

บทที่ 5

สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

การวิจัยเรื่องการพัฒนาแบบทดสอบวัดความสามารถในการใช้ภาษาอังกฤษ เป็นการวิจัยเพื่อพัฒนาแบบทดสอบฯ การรายงานผลในบทนี้มีดังนี้ คือ (5.1) สรุปผลการวิจัย (5.2) อภิปรายผล และ (5.3) ข้อเสนอแนะ

5.1 สรุปผลการวิจัย

วัตถุประสงค์ของการวิจัย โครงการวิจัยและพัฒนาแบบทดสอบวัดความสามารถในการอ่านภาษาอังกฤษออนไลน์ มีดังนี้

- (1) เพื่อพัฒนาแบบทดสอบวัดความสามารถในการอ่านภาษาอังกฤษออนไลน์
- (2) เพื่อตรวจสอบคุณภาพของแบบทดสอบวัดความสามารถในการอ่านภาษาอังกฤษออนไลน์
- (3) เพื่อศึกษาระดับความสามารถทางการอ่านภาษาอังกฤษของผู้ใช้บริการแบบทดสอบ
- (4) เพื่อวิเคราะห์ปัญหาหรือจุดอ่อนในการอ่านภาษาอังกฤษของผู้ใช้บริการแบบทดสอบ
- (5) เพื่อเสนอแนวทางการพัฒนาการอ่านภาษาอังกฤษของตนเองของผู้ใช้บริการแบบทดสอบ

การวิจัยและพัฒนาแบบทดสอบวัดความสามารถในการอ่านภาษาอังกฤษออนไลน์ มีผลที่ได้คือแบบทดสอบฯ จำนวน 3 ชุด ชุดละ 100 ข้อ แต่ละชุดแบ่งออกเป็นสามส่วนคือ ไวยากรณ์ คำศัพท์ และบทอ่าน นำเสนอออนไลน์ ณ เว็บไซต์ของมหาวิทยาลัยสุโขทัยธรรมาราช ผู้เข้าทดลองใช้แบบทดสอบฯ มีอายุต่ำกว่า 35 ปีมากที่สุด เป็นหญิงมากกว่าชาย จบปริญญาตรี และโท และกำลังศึกษาระดับปริญญาโท ร้อยละ 51.6 ความสามารถภาษาอังกฤษจากการประเมินตนเอง พอใช้มากที่สุด ระดับดี ร้อยละ 25.8 ผู้ทดลองใช้แบบทดสอบมีความพึงพอใจต่อแบบทดสอบในภาพรวมระดับมาก จำนวน 17 คน มากที่สุด คิดเป็นร้อยละ 54.80

การพัฒนาแบบทดสอบฯ มีการออกแบบแบบทดสอบที่อยู่บนพื้นฐานกรอบแนวคิดของกระบวนการอ่านและกระบวนการวัดผลความสามารถในการอ่านของ Khalifa & Weir (2009: 62) Bachman and Palmer (1996) เกณฑ์วัดระดับความสามารถของ CEFR การกำหนด test specifications ในส่วนไวยากรณ์กำหนดตาม รายละเอียดของ Grammar City & Guilds ESOL (www.cityandguilds.com) ซึ่งสามารถเทียบเคียงกับระดับของ CEFR คำศัพท์มาจากรายการคำศัพท์ทางวิชาการ (AWL list) รายละเอียดของบทอ่านระดับ B 1 B2 และ C 1 เป็นไปตามรายละเอียดของ

CEFR และ Unaldi (2009) โดยใช้การนำเสนอออนไลน์ที่เรียบง่าย และสามารถแสดงผลคะแนนรวมพร้อมคำอธิบายได้ มีกระบวนการและขั้นตอนที่ประกันคุณภาพของแบบทดสอบรายฉบับและรายข้อ มีการประเมินคุณภาพมี ความคิดเห็นและข้อเสนอแนะของผู้ทรงคุณวุฒิด้านภาษาและการวัดและประเมินผล มีการทดลองใช้ 2 ครั้ง มีการวิเคราะห์ข้อมูลทางสถิติรายข้อ และมีการปรับปรุงแก้ไขข้อสอบตามข้อมูลทางสถิติและข้อเสนอแนะของผู้ทรงคุณวุฒิและผู้ทดลองใช้แบบทดสอบ

5.2 อภิปรายผลการวิจัย

การอภิปรายผลการวิจัยนำเสนอตามวัตถุประสงค์ของการวิจัยดังต่อไปนี้

2.1 การพัฒนาแบบทดสอบวัดความสามารถในการอ่านภาษาอังกฤษออนไลน์

การอภิปรายผลเกี่ยวกับการพัฒนาแบบทดสอบฯ มีประเด็นที่สำคัญดังนี้

2.1.1 การออกแบบแบบทดสอบ: แนวคิด เกณฑ์กำหนดความสามารถและรายละเอียด

การพัฒนาแบบทดสอบฯ ได้จัดทำขึ้น โดยอิงกรอบแนวคิดที่ชัดเจนในเรื่องของทฤษฎีการอ่าน (Khalifa & Weir, 2009: 62) Bachman & Palmer (1996: 133-156) พิจารณาประโยชน์ของแบบทดสอบ ความหมายของความสามารถในการอ่าน (Alderson, 2000:169) กำหนดระดับความสามารถตามเกณฑ์มาตรฐานของ CEFR (www.uk.cambridge.org/elt p 69) กำหนดรายละเอียดแบบทดสอบรายข้อที่ชัดเจนด้านไวยากรณ์ คำศัพท์ บทอ่าน พัฒนาจำนวน 3 ชุด คู่ขนานกันเพื่อที่จะได้ทดสอบหลายครั้ง ในแบบทดสอบแต่ละชุดได้วัดความสามารถในการอ่าน (1) ระดับคำ (2) ระดับประโยค (3) ระดับข้อความต่อเนื่อง ดังรายละเอียดต่อไปนี้

