

การพัฒนาฐานข้อมูลภูมิศาสตร์ด้านคำศัพท์ในภาษาไทยถิ่น
ด้วยระบบสารสนเทศภูมิศาสตร์
Development of Geographical Database of Lexical items in Thai Dialects
Using Geographic Information System

ศิริวิไล อีระโรจนารัตน์
sirivilai.t@chula.ac.th
ม.ร.ว. กัลยา ดิงศกัณย์
kalaya.t@chula.ac.th

บทคัดย่อ

ภาษาทุกภาษาแปรไปตามถิ่นที่อยู่ของผู้พูด ภาษาไทยก็มีการแปรดังกล่าวเช่นเดียวกัน คำศัพท์และวรรณยุกต์เป็นองค์ประกอบหลักที่จำแนกภาษาไทยถิ่น วิทยาภาษาไทยถิ่นได้ก้าวหน้าขึ้นมากในช่วง 40 ปีที่ผ่านมา ผลของการศึกษาช่วยให้ความรู้ใหม่เกี่ยวกับการแปรของภาษาไทยตามถิ่นที่อยู่ของผู้พูด ในงานส่วนใหญ่การกระจายของลักษณะทางภาษาและตำแหน่งของแนวแบ่งเขตภูมิภาคทางภาษาถิ่นแสดงไว้ในรูปของตารางและ/หรือแผนที่ เท่าที่ผ่านมาแผนที่ภาษาไทยถิ่นไม่ได้สร้างขึ้นโดยใช้ระบบสารสนเทศภูมิศาสตร์ ทำให้ขาดความแม่นยำและความน่าเชื่อถือทั้งในเชิงตำแหน่งและแนวแบ่ง งานวิจัยนี้มีวัตถุประสงค์หลักคือการออกแบบพัฒนาระบบฐานข้อมูลภูมิศาสตร์คำศัพท์ภาษาไทยถิ่นในระดับตำบลทั่วประเทศด้วยระบบสารสนเทศภูมิศาสตร์ เพื่อใช้เป็นฐานข้อมูลหลักในการสร้างแผนที่แสดงการกระจายของข้อมูลคำศัพท์

พื้นที่ศึกษาของงานวิจัยนี้ครอบคลุมทุกจังหวัดในประเทศ ข้อมูลที่ใช้มาจากโครงการเก็บรวบรวมข้อมูลคำศัพท์ภาษาพื้นบ้าน ซึ่งสำนักงานคณะกรรมการวัฒนธรรมแห่งชาติร่วมกับคณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัยเป็นผู้ดำเนินการ การรวบรวมข้อมูลใช้วิธีส่งแบบสอบถามที่ประกอบด้วยหน่วยย่อย 170 หน่วยย่อยไปยังตำบลทั่วประเทศ งานวิจัยนี้ใช้หน่วยย่อยจำนวน 7 หน่วยย่อย ซึ่งได้เคยมีผู้วิเคราะห์แล้วว่าสามารถนำมาแยกภูมิภาคทางภาษาถิ่นได้ ขั้นตอนการวิจัยแบ่งออกเป็น 3 ขั้นตอนหลัก ได้แก่ การเตรียมข้อมูลก่อนการนำเข้า การนำเข้าและเชื่อมโยงข้อมูลคำศัพท์ภาษาถิ่นเข้ากับข้อมูลแผนที่ และการสร้างแผนที่ข้อมูลคำศัพท์ภาษาถิ่นทั้ง 7 หน่วยย่อย ผลการวิจัยแสดงให้เห็นอย่างชัดเจนถึงศักยภาพในการจัดการชุดข้อมูลคำศัพท์ภาษาถิ่นด้วยระบบสารสนเทศภูมิศาสตร์ นอกจากนี้ ผลลัพธ์การวิจัย ยังได้แก่ เส้นแบ่งเขตภาษา 7 แผ่นซึ่งสร้างจากคำศัพท์สำหรับหน่วยย่อย 7 หน่วยย่อย และสามารถอธิบายและสรุปได้เป็น 4 ประเด็นหลักดังนี้ (1) หน่วยย่อยส่วนใหญ่ใน 7 หน่วยย่อยซึ่งแสดงด้วยแผนที่เส้นแบ่งเขตภาษาสามารถแสดงเส้นแบ่งเขตทางภาษาถิ่นทั้ง 4 ถิ่นได้ ทั้งเขตภาษาไทยถิ่นเหนือ ภาษาไทยถิ่น

กลาง ภาษาไทยถิ่นอีสาน และภาษาไทยถิ่นใต้ (2) หน่วยยรรยงบางหน่วยยรรยงแสดงเส้นแบ่งเขตภูมิภาคทางภาษาถิ่นคล้ายคลึงกัน (3) ปรากฏพื้นที่ที่มีลักษณะเฉพาะที่มีการใช้ “ภาษาอื่น” ในประเทศไทยซึ่งมีความสอดคล้องกันใน 3 บริเวณหลัก คือ บริเวณภาคใต้ตอนล่าง บริเวณภาคอีสานตอนล่างและตอนใน และบางตำบลของจังหวัดกาญจนบุรี (4) เมื่อพิจารณาในภาพรวมเส้นแบ่งเขตภาษาของภาษาไทยถิ่นมีความสอดคล้องกับแนวแบ่งเขตภูมิภาคทางภูมิศาสตร์

คำสำคัญ : แนวแบ่งเขตภูมิภาคทางภาษาถิ่น แผนที่ภาษา ภาษาไทยถิ่น การแปรของ
คำศัพท์ หน่วยยรรยง เส้นแบ่งเขตภาษา ระบบสารสนเทศภูมิศาสตร์

Abstract

Every language varies from one region to another. Thai is no exception. Lexical items and tones are the two main features that differentiate Thai dialects. Over the past 40 years Thai dialectology has made much progress providing new knowledge on regional variation in this language. In most studies the distribution of linguistic characteristics and the location of linguistic borders are displayed as table and/or map presentation. Up to now GIS techniques have not been utilized in Thai dialect map creation. The location of areas and borders are thus inaccurate and unreliable. The main objective of this research is to design and develop the geographical database of lexical items in Thai dialects by using Geographic Information System in order to be used as a key database for the creation of lexical variation maps.

The area of study in this research covers every province in Thailand. The data used come from the “Collection of Words in Local Dialect Project” by the Office of the National Cultural Commission and the Faculty of Arts, Chulalongkorn University, where a questionnaire consisting of 170 semantic units was sent to every Tambon. In this research seven semantic units were selected from that project, all of which have been shown in previous studies to be effective in segregating the four main Thai dialects. There are three major steps in this study: data preparation and database design for lexical items, data entry and linkage, and map generation for all seven semantic units. Results clearly demonstrate the potential of GIS to organize a data set for Thai dialects. In addition, seven isogloss maps were created from the lexical items for seven semantic units and the following conclusions were reached: (1) Almost all

of the seven semantic units can be used to segregate the four main dialect regions of Northern Thai, Central Thai, Northeastern Thai, and Southern Thai. (2) Thai dialect maps created by some semantic units show similar location of dialect borders. (3) Three areas where other languages are spoken are identified – the low parts of southern Thailand, the low- and inner-part of northeastern Thailand, and some Tambons in Kanchanaburi. (4) In overview the isoglosses of Thai dialects correspond to geographical borders.

Keywords: dialect border; linguistic map; Thai dialect; lexical variation; semantic unit; isogloss, Geographic Information System

1. บทนำ: ภาษาถิ่นและการประยุกต์ใช้ระบบสารสนเทศภูมิศาสตร์สำหรับงานภาษาไทยถิ่น

ภาษาทุกภาษามีการแปรตามถิ่นที่อยู่ของผู้พูด ผู้ที่พูดภาษาเดียวกันแต่มีภูมิสำเนาต่างกันมักใช้ภาษาที่มีลักษณะแตกต่างกันในแง่ของเสียงพูด คำศัพท์และ/หรือโครงสร้างประโยค นักภาษาศาสตร์สนใจศึกษาภาษาถิ่นของภาษาต่างๆ กันอย่างกว้างขวาง การศึกษาวิจัยภาษาถิ่นอย่างเป็นระบบเริ่มขึ้นที่ประเทศเยอรมันเมื่อคริสต์ศตวรรษที่ 19 จากนั้นการศึกษาภาษาถิ่นก็ขยายวงออกไปในประเทศต่างๆ ในทวีปยุโรปและอเมริกา สำหรับประเทศไทยนั้นได้มีการศึกษาวิจัยภาษาถิ่นมาแล้วประมาณ 40 ปี (สมทรง บุรุษพัฒน์, 2543, Tingsabadh, 2538) ในปัจจุบันมีงานวิจัยเกี่ยวกับภาษาไทยถิ่นอยู่จำนวนมาก ซึ่งส่วนใหญ่เป็นวิทยานิพนธ์ระดับมหาบัณฑิต ภาษาไทยถิ่นที่สำคัญมี 4 ภาษา ได้แก่ ภาษาไทยถิ่นเหนือ ภาษาไทยถิ่นอีสาน ภาษาไทยถิ่นกลาง และภาษาไทยถิ่นใต้ แต่ละภาษายังแบ่งออกได้เป็นภาษาถิ่นย่อย เช่น ภาษาไทยถิ่นเหนือหรือคำเมืองมี 2 ภาษาถิ่นย่อย ได้แก่ ภาษาคำเมืองตะวันตกและภาษาคำเมืองตะวันออก (วันทนี พันธชาติ, 2529) เป็นต้น ลักษณะสำคัญทางภาษาที่จำแนกภาษาไทยถิ่นต่างๆ คือ คำศัพท์และวรรณยุกต์ ส่วนสระและพยัญชนะนั้นใช้จำแนกภาษาถิ่นได้ในบางพื้นที่ (เจริญขวัญ ธรรมประดิษฐ์, 2524 ฌลลิกามหาพูนทอง, 2539) โครงสร้างประโยคและคำลงท้ายจำแนกภาษาถิ่นได้บ้างแต่ไม่ได้มีผู้ศึกษาไว้มากนัก ในบรรดาเรื่องต่างๆ ที่สามารถศึกษาวิจัยได้นี้ การแปรของคำศัพท์เป็นเรื่องที่ได้มีผู้ศึกษาวิจัยไว้มากที่สุด เท่าที่ผ่านมาผลการวิจัยคำศัพท์ภาษาไทยถิ่นปรากฏใน 2 ลักษณะ คือ (1) รายการคำ ซึ่งแสดงคำในภาษาไทยถิ่นพร้อมความหมาย เช่น พจนานุกรมภาษาถิ่นใต้ (สถาบันทักษิณคดีศึกษา, 2530) หรือ เสนอเป็นหมวดหมู่ตามกลุ่มความหมาย (2) แผนที่ ซึ่งมีอยู่ 2 ประเภทใหญ่ๆ คือ แผนที่แสดงการกระจายของคำศัพท์ โดยมักใช้แผนที่ 1 แผ่นแสดงคำศัพท์ต่างๆ ที่มีความหมายเดียวกัน นักภาษาศาสตร์ใช้คำว่า “หน่วยอรรถ” แทนหน่วยความหมาย อาทิเช่น หน่วยอรรถ “พูด” ใช้คำศัพท์หลักภาษาไทยถิ่น

เหนือ “อู” และ ใช้คำศัพท์หลักภาษาไทยถิ่นอีสาน “เว้า” ดังนั้นในกรณีของแผนที่ประเภทนี้ แผนที่ 1 แผ่นจึงแสดง 1 หน่วยอรรถ มีแผนที่อีกประเภทหนึ่ง เป็นแผนที่ที่ใช้แสดงเส้นแบ่งเขตภาษา (isogloss map) และแผนที่ที่ใช้แสดงแนวแบ่งเขตภาษา (boundary map)¹ (ดูตัวอย่างใน สมทรง บุรุษพัฒน์, 2543) แผนที่ทั้งสองประเภทที่สร้างขึ้นเท่าที่ผ่านมา เป็นแผนที่ที่แผ่นกระดาษสร้างขึ้นอย่างหยาบ กำหนดจุดแสดงข้อมูลอย่างคร่าว ๆ ด้วยสายตา และใช้มือในการลากเส้นแบ่งเขตภาษาหรือแนวแบ่งเขตภูมิภาค ทำให้แผนที่ผลลัพธ์ขาดความน่าเชื่อถือ และยากต่อการตรวจสอบความถูกต้อง ทั้งในเชิงตำแหน่งและแนวแบ่ง นอกจากนี้ยังมีข้อจำกัดในการนำไปวิเคราะห์เปรียบเทียบหากทำงานกับฐานข้อมูลคำศัพท์ภาษาไทยถิ่นขนาดใหญ่ซึ่งมีความซับซ้อน

ระบบสารสนเทศภูมิศาสตร์ (Geographic Information System หรือ GIS) สามารถลดข้อจำกัดดังกล่าวได้ GIS เป็นระบบสารสนเทศที่มีลักษณะเด่นเฉพาะตัว โดยเฉพาะความสามารถในการเชื่อมโยงข้อมูลเชิงพื้นที่ (geographic data) เข้ากับข้อมูลเชิงบรรยาย (attribute data) ได้อย่างมีประสิทธิภาพ รวมทั้งความสามารถในการจัดเก็บและจัดการกับฐานข้อมูลขนาดใหญ่ ความสามารถในการค้นคืน วิเคราะห์ และแสดงผลข้อมูลเชิงพื้นที่ในรูปแบบของภาพและแผนที่เฉพาะเรื่อง เป็นการเพิ่มการรับรู้และความเข้าใจกับข้อมูลในพื้นที่ที่ทำการศึกษา การประยุกต์ใช้ระบบสารสนเทศภูมิศาสตร์ในการจัดเก็บข้อมูลภาษาถิ่น นอกจากจะเอื้อประโยชน์ในการจัดเก็บและการจัดการกับฐานข้อมูลซึ่งประกอบไปด้วยหน่วยอรรถที่มีความซับซ้อนในแง่ความหลากหลายของคำศัพท์ ยังช่วยในการแสดงผลข้อมูลในรูปแบบของแผนที่แสดงการกระจายของคำศัพท์ที่มีความชัดเจน สามารถระบุตำแหน่งได้จากแผนที่ว่าบริเวณใดใช้คำศัพท์ใด นอกจากนี้ GIS ยังสามารถนำมาใช้เป็นเครื่องมือในการวิเคราะห์และสร้างแผนที่แสดงเส้นหรือแนวแบ่งเขตภูมิภาคทางภาษาถิ่นให้สามารถกำหนดแนวกรอบที่ชัดเจน ด้วยวิธีการที่น่าเชื่อถือ และสามารถตรวจสอบได้

แม้ว่างานวิจัยเกี่ยวกับภาษาไทยถิ่นมีอยู่จำนวนมากในปัจจุบัน การนำระบบสารสนเทศภูมิศาสตร์มาประยุกต์ใช้ในการจัดการข้อมูลภาษาไทยถิ่นนับได้ว่าเป็นเรื่องใหม่ มีผลงานวิจัยการประยุกต์ใช้ในประเทศไทยไม่มากนัก งานวิจัยชิ้นนี้เป็นส่วนหนึ่งของชุด

¹ ความแตกต่างระหว่างแผนที่ที่ใช้แสดงเส้นแบ่งเขตภาษา (isogloss map) และแผนที่ที่ใช้แสดงแนวแบ่งเขตภาษา (boundary map) คือ แผนที่ที่ใช้แสดงเส้นแบ่งเขตภาษา สร้างเส้นสมมติลากเพื่อแบ่งบริเวณที่มีการใช้ภาษา เช่น เส้นสมมติที่ลากเพื่อใช้แบ่งเขตการกระจายของคำศัพท์ ในหน่วยอรรถหนึ่งๆ ในขณะที่ แผนที่ที่ใช้แสดงแนวแบ่งเขตภาษา คือ แผนที่ที่เกิดจากการนำแผนที่ที่ใช้แสดงเส้นแบ่งเขตภาษาหลายๆแผ่น ซึ่งอาจเป็นเส้นแบ่งเขตศัพท์ พยัญชนะ หรือสระ มาลากซ้อนทับกัน เพื่อสร้างและกำหนดแนวแบ่งเขตภาษา (สมทรง บุรุษพัฒน์, 2535)

งานวิจัยในโครงการเก็บรวบรวมข้อมูลคำศัพท์ภาษาพื้นบ้าน² จัดได้ว่าเป็นงานวิจัยนาร่อง โดยมีวัตถุประสงค์หลักในการออกแบบพัฒนาระบบฐานข้อมูลภูมิศาสตร์คำศัพท์ภาษาไทยถิ่นในระดับตำบลทั่วประเทศด้วยระบบสารสนเทศภูมิศาสตร์ เพื่อใช้เป็นฐานข้อมูลหลักในการสร้างแผนที่แสดงการกระจายของข้อมูลคำศัพท์ภาษาไทยถิ่น และการศึกษาเปรียบเทียบเส้นแบ่งเขตภาษาของภูมิภาคภาษาไทยถิ่น ระหว่างภาษาไทยถิ่นเหนือ ภาษาไทยถิ่นกลาง ภาษาไทยถิ่นอีสาน และภาษาไทยถิ่นใต้ ในงานวิจัยนาร่องนี้ ได้ทำการจัดสร้างฐานข้อมูลภูมิศาสตร์เพื่อใช้ในการสร้างแผนที่แสดงการกระจายของข้อมูลคำศัพท์จำนวน 7 หน่วยย่อย จากคำศัพท์ทั้งหมดของโครงการจำนวน 170 หน่วยย่อย ผลที่ได้จากงานวิจัยชิ้นนี้จะใช้เป็นแนวทางในการปรับปรุงและพัฒนาฐานข้อมูลภูมิศาสตร์ภาษาไทยถิ่นสำหรับประเทศไทยจำนวน 170 หน่วยย่อยต่อไป

2. วัตถุประสงค์ของการวิจัย

1. เพื่อจัดทำฐานข้อมูลภูมิศาสตร์ภาษาไทยถิ่นด้วยระบบสารสนเทศภูมิศาสตร์
2. เพื่อวิเคราะห์และสร้างแผนที่แสดงการกระจายของคำศัพท์จากข้อมูลคำศัพท์จำนวน 7 หน่วยย่อยจากฐานข้อมูลภูมิศาสตร์คำศัพท์ภาษาไทยถิ่น
3. เพื่อเปรียบเทียบความสอดคล้องของเส้นแบ่งเขตภาษาของภูมิภาคทางภาษาถิ่นในประเทศไทย จากแผนที่แสดงการกระจายของคำศัพท์

² โครงการเก็บรวบรวมข้อมูลคำศัพท์ภาษาพื้นบ้าน เป็นความร่วมมือระหว่างสำนักงานคณะกรรมการวัฒนธรรมแห่งชาติ กระทรวงวัฒนธรรม ศูนย์วิจัยความหลากหลายทางภาษา ภาควิชาภาษาศาสตร์ และ ภาควิชาภูมิศาสตร์ คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย มีเป้าหมายหลักในการบันทึกคำศัพท์ที่ปรากฏอยู่ในภาษาไทยถิ่นในปัจจุบันทั่วประเทศ ในโครงการนี้ ศูนย์วิจัยความหลากหลายทางภาษารับผิดชอบการสร้างแบบสอบถาม ซึ่งประกอบด้วยหน่วยย่อยจำนวน 170 หน่วยย่อย สร้างขึ้นโดยใช้ผลงานวิจัยคำศัพท์ภาษาไทยถิ่นในอดีตเป็นพื้นฐาน (วรรณพร ทองมาก 2525, ภาควิชาภาษาศาสตร์ มป., จรุง บุญพันธ์และวันทนี พันธ์ชาติ 2526, วันทนี พันธ์ชาติ 2526, วิจิตร ภาณุพงศ์ 2529, ชลิดา โรจนวัฒนวุฒิ 2533, สุนีย์พร เลิศกุลทานนท์ 2533) ส่วนสำนักงานคณะกรรมการวัฒนธรรมแห่งชาติรับผิดชอบการเก็บรวบรวมข้อมูลผ่านทางเครือข่ายสภาวัฒนธรรมทั่วประเทศ และการวางระบบการบันทึกข้อมูลด้วยคอมพิวเตอร์โดยใช้โปรแกรม Microsoft Access ภาควิชาภาษาศาสตร์รับผิดชอบการบันทึกข้อมูล วิเคราะห์และจัดกลุ่มคำศัพท์ภาษาถิ่นจากแบบสอบถามในระดับที่พัฒนาขึ้น และภาควิชาภูมิศาสตร์รับผิดชอบในส่วนของการจัดเก็บและจัดการฐานข้อมูลภาษาไทยถิ่นเชิงพื้นที่ด้วยระบบสารสนเทศภูมิศาสตร์ การวิเคราะห์และการนำเสนอผลด้วยแผนที่โครงการนี้เป็นโครงการต้นแบบของการสร้างเครือข่ายเพื่อศึกษาค้นคว้าวิจัยวัฒนธรรมไทย ผลที่ได้จากโครงการนี้จะประโยชน์แก่ผู้ที่สนใจภาษาไทยถิ่นและวัฒนธรรมศึกษา ตลอดจนผู้ที่รับผิดชอบการบริหารงานด้านวัฒนธรรมระดับจังหวัด ภาค และประเทศ

3. พื้นที่ศึกษาและกรอบการวิจัย

พื้นที่ศึกษาของงานวิจัยนี้ครอบคลุมข้อมูลระดับตำบลของทุกจังหวัดในประเทศไทย ยกเว้นเขตในจังหวัดกรุงเทพมหานคร³ ข้อมูลคำศัพท์ได้มาจากการจัดเก็บข้อมูลผ่านเครือข่ายวัฒนธรรมในปีพ.ศ. 2545 หน่วยย่อยที่เลือกมาศึกษาในงานวิจัยนี้มี 7 หน่วยย่อย ได้แก่ (1) “ภูเขา” (2) “กะลามาพร้าว” (3) “พุด” (4) “ตู” (5) “รีบ” (6) “เรอ” (7) “ฉลาด” การคัดเลือกใช้เกณฑ์ว่าหน่วยย่อยที่จะนำมาใช้ในงานวิจัยนี้ต้องมีศักยภาพที่จะแสดงเส้นแบ่งเขตภาษาของภูมิภาคทางภาษาไทยถิ่นหลักทั้ง 4 ถิ่น ได้แก่ ภาษาไทยถิ่นเหนือ ภาษาไทยถิ่นกลาง ภาษาไทยถิ่นอีสาน และภาษาไทยถิ่นใต้ หน่วยย่อยที่ใช้ต้องมีความหมายชัดเจนและมีความหมายตรงกันในภาษาไทยถิ่นทั้ง 4 ภาค โดยพิจารณาจากผลของงานวิจัยในอดีต สำหรับตำบลที่ผู้ตอบแบบสอบถามใช้คำศัพท์มากกว่า 1 คำ เพื่อแทนหน่วยย่อยหนึ่งๆ หรือ ใช้ภาษาอื่นนอกเหนือจากภาษาถิ่นไทยหลัก อยู่นอกเหนืองานวิจัยนี้

4. เครื่องมือที่ใช้ในการวิจัย

1. โปรแกรมระบบสารสนเทศภูมิศาสตร์ ARCVIEW 3.2
2. โปรแกรม Microsoft Excel

5. ขั้นตอนการทำวิจัย

ขั้นตอนที่ 1 : การเตรียมข้อมูลก่อนการนำเข้าระบบสารสนเทศภูมิศาสตร์

การเตรียมการเพื่อนำข้อมูลในแบบสอบถามมาเข้าระบบสารสนเทศภูมิศาสตร์ ประกอบด้วย การเตรียมข้อมูล 2 ชุด คือ

1.1 ข้อมูลแผนที่แสดงเขตการปกครองประเทศไทยในระดับตำบล

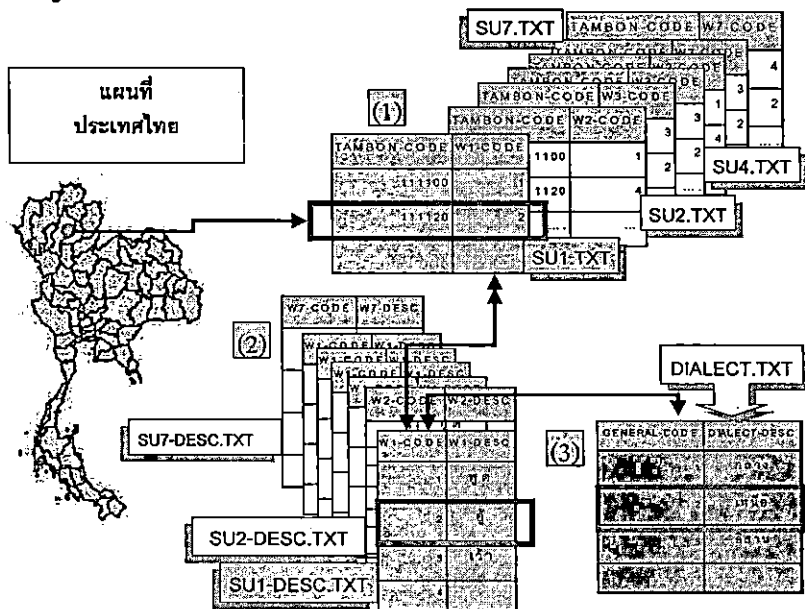
ข้อมูลดังกล่าวได้รับความอนุเคราะห์จากสถาบันสิ่งแวดล้อมไทย (Thailand Environment Institute, TEI) ในรูปข้อมูลแผนที่ฐานประเทศไทยเชิงตัวเลข (digital map) แสดงเขตการปกครองในระดับตำบล มาตราส่วน 1: 250,000 ข้อมูลเชิงพื้นที่ (Thailand on a Disc) ดังกล่าวได้มีการกำหนดรหัสตำบลไว้แล้ว และสามารถเชื่อมโยงกับตารางข้อมูลรหัสตำบลของกรมการปกครอง กระทรวงมหาดไทยได้ทันที

1.2 ข้อมูลคำศัพท์ภาษาไทยถิ่น จัดทำแฟ้มข้อมูลหลัก ดังต่อไปนี้ (ดูภาพที่1)

³ ในภาพกว้างกรุงเทพมหานครเป็นบริเวณที่ใช้ภาษาไทยถิ่นกลาง แต่นักภาษาศาสตร์จะไม่ศึกษาภาษาถิ่นในเขตกรุงเทพมหานคร เนื่องจากเป็นบริเวณที่มีประชากรหลากหลายทั้งผู้ที่เป็นชาวกรุงเทพฯแต่กำเนิด ผู้ที่อพยพย้ายถิ่นมาจากทั่วประเทศ และชาวต่างประเทศ อีกทั้งยังมีการใช้ภาษาไทยมาตรฐานอย่างกว้างขวางอีกด้วย (ในทางภาษาศาสตร์ ภาษามาตรฐานและภาษาถิ่นเป็นภาษาที่ไม่ได้อยู่ในมิติเดียวกัน)

(1) เพิ่มข้อมูลคำศัพท์สำหรับแต่ละหน่วยย่อย จำนวน 7 แฟ้ม (ตามจำนวนหน่วยย่อย) จากภาพตั้งชื่อแฟ้มว่า SU1.TXT ถึง SU7.TXT ตามลำดับ แฟ้มนี้เป็นแฟ้มที่ใช้เชื่อมข้อมูลคำศัพท์เข้ากับข้อมูลแผนที่แสดงเขตการปกครองในระดับตำบล ใช้เพื่อแสดงว่าตำบลใดใช้คำศัพท์ใด แต่ละแฟ้มประกอบด้วย ฟิลด์รหัสตำบล (TAMBON-CODE) และ ฟิลด์รหัสคำศัพท์ (W?-CODE⁴) เช่น แฟ้มข้อมูลคำศัพท์ของหน่วยย่อย “พูด” (SU1.TXT) ประกอบด้วยฟิลด์ตำบลที่มีรหัส “111120” และมีฟิลด์รหัสคำศัพท์ที่มีรหัส “2” ในแฟ้มข้อมูลนี้ รหัสตำบลจัดได้ว่าเป็นรหัสหลักที่ใช้ในการเชื่อมโยงข้อมูลในการแสดงคำศัพท์

(2) เพิ่มข้อมูลคำศัพท์ที่ใช้เรียกแต่ละหน่วยย่อย จำนวน 7 แฟ้ม (ตามจำนวนหน่วยย่อย) จากภาพตั้งชื่อแฟ้มว่า SU1-DESC.TXT ถึง SU7-DESC.TXT ตามลำดับ แต่ละแฟ้มประกอบด้วย ฟิลด์รหัสคำศัพท์ (W?-CODE⁵) และฟิลด์คำศัพท์ที่ใช้เรียก (SU?-DESC) เช่น แฟ้มข้อมูลคำศัพท์ของหน่วยย่อย “พูด” (SU1-DESC.TXT) รหัสคำศัพท์ “2” ใช้คำว่า “อุ้”



ภาพที่ 1 : โครงสร้างฐานข้อมูลภูมิศาสตร์คำศัพท์ภาษาไทยถิ่น
(โปรดดูภาพสีที่เว็บไซต์นี้ www.arts.chula.ac.th/~geography/faculty/Geo_sirivilai.htm)

(3) เพิ่มข้อมูลชื่อภาษาไทยถิ่น (Lookup Table for Thai Dialects) จำนวน 1 แฟ้ม ชื่อ DIALECT.TXT แฟ้มนี้ใช้เป็นแฟ้มอ้างอิงเพื่อแสดงรหัสคำหลักของภาษาไทยถิ่น

⁴ เครื่องหมาย ? ใน W?-CODE ใช้แทนตัวเลขของหน่วยย่อย 1 ถึง 7

⁵ เครื่องหมาย ? ใน W?-CODE ใช้แทนตัวเลขของหน่วยย่อย 1 ถึง 7

ที่กำหนดขึ้นในงานวิจัยนี้สำหรับทุกหน่วยย่อย เพิ่มประกอบด้วย ฟิรด์รหัสคำหลักของภาษาไทยถิ่นสำหรับทุกหน่วยย่อย (GENERAL-CODE) และฟิรด์ชื่อภาษาไทยถิ่น (DIALECT-DESC) เช่น รหัสคำหลัก “2” เป็น รหัสคำหลักของภาษาไทยถิ่นเหนือ

ความสัมพันธ์ของแฟ้มข้อมูลต่าง ๆ ในฐานข้อมูลภูมิศาสตร์ภาษาไทยถิ่นสามารถสรุปได้ดังนี้ (1) ข้อมูลแผนที่แสดงเขตการปกครองระดับตำบล มีความสัมพันธ์แบบหนึ่งต่อหนึ่ง (one-to-one relationship) กับ แฟ้มข้อมูลคำศัพท์ (2) แฟ้มข้อมูลคำศัพท์ มีความสัมพันธ์แบบพหุต่อหนึ่ง (many-to-one relationship) กับ แฟ้มข้อมูลคำศัพท์ที่ใช้เรียก และ (3) แฟ้มข้อมูลคำศัพท์ที่ใช้เรียก มีความสัมพันธ์แบบหนึ่งต่อหนึ่ง (one-to-one relationship) กับ แฟ้มข้อมูลชื่อภาษาไทยถิ่น

ขั้นตอนที่ 2 : การเก็บรวบรวมข้อมูลคำศัพท์ภาษาไทยถิ่น

ข้อมูลที่ใช้ในงานวิจัยนี้ได้มาจากโครงการเก็บรวบรวมข้อมูลคำศัพท์ภาษาพื้นบ้าน ซึ่งส่งแบบสอบถามจำนวน 7,156 ชุด ไปเก็บข้อมูลจากตำบลทั่วประเทศ ได้แบบสอบถามคืนมาและใช้ได้จำนวน 4,682 ชุด คิดเป็นร้อยละ 65.43 แบบสอบถามที่ใช้ไม่ได้มีจำนวน 490 ชุด คิดเป็นร้อยละ 6.85 สาเหตุที่ใช้ไม่ได้เกิดจากการกรอกข้อมูลไม่ครบถ้วน โดยเฉพาะข้อมูลเกี่ยวกับตำบล

ขั้นตอนที่ 3 : การจัดหมวดหมู่และกำหนดรหัสคำศัพท์ภาษาไทยถิ่น

ขั้นตอนนี้เป็นการออกแบบระบบฐานข้อมูลคำศัพท์ภาษาไทยถิ่นโดยอาศัยความร่วมมือระหว่างนักภาษาศาสตร์และนักภูมิศาสตร์ วัตถุประสงค์ของการดำเนินการคือ การจัดหมวดหมู่คำศัพท์ที่ปรากฏในข้อมูลและกำหนดรหัสสำหรับคำศัพท์ที่เป็นคำศัพท์หลักของแต่ละภาษาถิ่นเพื่อให้เห็นผลในลักษณะของแผนที่เฉพาะเรื่อง (Thematic map) ที่สามารถแสดงรูปแบบการกระจายของภาษาถิ่นได้ (ดูตัวอย่างคำศัพท์ในตารางที่ 1) หลังจากจัดหมวดหมู่และกำหนดรหัสคำศัพท์แล้ว จึงลงรหัสในแฟ้มต่าง ๆ ที่ได้จัดเตรียมไว้ในขั้นตอนที่ 1 (ดูภาพที่ 1)

ภาษาไทยถิ่น	คำศัพท์ที่ใช้เรียกหน่วยย่อย “ภูเข” ในท้องถิ่น รวบรวมจากผู้ตอบแบบสอบถาม
กลาง	ภูเขา
เหนือ	ดอย
อีสาน	ภู
ใต้	เขา
ภาษาอื่น	ควน ภูเขา บนเขา เขาเขี้ยว ภูเกด โฟเขา คูหรง โคก ชัว เนินสูง กะเจ้อ พนม ม่อน บลู เตอ

ตารางที่ 1 : การจัดหมวดหมู่คำศัพท์ที่ใช้เรียกหน่วยย่อย “ภูเข”

ขั้นตอนที่ 4 : การเชื่อมข้อมูลคำศัพท์เข้ากับข้อมูลเชิงพื้นที่

การเชื่อมข้อมูลคำศัพท์กับข้อมูลเชิงพื้นที่ที่ใช้ฟิลเตอร์หีสดาบลในการเชื่อมโยงเนื่องจากข้อมูลเชิงพื้นที่แสดงเขตการปกครองในระดับตำบลได้มีการกำหนดรหัสตำบลไว้แล้ว จึงเป็นการเชื่อมแผนที่เขตการปกครองในระดับตำบลเข้ากับแฟ้มข้อมูลคำศัพท์ ซึ่งเก็บอยู่ในโปรแกรมสำเร็จรูป Microsoft Excel โดยใช้ลักษณะการเชื่อมโยงแบบชั่วคราว (temporary link) การเชื่อมโยงในลักษณะนี้ช่วยลดขนาดฐานข้อมูลเชิงพื้นที่ และง่ายต่อการปรับแก้แฟ้มข้อมูลคำศัพท์ของแต่ละหน่วยย่อย เนื่องจากสามารถปรับแก้ได้โดยตรงใน Microsoft Excel โดยนักภาษาศาสตร์

ขั้นตอนที่ 5 : การตรวจสอบความถูกต้องแผนที่ข้อมูลคำศัพท์

หลังจากที่ทำการเชื่อมโยงข้อมูลคำศัพท์กับข้อมูลเชิงพื้นที่แล้ว ได้กำหนดสัญลักษณ์สี (color symbol) เพื่อดูการกระจายของคำศัพท์ในพื้นที่ ทั้งนี้เพื่อตรวจสอบความถูกต้องในการกรอกข้อมูลของผู้ตอบแบบสอบถามและเป็นการตรวจสอบความถูกต้องในการลงรหัสคำศัพท์ในแฟ้มข้อมูลอีกด้วย ขั้นตอนนี้อาจกล่าวได้ว่ามีประโยชน์อย่างยิ่งสำหรับนักภาษาศาสตร์ เนื่องจากในอดีตที่ผ่านมา ข้อมูลคำศัพท์เก็บอยู่ในตารางข้อมูล ทำให้ยากต่อการตรวจสอบหรือใช้เวลานานในการตรวจสอบ การนำระบบสารสนเทศภูมิศาสตร์มาช่วยเพิ่มประสิทธิภาพทั้งในแง่ความถูกต้องและความรวดเร็วในการตรวจสอบข้อมูลจากแผนที่ ในกรณีที่นักภาษาศาสตร์พบบริเวณที่ผิดพลาดอันเนื่องมาจากการกรอกข้อมูลผิด ก็สามารถปรับแก้แฟ้มข้อมูลคำศัพท์ได้ทันทีและย้อนกลับไปทำการเชื่อมโยงข้อมูลในขั้นตอนที่ 4 อีกครั้ง

ขั้นตอนที่ 6 : การผลิตแผนที่ผลลัพธ์

แผนที่ผลลัพธ์ ได้แก่ แผนที่แสดงข้อมูลคำศัพท์ภาษาไทยถิ่น 7 หน่วยย่อย ได้แก่ แผนที่ของหน่วยย่อย (1) “ภูเขา” (2) “กะลามาพร้าว” (3) “พุด” (4) “คู” (5) “ริบ” (6) “เรือ” และ (7) “ฉลาด” (ดูตัวอย่างแผนที่ผลลัพธ์โครงการ ในภาพที่ 2) มีการกำหนดคำศัพท์หลักของภาษาไทยถิ่นแต่ละภาษาและกำหนดคำศัพท์อื่นๆ เป็นคำศัพท์ของภาษาอื่น (ดูตารางที่ 2) การกำหนดสัญลักษณ์สี (color symbol) และคำอธิบายสัญลักษณ์ (legend) ใช้การระบายสี (shade) ในระดับตำบลเพื่อแทนข้อมูลคำศัพท์หนึ่ง ๆ ที่ใช้ในตำบลนั้น ๆ ทั้งนี้ เพื่อให้เป็นมาตรฐานเดียวกันสำหรับแผนที่ทั้งหมดในงานวิจัยนี้ (ดูตารางที่ 3)

ความหมายคำ	คำศัพท์หลัก ภาษาไทยถิ่น กลาง	คำศัพท์หลัก ภาษาไทยถิ่น เหนือ	คำศัพท์หลัก ภาษาไทยถิ่น อีสาน	คำศัพท์หลัก ภาษาไทยถิ่น ใต้	คำศัพท์ ภาษาอื่น
1. “ภูเขา”	ภูเขา	ดอย	ภู	เขา	อื่น ๆ
2. “กะลามะพร้าว”	กะลา	กะโหล่ง	กะโปะ	พรก	อื่น ๆ
3. “พูด”	พูด	อู้	เว้า	แล้ง	อื่น ๆ
4. “ดู”	ดู	ผ่อ	เบิง	แล	อื่น ๆ
5. “รีบ”	รีบ	ฟั่ง	ฟ้าว	แซบ	อื่น ๆ
6. “เรอ”	เรอ	เออะ	เอี่ยม	เอย	อื่น ๆ
7. “ฉลาด”	ฉลาด	หลืวก	ป่อง	ปัญญา	อื่น ๆ

ตารางที่ 2 : ตารางการอ่านแผนที่คำศัพท์ภาษาไทยถิ่นทั้ง 7 หน่วยย่อย

6. ผลลัพธ์และการอภิปรายผล

ผลลัพธ์ของงานวิจัยชิ้นนี้ ได้แก่ แผนที่ข้อมูลคำศัพท์ภาษาไทยถิ่น 7 แผ่น ซึ่งสร้างจากจำนวนหน่วยย่อย 7 หน่วยย่อย คือ (1) “ภูเขา” (2) “กะลามะพร้าว” (3) “พูด” (4) “ดู” (5) “รีบ” (6) “เรอ” และ (7) “ฉลาด” แผนที่เหล่านี้แสดงลักษณะการกระจายของการใช้ภาษาในเชิงพื้นที่ (spatial distribution of language) ผลการวิจัยมีข้อน่าสังเกต สรุปได้เป็นประเด็นหลัก ๆ ดังนี้

ประเด็นที่ 1 : ข้อมูลคำศัพท์ในทั้ง 7 หน่วยย่อย สามารถนำมาใช้แสดงเส้นแบ่งเขตภาษา (isogloss) ของภูมิภาคทางภาษาถิ่นได้หรือไม่

จากการพิจารณาผลลัพธ์แผนที่ข้อมูลคำศัพท์ภาษาไทยถิ่นทั้ง 7 หน่วยย่อย อาจสรุปได้ว่า หน่วยย่อยเกือบทุกหน่วยย่อยที่นำมาใช้ในงานวิจัย (ยกเว้นหน่วยย่อย “ภูเขา”) สามารถแสดงเส้นแบ่งเขตภาษาของภูมิภาคทางภาษาถิ่นระหว่าง 4 ถิ่นได้ ทั้งภาษาไทยถิ่นเหนือ ภาษาไทยถิ่นกลาง ภาษาไทยถิ่นอีสาน และภาษาไทยถิ่นใต้⁶ โดยคำศัพท์ที่ใช้เรียกหน่วยย่อยดังกล่าว เมื่อนำมาแสดงด้วยการให้สี (shade) ในแผนที่ การกระจายตัวของคำศัพท์ในทางพื้นที่มีรูปแบบการเกาะกลุ่ม (cluster) และสามารถลากเส้นแบ่งเขตภาษาของภูมิภาคทางภาษาถิ่นได้ (ดูตารางที่ 4) อย่างไรก็ตาม สำหรับกรณีของหน่วยย่อย “ภูเขา” มีลักษณะการปะปนกันของการใช้คำศัพท์ระหว่างคำศัพท์หลักของภาษาไทยถิ่นกลาง (“ภูเขา”) ภาษาไทยถิ่นใต้ (“เขา”) และภาษาไทยถิ่นอีสาน (“ภู”) ทำให้ไม่สามารถ

