

การศึกษาการปรากฏร่วมจำเพาะของคำเนื้อหาที่มีความถี่สูงใน OUTSIDE WORD LIST (OWL)
โดยใช้คลังข้อมูลภาษาจากบทความวิจัยที่เกี่ยวข้องกับสัตว์ทดลอง

A Corpus-Based Study on High-Frequency Content Words and Collocations in the
Outside Word List (OWL) in Laboratory Animal Research Articles

วิรดา ปัญจนาท
ทรงศรี สรณสถาพร

บทคัดย่อ

งานวิจัยชิ้นนี้มีวัตถุประสงค์เพื่อศึกษาการปรากฏร่วมจำเพาะ (Collocation) ของคำเนื้อหา (Content word) ที่มีความถี่สูงที่สุดใน Outside Word List (OWL) โดยใช้บทความวิจัยที่เกี่ยวข้องกับสัตว์ทดลอง เป็นแหล่งของคลังข้อมูล ซึ่งประกอบด้วย 100 บทความวิจัยที่ได้รับการตีพิมพ์ในวารสารสัตว์ทดลองของ Institute for Laboratory Animal Research (ILAR) ระหว่างปี 2553 ถึง 2557 โดยวิธีการสุ่มตัวอย่างแบบชั้นภูมิและการสุ่มตัวอย่างแบบง่าย คลังข้อมูลในการศึกษานี้มีจำนวนคำศัพท์ทั้งหมด 555,526 คำ เครื่องมือที่ใช้ในการวิเคราะห์คือ โปรแกรม AntConc และเครื่องมือที่ใช้ในจำแนกคำศัพท์คือ VocabProfile ข้อมูลจากรายการคำศัพท์ที่ได้รับการวิเคราะห์โดยการหาความถี่ ร้อยละ และควอไทล์ งานวิจัยนี้พบว่า อัตราส่วนของ OWL คิดเป็นร้อยละ 65.87 โดยคำที่มีความถี่สูงคือคำที่มีความถี่อยู่ในควอไทล์ที่ 1 ซึ่งคือคำจากลำดับที่ 1 ถึง 3,055 คำศัพท์ที่มีความถี่สูงที่สุดคือคำว่า zebrafish ซึ่งการปรากฏร่วมจำเพาะของ zebrafish ส่วนมากอยู่ในลักษณะคำนามประสม (compound noun)

คำสำคัญ: การปรากฏร่วมจำเพาะ, คลังข้อมูล, คำเนื้อหา, สัตว์ทดลอง, Outside Word List (OWL)

Abstract

The present study was designed to investigate the collocations using the highest-frequency content words in the Outside Word List (OWL). The source of the corpus was 555,526 running words selected by stratified random sampling from 100 research articles published in the Institute of Laboratory Animal Research (ILAR) journal between 2010 and 2014. AntConc (3.4.4) was used to determine word frequency and create a laboratory animal science word list. Then VocabProfile was used to separate these words from the OWL. The data were analyzed and displayed as percentages, frequencies, and quartiles. The results show

that the text coverage for OWL amounted to 65.87%. The high frequency words in OWL were the words belonging to the first quartile, or those ranking from 1st to 3055th. The highest-frequency word was ‘zebrafish’, which collocated most frequently with another noun to form a compound noun.

Keywords: collocation, content word, corpus, laboratory animal, Outside Word List (OWL)

บทนำ

ภาษาอังกฤษได้รับการยอมรับเป็นภาษาสากลที่ใช้ในการติดต่อสื่อสารในระดับนานาชาติ (Kirkpatrick, 2012; Sharifian, 2009) ดังนั้นภาษาอังกฤษจึงมีบทบาทสำคัญต่อการสื่อสารในฐานะภาษาที่สอง (English as a Second Language) และ ภาษาต่างประเทศ (English as a Foreign Language) รวมถึงการทำงาน (English for Occupational Purposes) การเรียนเพื่อวิชาการ (English for Academic Purposes) และเพื่อวัตถุประสงค์เฉพาะ (English for Specific Purposes) (Montgomery, 2013; Basturkmen, 2010) โดยเฉพาะการเรียนการสอนในสาขาวิทยาศาสตร์และเทคโนโลยี (English for Science and Technology) ที่ต้องอาศัยความรู้ความเจริญก้าวหน้าทางวิทยาศาสตร์จากต่างประเทศเพื่อศึกษาและทำการวิจัย นักศึกษาและผู้วิจัยจึงจำเป็นต้องอ่านบทความหรือวารสารจากต้นฉบับภาษาอังกฤษ ดังนั้นเพื่อให้เข้าใจข้อมูลจากต้นฉบับได้อย่างถูกต้อง ผู้อ่านจะต้องมีความรู้ภาษาอังกฤษเป็นอย่างดี นอกจากนี้การเผยแพร่ผลงานวิจัยในระดับนานาชาติจำเป็นต้องใช้อย่างยิ่งที่ต้องใช้ภาษาอังกฤษในการสื่อสาร โดยเฉพาะการเขียนบทความและการนำเสนอผลงาน (Belcher, Johns & Paltridge, 2011)

ความรู้ด้านคำศัพท์เป็นพื้นฐานสำคัญในการสร้างประโยคเพื่อให้ได้ใจความสำคัญครบสมบูรณ์ (Folse, 2010; Laufer & Kalovsky, 2010; Nation, 2001) อย่างไรก็ตาม กลุ่มคำศัพท์ในแต่ละสาขาวิชามีความแตกต่างกัน จึงทำให้คำศัพท์ที่เรียนทั่วไปไม่เพียงพอต่อการอ่านบทความหรือวารสารเฉพาะทางที่ใช้เป็นแหล่งศึกษาค้นคว้าในระดับมหาวิทยาลัยและในการทำงานวิจัย เพราะบทความเหล่านั้นมักใช้คำศัพท์เฉพาะทางหรือคำศัพท์เทคนิคในการเขียน ดังนั้นคำศัพท์เฉพาะจึงจำเป็นต้องการเรียนภาษาอังกฤษเพื่อวัตถุประสงค์เฉพาะ (English for Specific Purposes)

จากการสัมภาษณ์ผู้อำนวยการศูนย์วิจัยสัตว์ทดลองแห่งชาติ มหาวิทยาลัยมหิดล เมื่อวันที่ 20 มีนาคม 2558 พบว่านักศึกษาและนักวิจัยส่วนใหญ่ไม่สามารถอ่านและเขียนบทความงานวิจัยภาษาอังกฤษได้ เนื่องจากขาดความรู้ด้านคำศัพท์เฉพาะทางด้านสัตว์ทดลอง ประกอบกับไม่มีการจัดทำหลักสูตรการเรียนการสอนภาษาอังกฤษเฉพาะทางด้านสาขาสัตว์ทดลอง นอกจากนี้จากการค้นหาคำศัพท์เกี่ยวกับคลังข้อมูลภาษาด้านสัตว์ทดลองในฐานข้อมูล ScienceDirect และ Scopus เมื่อวันที่ 1 พฤษภาคม 2558 โดยใช้คำสำคัญ ‘คลังข้อมูล’ และ ‘สัตว์ทดลอง’ พบว่า มีการศึกษาเกี่ยวกับคลังข้อมูลภาษาในหลายสาขาวิชาเช่น ด้านธุรกิจ

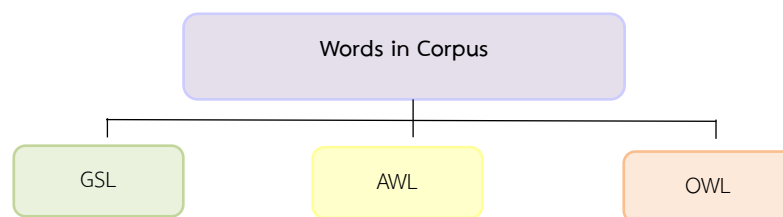
(Walker, 2011) การเกษตร (Martinez, Beck, & Panza, 2009) เคมี (Valipouri, & Nassaji, 2013) การแพทย์ (Staples, 2015) และสิ่งแวดล้อม (Liu & Han, 2015) แต่ไม่พบข้อมูลที่ศึกษาด้านสัตว์ทดลอง ดังนั้นงานวิจัยชิ้นนี้จึงมีวัตถุประสงค์ดังนี้