การสร้างแบบทดสอบฯ ในส่วนของไวยากรณ์ (Grammar) ได้กำหนด test specifications โดยเลือกประเด็นทางไวยากรณ์ตามระดับ B1 B2 C1 โดยใช้ ตาราง ประเด็นทางไวยากรณ์ ได้เลือกจากเกณฑ์ของ Grammar City & Guilds ESOL 3 ระดับคือ B1 (Achiever) B2 (Communicator) C1 (Expert) (ระดับละ 10 ข้อ รวม 30 ข้อ) (www.cityandguilds.com) ในแบบทดสอบฯ แต่ละชุดจะมีลักษณะคู่ขนานกัน

การสร้างแบบทดสอบฯ ในส่วนของคำศัพท์ (vocabulary) คำศัพท์ที่เลือกมาทดสอบมาจากรายการคำศัพท์ทางวิชาการ (AWL) รวบรวมโดย Coxhead ณ มหาวิทยาลัย Victoria University of Wellington ประเทศนิวซีแลนด์ ซึ่งเป็นที่ยอมรับในวงวิชาการ การเลือกใช้วัดของผู้วิจัยซึ่งมีประสบการณ์ในการอ่านเรื่องทั่วไปบทความ รายงานทางวิชาการ

การสร้างแบบทดสอบฯ ในส่วนของบทอ่าน (Reading) ได้กำหนด test specifications ตามระดับ B1 B2 C1 โดยได้การกำหนดพื้นฐานความรู้ระดับคำและประโยคโดยเน้นที่เรื่องการใช้คำศัพท์ในประโยค การใช้ประเด็นไวยากรณ์ให้ถูกต้องในประโยค เป็นพื้นฐานให้สามารถบอกความหมายของบทอ่านด้วยวิธี Bottom up จากล่างสู่บนได้

ในระดับคำ การที่มีกรอบแนวคิดที่ชัดเจนในการพัฒนาความสามารถในการอ่าน และการกำหนดรายละเอียดของข้อสอบอย่างชัดเจน (Khalifa & Weir, 2009: 62) พิจารณาจากระดับล่างสุดคือ visual input ระดับคำ lexical form การรู้จักคำในระดับ orthography phonology morphology meaning word class การเข้าถึงคำศัพท์ ประเภทของคำ (word class)

ในระดับประโยค ประกอบด้วย การแจงองค์ประกอบของประโยค (syntactic parsing) การสร้างความหมายระดับประโยคย่อย และระดับประโยค การอนุมาน การสร้างภาพในใจคือบูรณาการข้อมูลใหม่เข้ากับสิ่งที่ได้อ่าน ซึ่งจุดนี้ ต้องการความรู้ทั่วไปในโลกประสบการณ์ผู้อ่าน ความรู้เกี่ยวกับหัวข้อเรื่องที่อ่าน

ส่วนการสร้าง ความหมายระดับข้อความต่อเนื่อง และระหว่างข้อความต่อเนื่อง ต้องการความรู้เรื่องโครงสร้างของบทอ่าน ประเภท และ rhetorical tasks คือลักษณะของคำถาม รวมถึงวิธีการอ่าน ระดับภาพรวม การอ่านเพื่อใจความหลัก และรายละเอียด การตั้งเป้าหมายการอ่าน การตรวจสอบเป้าหมาย การปรับแก้ ผู้วิจัยได้ออกแบบแบบทดสอบฯ ตามแนวทางนี้ โดยเฉพาะอย่างยิ่งการวัดระดับความยากง่ายของบทอ่านที่รอบด้านของ (Unaldi 2009) ในด้าน บริบท (การตอบ ความยาวของบทอ่าน) discourse mode (ประเภทของบทอ่าน กิจกรรม รูปแบบการเรียงความ ความชัดเจนของบทอ่าน) structural resources (การใช้คำ จำนวนคำต่อประโยค ค่าเฉลี่ยระดับ Flesch-Kincaid Grade level ความซับซ้อนของโครงสร้างประโยค, ความเชื่อมโยงในบทอ่าน (cohesion), แหล่งของคำศัพท์ (K1,2,3, AWL), ธรรมชาติของข้อมูล, ความรู้เกี่ยวกับเนื้อหา ด้านพุทธิพิสัย (ประเภทของการอ่าน ระดับของบทอ่าน --ดูรายละเอียดที่ภาคผนวก 3.7 และ 3.8) รวมถึงการใช้โปรแกรม lexitutor เพื่อตรวจสอบระดับความยากง่ายของคำศัพท์ (“VocabProfile Home” Retrieved from <http://www.lexitutor.ca/vp/> April 20, 2010)

2.1.2 ผลการวิเคราะห์ห้บทอ่านที่สร้างขึ้นในกรอบของ test specifications

ในส่วนบทอ่านซึ่งมีขั้นตอนและกระบวนการที่ซับซ้อน กำหนด test specifications ตามรายละเอียดหัวข้อต่าง ๆ ทำให้สามารถได้บทอ่านที่มีระดับความยากง่ายตามระดับที่กำหนด คือ B1 B2 และ C1 และอยู่ในขอบข่ายของโครงร่างรายละเอียดที่กำหนดไว้ ในระดับต่างๆ ตามที่ได้แสดงผลในตารางที่ 4.5 4.6 และ 4.7 ในบทที่ 4 สามารถสรุปได้ว่าการกำหนดเกณฑ์เป็นแนวทางช่วยให้หาบทอ่านที่มีความเหมาะสมได้ใกล้เคียงเกณฑ์ที่สุดโดยมีทฤษฎีการอ่านรองรับ แม้จะไม่สามารถปรับบทอ่านให้ได้ตรงตามทุกประเด็นก็ตามหากอยู่ในกรอบแนวคิดและเกณฑ์ก็จะทำให้การกำหนดบทอ่านมีเป้าหมาย และสามารถวัดระดับความสามารถได้ตามที่ต้องการ โดยจะอภิปรายรายละเอียดแต่ละระดับดังนี้

การเลือกหัวข้อ (topic) ของบทอ่าน นอกจากภายในกรอบของ test specifications ในประเด็นเรื่องความรู้ทางเนื้อหาแล้ว ผู้วิจัยเลือกจากเรื่องที่ใช้แบบทดสอบน่าจะสนใจ โดยที่กลุ่มผู้ที่คาดว่าจะโดยที่กลุ่มผู้ที่คาดว่าจะใช้แบบทดสอบคือนักศึกษาระดับบัณฑิตศึกษาหัวข้อเรื่องจึงอยู่ในแวดวง

วิชาการ

ระดับ B 1 ส่วนที่กำหนดเหมือนกันคือ response methods การทำตอบ เช่น multiple choice multiple matching จำนวนคำแตกต่างกันบ้างภายใน 500 คำ ยกเว้นบทอ่านในแบบทดสอบฯ ชุดที่ 3 ที่มีจำนวนคำ 758 คำ หากแต่เป็นการเล่าเรื่องราวส่วนบุคคล มีความต่อเนื่องทางเนื้อหาที่ต้องคงไว้ ไม่สามารถตัดออกได้ ประเภทของบทอ่านเป็นความเรียงให้ข้อมูล และเรื่องราวส่วนบุคคล ลักษณะภาษาเป็นการเล่าเรื่อง พรรณนา บอกขั้นตอน รูปแบบของเรียงความเป็นการบรรยาย แบ่งประเภท ลักษณะของโครงสร้างเนื้อหาที่มีความชัดเจนมาก ในส่วนของ structural resources จำนวนคำต่อประโยค 18, 15.7, และ 20.5 เกินที่กำหนด (15) ไปบ้างโดยเฉพาะบทอ่านชุดที่ 3 ซึ่งเป็นการกล่าวถึงเรื่องราวส่วนบุคคล ในส่วนระดับค่าเฉลี่ย Flesch-Kincaid Grade level ซึ่งกำหนดไว้ที่ 8 มีค่าใกล้เคียง บทอ่านชุดที่ 1 ค่าต่างไป (10, 8.6 และ 8.2) ความซับซ้อนของโครงสร้างประโยคทั้งสามชุด ส่วนใหญ่เป็น simple sentences และมีการใช้ subordinate clauses in ระดับ Preliminary English Test) ส่วน ความเชื่อมโยงของบทอ่าน (Cohesion) ทั้งสามชุดมีลักษณะชัดเจน (explicit) แหล่งของคำศัพท์ใกล้เคียงกับที่กำหนดโดยคำศัพท์วิชาการในส่วนนี้กำหนดไว้ 2.5 % ชุดที่ 1 และ 2 เนื่องจากมีเนื้อหาที่เป็นเรื่อง ข้อมูลจึงมีคำศัพท์ทางวิชาการที่วิเคราะห์ห้อออกมาสูงกว่า ชุดที่ 1 AWL 6.32% ชุดที่ 2 AWL 8.51 แต่เป็นศัพท์ที่ใช้ทับศัพท์ในภาษาไทยอยู่แล้ว จึงได้นำมาใช้ ชุดที่ 3 เป็นเรื่องส่วนบุคคลจึงมีศัพท์วิชาการ น้อย 1.81% ลักษณะธรรมชาติของข้อมูลทั้งสามชุดมีลักษณะเป็นรูปธรรม ความรู้เกี่ยวกับเนื้อหา ทั้งสามชุด ไม่จำเป็น ประเด็นทาง cognitive ประเภทของการอ่านทั้งสามชุดใช้ลักษณะเดียวกันคือ main point/multiple matching/relevant point/vocabulary in context/pronoun referents ระดับของบทอ่านทั้งสามชุด เป็นระดับคำ ประโยคและระหว่างประโยค

ระดับ B2 มีบทอ่านชุดละ 2 บท ทั้งสามชุด รูปแบบการตอบ เป็นแบบมีตัวเลือก และ แบบจับคู่ ที่มีตัวเลือกเกิน (multiple matching) เหมือนกัน จำนวนคำอยู่ในเกณฑ์ ไม่เกิน 750 คำ ในส่วนของ ประเภทผลงาน (genre) ทั้ง 3 ชุดเป็นบทความจากหนังสือพิมพ์ ที่มีลักษณะเป็นรายงานวิชาการ แหล่งข้อมูล เป็นงานเขียนแบบความเรียง ลักษณะของบทความ ทั้ง 3 ชุด เป็นการเล่าเรื่อง พรรณนา โครงสร้างของบทอ่าน ทั้ง 3 ชุด มีหลายลักษณะคือ ให้คำจำกัดความ บรรยาย ขยาย ความ วาดภาพ เปรียบเหมือนและเปรียบเทียบ จัดประเภท ปัญหาและการแก้ปัญหาให้เหตุผลหลักการ (justify) เนื่องจากเป็นบทความเกี่ยวกับสภาพแวดล้อม วิทยาศาสตร์ คอมพิวเตอร์ ปัญหาเรื่องการทำงาน มีเนื้อหาที่เกี่ยวข้องกับความสนใจในสังคม ลักษณะภาษาจึงมีหลากหลาย ลักษณะความชัดเจนของโครงสร้างของบทอ่านมีลักษณะ explicit น้อยลง กล่าวคือต้องใช้การอ่านแบบตีความ

เรื่องของ structural resources เฉลี่ยจำนวนคำ ต่อประโยคเกณฑ์กำหนดไม่เกิน 18 บทอ่าน ทั้ง 3 ชุดจะเกินเกณฑ์ไปบ้าง (ชุดที่ 1 25.6, 17.9 ชุดที่ 2 20.6, 24 ชุดที่ 3 22.1, 25.7) Flesch-Kincaid Grade level average กำหนดระดับ 12 ทั้ง 3 ชุดอยู่ในเกณฑ์ที่กำหนด (ชุดที่ 1 12.9, 12.5 ชุดที่ 2 12.4, 11.9 ชุดที่ 3 11.6, 11.6) ความซับซ้อนของโครงสร้างประโยคทั้งสามชุด ส่วนใหญ่มีความ

หลากหลายของประโยค ใช้ compound sentences ส่วน Cohesion (ความเชื่อมโยงของบทอ่าน) ทั้งสามชุดมีความสัมพันธ์ระหว่างส่วนต่าง ๆ ของบทความอาจจะไม่ได้แสดงไว้อย่างชัดเจนมาก ผู้อ่านต้องหาความหมาย แหล่งของคำศัพท์ กำหนดคำศัพท์วิชาการ ที่ 3.3 % ตามที่วิเคราะห์ค่าของศัพท์วิชาการมีสูงมาก (ชุดที่ 1 AWL 6.71% AWL 3.46% ชุดที่ 2 บทอ่านที่ 1 AWL 7.28% บทอ่านที่ 2 K1 AWL 8.62% ชุดที่ 3 บทอ่านที่ 1 AWL 4.01% บทอ่านที่ 2 AWL 4.53%) อย่างไรก็ตามเพราะเป็นแบบทดสอบที่ มุ่งให้นักศึกษาระดับบัณฑิตศึกษาใช้ โดยที่เนื้อหาและองค์ประกอบอื่นเหมาะสมจึงนำมาใช้ ลักษณะธรรมชาติของข้อมูลทั้งสามชุดมีลักษณะเป็นรูปธรรม และนามธรรม ความรู้เกี่ยวกับเนื้อหา ทั้งสามชุด มีลักษณะเฉพาะแต่ผู้อ่านทั่วไปก็สามารถเข้าถึงได้ มีส่วนที่เป็นการอภิปรายเรื่องทางวิชาการ เทคนิค ประเด็นทาง cognitive ประเภทของการอ่านทั้งสามชุดใช้ลักษณะเดียวกันคือการอ่านอย่างรวดเร็วเพื่อหาข้อมูลที่เกี่ยวข้องโดยใช้ความรู้ทางโครงสร้างภาษา การเข้าใจใจความหลักและรายละเอียด ระดับของบทอ่านทั้งสามชุด เป็นระดับ ประโยคและระหว่างประโยค

ระดับ C1 รูปแบบการตอบ เป็น multiple choice และ summary จำนวนคำอยู่ในเกณฑ์ที่กำหนด คือไม่เกิน 750 คำ (679, 738, 651) ในส่วนของ genre (ประเภทผลงาน) เป็นงานเขียนแบบความเรียง ลักษณะภาษาในบทความ ทั้งสามชุดเป็นแบบ descriptive, narrative, expository, argumentative, instructive มีลักษณะโต้แย้งและคำสั่งด้วย โครงสร้างของบทอ่านทั้งสามชุด มีลักษณะหลากหลาย คือ define, describe, elaborate, illustrate, compare and contrast, classify, cause and effect, problem and solution, justify ลักษณะความชัดเจนของโครงสร้างของบทอ่านทั้ง 3 ชุด มีลักษณะไม่ได้แสดงโครงสร้างอย่างชัดเจน

ในส่วน structural resources เฉลี่ยจำนวนคำ ต่อประโยคเกินเกณฑ์บ้าง (เกณฑ์ 19) ชุดที่ 1 30 คำ ชุดที่ 2 23.6 คำ ชุดที่ 3 23.6 คำ Flesch-Kincaid Grade level average เกณฑ์ 12 (ชุดที่ 1 15.7 ชุดที่ 2 15.7 ชุดที่ 3 15.8) ความซับซ้อนของโครงสร้างประโยคทั้งสามชุด เป็นประโยค complex sentences ใช้ modals, ellipsis, range of pronouns and adverbials Cohesion ของทั้งสามชุดมีลักษณะที่แสดงนัยต้องอนุมานจากบทอ่าน แหล่งของคำศัพท์ วิชาการเกณฑ์กำหนดไว้ 4.6% บทอ่านชุดที่ 1 AWL 6.06 % ชุดที่ 2 AWL 17.03% ชุดที่ 3 AWL 15.41% เนื่องจากเป็นบทความจากวารสารทางวิชาการที่ยกมาโดยมิได้ดัดแปลง ลักษณะธรรมชาติของข้อมูลทั้งสามชุดมีลักษณะเป็นนามธรรมบ้าง ความรู้เกี่ยวกับเนื้อหา ทั้งสามชุด อาจต้องเข้าใจความรู้ทางเนื้อหาเฉพาะอาชีพหรือทางวิชาการ

ประเด็นทาง cognitive ประเภทของการอ่านทั้งสามชุดใช้ลักษณะเดียวกันคือวิธีการอ่านอย่างรวดเร็วเพื่อหาข้อมูลที่เกี่ยวข้องโดยใช้ประโยชน์จากลักษณะของบทอ่าน (Search read quickly to locate relevant information using textual feature) เข้าใจใจความหลักและรายละเอียด เข้าใจความสัมพันธ์ระหว่างใจความหลักและรายละเอียด อนุมานความหมายที่ไม่ได้บอกอย่างชัดเจนในบทอ่าน เข้าใจความสัมพันธ์ระหว่างความคิดที่แตกต่างในบทอ่าน ระดับของบทอ่านทั้งสามชุด เป็นระดับ ระหว่าง

ประโยชน์ และระดับระหว่างบทอ่าน

2.1.3 รูปแบบของแบบทดสอบออนไลน์

ในการสร้างแบบทดสอบทดสอบออนไลน์ ได้จัดรูปแบบที่เรียบง่ายเหมาะกับการวัด และประเมินผล ไม่มีส่วนที่ทำให้ไขว้เขว สามารถทบทวนคำตอบได้จนกว่าจะส่ง (submit) มีกรอบ คะแนนแสดงผล การรายงาน คะแนนมีคำอธิบายในรายละเอียด และแนะนำการศึกษาเพิ่มเติม เมื่อพิจารณาเกณฑ์การวัดแบบทดสอบทางคอมพิวเตอร์ของ Fulcher (2003: 384-408) ใน Chapelle and Douglas (2006: 84-85) มีประเด็นต่าง ๆ คือ ลักษณะของฮาร์ดแวร์และซอฟต์แวร์ ระบบนำทาง (Navigation) คำศัพท์ (Terminology) การจัดรูปแบบหน้า (Page layout) ตัวบท (Text) สี (Text color) เครื่องมือ (Tool bars and control) สัญลักษณ์ (Icons and graphics) ความช่วยเหลือ (Help facilities) ประเภทข้อสอบ (Item types) รูปแบบการเขียนคำตอบ (Forms for writing and short answer tasks) การให้ผลย้อนกลับ (Feedback)

ในส่วนของการพัฒนาแบบทดสอบฯ นี้ได้ใช้โปรแกรมสำเร็จรูปซึ่งสามารถเก็บคะแนนผู้เข้าสอบรายบุคคลได้ มีระบบนำทางที่ชัดเจน ง่ายไม่ซับซ้อน การจัดรูปแบบอ่านง่าย ตัวบทอ่านง่าย ไม่มีสีประกอบเนื่องจากเป็นแบบทดสอบฯ ไม่ต้องการการมีปฏิสัมพันธ์กับหน้าจอมาก ต้องการสมาธิในการอ่าน มีข้อจำกัดคือต้องเลื่อนหน้าจอเนื้อหามบทอ่าน ระบบเฉลยโดยสัญลักษณ์เข้าใจยาก จึงต้องระบุคำตอบที่ถูกต้องในส่วนของผลย้อนกลับ อีกครั้ง ไม่มีภาพประกอบเพราะต้องการสมาธิในการอ่าน เป็นข้อสอบปรนัยหลายรูปแบบ โดยอัตโนมัติ มีการเขียนตอบแต่เนื่องจากโปรแกรมไม่สามารถให้คะแนนได้ ผู้วิจัยต้องการให้คิดคะแนนอัตโนมัติ จึงได้ตัดออก

2.1.4 คำอธิบายประเด็นทางภาษาในแบบทดสอบฯ

ส่วนของไวยากรณ์ (Grammar) อธิบายรายชื่อตามประเด็นไวยากรณ์ ส่วนคำศัพท์ (Vocabulary) ให้รายการคำศัพท์วิชาการ (AWL) และวิธีการศึกษาคำศัพท์ ฝึกฝนกิจกรรมจากเว็บไซต์ การอ่าน (Reading) เขียนคำอธิบายหลักการอ่านตามแนวคำถามในแบบทดสอบ เช่น การตอบคำถาม main idea pronoun reference เนื้อหารายละเอียด และให้รายการวิธีอ่านและแบบฝึกหัดการอ่านเพิ่มเติม จากเว็บไซต์ การอธิบายคำถามรายชื่อของการอ่านให้เป็นลักษณะคำถามแบบเดียวกัน เช่นการหาใจความหลัก ซึ่งหากอธิบายเป็นข้อเขียนจะยาวมาก ในโอกาสต่อไปอาจพัฒนาเป็นคำอธิบายที่มีเสียงประกอบ

ในด้านการออกแบบแบบทดสอบในบริบทของไทยจากงานวิจัยเท่าที่ผู้วิจัยพบ ที่เกี่ยวข้องกับการออกแบบแบบทดสอบฯ ในโครงการวิจัยนี้ กล่าวคือในด้านการกำหนดบทอ่าน ผู้วิจัยได้ศึกษาประเด็นเรื่องหัวข้อของบทอ่านกับความสามารถในการอ่านภาษาอังกฤษมีความสัมพันธ์กัน ในการกำหนดบทอ่านผู้วิจัยได้คำนึงหัวข้อของบทอ่านกับลักษณะของผู้เข้าสอบที่เป็นผู้ต้องการเข้าศึกษาระดับ

บัณฑิตศึกษา (Vongpadungkiat, 2006)

ในด้านกาให้คำอธิบายและความรู้เกี่ยวกับโครงสร้างของบทอ่าน หรือประเภทของบทอ่าน มีลักษณะที่สอดคล้องกับผลวิจัยเรื่องความสัมพันธ์ระหว่างความรู้พื้นฐานของผู้เข้าสอบกับลักษณะของบทอ่าน (Vongpumivitch 2004) รวมถึงลำดับขั้นการเรียนรู้การอ่านเข้าใจความภาษาอังกฤษ (กรองแก้ว วรรณสูตร 2541)

ส่วนผลการวิจัยเกี่ยวกับทัศนคติ ความเครียดความคุ้นเคย กับความสามารถในการอ่านภาษาอังกฤษด้วยความเข้าใจด้วยแบบทดสอบคอมพิวเตอร์ ของผู้ทดลองใช้แบบทดสอบไม่เป็นปัญหา เช่นที่แสดงในงานของ Kiratibodee (2005) เนื่องจากผู้ใช้แบบทดสอบเป็นผู้ที่ต้องการศึกษาต่อระดับบัณฑิตศึกษาที่มีความคุ้นเคยกับการใช้คอมพิวเตอร์อยู่แล้ว และไม่มีความเครียดเพราะเป็นการสอบเพื่อวัดระดับความสามารถ มิใช่การตัดสินผลเพื่อนำไปใช้อย่างอื่น และส่วนสุดท้ายคือนำผลการวิจัยของการใช้กลวิธีการเรียนกับการทำแบบทดสอบ (Phakiti 2008) ไปประยุกต์ใช้ต่อการออกแบบแบบทดสอบและการแนะนำการฝึกฝนทักษะการอ่านภาษาอังกฤษ

กระบวนการ ออกแบบแบบทดสอบ: แนวคิด เกณฑ์กำหนดความสามารถและรายละเอียด
เกณฑ์การสร้างบทอ่าน ในกรอบของ test specifications รูปแบบแบบทดสอบออนไลน์ นับได้ว่าการ
พัฒนาแบบทดสอบมีความหนักแน่นทางทฤษฎี ครอบคลุมประเด็นต่าง ๆ อย่างรอบด้าน สมควร
จัดเป็นขั้นตอนและแนวทางที่ดีในการพัฒนาแบบทดสอบ แบบทดสอบฯนี้ยังมีคุณภาพ สามารถเป็น
ทางเลือกหนึ่งในจำนวนแบบทดสอบการอ่านภาษาอังกฤษออนไลน์จำนวนมากที่นำเสนอทางเว็บไซต์
ปัจจุบัน โดยเฉพาะอย่างยิ่งได้มุ่งเน้นที่ระดับบัณฑิตศึกษาเพื่อเตรียมความพร้อมในการอ่านงาน
วิชาการ

2.2 การตรวจสอบคุณภาพของแบบทดสอบวัดความสามารถในการอ่านภาษาอังกฤษออนไลน์

งานวิจัยนี้เน้นที่กระบวนการพัฒนาแบบทดสอบฯความสามารถในการอ่านภาษาอังกฤษการตรวจสอบคุณภาพของแบบทดสอบฯจึงมีความสำคัญยิ่ง

การตรวจสอบคุณภาพของแบบทดสอบฯ มีหลายขั้นตอน โดยอยู่บนพื้นฐานของกรอบแนวคิดเรื่องการอ่าน นับตั้งแต่ (1) การสร้างแบบทดสอบที่มีกรอบแนวคิดชัดเจน มีเกณฑ์ชัดเจน มีรายละเอียดชัดเจน (2) ให้ผู้ทรงคุณวุฒิทั้งด้านภาษา เจ้าของภาษา และด้านการวัดและประเมินผลพิจารณาให้ข้อเสนอแนะ ตรวจสอบ (3) มีการทดลองใช้ เบื้องต้นกับกลุ่มที่ไม่ใช่กลุ่มผู้ให้ข้อมูลจำนวน 5 คน เพื่อปรับแก้ภาษา กำหนดระยะเวลาประมาณในการสอบ (4) จัดทำแบบทดสอบออนไลน์ (5) ทดลองใช้จำนวน 31 คน (8) หากคุณภาพข้อสอบทั้งฉบับ ใช้สูตรการคำนวณของ *Kuder Richardson 20* (KR-20) ค่าความเที่ยงของแบบทดสอบชุดที่ 1 คือ $KR\ 20 = .86$ แบบทดสอบชุดที่ 2 คือ $KR\ 20 = .88$ แบบทดสอบชุดที่ 3 คือ $KR\ 20 = .61$ ซึ่งแสดงให้เห็นว่ามีค่าความเที่ยงสูง ส่วนแบบทดสอบชุดที่ 3 ค่าความเที่ยงน้อยกว่าชุดที่ 1 และ 2 อาจเนื่องมาจากเป็นการทำแบบทดสอบชุดสุดท้ายที่ทำต่อเนื่องกัน

ผู้สอบอาจกล้า (9) หากคุณภาพข้อสอบรายข้อ ทุกข้อ เนื่องจากข้อสอบมีหลายลักษณะคือปรนัย ตัวเลือก 4 ตัว 3 ตัว จับคู่ และแบบปลายเปิด จึงใช้วิธีวิเคราะห์ส่วนปรนัยตัวเลือก (9.1) โดยจำแนกเป็น ข้อที่ใช้ได้ ข้อที่ไม่มีอำนาจจำแนก ยาก ไป ง่ายไป โดยกำหนดเกณฑ์ดังนี้คือ ข้อสอบใช้ได้ มีค่า $.20 < p \leq .80$ ข้อสอบไม่มีอำนาจจำแนก มีค่า $.20 < p \leq .80$ และ $r < .20$ ข้อสอบยากเกินไป มีค่า $p < .20$ ข้อสอบง่ายไป มีค่า $p > .80$ (9.2) ส่วนจับคู่ใช้การวิเคราะห์คะแนนเฉลี่ย (10) ในการปรับปรุงข้อสอบรายข้อ ใช้กรอบแนวคิดเรื่องการอ่าน ระดับเกณฑ์ความสามารถที่กำหนด รายละเอียดแบบทดสอบเป็นพื้นฐาน พิจารณาค่า p_r ข้อที่ใช้ได้ ข้อที่ไม่มีอำนาจจำแนก ยาก ไป ง่ายไป ประกอบกับร้อยละของคนที่ตอบถูก และการวิเคราะห์ตัวเลือกรายตัว เช่น ลวงคนเก่งมากกว่าคนอ่อนในกลุ่ม ลวงคนได้น้อยควรปรับ ยากไป ง่ายไป เมื่อปรับปรุงแล้วให้เจ้าของภาษาตรวจสอบอีกครั้งหนึ่ง อย่างไรก็ตามแม้ข้อมูลทางสถิติอาจจะต้องปรับแต่ถ้าเป็นลักษณะของข้อคำถามที่ต้องถามและมีตัวเลือกเป็นแนวนั้น เช่นเรื่อง if clause จะต้องมิตัวเลือกเป็นรูปคำกริยาต่างๆ แม้ตัวเลือกบางข้อผลทางสถิติบอกว่าลวงคนเก่งมากกว่าคนอ่อนก็ต้องคงไว้ หมายความว่าคนเก่งในกลุ่มที่ทดสอบนี้ยังเข้าใจหลักภาษาไม่ชัดเจน เป็นประเด็นที่ควรนำไปสอนนักศึกษาต่อไป

อย่างไรก็ตามประเด็นเรื่องความตรง (validity) ของแบบทดสอบ Messick (1989 อ้างใน Khalifa & Weir 2009) ได้กล่าวว่าความตรง เป็นเรื่องของระดับมากกว่าจะมีทั้งหมดหรือไม่มีเลย ควรจะพิจารณาในลักษณะสัมพัทธ์มากกว่าตายตัว

จากกระบวนการตรวจสอบคุณภาพแบบทดสอบฯข้างต้นจึงกล่าวได้ว่าแบบทดสอบฯมีคุณภาพตามเกณฑ์ที่กำหนด