⁶ ผลลัพธ์ในงานวิจัยนี้มีข้อจำกัดเนื่องจากข้อมูลที่ใช้ในการวิเคราะห์มาจากร้อยละ 65.43 ของตำบลทั่วประเทศ ขณะนี้สำนักงานคณะกรรมการวัฒนธรรมแห่งชาติและคณะอักษรศาสตร์กำลังเก็บรวบรวมข้อมูลเพิ่มเติมในอนาคตเมื่อได้ข้อมูลเพิ่มขึ้น เส้นแบ่งเขตภาษาของภูมิภาคทางภาษาถิ่นอาจเปลี่ยนแปลงได้

กำหนดเส้นแบ่งเขตภาษาของภูมิภาคทางภาษาถิ่นได้ เมื่อพิจารณาในแง่นี้ อาจกล่าวได้ว่าหน่วยย่อย “ภูเขา” ไม่เหมาะที่จะนำมาใช้ศึกษาเส้นแบ่งเขตภาษาของภูมิภาคทางภาษาถิ่น

คำศัพท์ ภาษาไทยถิ่น	แทนด้วย สัญลักษณ์สี	หมายเหตุ
กลาง	เหลือง (yellow)	คำศัพท์หลักของภาษาถิ่นกลางที่ระบุไว้ในตารางที่ 2
เหนือ	แดง (red)	คำศัพท์หลักของภาษาถิ่นเหนือที่ระบุไว้ในตารางที่ 2
อีสาน	น้ำเงิน (blue)	คำศัพท์หลักของภาษาถิ่นอีสานที่ระบุไว้ในตารางที่ 2
ใต้	ม่วง (magenta)	คำศัพท์หลักของภาษาถิ่นใต้ที่ระบุไว้ในตารางที่ 2
ภาษาอื่น	เทา (gray)	ผู้ตอบแบบสอบถามใช้คำศัพท์อื่นนอกเหนือจากคำศัพท์หลักของภาษาถิ่นทั้ง 4 ภาษาที่ระบุไว้ในตารางที่ 2
“ผสม 2 คำ”	ส้มอ่อน (light orange)	ผู้ตอบแบบสอบถามใช้คำศัพท์หลักของภาษาถิ่น 2 ภาษาที่ระบุไว้ในตารางที่ 2
“ผสม 3 คำ”	เขียวอ่อน (light green)	ผู้ตอบแบบสอบถามใช้คำศัพท์หลักของภาษาถิ่น 3 ภาษาที่ระบุไว้ในตารางที่ 2
“ผสม 4 คำ”	ฟ้าอ่อน (light cyan)	ผู้ตอบแบบสอบถามใช้คำศัพท์หลักของภาษาถิ่น 4 ภาษาที่ระบุไว้ในตารางที่ 2
“ไม่มีข้อมูล”	ขาว (white)	ไม่มีข้อมูลในแฟ้มข้อมูล

ตารางที่ 3 : ตารางการอ่านสัญลักษณ์สีและคำอธิบายสัญลักษณ์แผนที่คำศัพท์ภาษาไทยถิ่นทั้ง 7 หน่วยย่อย

ประเด็นที่ 2 : การกระจายของคำศัพท์ใน 7 หน่วยย่อยแสดงเส้นแบ่งเขตภาษาของภูมิภาคทางภาษาถิ่นสอดคล้องกันหรือไม่

เมื่อเปรียบเทียบแผนที่ผลลัพธ์ของทั้ง 7 หน่วยย่อย สรุปได้ว่าแม้หน่วยย่อยเกือบทุกหน่วยย่อยจะสามารถแสดงเส้นแบ่งเขตภาษาของภูมิภาคภาษาถิ่นได้ชัดเจน แต่เส้นแบ่งเหล่านี้มีความคล้ายคลึงและแตกต่างกัน สรุปได้ดังนี้

2.1 ภาษาไทยถิ่นเหนือ : แผนที่ผลลัพธ์ของทั้ง 7 หน่วยย่อยสามารถแสดงเส้นแบ่งเขตภาษาของภาษาไทยถิ่นเหนือได้คล้ายคลึงกัน ตัวอย่างในภาพที่ 4 แสดงความสอดคล้องของเส้นแบ่งเขตภาษาระหว่างหน่วยย่อย “ภูเขา” “กะลามาพะรว้า” “พูด” และ “ฉลาด”

2.2 ภาษาไทยถิ่นกลาง : เฉพาะแผนที่ผลลัพธ์ของหน่วยย่อย “ดู” “rieb” และ “เรอ” สามารถแสดงเส้นแบ่งเขตภาษาของภาษาไทยถิ่นกลางได้คล้ายคลึงกัน ตัวอย่างในภาพที่ 5 แสดงความสอดคล้องของเส้นแบ่งเขตภาษาระหว่างหน่วยย่อย “ดู” และ “rieb”

2.3 ภาษาไทยถิ่นอีสาน : เฉพาะแผนที่ผลลัพธ์ของหน่วยย่อย “พูด” “ดู” และ “rieb” สามารถแสดงเส้นแบ่งเขตภาษาของภาษาไทยถิ่นอีสานได้คล้ายคลึงกัน ตัวอย่างในภาพที่ 6 แสดงความสอดคล้องของเส้นแบ่งเขตภาษาระหว่างหน่วยย่อย “ดู” และ “rieb”

2.4 ภาษาไทยถิ่นใต้ : เส้นแบ่งเขตภาษาของภาษาไทยถิ่นใต้มีลักษณะคล้ายคลึงกัน แบ่งออกได้เป็น 2 กลุ่ม (ดูภาพที่ 7) ดังนี้

กลุ่มที่ 1 ในแผนที่ผลลัพธ์ของหน่วยอรรถ “ดู” “รีบ” และ “เรอ” มีเส้นแบ่งเขตภาษาระหว่างภาษาถิ่นกลางและภาษาถิ่นใต้ที่แนวจังหวัดประจวบคีรีขันธ์ต่อกับจังหวัดชุมพร

กลุ่มที่ 2 ในแผนที่ผลลัพธ์ของหน่วยอรรถ “พูด” และ “ฉลาด” มีเส้นแบ่งเขตภาษาระหว่างภาษาถิ่นกลางและภาษาถิ่นใต้ที่แนวจังหวัดระนองต่อกับจังหวัดสุราษฎร์ธานี

ความหมายคำ	ความชัดเจนในการกำหนดเส้นแบ่งเขตภาษา			
	ภาษาไทยถิ่น กลาง	ภาษาไทยถิ่น เหนือ	ภาษาไทยถิ่น อีสาน	ภาษาไทยถิ่น ใต้
1. “ภูเขา”	×	✓	×	×
2. “กะลามะพร้าว”	✓	✓	✓	✓
3. “พูด”	✓	×	✓	×
4. “ดู”	×	✓	✓	✓
5. “รีบ”	✓	✓	✓	✓
6. “เรอ”	✓	✓	✓	✓
7. “ฉลาด”	✓	✓	✓	✓

ตารางที่ 4 : ความชัดเจนในการกำหนดเส้นแบ่งเขตภาษาของภูมิภาคภาษาถิ่นของคำศัพท์ทั้ง 7 คำ

ประเด็นที่ 3 : การปะปนของการใช้ภาษาอื่นในบริเวณภาคต่าง ๆ ในประเทศไทย

“ภาษาอื่น” ในงานวิจัยครั้งนี้ หมายถึงภาษาถิ่นที่ใช้พูดกันเฉพาะในท้องถิ่นของตน หรือ ภาษาของชนกลุ่มน้อย ภาษาเหล่านี้ใช้คำศัพท์ที่แตกต่างจากคำศัพท์หลักของภาษาถิ่น 4 ภาษาที่กำหนดไว้ในตารางที่ 2 เมื่อพิจารณาผลการวิจัยพบว่าพื้นที่ที่ปรากฏการใช้ “ภาษาอื่น” มีความสอดคล้องกันในเชิงตำแหน่งที่ตั้ง (location) และ รูปแบบ (pattern) การกระจุกตัว (cluster) ในบริเวณต่าง ๆ (ดูภาพที่ 8) สรุปได้ดังนี้

- บริเวณภาคอีสานตอนล่าง (บางส่วนของจังหวัดบุรีรัมย์ จังหวัดสุรินทร์ และ จังหวัดศรีสะเกษ) ส่วนที่ติดกับประเทศกัมพูชา นอกจากนี้ ในบางหน่วยอรรถแผนที่แสดงให้เห็นการใช้ “ภาษาอื่น” รุกเข้ามาถึงอีสานตอนใน
- บริเวณภาคใต้ตอนล่าง (จังหวัดยะลา และ จังหวัดปัตตานี) ส่วนที่ติดกับประเทศมาเลเซีย

⁷ ภาษาอื่น ในแผนที่ผลลัพธ์ ใช้คำอธิบายสัญลักษณ์ว่า “อื่น ๆ”

ประเด็นที่ 4 : ความสอดคล้องของแนวแบ่งเขตภูมิภาคทางภูมิศาสตร์กับเส้นแบ่งเขตภาษาของภูมิภาคทางภาษาถิ่น^๘

ภูมิภาคทางภูมิศาสตร์ แบ่งประเทศไทยออกเป็นภูมิภาคใหญ่ ๆ 6 ภูมิภาค ตามลักษณะภูมิประเทศ ได้แก่ ภาคเหนือ ภาคกลาง ภาคตะวันตก ภาคอีสาน ภาคใต้ และ ภาคตะวันออก (ดูภาพที่ 2) ในขณะที่ภูมิภาคทางภาษาถิ่นไทยประกอบด้วย 4 ภาค คือ ภูมิภาคภาษาไทยถิ่นเหนือ ภาษาไทยถิ่นกลาง ภาษาไทยถิ่นอีสาน และภาษาไทยถิ่นใต้

เมื่อนำแผนที่แสดงแนวแบ่งเขตภูมิภาคทางภูมิศาสตร์มาซ้อนทับกับแผนที่แสดงเส้นแบ่งเขตภาษาจากข้อมูลคำศัพท์ภาษาของแต่ละหน่วยย่อย โดยพิจารณาในภาพรวม (ดูภาพที่ 9) มีข้อสังเกตดังนี้

1. ภาคเหนือ ภาคอีสานและภาคใต้มีความสอดคล้องของแนวแบ่งเขตภูมิภาคทางภูมิศาสตร์กับเส้นแบ่งเขตภาษาของภูมิภาคทางภาษาถิ่น ยกเว้นในบริเวณ

- พื้นที่ตอนบนและตอนกลางของจังหวัดตาก^๙ ใช้คำศัพท์ภาษาถิ่นเหนือ
- พื้นที่ส่วนใหญ่ของจังหวัดนครราชสีมา^๙ ใช้คำศัพท์ภาษาถิ่นกลาง

2. ภาษาถิ่นกลางใช้กันทั่วไปในภาคกลาง ภาคตะวันตก และภาคตะวันออกของประเทศ อย่างไรก็ตาม พื้นที่แถบตะวันออกของภาคตะวันออกในจังหวัดสระแก้วและปราจีนบุรีใช้ภาษาถิ่นอีสานหรือภาษาอื่น