1. เพื่อศึกษาคำเนื้อหา (content word) ที่มีความถี่สูงที่ปรากฏใน Outside Word List (OWL)
2. เพื่อศึกษาการปรากฏร่วมจำเพาะ (collocation) ของคำเนื้อหาที่มีความถี่สูงสุดใน Outside Word List (OWL)

ทฤษฎีและวรรณกรรมที่เกี่ยวข้อง

คลังข้อมูลภาษา (corpus) คือ การรวบรวมภาษาเขียน หรือ ภาษาพูด ที่ใช้ในสถานการณ์จริงจากแหล่งข้อมูลต่างๆ เช่น บทความ หนังสือพิมพ์ ตำราเรียน สุนทรพจน์ เป็นต้น (Biber, Conrad & Reppen, 2000; Hunston, 2002) ดังนั้น นักภาษาศาสตร์จึงใช้คลังข้อมูลภาษาในการรวบรวมและวิเคราะห์ คำศัพท์ โครงสร้างประโยค หรือ รูปแบบบทความที่ใช้ในสาขาวิชาเฉพาะ (Chung & Nation, 2004; Hunston, 2011) คลังข้อมูลภาษาจึงมีประโยชน์ต่อการเรียนการสอนในฐานะที่เป็นแหล่งอ้างอิงที่น่าเชื่อถือ เพราะรวบรวมจากข้อมูลที่ใช้ในสถานการณ์จริง(authentic materials) นอกจากนี้คลังข้อมูลภาษายังใช้เป็นแหล่งข้อมูลอ้างอิงในการทำพจนานุกรมและพัฒนาบทเรียน (Hunston, 2011) โดยเฉพาะการพัฒนาบทเรียนภาษาอังกฤษเพื่อวัตถุประสงค์เฉพาะ (English for Specific Purposes) (Bruce, 2011; Flowerdew & Peacock, 2001; Hutchinson & Waters, 1987; Jolly & Bolitho, 2011; Tomlinson, 2012) ที่ผู้สอนรวบรวมและวิเคราะห์ คำศัพท์ โครงสร้างประโยค หรือตัวอย่างประโยคที่ใช้ในบริบทจำเพาะได้จากคลังข้อมูลภาษา เช่น บทอ่านที่ใช้ในหนังสือ *Reading and Writing for First Year Medical Students ของมหาวิทยาลัยมหิดล (2004) หรือ ตัวอย่างบทคัดย่อและประโยคในหนังสือ การอ่านและเขียนภาษาอังกฤษอย่างมีประสิทธิภาพ* เรียบเรียงโดย ทรงศรี สรณสถาพร (2557) ยิ่งไปกว่านี้ ปัจจุบันวิทยาศาสตร์และเทคโนโลยีมีการพัฒนาและก้าวหน้าอย่างมาก นักวิทยาศาสตร์และเทคโนโลยีใช้ภาษาอังกฤษในการเผยแพร่ผลงานวิจัยวิชาการซึ่งมีเนื้อหาจำนวนมากในฐานข้อมูลอิเล็กทรอนิกส์ต่างๆ การศึกษาคำเนื้อหา (content word) จะช่วยให้ทราบแนวโน้มหรือประเด็นที่นักวิทยาศาสตร์และเทคโนโลยีทำการศึกษาวิจัยในแต่ละช่วงเวลา ดังนั้น จึงกล่าวได้ว่า การศึกษาคคลังข้อมูลภาษาเป็นประโยชน์ต่อการวิจัยทั้งทางด้านภาษาศาสตร์ การเรียนการสอนภาษาอังกฤษ และวิทยาศาสตร์

คำศัพท์ในคลังข้อมูลภาษาประกอบด้วย 3 ประเภทหลักๆ คือ General Service List (GSL), Academic Word List (AWL) และ Outside Word List (OWL) ดังแผนภาพที่ 1



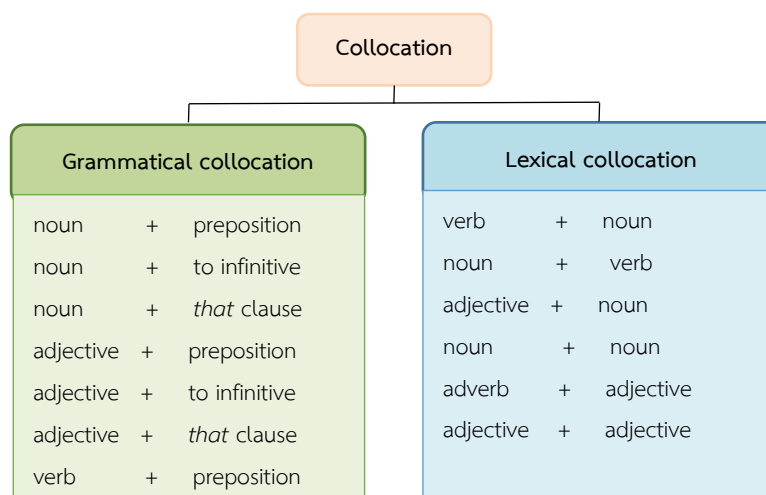
แผนภาพที่ 1 รายการคำศัพท์ในคลังข้อมูลภาษา

1. General Service List (GSL) คือ รายการคำศัพท์ที่มีความถี่สูง 2,000 คำแรกมักใช้ในบทความทั่วไป (Brown, 2013)
2. Academic Word List (AWL) คือ คำศัพท์ใช้ในบริบททางวิชาการ (Brown, 2013)
3. Outside Word List (OWL) คือ คำศัพท์เฉพาะหรือคำศัพท์เทคนิค (Technical terms) ที่ไม่ปรากฏใน GSL หรือ AWL การแยกคำศัพท์เฉพาะออกจากคำศัพท์ที่ใช้ทั่วไป สามารถทำได้โดยการแบ่งคำศัพท์ในบทความออกเป็น GSL และ AWL คำศัพท์ส่วนที่อยู่นอกเหนือจากสองรายการนี้คือ OWL ซึ่งเป็นกลุ่มคำศัพท์ที่ใช้เฉพาะในสาขาวิชานั้นๆ (Coxhead & Hirsh, 2007; Coxhead, 2000)

รายการคำศัพท์ที่ใช้ศึกษาในงานวิจัยนี้ อ้างอิงจาก Brown (2013) เนื่องจากเป็นรายการคำศัพท์ที่ได้รับการปรับปรุงใหม่ ทำให้มีความน่าเชื่อถือ เนื่องจากรายการคำศัพท์ของ Brown มีคำศัพท์เพิ่มขึ้นร้อยละ 5 จาก GSL ของ West (1953) และ AWL ของ Coxhead (2000) ปริมาณคำศัพท์ที่เพิ่มขึ้นถือว่าการขยายฐานคำศัพท์ ทำให้ผู้เรียนมีวงคำศัพท์ที่นำศึกษาเรียนรู้เพิ่มขึ้น และจะส่งผลให้ผู้เรียนที่สนใจศึกษาค้นคว้าเพิ่มเติมมีความรู้ทางคำศัพท์มากขึ้น

นอกจากนั้น Fries (1952) แบ่งคำในภาษาอังกฤษออกเป็น 2 ประเภท คือ คำเนื้อหา (content word) และคำแสดงหน้าที่ (function word) คำเนื้อหามีหน้าที่บอกความหมายของคำ ส่วนคำแสดงหน้าที่คือคำที่ทำหน้าที่แสดงความสัมพันธ์ทางไวยากรณ์กับคำอื่นๆ ในการศึกษาครั้งนี้ ผู้วิจัยจะศึกษาเฉพาะแสดงคำเนื้อหา เพราะคำเนื้อหาเป็นส่วนที่ส่วนที่ให้ความสำคัญของถ้อยความ และสามารถเพิ่มความรู้ทางคำศัพท์ของผู้เรียนได้ (Fries, 1952; Murphy, 2010)

นอกจากความรู้ด้านคำศัพท์แล้ว การปรากฏจำเพาะของคำศัพท์ (Collocation) ยังมีความสำคัญในการช่วยสร้างประโยคภาษาอังกฤษให้ใกล้เคียงกับเจ้าของภาษา (Hill, 2000; Nation, 2001; Durrant & Schmitt, 2010) เพราะ Collocation เกี่ยวข้องกับความสัมพันธ์ของคำในระดับ วลี และประโยค Collocation แบ่งเป็น 2 ประเภทหลักๆ ได้แก่ Grammatical collocation และ Lexical collocation (Benson, 1985; Biskup, 1992; Bahns, 1993) ดังแผนภาพที่ 2



แผนภาพที่ 2 ประเภทของ Collocation

Grammatical collocation คือ การปรากฏร่วมจำเพาะระหว่าง คำนาม คำกริยา หรือ คำคุณศัพท์ กับ คำบุพบท to infinitive หรือ that clause

Lexical collocation คือ การปรากฏร่วมจำเพาะระหว่างคำนาม คำกริยา หรือ คำคุณศัพท์ด้วยกันเอง

ในการศึกษาครั้งนี้จะศึกษาเฉพาะ lexical collocation เพื่อให้สอดคล้องกับวัตถุประสงค์ของงานวิจัย ซึ่งเน้นศึกษาเฉพาะคำเนื้อหา นอกจากนี้การปรากฏร่วมของคำเนื้อหาหรือ lexical collocation สามารถทำให้เกิดคำศัพท์ใหม่หรือความหมายใหม่ขึ้นได้ (Fries, 1952)

การศึกษา lexical collocation จะดูจากไวยากรณ์ภาษาอังกฤษในแง่ของไวยากรณ์หน้าที่ (functional grammar) กล่าวคือ ศึกษาโดยการจำแนกชนิดคำ (part of speech) ออกตามหน้าที่ที่ใช้ในบริบทเช่น คำนาม คำกริยา คำคุณศัพท์ คำกริยาวิเศษณ์ (Halliday & Matthiessen, 2014) นอกจากนี้ชนิดของคำมีความสำคัญในการเรียงประโยคให้มีใจความสมบูรณ์ เพราะชนิดของคำจะบอกหน้าที่ต่างๆในประโยค เช่น คำนี้ทำหน้าที่เป็นคำนาม และแสดงหน้าที่เป็นประธานหรือกรรมในประโยค ดังนั้นเพื่อให้ผู้อ่านเข้าใจความหมายของประโยคได้อย่างถ่องแท้ ผู้อ่านจำเป็นต้องทราบหน้าที่และชนิดของคำด้วย ทั้งนี้งานวิจัยชิ้นนี้จะเน้นศึกษาในกลุ่มคำนาม เนื่องจากคำนามเป็นหน่วยคำที่มีความสำคัญที่สุดในการสร้างประโยค และเป็นหน่วยคำที่สื่อความหมายหลักของประโยคมากกว่าคำชนิดอื่นเช่น คำคุณศัพท์ หรือ คำวิเศษณ์ (Downing & Locke, 1992) นอกจากนี้จากการศึกษาเบื้องต้นพบว่า ลักษณะของ lexical collocation มักเกิดในลักษณะการปรากฏร่วมของคำนามในรูปแบบคำนามประสมมากที่สุด อีกทั้งตำแหน่งของคำนามหลักและคำนามรองที่ประกอบเป็นคำนามประสม มีการเรียงลำดับคำที่แตกต่างจากภาษาไทย กล่าวคือคำนามประสมในภาษาไทยมีการอ่านและให้ความหมายเรียงจากซ้ายไปขวา แต่คำนามหลักในภาษาอังกฤษจะอยู่ด้านหลัง คำนามรองจะอยู่ด้านหน้า ยกตัวอย่างเช่น คำว่า ‘รถโรงเรียน’

ภาษาไทย	รถโรงเรียน	มาจาก	รถ (คำนาม)	รวมกับ	โรงเรียน (คำนาม)
ภาษาอังกฤษ	school bus	มาจาก	school (คำนาม)	รวมกับ	bus (คำนาม)

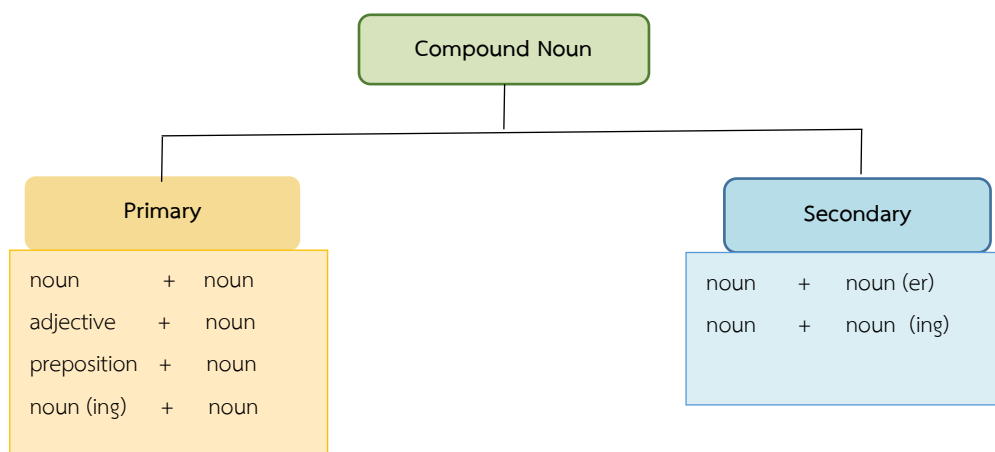
ฉะนั้นคำนามประสมภาษาอังกฤษจะต้องแปลจากขวามาซ้ายจึงจะได้ความหมาย คำนามประสมยังมีประโยชน์ทางภาษาศาสตร์ดังนี้

1. การใช้คำนามประสมทำให้ประโยคมีใจความกระชับขึ้น และยังหลีกเลี่ยงการใช้คำกริยาซ้ำซ้อน ทำให้ผู้เขียนไม่ต้องอธิบายความยาวขึ้น กล่าวคือ คำนามประสมเป็นการใช้ nominalization อย่างหนึ่ง (Quirk, 1972; Baker, 2003) เช่น
 ‘the woman who works for living’ เรียกว่า ‘working woman’
 ‘the bag which is used for shopping’ เรียกว่า ‘shopping bag’
2. การรวมกันของคำนามประสมสามารถสร้างคำใหม่ขึ้นได้ เนื่องจากคำนามเป็นคำเนื้อหา (content word) มีลักษณะเป็นคำชนิดเปิด (open class) มีคำศัพท์จำนวนมาก ต่างจากคำหน้าที่ (function word) ซึ่งมีลักษณะเป็นคำชนิดปิด (close class) และมีจำนวนคำศัพท์ตายตัว (Fries, 1952) ดังนั้นเมื่อนำคำเนื้อหาหรือคำนามมาประสมกัน จึงทำให้เกิดคำใหม่เพิ่มขึ้นตามการพัฒนาของสาขานั้นๆ ได้ (Diarmuid, 2008) เช่น ice cream, swimming pool, bus stop
3. คำนามประสมทำให้คำสองคำมีความหมายสำคัญเท่าเทียมกัน (McCarthy, 2002; Baker, 2003) โดยทั่วไปคำคุณศัพท์ภาษาอังกฤษขยายจะทำให้หน้าที่ขยายคำนาม ดังนั้นใจความหลักจะอยู่ที่คำนาม แต่การใช้คำนามประสมหรือคำนามขยายคำนาม จะทำให้ใจความหลักอยู่ที่คำนามทั้งสองคำ กล่าวคือ หากละคำนามตัวใดตัวหนึ่งไป จะทำให้ความหมายผิดเพี้ยนหรือไม่สมบูรณ์ ประกอบกับคำนามภาษาอังกฤษบางคำ ไม่สามารถเติมปัจจัย (suffix) เพื่อเปลี่ยนรูปเป็นคำคุณศัพท์ได้ จึงต้องใช้คำนามทำหน้าที่ขยายคำนาม คำนามประสมเหล่านี้จะเรียกว่า root compound (Lieber, 2010)

ดังนั้นเพื่อเป็นประโยชน์ต่อผู้เรียนไทยที่เรียนภาษาอังกฤษในฐานะภาษาที่สองและภาษาต่างประเทศ ให้ตระหนักถึงความแตกต่างของการเรียงลำดับคำในภาษาอังกฤษ และความสำคัญของการสร้างคำนามประสม งานวิจัยชิ้นนี้จึงวิเคราะห์ประเภทของคำนามประสมและตำแหน่งของคำนามประสมในประโยคของบทความ สาขาสัตว์ทดลองมาจากวารสาร Institute of Laboratory Animal Research Articles journal (ILAR)

ประเภทของคำนามประสม

ประเภทของคำนามประสม แบ่งได้ออกเป็น 2 ประเภท ได้แก่ คำนามประสมประเภทหลัก (primary compound) และ คำนามประสมประเภทรอง (secondary compound) (McCarthy, 2002) ดังแผนภาพที่ 3

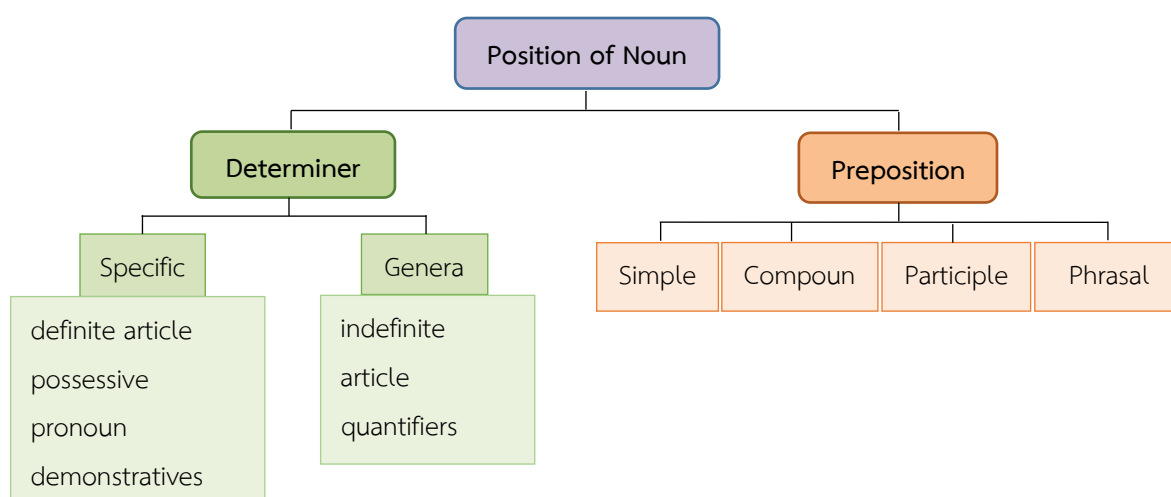


แผนภาพที่ 3 ประเภทของคำนามประสม

1. คำนามประสมประเภทหลัก คือ คำนามประสมที่มีส่วนหลักเป็นคำนาม และมีส่วนขยายเป็นคำนาม คำคุณศัพท์ คำบุพบท และ คำกริยาที่เติมหน่วยคำอนุพันธ์สร้างคำนาม -ing
2. คำนามประสมประเภทรอง คือ คำนามประสมที่มีส่วนหลักมาจากคำกริยาที่เติมหน่วยคำอนุพันธ์สร้างคำนาม -ing และ -er

ตำแหน่งของคำนามประสม

คำนามประสมมักนำหน้าด้วย คำกำกับคำนาม (determiner) และ คำบุพบท (preposition) ดูแผนภาพที่ 4



แผนภาพที่ 4 ตำแหน่งของคำนามประสม

1. คำกำกับหน้านาม (determiner) ทำหน้าที่กำกับคำนามและช่วยบ่งชี้ว่าคำนามนั้นเป็นคำนามทั่วไป หรือชี้เฉพาะเจาะจง
2. คำบุพบท (preposition) ใช้แสดงสถานที่ ตำแหน่ง การเคลื่อนไหว ทิศทาง เวลา ลักษณะ และความสัมพันธ์ คำบุพบทสามารถเป็นคำเดี่ยวโดดๆ (simple preposition) คำผสม (compound preposition) คำที่ลงท้ายด้วย -ing หรือ เป็นกลุ่มคำวลี (phrasal preposition)

ระเบียบและวิธีการวิจัย

ระเบียบวิธีการวิจัยในการศึกษาคำนี้ประกอบด้วย แหล่งของคลังข้อมูลภาษา เครื่องมือวิจัย ขั้นตอนในการดำเนินการวิจัย และสถิติที่ใช้ในการวิเคราะห์ ได้อธิบายไว้ในส่วนนี้

แหล่งของคลังข้อมูลภาษา

แหล่งของคลังข้อมูลภาษาที่นำมาศึกษาคำศัพท์เฉพาะของสาขาสัตว์ทดลองมาจากวารสาร Institute of Laboratory Animal Research Articles journal (ILAR) ซึ่งประกอบไปด้วยบทความ 100 ฉบับ จากทั้งหมด 160 บทความ ระหว่างปี 2553 ถึง 2557 โดยวิธีการสุ่มตัวอย่างแบบชั้นภูมิและการสุ่มตัวอย่างแบบง่าย โดยมีคำศัพท์ในคลังข้อมูลทั้งหมด 555,526 คำศัพท์ วารสาร ILAR มีความน่าเชื่อถือ และได้มาตรฐานตามเกณฑ์ดังต่อไปนี้

1. Beall's List

Beall's list คือ รายชื่อของวารสารที่เข้าข่ายล่อหลวงเพื่อแสวงหากำไร เป็นวารสารที่ไม่ได้มาตรฐาน ซึ่งจากรายชื่อของ Beall ไม่พบว่ามีการวารสาร ILAR ปรากฏอยู่

2. Peer review

Peer review คือ การตรวจทานจากผู้ทรงคุณวุฒิ ซึ่งบทความในวารสาร ILAR ต้องผ่านการตรวจทานจากผู้ทรงคุณวุฒิก่อนตีพิมพ์

3. SCImago Journal & Country Rank (SJR)

การจัดอันดับวารสารในระดับนานาชาติของ SJR แสดงให้เห็นว่าวารสาร ILAR อยู่ในระดับต้นๆ (Q1 และ Q2) ของสาขาวิชาสัตววิทยา ชีวเคมี พันธุศาสตร์ และอนุชีววิทยา

เครื่องมือวิจัย

โปรแกรม AntConc (3.4.3) VocabProfile และ Microsoft Excel 2013 เป็นเครื่องมือที่ใช้ในการวิจัยครั้งนี้

1. AntConc

AntConc เวอร์ชัน 3.4.3 พัฒนาโดย Anthony (2014) ใช้รวบรวมและแสดงตรรกะคำในบริบท (concordance) และใช้ในการวิเคราะห์คำศัพท์ในคลังข้อมูลภาษา นอกจากนี้ AntConc สามารถสร้างรายการคำศัพท์ โดยการจัดเรียงตามลำดับตามความถี่ นักวิจัยภาษาศาสตร์ได้นำโปรแกรม AntConc มาใช้ในการวิเคราะห์ภาษาทั้งในระดับคำและประโยคอย่างแพร่หลายเพราะความสะดวกและง่ายต่อการใช้งาน Froehlich (2015) กล่าวว่า AntConc ยังสามารถแสดงเฉพาะคำหรือประโยคที่ต้องการวิเคราะห์ เพื่อหาความถี่หรือโครงสร้างที่คำนั้นๆปรากฏรวม

2. VocabProfile

VocabProfile พัฒนาโดย Cobb (2003) เพื่อใช้ในการจำแนกรายการคำศัพท์ออกเป็น GSL, AWL และ OWL Laufer และ Nation (1995) ได้ทดสอบความเที่ยงตรงของโปรแกรม VocabProfile พบว่าโปรแกรมนี้มีประโยชน์ต่อการเรียนรู้คำศัพท์เพราะมีฐานข้อมูลคำศัพท์ที่เหมาะสมกับระดับความรู้ของผู้เรียน ตั้งแต่ระดับเริ่มต้นจนถึงระดับสูง นอกจากนี้ฐานข้อมูลคำศัพท์ของ VocabProfile ยังครอบคลุมรายการคำศัพท์ใน AWL ที่จำเป็นต่อผู้เรียน (Cobb & Horst, 2001) ฐานข้อมูลคำศัพท์ของ VocabProfile ได้ปรับปรุงโดยอ้างอิงจากรายการคำศัพท์ของ Browne (2013)

3. Microsoft Excel 2013

Microsoft Excel 2013 ใช้ในการเก็บบันทึกข้อมูล สำหรับรวบรวม คำนวณและจัดเรียงข้อมูลในรูปแบบตาราง อีกทั้งสามารถจัดทำกราฟ แผนภูมิเพื่อแสดงผลของข้อมูล

ขั้นตอนในการดำเนินงานวิจัย

1. ดาวน์โหลดบทความ 100 บทความที่ได้จากการสุ่มตัวอย่างแบบง่าย และแปลงไฟล์เป็น text file ในขั้นตอนนี้คลังคำศัพท์จะไม่รวมรูปภาพและคำบรรยายใต้ภาพ ตาราง บรรณานุกรม เชิงอรรถ และ กิตติกรรมประกาศในบทความวิจัย
2. สร้างรายการคำศัพท์จากบทความวิจัยสัปดาห์ทดลอง โดยใช้ฟังก์ชัน Word List ในโปรแกรม AntConc เนื่องจากขนาดคลังข้อมูลที่ใช้ในการศึกษานี้มีทั้งหมด 555,526 คำ จึงไม่สามารถใช้โปรแกรม VocabProfile วิเคราะห์ข้อมูลได้ เพราะตัวโปรแกรมจำกัดจำนวนบทความและคำในการวิเคราะห์ไม่เกิน 400,000 คำ ดังนั้นผู้วิจัยจึงใช้โปรแกรม AntConc เพราะสามารถจุข้อมูลได้มากกว่า และสามารถแสดงความถี่ของคำได้ หลังจากนั้นจึงเอาเฉพาะคำเนื้อหาที่ได้จากรายการคำศัพท์นี้ไปจำแนกต่อไป
3. จำแนกรายการคำศัพท์ออกเป็น GSL, AWL และ OWL โดยใช้โปรแกรม VocabProfile และเลือกการจำแนกกลุ่มคำศัพท์ดังกล่าว โดยใช้ฟังก์ชัน New General Service List (NGSL) และ New

Academic Word List (NAWL) ซึ่งฟังก์ชันนี้เป็นรายการคำศัพท์ที่ได้รับการปรับปรุงใหม่โดย Browne (2013)

4. หาคำที่มีความถี่สูงใน OWL โดยวิธีการเรียงลำดับคำที่มีความถี่สูงที่สุดไปน้อยที่สุด จากนั้นใช้ควอร์ไทล์แบ่งคำศัพท์ออกเป็น 4 กลุ่ม ซึ่งคำศัพท์ที่อยู่ในควอร์ไทล์ที่ 1 คือคำศัพท์ที่มีความถี่สูง
5. นำคำที่มีความถี่สูงที่สุดมาวิเคราะห์ Lexical collocation โดยใช้ฟังก์ชัน Concordance ในโปรแกรม AntConc

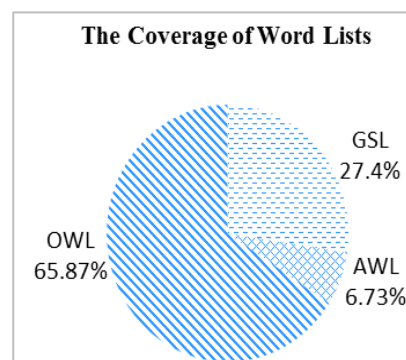
การวิเคราะห์ข้อมูล

สถิติที่ใช้ในการข้อมูลวิเคราะห์จากรายการคำศัพท์ คือ ความถี่ ร้อยละ และ ควอร์ไทล์

ผลการวิจัย

1. คำที่มีความถี่สูงใน OWL

จากการศึกษาพบว่าอัตราส่วนของคำศัพท์ที่ปรากฏใน OWL มีอัตราส่วนสูงที่สุด โดยมีอัตราส่วนครอบคลุมเกินครึ่งของคลังข้อมูลภาษา อัตราส่วนของ GSL ครอบคลุมรองมาเป็นอันดับสอง และอัตราส่วนของ AWL ครอบคลุมคลังข้อมูลภาษาน้อยที่สุด (ดูกราฟที่ 1) สาเหตุที่อัตราส่วนของ AWL ลดลง ทั้งๆที่เป็นคำศัพท์ที่รวบรวมจากบทความวิชาการเป็นเพราะ แหล่งของคลังข้อมูลเก็บจากบทความเฉพาะทางของสัตว์ทดลอง ซึ่งเป็นเพียงสาขาหนึ่งของวิทยาศาสตร์ และเก็บจากแหล่งข้อมูลเดียวคือ บทความวิชาการ จึงทำให้เป็นคลังข้อมูลเฉพาะทาง และคำศัพท์ส่วนใหญ่เป็นคำศัพท์เฉพาะ คลังข้อมูลที่ศึกษานี้ จึงไม่ใหญ่พอที่จะทำให้เกิดคำศัพท์ที่หลากหลาย หรือคำศัพท์ในกลุ่มของ AWL ที่ใช้ในสาขาอื่นๆ ดังนั้นคำศัพท์ที่มีความถี่สูงในงานวิจัยชิ้นนี้จะคำนวณจากคำศัพท์ในกลุ่ม OWL เพราะ OWL เป็นรายการคำศัพท์ที่ครอบคลุมเนื้อหาส่วนมากในคลังข้อมูลภาษาของการศึกษาครั้งนี้ และเป็นคำศัพท์เฉพาะที่ปรากฏอยู่ในสาขาสัตว์ทดลอง (ดูตารางที่ 1)

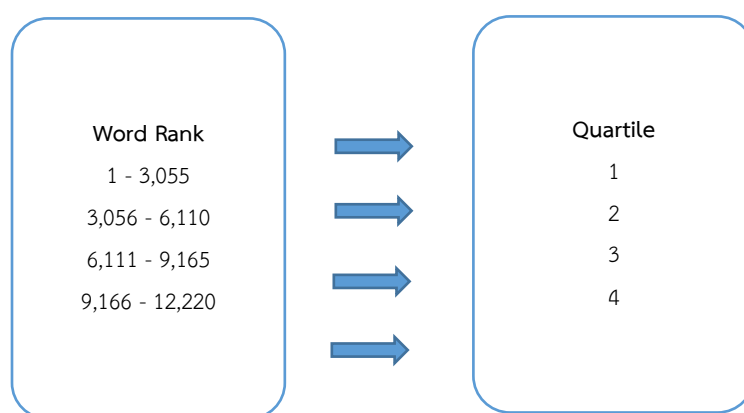


กราฟที่ 1 อัตราส่วนของคำศัพท์ในคลังข้อมูลภาษาสัตว์ทดลอง

ตารางที่ 1 คำที่มีความถี่สูงสุดใน OWL 50 คำแรก

Rank	f	Word	Rank	f	Word	Rank	f	Word	Rank	f	Word
1	675	zebrafish	14	188	macaques	27	160	bartonella	40	134	plasma
2	533	methylation	15	186	rodent	28	159	vervet	41	128	anesthesia
3	435	epigenetic	16	185	primates	29	158	transgenic	42	128	baboons
4	372	cocaine	17	185	swine	30	156	activation	43	128	lifespan
5	284	genome	18	177	ketamine	31	154	investigators	44	121	susceptibility
6	254	canine	19	177	viral	32	147	promoter	45	120	intake
7	237	nicotine	20	175	invertebrates	33	145	extinction	46	119	phenotype
8	219	rodents	21	175	nonhuman	34	145	obesity	47	118	addiction
9	216	histone	22	170	genomic	35	143	metabolic	48	118	biomedical
10	206	wildlife	23	168	pet	36	142	therapeutic	49	117	cage
11	198	rhesus	24	165	influenza	37	141	alterations	50	115	imprinted
12	196	neuropathy	25	162	lesions	38	141	diabetes			
13	195	veterinary	26	161	primate	39	141	insulin			

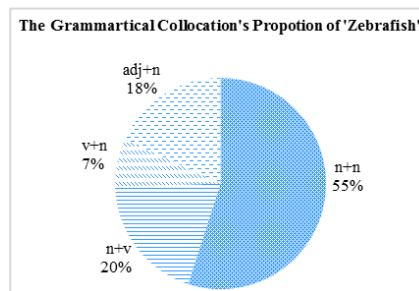
จากการศึกษาพบว่า OWL มีคำศัพท์ทั้งหมด 12,220 คำ ดังนั้นคำศัพท์ที่อยู่ในช่วง ควอร์ไทล์ที่ 1 (ลำดับที่ 1 ถึง 3,055) จึงเป็นคำศัพท์ที่มีความถี่สูง (รูปภาพที่ 1)



รูปภาพที่ 1 การแบ่งควอร์ไทล์

2. Lexical collocation ของ คำที่ปรากฏสูงที่สุดใน OWL

จากตารางที่ 1 แสดงให้เห็นว่าคำที่มีความถี่สูงสุดคือคำว่า zebrafish ดังนั้น zebrafish จึงถูกนำมาวิเคราะห์เพื่อหา lexical collocation ซึ่งผลการวิจัยพบว่า การปรากฏร่วมจำเพาะของ zebrafish ที่พบมากที่สุดคือ ปรากฏร่วมกับคำนาม (ดูกราฟที่ 2) การปรากฏร่วมจำเพาะของคำนามกับคำนามนี้เกิดมากเกินครึ่งของการปรากฏร่วมจำเพาะทั้งหมด ซึ่งส่วนมากอยู่ในลักษณะคำนามประสมประเภทหลัก (primary compound noun) (ดูตารางที่ 2) และ เกิดตามหลังคำบุพบท (preposition) (ดูตารางที่ 3)



กราฟที่ 2 อัตราส่วน Lexical collocation

ตารางที่ 2 ประเภทคำนามประสมของคำที่ปรากฏสูงที่สุดใน OWL (zebrafish)

Type of Compound Noun		Examples	Total (%)
Primary	n+n	1. The present review focuses specifically on the transmission and control of microsporidia in zebrafish facilities . 2. Certain other infectious agents that often cause profound hyperplasia in zebrafish tissues , such as Piscinoodinium pillulare in the gill, have not acted as a tumor promoter in any carcinogen experiments that we have conducted.	96.88%
	n(ing)+n	1. It is important to note that for all of the following performance-based approaches for ethanizing zebrafish by submersion in ice water (2-4°C), the fish do not come into direct contact with the ice.	0.67%
Secondary	n+n(ing)	1. One of the major concerns related to zebrafish breeding and line propagation in the laboratory is preservation of the vigor of a line. 2. Goolish and colleagues have evaluated chamber volume requirements for zebrafish spawning .	2.43%

ตารางที่ 3 ตำแหน่งของคำนามประสมของคำที่ปรากฏสูงสุดใน OWL (*zebrafish*)

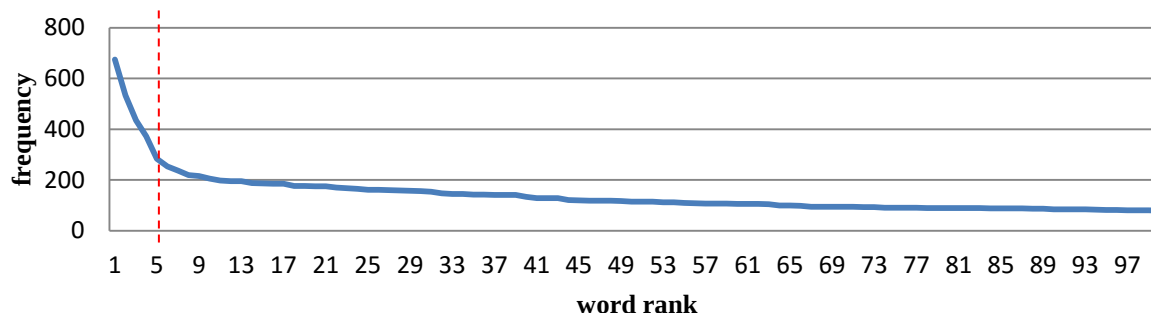
Position of Compound Noun		Examples	Total (%)
Determiner	general	1. The Sinhuber Aquatic Research Laboratory at Oregon State University has established <u>an</u> SPF zebrafish facility in which <i>P. neurophilia</i> has been eradicated, but mycobacteriosis is occasionally diagnosed by histopathology in clinically normal sentinel fish or 1-year-old retired brood stock at this facility. 2. Another microsporidium, <i>Pleistophora hypohessobryconis</i>, has recently been detected in <u>a few</u> zebrafish facilities.	10.76%
	specific	1. During the past decade, <u>the</u> zebrafish has emerged as a leading model for mechanistic cancer research because of its sophisticated genetic and genomic resources, its tractability for tissue targeting of transgene expression, its efficiency for forward genetic approaches to cancer model development, and its cost effectiveness for enhancer and suppressor screens once a cancer model is established. 2. Zebrafish aquaculture housing systems combining design principles from industrial aquaculture, laboratory rodent housing, and research genetics are commercially available from a number of sources, and an increasing number of academic research institutions are constructing large centralized facilities to support <u>their</u> growing zebrafish research programs. 3. The lengthy incubation time and specialized medium may contribute to the underdiagnosis of <u>this</u> important zebrafish pathogen.	24.22%
Preposition	simple	1. Occurrence of identical strains of <i>Mycobacterium</i> in both fish and biofilms <u>in</u> zebrafish systems suggests transmission can occur when fish feed on infected tissues or tank detritus containing mycobacteria. 2. To the best of our knowledge, there are no documented cases of human mycobacteriosis associated <u>with</u> zebrafish handling.	60.54%
	compound	1. Xenografts of human tumors <u>into</u> zebrafish embryos and zebrafish cancer models allowing tumor induction very early in life provide a cost-effective, high-throughput system for drug discovery.	4.48%

อภิปรายผลการวิจัย

1. คำที่มีความถี่สูงใน OWL

งานวิจัยครั้งนี้ให้ความสำคัญกับกลุ่มคำศัพท์ที่มีความถี่สูงใน OWL เพราะเป็นคำศัพท์ที่ปรากฏถี่ในสาขาวิชาเฉพาะทาง Nation (2001) กล่าวว่า คำศัพท์ที่มีความถี่สูงจะต้องปรากฏประมาณร้อยละ 80 ของคำศัพท์ทั้งหมด แต่เนื่องจาก OWL เป็นคำศัพท์เฉพาะทางที่ไม่ได้มีความถี่สูงหรือปรากฏบ่อยเท่าคำศัพท์ทั่วไป ดังนั้นผลของ OWL ในการศึกษาครั้งนี้จึงปรากฏไม่ถึงร้อยละ 80 ด้วยเหตุนี้ควอร์ไทล์จึงใช้เป็นเกณฑ์ในการพิจารณาหาคำที่มีความถี่สูง จากการพล็อตกราฟ ดังกราฟที่ 3 แสดงให้เห็นว่า กราฟของห้าอันดับคำแรกที่มีความถี่สูงสุดลดลงอย่างมาก ซึ่งต่างจากคำอื่นๆที่ค่อยๆลดลงทีละน้อย สิ่งนี้ชี้ให้เห็นว่าห้าคำแรกนี้ (zebrafish, methylation, epigenetic, cocaine, และ genome) ปรากฏถี่มากกว่าคำอื่นๆในงานวิจัยสาขาสัตว์ทดลอง

The Frequency of the Top 100 OWL



กราฟที่ 3 ความถี่ของ OWL 100 คำแรก

2. Lexical collocation (noun + noun)

งานวิจัยชิ้นนี้มีวัตถุประสงค์เพื่อหาการปรากฏร่วมจำเพาะ ซึ่งผลการวิจัยพบว่า lexical collocation ของคำศัพท์ 'zebrafish' ปรากฏคู่กับคำนามด้วยความถี่สูงสุด คิดเป็นร้อยละ 55 ดังนั้นการเกิดร่วมกันของคำนามกับคำนามจึงทำให้เกิดคำนามประสมขึ้น ซึ่งคำนามประสมที่ปรากฏส่วนมากเป็นคำนามประสมประเภทหลัก กล่าวคือเป็นคำนามประสมที่มีความหมายมาจากคำนามหลักและคำนามที่ขยาย (McCarthy, 2002) ยกตัวอย่างเช่น zebrafish model, zebrafish line และ zebrafish system นอกจากนี้ zebrafish ยังทำหน้าที่คล้ายคำคุณศัพท์ขยายคำนามหลักที่ปรากฏร่วม ซึ่งผลงานวิจัยครั้งนี้แสดงให้เห็นว่า นักวิทยาศาสตร์ใช้คำนามประสมมากที่สุด คือ ประมาณครึ่งหนึ่งของการปรากฏร่วมจำเพาะทั้งหมด การลดรูปประโยคโดยการสร้างคำนามหรือนามวลีด้วยวิธีการ Nominalization เป็นวิธีหนึ่งที่ใช้ในการเขียนเชิงวิชาการ ซึ่งการวิจัยครั้งนี้พบการใช้ Nominalization จำนวน 55 คำ (19.1%) ของการเกิดร่วมจำเพาะแบบคำนามประสมทั้งหมด และสอดคล้องกับ Fries (1952) Quirk (1972) McCarthy (2002) Baker (2003) Diarmuid (2008) และ Lieber

(2010) ที่กล่าวว่า Nominalization เป็นการทำให้ประโยคกระชับขึ้นและถือว่าเป็นประเภทหนึ่งของคำนามประสม นอกจากนี้คำนามประสมเป็นการสร้างคำใหม่โดยการนำคำนามกับคำนามมารวมกัน ทำให้คำสองคำมีความสำคัญเท่ากัน ทั้งนี้ การที่นักวิทยาศาสตร์ใช้คำนามประสมมากที่สุดอาจอธิบายได้ดังต่อไปนี้

1. งานเขียนครั้งนี้เป็นการเขียนเชิงวิชาการ วารสารจำกัดจำนวนคำในการส่งงานวิจัยเพื่อตีพิมพ์ ดังนั้นนักวิทยาศาสตร์จึงเลือกใช้คำนามประสม เพื่อสามารถเขียนงานวิจัยและอธิบายคำศัพท์เฉพาะได้อย่างกระชับ รัดกุม แต่ให้ความหมายตรงตามที่นักวิทยาศาสตร์ต้องการสื่อออกไปถึงผู้อ่าน ยกตัวอย่างเช่น zebrafish sedation, zebrafish development, zebrafish user, euthanizing zebrafish, zebrafish study, zebrafish reproduction และ zebrafish application
2. การสร้างคำนามประสมบางคำไม่สามารถเติมปัจจัย (suffix) เพื่อเปลี่ยนคำนามให้เป็นคำคุณศัพท์ได้ โดยเฉพาะคำศัพท์เทคนิคหรือคำศัพท์เฉพาะที่ใช้ในสาขาสัตว์ทดลอง เช่น zebrafish ดังนั้นนักวิทยาศาสตร์จึงจำเป็นต้องใช้คำนามทำหน้าที่ขยายคำนาม เป็นการใช้คำนามประสมสร้างคำใหม่ เช่น zebrafish housing, zebrafish mate และ zebrafish line
3. นักวิทยาศาสตร์ต้องการสื่อความหมายของคำสองคำให้มีศักดิ์ศรีเท่ากันหรือมีความสำคัญเท่ากัน นักวิทยาศาสตร์จึงใช้คำนามสองคำทำเป็นคำนามประสม เช่น zebrafish colony, zebrafish adult, zebrafish male และ zebrafish facility

ผลการวิจัยยังพบว่า กลุ่มคำนามประสมเหล่านี้ทำหน้าที่ด้วยคำกำกับคำนาม (determiner) คิดเป็นร้อยละ 24.22 ส่วนใหญ่กลุ่มคำนามที่พบจะใช้คำกำกับคำนามแบบชี้เฉพาะเจาะจงในรูป definite article หรือ the ซึ่งสอดคล้องกับรายการคำศัพท์ GSL ของ Browne (2013) ที่พบว่ามีการใช้ the สูงเป็นอันดับหนึ่งในส่วนของกลุ่มคำนามที่ตามหลังด้วยคำบุพบท คิดเป็นร้อยละ 69.02 เมื่อกลุ่มคำนามทำหน้าที่ด้วยคำบุพบทจะทำให้กลุ่มคำนามนั้นอยู่ในรูปบุพบทวลี (prepositional phrase) ผลการวิจัยตำแหน่งของคำนามประสมนี้ จึงมีประโยชน์ต่อการเรียนการสอนภาษาอังกฤษ ทำให้ผู้สอนและผู้เรียนตระหนักถึงวิธีการใช้คำนามประสมสร้างประโยคได้อย่างถูกต้อง โดยครูผู้สอนควรให้ความสำคัญในการสอนเรื่องคำกำกับคำนามแบบชี้เฉพาะเจาะจง และกลุ่มคำนามในรูปบุพบทวลีควบคู่กับการสอนเรื่องคำนามประสม

สรุปผลการวิจัย ข้อจำกัด และข้อเสนอแนะ

สรุปผลการวิจัย

จากการทบทวนวรรณกรรมพบว่าไม่มีการศึกษาด้านคลังภาษาศาสตร์ของสาขาสัตวทดลอง งานวิจัยชิ้นนี้จึงเป็นการสร้างองค์ความรู้ใหม่ทางด้านคำศัพท์ในสาขาวิชาสัตวทดลอง โดยผู้วิจัยได้รวบรวมรายการคำศัพท์ OWL ซึ่งเป็นรายการคำศัพท์เฉพาะที่ใช้ในสาขาวิชาสัตวทดลอง ผู้วิจัยได้ใช้รายการคำศัพท์ GSL และ

AWL ของ Browne (2013) ซึ่งเป็นรายการคำศัพท์ที่ปรับปรุงขึ้นใหม่ในการอ้างอิงเพื่อจำแนกคำศัพท์ในงานวิจัยชิ้นนี้ จากการวิจัยพบว่า หัวอันดับแรกได้แก่คำว่า zebrafish, methylation, epigenetic, cocaine, และ genome มีความถี่สูงโดดเด่นกว่าคำอื่นในคลังข้อมูลสัตว์ทดลอง นอกจากนี้ผู้วิจัยยังได้วิเคราะห์ Lexical collocation ของคำที่ปรากฏสูงที่สุดซึ่งอยู่ในกลุ่มคำนามที่มีลักษณะเป็นคำนามประสม เพื่อเป็นประโยชน์ต่อนักศึกษาและนักวิจัยที่ต้องอ่านและเขียนงานวิจัยเป็นภาษาอังกฤษ

ข้อจำกัดและข้อเสนอแนะในการศึกษาครั้งต่อไป

ข้อจำกัดการวิจัยครั้งนี้ คือ ผู้วิจัยใช้วิธีการสุ่มในการเลือกบทความและใช้ความถี่ของคำที่ปรากฏร่วมเฉพาะในการคัดเลือกคำที่นำมาวิเคราะห์ทำให้ทราบแนวโน้มหรือประเด็นที่นักวิทยาศาสตร์สาขาสัตว์ทดลองทำการศึกษาวิจัย ในการศึกษาครั้งต่อไป อาจทำการศึกษาเพิ่มเติมดังนี้