2.3. ระดับความสามารถทางการอ่านภาษาอังกฤษของผู้เข้าทดลองสอบ

ในส่วนผลของการทดลองสอบมีดังนี้

แบบทดสอบชุดที่ 1 มีผู้ได้คะแนนระดับ C1 คือ (effective operational proficiency) ใช้ภาษาเพื่อวัตถุประสงค์ใช้งานได้อย่างมีประสิทธิภาพจำนวน 3 คน ระดับ B2 (independent user ระดับเชี่ยวชาญ) จำนวน 12 คน ระดับ B1 (independent user ระดับเริ่มต้น) จำนวน 16 คน

แบบทดสอบชุดที่ 2 มีผู้ได้คะแนนระดับ C1 จำนวน 3 คน ระดับ B2 (จำนวน 14 คน ระดับ B1 จำนวน 14 คน

ในแบบทดสอบชุดที่ 3 มีผู้ได้คะแนนระดับ C1 จำนวน 2 คน ระดับ B2 จำนวน 11 คน ระดับ B1 จำนวน 18 คน

จากผลคะแนนดังกล่าวนี้บ่งชี้ว่าสอดคล้องความเป็นจริงที่คาดการณ์จากข้อมูลระดับความรู้พื้นฐานของผู้ทดลองใช้แบบทดสอบฯ ระดับปริญญาตรี กำลังศึกษาต่อระดับปริญญาโท และจากการที่ผู้ทดลองใช้ประเมินความสามารถทางภาษาอังกฤษของตนเองในระดับพอใช้มากที่สุด และระดับดีรองลงมา ผลคะแนนที่ระดับ B1 B2 (independent user) มีจำนวนมากกว่า ระดับ C1 (effective

operational proficiency) ซึ่งมีเพียง 2-3 คน นับว่าโดดเด่นออกมาจากกลุ่ม นอกจากนี้เมื่อพิจารณาประกอบกับด้านอื่น ๆ ยังแสดงให้เห็นว่าผลคะแนนสะท้อนความสามารถในการอ่าน 3 ระดับตามที่กำหนดไว้ อย่างไรก็ตามการกำหนดระดับความสามารถดังกล่าวเป็นเพียงเป้าหมาย ยังต้องมีองค์ประกอบด้านอื่น ๆ ที่มีผลกระทบต่อทดสอบด้วย เช่น ลักษณะของผู้เข้าสอบ การบริหารจัดการการสอบ การให้คะแนน (Shaw & Weir 2007) เช่นในการทดลองสอบนี้ ตามแบบสอบถามข้อมูลพื้นฐานและความเห็นทั่วไป หลายรายระบุว่าเครื่องปรับอากาศเย็นเกินไป อันอาจมีผลต่อการทำแบบทดสอบ อย่างไรก็ตามในงานวิจัยนี้ยังมิได้ครอบคลุมรายละเอียดดังกล่าว

2.4 การวิเคราะห์ปัญหาหรือจุดอ่อนในการอ่านภาษาอังกฤษของผู้ใช้บริการแบบทดสอบ

วัดความสามารถในการอ่านภาษาอังกฤษ

การวิเคราะห์ปัญหาหรือจุดอ่อนในการอ่านภาษาอังกฤษของผู้ใช้บริการแบบทดสอบวัดความสามารถในการอ่านภาษาอังกฤษ พิจารณาจากผลคะแนน ผลการวิเคราะห์ข้อคำถามที่ผู้ทดลองสอบ ค่าเฉลี่ย ร้อยละที่ตอบถูก ตอบผิด ข้อที่ไม่มีอำนาจจำแนก ยากไป ง่ายไป

กล่าวโดยสรุปผู้ใช้แบบทดสอบมีความสามารถทางไวยากรณ์พื้นฐานระดับที่ใช้ได้ แต่ยังมีประเด็นที่ซับซ้อน หรือต้องการความเข้าใจละเอียดลึกซึ้ง และรูปแบบที่พลิกแพลงไปจากตัวอย่างกฎเกณฑ์ของไวยากรณ์ที่ศึกษาทั่วไปยังทำไม่ได้มากนัก

ในส่วนคำศัพท์ ส่วนที่เป็นคำศัพท์ความหมายโดค จับคู่ คำศัพท์เดิมคำคำศัพท์แทนที่ คำที่ขีดเส้นใต้ คำศัพท์ในบริบท ส่วนที่เป็นประโยคทำได้ดีกว่า ส่วนที่เป็นบริบท เพราะต้องเข้าใจความหมายเนื้อหา

ในส่วนการอ่าน จากผลการวิเคราะห์การอ่านพบว่าคำถามที่มีปัญหาคือลักษณะคำถามในระดับ C1 ที่ต้องการความรู้เกี่ยวกับการเชื่อมโยงข้อความในบทอ่าน ระดับรายละเอียด ข้อความที่เกี่ยวข้อง การสรุป ใจความหลัก เข้าใจความสัมพันธ์ระหว่างใจความหลักกับรายละเอียด ซึ่งเป็นทักษะการอ่านขั้นสูง ประกอบกับเนื้อหาของบทอ่านจะมีระดับคำศัพท์ที่ยากกว่าระดับอื่น ต้องการความรู้เฉพาะทางเนื้อหาบ้าง ในส่วนคำถามระดับ B2 และ B1 การหาใจความหลัก รายละเอียด คำศัพท์จากบทอ่าน คำสรรพนามอ้างอิง

ข้อคำถามลักษณะนี้เป็นลักษณะแบบทดสอบการหาใจความหลักในย่อหน้าต่างๆ ในบทอ่าน มีรูปแบบคำถามหลากหลาย โดยเป็นการเลือก topic sentence, สรุปใจความออกมาใหม่ ระบุชื่อคนที่เกี่ยวข้องในย่อหน้า โดยตัวเลือกว่ากับจำนวนย่อหน้าจะไม่ตรงกัน (multiple matching) ข้อมูลโดยรวมตอบได้ถูกต้องเกินร้อยละ 50 ในแบบทดสอบชุดที่ 2 ทำได้ถึงร้อยละ 90 ส่วนที่ทำคะแนนได้น้อยจะเป็นข้อคำถามจากบทอ่านระดับ C1 ซึ่งมีเนื้อหาสูงกว่าระดับอื่น