7. ปัญหาและอุปสรรค

1. ปัญหาในการเก็บรวบรวม นำเข้าข้อมูลแบบสอบถาม และวิเคราะห์คำศัพท์ในโครงการเก็บรวบรวมข้อมูลคำศัพท์ภาษาพื้นบ้าน

เนื่องจากโครงการเก็บรวบรวมข้อมูลคำศัพท์ภาษาพื้นบ้านเป็นโครงการขนาดใหญ่ ครอบคลุมการเก็บรวบรวมและนำเข้าข้อมูลในระดับตำบลทั่วประเทศ ผ่านทางเครือข่ายวัฒนธรรม จากสภาวัฒนธรรมประจำจังหวัด ผ่านยังอำเภอ จนถึง ระดับตำบล ปัญหาข้อแรก คือ โครงการใช้ระยะเวลานาน ในการส่งแบบสอบถามและได้รับข้อมูลตอบกลับจากระดับตำบลทั่วประเทศ พื้นที่บางแห่งไม่ได้รับข้อมูลกลับเนื่องจากอยู่ในพื้นที่ห่างไกลและไม่สามารถเข้าถึง โดยเฉพาะตำบลที่อยู่บริเวณตะเข็บรอยต่อระหว่างประเทศ

⁸ ตามลักษณะภูมิภาคทางภูมิศาสตร์ จังหวัดตากจัดอยู่ในภาคตะวันตก

⁹ ตามภูมิภาคทางภูมิศาสตร์ จังหวัดนครราชสีมาจัดอยู่ในภาคอีสาน แต่ภาษาไทยถิ่นโคราชมีลักษณะพิเศษคือเป็นภาษาลูกผสม (hybrid language) ระหว่างภาษาไทยถิ่นกลางกับภาษาไทยถิ่นอีสาน คำศัพท์จำนวนมากในภาษาไทยถิ่นโคราชใช้คำศัพท์ในภาษาไทยถิ่นกลาง ส่วนระบบเสียง โดยเฉพาะอย่างยิ่งระบบวรรณยุกต์ มีลักษณะผสมระหว่างภาษาไทยถิ่น 2 ภาษาดังกล่าว

ปัญหาอีกประการหนึ่งคือ แบบสอบถามที่ตอบกลับมามีบางส่วนไม่สามารถสะท้อนถึงคำศัพท์ภาษาถิ่นที่ใช้ในบริเวณนั้นได้จริง ทั้งนี้เนื่องจาก

- การไม่สามารถประเมินผู้ตอบแบบสอบถามได้ว่าเป็นคนในพื้นที่นั้นจริง ๆ
- ผู้ตอบแบบสอบถามบางส่วนไม่กรอกประวัติส่วนตัวและที่อยู่ของพื้นที่ที่กรอกข้อมูลทำให้ไม่สามารถลงรหัสตำบลเชื่อมโยงกับแผนที่ได้
- เสียงของคำเป็นปัญหาหลักที่ทำให้ผู้ตอบแบบสอบถามไม่แน่ใจว่าควรใช้ตัวเลือกใด เป็นผลให้ต้องใช้เวลานานในการจัดหมวดหมู่และกำหนดรหัสข้อมูล ยกตัวอย่างเช่น

คำศัพท์คำว่า “พูด” มีตัวเลือก ก.พูด ข.อู๋ ค.เว้า ง.แล้ง จ.....

หากออกเสียงว่า “แหลง” ผู้ตอบแบบสอบถามอาจตอบแบบสอบถาม ดังนี้

คำศัพท์คำว่า “พูด” มีตัวเลือก ก.พูด ข.อู๋ ค.เว้า ☒ง.แล้ง จ.....

ซึ่งเป็นการตอบตามการเขียน หรือ

คำศัพท์คำว่า “พูด” มีตัวเลือก ก.พูด ข.อู๋ ค.เว้า ง.แล้ง ☒จ.แหลง..

ซึ่งเป็นการตอบตามเสียงพูด

นอกเหนือจากปัญหาเกี่ยวกับการเก็บข้อมูลแล้ว ปัญหาหลักอีกสองประการหนึ่งได้แก่ ปัญหาในการวิเคราะห์คำศัพท์และปัญหาในการลงรหัสข้อมูล ปัญหาแรก สืบเนื่องมาจาก ผู้ตอบแบบสอบถามในบางตำบลตอบตัวเลือกมากกว่าหนึ่งคำศัพท์ ทำให้ต้องใช้ระยะเวลาในการวิเคราะห์และตัดสินใจกำหนดรหัสในฐานข้อมูล ส่วนปัญหาที่สองเป็นปัญหาในเรื่องความผิดพลาดในการลงรหัสข้อมูลผิดของผู้นำเข้าสู่ข้อมูลในระบบคอมพิวเตอร์ ดังนั้นเมื่อทำการแสดง (plot) ข้อมูลคำศัพท์เหล่านั้นในทางพื้นที่ ข้อมูลเหล่านี้จึงมีลักษณะผิดปรกติมากในแผนที่ เช่น พบภาษาถิ่นเหนือในบริเวณพื้นที่ภาคใต้ เป็นต้น ทำให้พื้นที่บริเวณดังกล่าวต้องทำการตรวจสอบการลงรหัสข้อมูลหรือในกรณีที่ข้อมูลผิดจริงต้องส่งแบบสอบถามไปให้กรอกใหม่อีกครั้ง

2. ปัญหาในการเชื่อมโยงข้อมูล

เนื่องจากข้อมูลจำนวนตำบลและอำเภอของข้อมูลแผนที่แสดงเขตการปกครองในระดับตำบลของสถาบันสิ่งแวดล้อมไทยต่างจากจำนวนตำบลและอำเภอของกระทรวงมหาดไทยที่ใช้ในการบันทึกข้อมูล นอกจากนี้ รหัสตำบลและอำเภอของ Thailand on Disc ในบางบริเวณไม่ตรงกับรหัสตำบลและอำเภอของกระทรวงมหาดไทย ทั้งนี้เป็นผลสืบเนื่องมาจากความไม่สอดคล้องของข้อมูลทั้งสองแหล่ง ทำให้การเชื่อมโยงข้อมูลที่ได้จากแบบสอบถามและจากแผนที่แสดงเขตการปกครองใช้เวลานานมากในการตรวจสอบ การเชื่อมโยงข้อมูลทำได้เฉพาะบริเวณที่ตำบลจากทั้งสองแหล่งข้อมูลตรงกันเท่านั้น

ปัญหาดังกล่าวทำให้ไม่สามารถแสดงภาพการกระจายของข้อมูลคำศัพท์ในแผนที่ได้อย่างครบถ้วนโดยในแผนที่จะเห็นบางบริเวณเป็นสีขาว แสดงคำอธิบายสัญลักษณ์ “ไม่มีข้อมูล”

8. ข้อสรุปและงานวิจัยในอนาคต

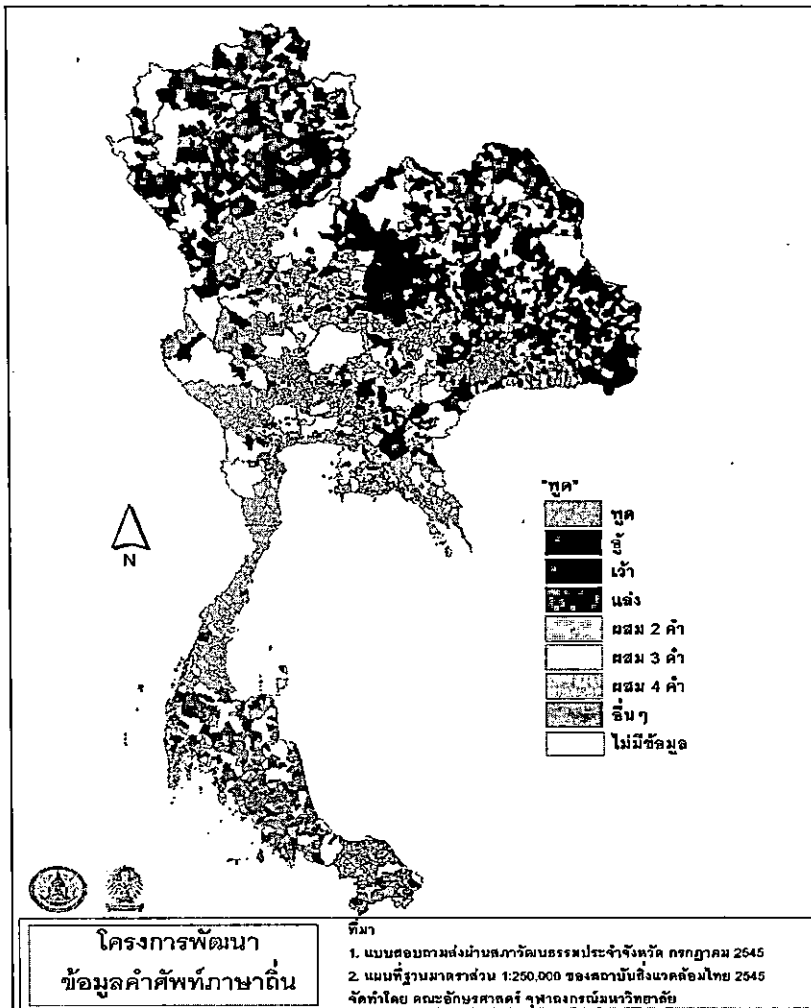
งานวิจัยข้อมูลคำศัพท์ภาษาไทยถิ่นเก่าที่ผ่านมา มักสร้างและแสดงข้อมูลคำศัพท์ในเชิงพื้นที่ด้วยมือในรูปของแผนที่แผ่นกระดาษอย่างหนา มีกำหนดจุดแสดงตำแหน่งข้อมูลอย่างคร่าว ๆ ยากต่อการแสดงเส้นหรือแนวแบ่งเขตภาษาไทยถิ่นที่ชัดเจน แผนที่ผลลัพธ์ขาดความน่าเชื่อถือ ยากต่อการตรวจสอบความถูกต้องและการปรับปรุงแก้ไขข้อมูลในอนาคต ในงานวิจัยชิ้นนี้ ได้มีการนำระบบสารสนเทศภูมิศาสตร์มาใช้ในการจัดเก็บและแสดงผลการจัดทำแผนที่ข้อมูลคำศัพท์ภาษาไทยถิ่น นับเป็นงานวิจัยนำร่องในการพัฒนาฐานข้อมูลภูมิศาสตร์คำศัพท์ภาษาไทยถิ่นสำหรับประเทศไทยซึ่งเก็บและแสดงผลข้อมูลครอบคลุมพื้นที่ในระดับตำบลทั่วประเทศ ผลงานวิจัยชี้ให้เห็นศักยภาพของการนำข้อมูลในฐานข้อมูลภูมิศาสตร์คำศัพท์มาใช้ในการสร้างแผนที่การกระจายของข้อมูลคำศัพท์และแนวทางในการสร้างแนวแบ่งเขตภูมิภาคทางภาษาถิ่นอย่างมีระบบ และเป็นประโยชน์อย่างยิ่งสำหรับนักภาษาศาสตร์ทั้งในเชิงการแสดงผลการกระจายของภาษาถิ่น และการเพิ่มประสิทธิภาพในการตรวจสอบความถูกต้องของข้อมูลภาษาถิ่นจากแผนที่ ทั้งนี้ผลประโยชน์ความสำเร็จของโครงการ ปัญหาและอุปสรรคในการดำเนินงานวิจัย จะเป็นประโยชน์อย่างยิ่งในการปรับปรุงฐานข้อมูลและแผนที่ข้อมูลภูมิศาสตร์คำศัพท์ภาษาไทยถิ่นทั้ง 170 หน่วยย่อยให้สำเร็จต่อไป นอกจากนี้ ในงานวิจัยในอนาคต ฐานข้อมูลภูมิศาสตร์คำศัพท์ภาษาถิ่นที่จัดทำขึ้น สามารถนำไปใช้ศึกษาและวิเคราะห์ร่วมกับปัจจัยและตัวแปรด้านต่าง ๆ ที่ก่อให้เกิดความแตกต่างของภาษาไทยถิ่น เช่น ปัจจัยลักษณะภูมิประเทศและตัวแปรทางสังคม อาทิเช่น อายุ ระดับการศึกษา เพศ ทัศนคติ