1. สร้างคลังข้อมูลที่ใหญ่ขึ้นและให้ความหลากหลายทางวิธภาษา (variety) โดยครอบคลุมทุกประเภทตัวบท (text types) และทุกทำเนียบภาษา (register) เช่น การนำเสนอผลงานในการประชุมวิชาการ และ หนังสือเรียน
2. พิจารณาคำในการศึกษาให้เฉพาะเจาะจง เช่น ข้อจำกัดการเกิดร่วมกัน (selectional restriction) ระดับขั้นของความเป็นธรรมชาติ (degree of naturalness) หรือ คำที่เป็นปัญหาหรือเป็นประโยชน์ต่อนักวิทยาศาสตร์สาขาสัตว์ทดลอง ในการพิจารณาคำที่จะศึกษานี้ ผู้วิจัยอาจเลือกคำปรากฏร่วมที่มีความเฉพาะเจาะจงกับคำแต่ละคำหรือตัวบทเนื่องจากคลังข้อมูลในการศึกษาครั้งนี้เป็นคลังข้อมูลเฉพาะ (specialized corpus)
3. ทำการวิเคราะห์ภาษาของคำศัพท์เพิ่มเติม เช่น หน้าที่ในการสื่อสาร (communicative function) และ อรรถภาค (Move analysis)

ข้อเสนอแนะในการวิจัย ผู้สอนวิชาการอ่านหรือภาษาอังกฤษเชิงวิชาการ ควรชี้ให้นักศึกษาเห็นความสำคัญและประโยชน์ของคำนามประสม อธิบายและ ฝึกการสร้างคำนามผสม โดยเฉพาะ nominalization เพื่อให้นักศึกษาจะได้เข้าใจ มีความรู้ ทักษะ nominalization แล้วนำไปใช้ในการอ่านและเขียนภาษาอังกฤษต่อไป

เอกสารอ้างอิง

- ทรงศรี สรณสถาพร. (2557). *การอ่านและเขียนภาษาอังกฤษอย่างมีประสิทธิภาพ*. กรุงเทพฯ : สำนักพิมพ์วังอักษร.
- Anthony, L. (2014). AntConc (Version 3.4.3) [Computer Software]. Tokyo, Japan: Waseda University. Available from <http://www.laurenceanthony.net>
- Bahns, J. (1993). Lexical collocations: A contrastive view. *ELT Journal*, 47(1), 56-63.
- Baker, M. (2003). *Lexical categories: verbs, nouns, and adjectives*. Cambridge: Cambridge University Press.
- Basturkmen, H. (2010). *Developing courses in English for specific purposes*. Basingstoke: Palgrave Macmillan.
- Belcher. D., Johns. A., & Paltridge. B. (2011). *New directions in English for specific purposes Research*. Michigan: The University of Michigan Press.
- Benson, M. (1985). *Collocations and idioms*. In R. Ilson (Ed.), *Dictionaries, lexicography and language learning*. British Council: Pergamon Press.
- Biber, D., Conrad, S., & Reppen, R. (2000). *Corpus linguistics*. Cambridge: Cambridge University Press.
- Biskup, D. (1992). Li influence on learners' renderings of English collocation: A Polish/German empirical study. In P.J.L. Arnauld, & H. Bejoint (Eds.), *Vocabulary and applied linguistics*. London: Macmillan.
- Browne, C. et al. (2013). The new general service list: Celebrating 60 years of vocabulary learning. *The Language Teacher*, 37 (4), 13-16.
- Bruce, I. (2011). *Theory and concepts of English for academic purposes*. New York, NY: Palgrave Macmillan.

Chung, T., & Nation, P. (2004). Identifying Technical Vocabulary. *System*, 32(2), 251-263.

Cobb, T. & Horst, M. (2001). Growing academic vocabulary with a collaborative online database. In B. Morrison, D. Gardner, K. Koebeke, & M. Spratt (Eds.), *LT Perspectives on IT & Multimedia* (pp. 189-226). Hong Kong: Polytechnic University Press.

Coxhead, A. & Hirsh, D. (2007). A pilot science word list for EAP. *Revue Française de linguistique appliquée xii*, 2, 65-78.

Coxhead, A. (2000). A New Academic Word List, *TESOL Quarterly*, 34(2), 213-238.

Department of Foreign Language. (2004). *Reading and writing for first year medical students, Faculty of Medicine, Mahidol University*.

Diarmuid, S. (2008). *Learning compound noun semantics*. Cambridge: Cambridge University Press.

Downing, A. & Locke, P. (1992). *University course in English grammar*. London: Prentice Hall International.

Durrant, P., & Schmitt, N. (2010). Adult learners' retention of collocations from exposure. *Second Language Research*, 26(2), 163-188.

Flowerdew, J. A. & Peacock, M. (2001). *Research perspective on English for academic purposes*. Cambridge: Cambridge University Press.

Folse, K. (2010). Is explicit vocabulary focus the reading teacher's job?. *Reading in a Foreign Language*, 22, 139-160.

Fries, C. (1952). *The structure of English*. New York: Harcourt Brace.

Froehlich, H. (2015). Corpus analysis with Antconc. *Programming Historian*. Retrieved from <http://programminghistorian.org/lessons/corpus-analysis-with-antconc.html>

- Halliday, M. & Matthiessen, C. (2014). *Halliday's introduction to functional grammar* (4th ed.). New York: Routledge.
- Hill, J. (2000). Teaching collocation. In M. Lewis (Ed.), *Revising priorities: From grammatical failure to collocational success* (pp. 47-69). Hove: Language Teaching Publications.
- Hunston, S. (2011). *Corpus approaches to evaluation*. New York: Routledge.
- Huston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.
- Hutchinson, T. & Waters, A. (1987). *English for specific purposes: A learning-centred approach*. Cambridge: Cambridge University Press.
- Jolly, D., & Bolitho, R. (2011). The process of materials writing. In B. Tomlinson (Ed.), *Materials development and language teaching*. Cambridge: Cambridge University Press.
- Kirkpatrick, A. (2012). English as an international language in Asia: Implications for language education. In A. Kirkpatrick & A. Sussex (Eds.), *English as an international language in Asia: Implications for language education* (pp. 29-44). London: Springer.
- Lieber, R. (2010). *Introducing morphology*. Cambridge: Cambridge University Press.
- Liu, J., & Han, L. (2015). A corpus-based environmental academic word list building and its validity test. *English for Specific Purposes*, 39, 1-11.
- Martinez, I., Beck, S. & Panza, C. (2009). Academic vocabulary in agriculture research articles: A corpus-based study. *English for Specific Purposes*, 28, 183-198.
- McCarthy, A. (2002). *An introduction to English morphology*. Edinburgh: Edinburgh University Press.
- Montgomery, S., & Crystal, D. (2013). *Does Science need a global language? : English and the future of research*. Chicago: University Press.

Murphy, M. (2010). *Lexical meaning*. Cambridge: Cambridge University Press.

Nation, P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Quirk, R. (1972). *A Grammar of contemporary English*. London: Longman.

Sharifian, F. (2009). *English as an international language: Perspectives and pedagogical issues*. UK: Multilingual Matters.

Staples, S. (2015). Examining the linguistic needs of internationally educated nurses: A corpus-based study of lexico-grammatical features in nurse–patient interactions. *English for Specific Purposes*, 37, 122-136.

Tomlinson, B. (2012). *Materials development for language learning and teaching*. Cambridge: Cambridge University Press.

Valipouri, L., & Nassaji, H. (2013). A corpus-based study of academic vocabulary in chemistry research articles. *English for Academic Purposes*, 12(4), 248-263.

Walker, C. (2011). How a corpus-based study of the factors which influence collocation can help in the teaching of business English. *English for Specific Purposes*, 30(2), 101-112.

West, M. (1953). *A general service list of English words*. London: Longman, Green & Co.