กล่าวโดยสรุปปัญหาของผู้ใช้แบบทดสอบคือความเข้าใจโครงสร้างทางภาษาที่ซับซ้อน วงคำศัพท์ และการใช้คำศัพท์ในบริบท การอ่านในระดับของการอนุมานข้อมูล แสดงนัย

2.5 แนวทางการพัฒนาการอ่านภาษาอังกฤษของตนเองของผู้ใช้บริการแบบทดสอบ

แนวทางการพัฒนาการอ่านภาษาอังกฤษของตนเอง สามารถพิจารณาจากผลการวิเคราะห์ความสามารถในการอ่านที่ได้กล่าวมาในข้อ 2.4 ผู้ใช้แบบทดสอบสามารถปรับปรุงวิธีการเรียนภาษาอังกฤษด้วยตนเองโดยการฝึกฝนเพิ่มเติมโดยใช้การคิดวิเคราะห์เป็นหลัก และฝึกฝนโดยการอ่านบริบทของบทอ่านที่อยู่ในระดับของการเรียนระดับบัณฑิตศึกษา ซึ่งเป็นการอ่านแบบวิเคราะห์วิพากษ์วิจารณ์ โดยในขั้นต้นต้องเข้าใจความหมายของข้อความตามจริง ความหมายแสดงนัย ที่สำคัญอย่างยิ่งคือควรศึกษาวงคำศัพท์วิชาการในสาขาวิชาที่ศึกษา โครงสร้างประโยคที่ซับซ้อน ให้ฝึกฝนการอ่านให้มาก

ผลการวิเคราะห์ความสามารถในการอ่านโดยรวมของกลุ่มผู้ใช้แบบทดสอบ ในด้าน ความเข้าใจโครงสร้างทางภาษาที่ซับซ้อน วงคำศัพท์ และการใช้คำศัพท์ในบริบท การอ่านในระดับของการอนุมานข้อมูล แสดงนัย ผู้ใช้แบบทดสอบสามารถปรับปรุงการเรียนภาษาอังกฤษด้วยตนเองโดยพิจารณาในมุมมองของกลวิธีการเรียนทางพุทธิพิสัย อภิปัญญา สังคมและจิตใจ (O'Malley and Chamot 1990; Chamot 2004) และ กลวิธีการเรียนทางตรง หรือ ทางอ้อม (Oxford 1990) ซึ่งเป็นแนวทางการอ่านเพื่อความเข้าใจที่ครอบคลุมรอบด้านนับแต่กระบวนการคิด การบริหารจัดการ และการจัดการกับความรู้สึกร่างกายของตน ในรายละเอียดของแต่ละด้าน นอกจากนี้อาจศึกษาเทคนิควิธีการทำแบบทดสอบ (Bachman & Cohen 1998) เพื่อใช้ประกอบในการทำแบบทดสอบที่ต้องนำไปตัดสินใจเรื่องต่างๆ ซึ่งกลวิธีการเรียนทางพุทธิพิสัย เป็นสิ่งสำคัญที่สุด รายละเอียดของกลวิธีการเรียนเหล่านี้แสดงอยู่ในผลของการวิจัยแล้ว

5.3 ข้อเสนอแนะ

5.3.1 ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1. ควรให้บริการแบบทดสอบวัดความสามารถในการใช้ภาษาอังกฤษออนไลน์แก่นักศึกษาระดับบัณฑิตศึกษา และผู้สนใจทั่วไปที่ต้องการประเมินผลความสามารถในการอ่านโดยให้คำแนะนำเพิ่มเติม

2. ควรจัดฝึกอบรมวิธีการสร้างแบบทดสอบวัดความสามารถในการอ่านในรายละเอียดและขั้นตอนรวมถึงเครื่องมือการวัดความสามารถในการอ่าน แก่ครู อาจารย์และนักวิชาการด้านภาษาอังกฤษเพื่อนำไปพัฒนาแบบทดสอบการอ่าน โดยเฉพาะอย่างยิ่งชุดวิชาด้านการอ่านภาษาอังกฤษของมหาวิทยาลัยสุโขทัยธรรมมาธิราช

3. ควรจัดฝึกอบรมและสัมมนา กำหนดเกณฑ์ระดับความสามารถในการอ่านภาษาอังกฤษของนักศึกษาไทยเพื่อเป็นมาตรฐาน

5.3.2 ข้อเสนอแนะในการวิจัยครั้งต่อไป

1. ควรศึกษาวิจัยองค์ประกอบของกรอบแนวคิดเรื่องความน่าเชื่อถือในการทดสอบการอ่าน ในด้านของผู้เข้าสอบ (test taker characteristics) บริบท (context validity) โดยเฉพาะอย่างยิ่งเรื่องพุทธิพิสัย (cognitive validity) ซึ่งรวมถึงเป้าหมายการอ่าน การรู้จักคำ การเข้าถึงแหล่งข้อมูล คำศัพท์ การแจ่งส่วนประโยค (sentence parsing) การสร้างความหมาย การอนุมาน การสร้างภาพต้นแบบ การสร้างความเชื่อมโยงระดับบทอ่าน การสร้างความเชื่อมโยง ระหว่างบทอ่าน และการทดสอบความเข้าใจ (Khalifa & Weir 2009)

2. ควรศึกษาวิจัยการจัดระดับความสามารถในการอ่านที่พึงประสงค์ของนักศึกษาไทยในระดับประเทศ ในลักษณะของกรอบมาตรฐานความสามารถ (framework of references) เช่นที่ Council of Europe ได้จัดทำ

3. ควรศึกษาวิจัยเรื่องความสัมพันธ์ระหว่างกลวิธีการอ่านและการทำแบบทดสอบการอ่าน

4. ควรวิจัยและพัฒนาแบบทดสอบวัดความสามารถในการอ่านภาษาอังกฤษออนไลน์ ที่มีคำอธิบายเป็นเสียงประกอบเนื้อหาเพื่อช่วยให้เข้าใจง่ายขึ้น