กิตติกรรมประกาศ

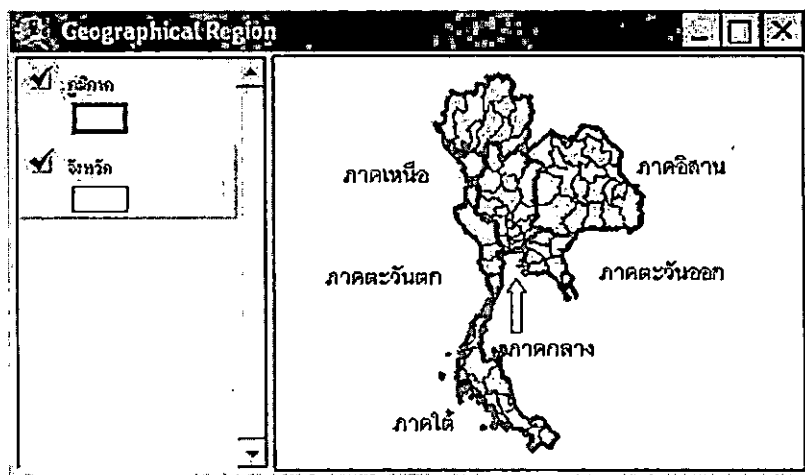
ผู้วิจัยขอขอบคุณสถาบันสิ่งแวดล้อมไทยที่ให้ความอนุเคราะห์ข้อมูลแผนที่ฐานประเทศไทย

บรรณานุกรม

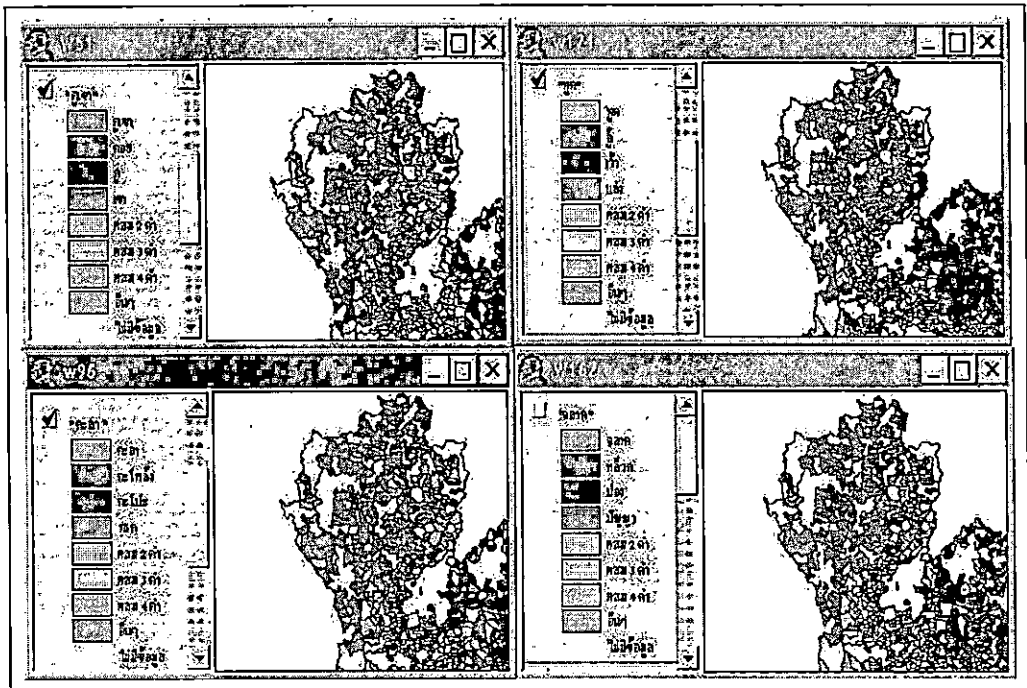
- จรูญ บุญพันธ์ และ วันทนี พันธ์ชาติ. 2526. “การศึกษาการกระจายของคำศัพท์ในประเทศไทย: โครงการทดลอง.” ศาสตรแห่งภาษา จ.3 (มิถุนายน 2526). 27-61.
- เจริญขวัญ ธรรมประดิษฐ์. 2524. การใช้ลักษณะทางสัทศาสตร์ของสระสูงในการแบ่งเขตภาษาถิ่นในจังหวัดตรัง กระบี่ พังงา และภูเก็ต. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต ภาควิชาภาษาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- ชลิตา โรจนวัฒนวุฒิ. 2533. ภูมิศาสตร์คำศัพท์ภาษาไทย 4 ถิ่น: การศึกษาเชิงจำกัด. จุฬาลงกรณ์มหาวิทยาลัย. อัดสำเนาเย็บเล่ม.
- ฉลิลิกา มหาพูนทอง. 2539. การกำหนดแนวแบ่งเขตภาษาไทยถิ่นใต้ตะวันออกกับภาษาไทยถิ่นใต้ตะวันตกโดยใช้พยางค์ที่มีสระเสียงยาวกับพยัญชนะท้าย /x/ หรือ /ʔ/. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต ภาควิชาภาษาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
- ทักษิณคดีศึกษา, สถาบัน มหาวิทยาลัยศรีนครินทรวิโรฒ. 2530. พจนานุกรมภาษาถิ่นใต้ พุทธศักราช 2525. พิมพ์ครั้งที่ 3. กรุงเทพมหานคร: กรุงเทพมหานครพิมพ์.
- ภาษาศาสตร์, ภาควิชา. มป. โครงการวิจัยภาษาไทยถิ่นอีสาน. (ไม่ได้ตีพิมพ์).
- วรรณพร ทองมาก. 2526. แนวแบ่งเขตภาษาไทยถิ่นกลางกับภาษาไทยถิ่นใต้โดยใช้ศัพท์เป็นเกณฑ์. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต ภาควิชาภาษาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- วันทนี พันธ์ชาติ. 2529. ภาษาถิ่นย่อยของคำเมือง: การศึกษาศัพท์. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต ภาควิชาภาษาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- วิจิตร ภาณุพงศ์. 2529. รายงานการวิจัยภูมิศาสตร์คำศัพท์ในจังหวัดนครราชสีมา: โครงการนำร่อง. กรุงเทพฯ: โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.
- สมทรง บุรุษพัฒน์. 2535. ภูมิศาสตร์คำศัพท์ภาษาไทยถิ่นในจังหวัดเพชรบูรณ์. สถาบันวิจัยภาษาและวัฒนธรรมเพื่อพัฒนาชนบท มหาวิทยาลัยมหิดล.
- สมทรง บุรุษพัฒน์. 2543. ภูมิศาสตร์ภาษาถิ่น. สถาบันวิจัยภาษาและวัฒนธรรมเพื่อพัฒนาชนบท มหาวิทยาลัยมหิดล.
- สุนีย์พร เลิศกุลทนนท์. 2534. ภูมิศาสตร์คำศัพท์ภาษาไทยถิ่นใต้จังหวัดพังงาและกระบี่. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต ภาควิชาภาษาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- Tingsabadh, M.R. Kalaya. 1985. “Thai Dialectology up to the Year 1984.” Science of Language Papers 5. (August 1985). 75-102.



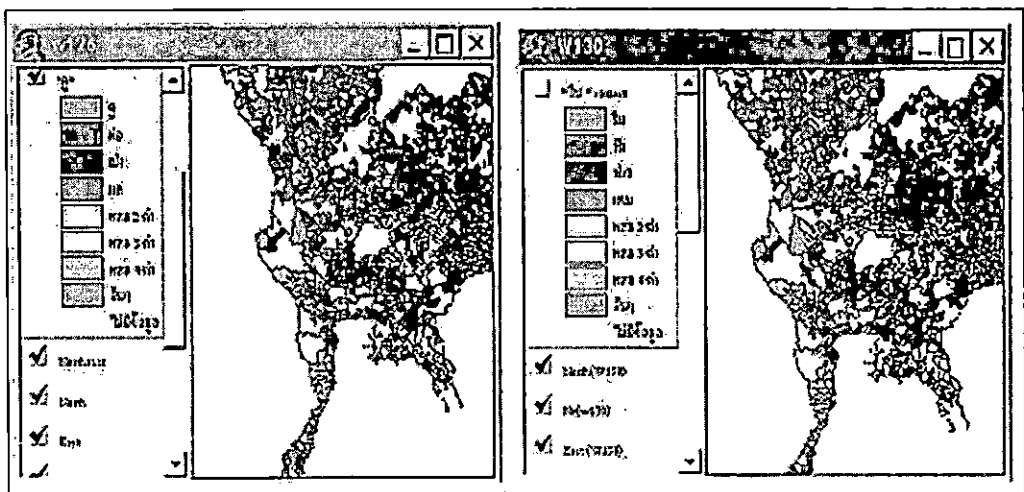
ภาพที่ 2 : ผลลัพธ์โครงการ แผนที่ข้อมูลคำศัพท์ภาษาไทยถิ่นของหน่วยบรรด “พูด”



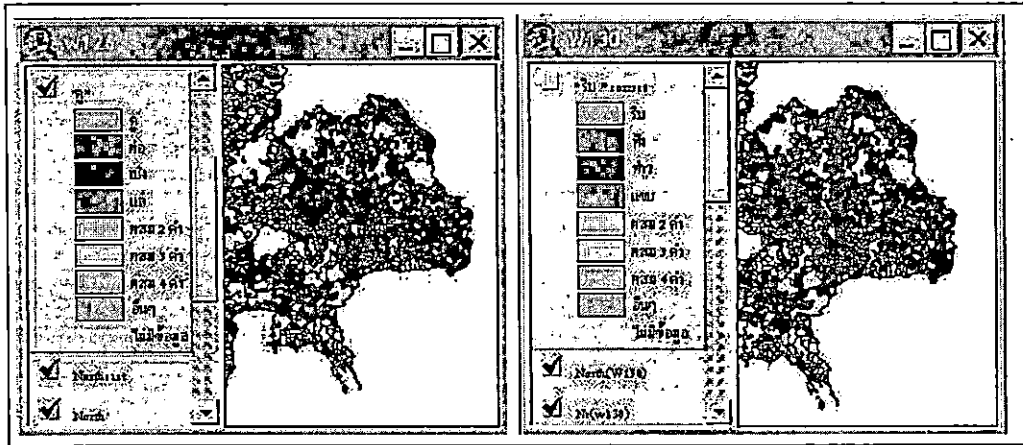
ภาพที่ 3 : แผนที่ประเทศไทย 76 จังหวัด



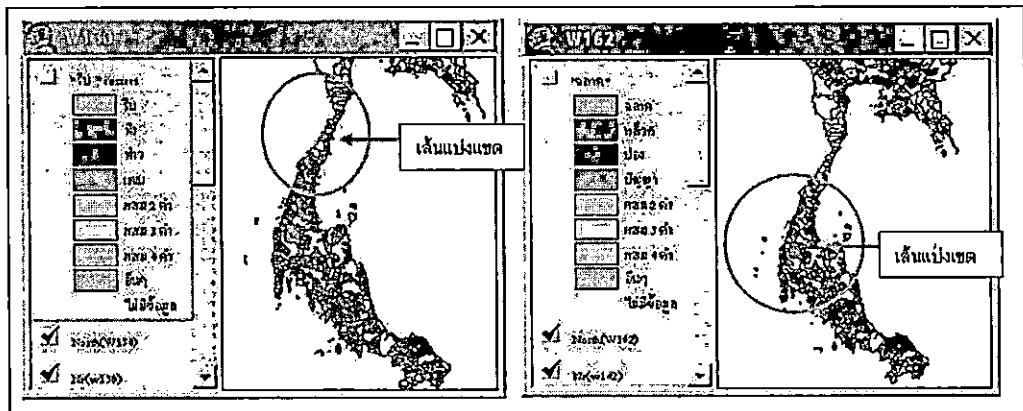
ภาพที่ 4 : แผนที่ข้อมูลคำศัพท์ภาษาถิ่นเหนือ แสดงความคล้ายคลึงของเส้นแบ่งเขตภาษาของภาษาถิ่นเหนือ โดยเปรียบเทียบระหว่างหน่วยย่อย “ภูเขา” (บนซ้าย) “กะลามาพร้าว” (ล่างซ้าย) “พุด” (บนขวา) และ “ฉลาด” (ล่างขวา)



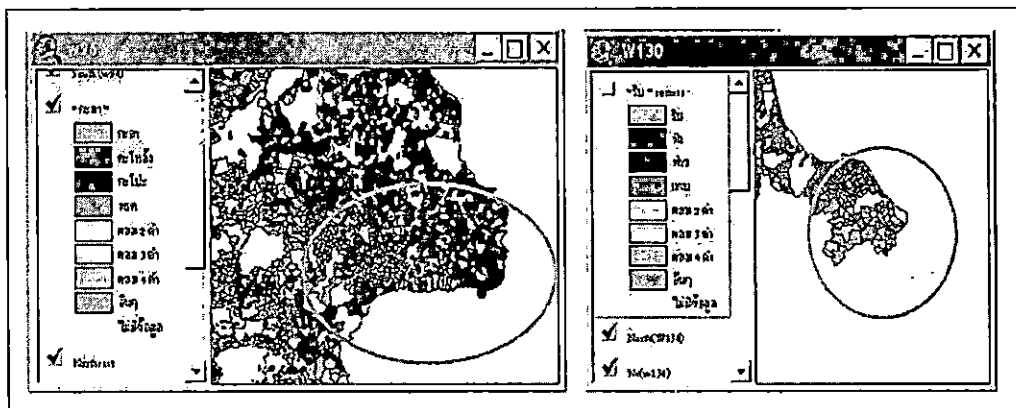
ภาพที่ 5 : แผนที่ข้อมูลคำศัพท์ภาษาถิ่นกลาง แสดงความคล้ายคลึงของเส้นแบ่งเขตภาษาของภาษาถิ่นกลาง โดยเปรียบเทียบระหว่างหน่วยย่อย “ตู” (ซ้าย) และ “รีบ” (ขวา)



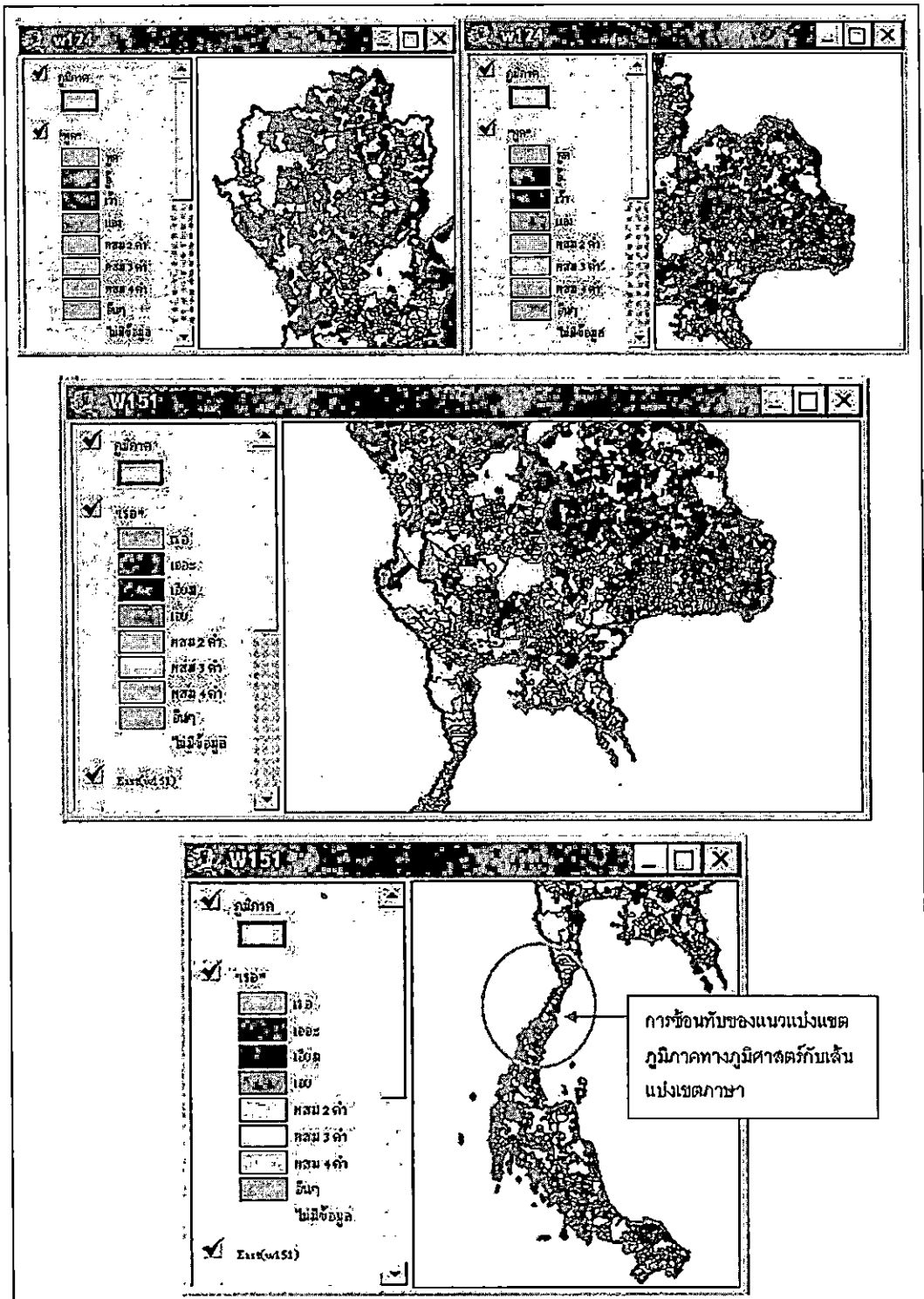
ภาพที่ 6 : แผนที่ข้อมูลคำศัพท์ภาษาถิ่นอีสาน แสดงความคล้ายคลึงของเส้นแบ่งเขตภาษาของภาษาถิ่นอีสาน โดยเปรียบเทียบระหว่างหน่วยวรรด “ดู” (ซ้าย) และ “ริบ” (ขวา)



ภาพที่ 7 : แผนที่ข้อมูลคำศัพท์ภาษาถิ่นใต้ แสดงความแตกต่างของเส้นแบ่งเขตภาษาระหว่างภาษาถิ่นกลางและภาษาถิ่นใต้ โดยเปรียบเทียบระหว่างหน่วยวรรด “ริบ” (ซ้าย) และ “ลาด” (ขวา)



ภาพที่ 8 : แผนที่ข้อมูลคำศัพท์ภาษาถิ่นแสดงสอดคล้องของรูปแบบการเกาะกลุ่มและตำแหน่งที่ตั้งที่มีการใช้ภาษาในบริเวณภาคอีสานตอนล่าง (บน) บริเวณภาคใต้ตอนล่าง (ล่างซ้าย) บริเวณตำบลที่ติดอยู่กับชายแดนด้านตะวันตกของจังหวัดกาญจนบุรี (ขวา)



ภาพที่ 9 : แผนที่แสดงแนวแบ่งเขตภูมิภาคทางภูมิศาสตร์ซ้อนทับกับแผนที่คำศัพท์ภูมิภาคทางภาษาถิ่น บริเวณภาคเหนือ (บนซ้าย) ภาคอีสาน (บนขวา) ภาคกลาง - ภาคตะวันตก - ภาคตะวันออก (กลาง) และ ภาคใต้ (ล่าง)