดับคำผิด		nei8	nei?	neither?
		ag4	ag.	again.
		ei4	ei.	either.
		*4	ch.	child.
		,s-call$	S-called	So-called

ประวัติผู้เขียนวิทยานิพนธ์

นายวีรชัย อำพรไพบูลย์ เกิดวันที่ 23 มิถุนายน พ.ศ. 2512 ที่กรุงเทพมหานคร สำเร็จการศึกษาศิลปศาสตรบัณฑิต สาขาภาษาอังกฤษ จากคณะศิลปศาสตร์ มหาวิทยาลัยธรรมศาสตร์ ในปีการศึกษา 2536 สำเร็จการศึกษาอักษรศาสตรมหาบัณฑิต สาขาภาษาศาสตร์ จากคณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2543 และเข้าศึกษาต่อในหลักสูตรอักษรศาสตรดุษฎีบัณฑิต สาขาภาษาศาสตร์ ที่คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย เมื่อ พ.ศ. 2556 ปัจจุบันเป็นอาจารย์ประจำหลักสูตรภาษาอังกฤษธุรกิจ คณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยราชภัฏธนบุรี