

รายการคำศัพท์ที่มีความถี่สูงสุดในบทความปริทัศน์สาขาสัตวทดลอง High Frequency Word List in Laboratory Animal Review Articles

ทรงศรี สรณสถาพร

สิริวิชัย ธรรมพานิช

บทคัดย่อ

การศึกษานี้มีวัตถุประสงค์เพื่อ (1) เปรียบเทียบความถี่ของ “คำเนื้อหา” และ “คำไวยากรณ์” ที่ปรากฏในรายการคำศัพท์ที่มีความถี่สูงสุดในคลังข้อมูลภาษาสาขาสัตวทดลองและ (2) เพื่อค้นหาและรายงานคำเนื้อหา (content words) 100 คำแรกที่มีความถี่สูงสุดในบทความวิจัยสัตวทดลอง ตัวอย่างที่ใช้ในการศึกษาคือ เนื้อหาของบทความปริทัศน์สัตวทดลองในวารสาร ILAR จำนวน 160 บทความ ย้อนหลัง 5 ปี ตั้งแต่ปี พ.ศ. 2553 ถึงปี พ.ศ. 2557 ใช้โปรแกรม AntConc เวอร์ชัน 3.4.4 (Anthony, 2014) วิเคราะห์หาความถี่ของ “คำเนื้อหา” และ “คำไวยากรณ์” เกณฑ์การพิจารณา รายการคำศัพท์ที่มีความถี่สูงสุดคือ รายการคำศัพท์ชุดนั้นจะต้องมีจำนวนร้อยละ 80 ของจำนวนคำศัพท์ทั้งหมด (Nation, 2001) สถิติที่ใช้ในการวิเคราะห์ข้อมูลคือ ค่าความถี่และร้อยละ คลังข้อมูลภาษาสาขาสัตวทดลองประกอบด้วย word tokens จำนวน 840,773 คำ word types มีจำนวน 22,434 คำ คำศัพท์ที่ปรากฏสูงสุด 2,000 คำแรกคิดเป็นร้อยละ 83.65 ของคำทั้งหมด และคำเนื้อหาที่ปรากฏใน 2,000 คำแรกคิดเป็นร้อยละ 40.13 ผลการวิจัยพบว่า (1) ความถี่ของ “คำเนื้อหา” (55.85%) ที่ปรากฏในคลังข้อมูลภาษาของบทความปริทัศน์สัตวทดลองสูงกว่า ความถี่ของ “คำไวยากรณ์” (44.15%) (2) คำเนื้อหา (content words) 100 คำแรกที่มีความถี่สูงสุดในบทความวิจัยสัตวทดลองเป็นคำศัพท์พื้นฐาน เช่น studies ($f = 3031, 0.36\%$) และ animal ($f = 2942, 0.35\%$)

คำสำคัญ: คำที่มีความถี่สูงสุด คำเนื้อหา คำไวยากรณ์ สัตวทดลอง

Abstract

This study aimed (1) to compare the frequency of content words and function words in a corpus of laboratory animal journals review articles (LARA), and (2) to explore and report the first 100 content words with the highest frequency the LARA corpus. The corpus of LARA included 160 review articles spanning the years 2010 to 2014 from the *Journal of the Institute for Laboratory Animal Research* (ILAR). The concordance software AntConc Version 3.4.4 (Anthony, 2014) was the research instrument used to investigate the high frequency of content words and function words. The criterion of investigating the high frequency word list was calculated as 80% of the running words (Nation, 2001). The frequency and percentage were used for data analysis. The LARA corpus included 840,773 word tokens and 22,434 word types. The first 2,000 high frequency words were calculated at 83.65%

of the running words which consisted of the content words (40.13%). The results revealed that: (1) the frequency of content words (55.85%) was higher than that of function words (44.15%) in the LARA corpus, and (2) the first 100 content words with the highest frequency in the LARA corpus were basic words such as “studies” ($f = 3031$, 0.36%) and “animal” ($f = 2942$, 0.35%).

Keywords: high frequency words, content words, function words, laboratory animal, corpus linguistics

1. บทนำ

คำศัพท์ภาษาอังกฤษมีบทบาทสำคัญต่อผู้เรียนภาษาอังกฤษในฐานะภาษาที่สอง (L2 learner) หรือ ภาษาต่างประเทศ (English as a Foreign Language: EFL Learner) เพื่อใช้ในการศึกษาผลงานทางวิชาการ (Nation, 2001; Gass & Selinker, 2008) ผู้เรียนจำเป็นต้องอ่านและเขียนบทความทางวิชาการเป็นภาษาอังกฤษ โดยเฉพาะผู้เรียนในสาขาวิทยาศาสตร์ที่เรียนภาษาอังกฤษเป็นภาษาที่สองหรือ ภาษาต่างประเทศจำเป็นต้องมีความรู้พื้นฐานเกี่ยวกับภาษาอังกฤษ กล่าวคือ ในศตวรรษที่ 20 บทความวิจัยหรือหนังสือเรียนในสาขาวิทยาศาสตร์ธรรมชาติ (natural science) มากกว่าร้อยละ 90 ตีพิมพ์เป็นภาษาอังกฤษ (Hamel, 2007) นอกจากนี้ ภาษาอังกฤษยังเป็นภาษากลาง (lingua franca) ในกลุ่มผู้ศึกษางานทางวิทยาศาสตร์ เนื่องจากว่าวารสารในฐานข้อมูล Scopus ประมาณร้อยละ 80 ตีพิมพ์เป็นภาษาอังกฤษ กล่าวคือ มากกว่า 210,000 บทความวิชาการจาก 239 ประเทศในฐานข้อมูล Scopus ตีพิมพ์เป็นภาษาอังกฤษ (Van, 2012) เพราะฉะนั้น ผู้เรียนในสาขาวิทยาศาสตร์ที่เรียนภาษาอังกฤษเป็นภาษาที่สองจึงจะต้องเข้าใจคำศัพท์ภาษาอังกฤษ (vocabulary) รูปแบบประโยค (sentence pattern) และ/หรือ รูปแบบทางวาทกรรม (discourse pattern) เพื่อว่าผู้เรียนจะได้เข้าใจบทความวิจัยหรือเนื้อหาทางวิทยาศาสตร์ที่ตีพิมพ์เป็นภาษาอังกฤษได้อย่างถ่องแท้ (Nation, 2001; Paltridge & Starfield, 2013) ที่สำคัญไปกว่านั้น คำศัพท์ภาษาอังกฤษทางด้านวิทยาศาสตร์เป็นประเด็นหนึ่งที่สำคัญในการเรียนการสอนภาษาอังกฤษเพื่อจุดมุ่งหมายเฉพาะ (English for specific purposes หรือ ESP) คำศัพท์ในกลุ่มนี้จึงเป็นปัจจัยพื้นฐานสำหรับผู้เรียนภาษาอังกฤษในฐานะภาษาที่สองหรือภาษาต่างประเทศในสาขาเฉพาะ (Trimble, 1985; Hutchison & Waters, 1987; Paltridge & Starfield, 2013) และยังส่งผลให้ผู้เรียนเข้าใจและใช้คำศัพท์ภาษาอังกฤษได้อย่างเหมาะสมกับสาขาเฉพาะกลุ่ม (Paltridge & Starfield, 2013) ในการศึกษาคำศัพท์ในบทความวิชาการในสาขาวิทยาศาสตร์จึงมีความจำเป็นต่อผู้เรียนภาษาอังกฤษในฐานะภาษาที่สองและมีผู้วิจัยหลายท่านให้ความสนใจกับการศึกษาคำศัพท์โดยการใช้คลังข้อมูลภาษา (corpus) มากขึ้น (Valipouri & Nassji, 2013)

ความถี่ของคำ (word frequency) เป็นเกณฑ์ที่สำคัญในการพิจารณาคำศัพท์ ในคลังข้อมูลภาษา เพราะคำศัพท์ที่ต้องใช้ในรูปแบบที่ต่างกันจะขึ้นอยู่กับบริบทการใช้คำศัพท์นั้น ๆ (Schmitt, 2000) คำศัพท์เหล่านี้ปรากฏขึ้นทั้งในตัวบทภาษาพูดและภาษาเขียน (spoken and written text) ผู้เรียนภาษาอังกฤษในฐานะภาษาที่สองหรือภาษาต่างประเทศ จึงต้องมีคลังข้อมูลคำศัพท์ที่มากพอ (Nation, 2001) นอกจากนี้ผู้สอนควรสอนรายการคำที่มีความถี่สูงสุด (High Frequency Word) ให้กับผู้เรียนก่อน รายการคำศัพท์อื่น ๆ เช่น รายการคำศัพท์ทางวิชาการ (Academic Word List: AWL) หรือรายการคำศัพท์ที่ไม่ได้อยู่ในรายการคำศัพท์ทางวิชาการ (Outside Word List: OWL) ซึ่งนักภาษาศาสตร์จัดรายการคำศัพท์กลุ่ม OWL นี้เป็นรายการคำศัพท์เฉพาะ (technical term)

Nation (2001) ได้กล่าวไว้ว่า คำที่มีความถี่สูงสุดรวมคำศัพท์ทั้งสองประเภทใหญ่ ๆ ไว้กล่าวคือ content word และ function word คำว่า “content word” ได้บัญญัติเป็นคำภาษาไทยไว้หลายคำ ได้แก่ “คำบอกเนื้อหา” (น้ำทิพย์, 2551; ดิยูและคณะ, 2558) “คำแสดงเนื้อหา” (นันทนา, 2555) และ “คำเนื้อหา” (ณัฐมา, 2554; ราชบัณฑิต, 2557) ดังนั้นในบทความฉบับนี้จะยึดตามราชบัณฑิตคือ “คำเนื้อหา” ส่วนคำว่า “function word” ได้บัญญัติเป็นคำภาษาไทยไว้ว่า “คำไวยากรณ์” (ราชบัณฑิต, 2546; น้ำทิพย์, 2551; ณัฐมา, 2554; นันทนา, 2555; ดิยูและคณะ, 2558) ดังนั้นในบทความฉบับนี้จะใช้ศัพท์บัญญัติดังกล่าว ในทางภาษาศาสตร์ คำเนื้อหา หมายถึง คำที่ให้ความหมาย อาจสื่อความถึงสิ่งของ อากาศ การกระทำ บุคคล สถานที่ สภาวะและการกระทำ เป็นต้น (Kantamba & Stonham, 2006; Murphy, 2010; ณัฐมา, 2554; ดิยูและคณะ, 2558) ผู้ใช้ภาษาสามารถสร้างคำเนื้อหาเป็นคำใหม่ได้และเกิดความหมายใหม่ขึ้นได้ (ณัฐมา, 2554) ดังนั้น คำเนื้อหามักจะเป็นหน่วยคำอิสระเทียบเท่ากับคำหนึ่งคำ และจัดอยู่ในหมวดคำนาม คำกริยา คำคุณศัพท์ และคำวิเศษณ์ (Murphy, 2010; ณัฐมา, 2554; ดิยูและคณะ, 2558) เช่น bag, lesson, person, put, และ suggest ส่วนคำไวยากรณ์ (function word) หมายถึง คำที่มีหน้าที่แสดงความสัมพันธ์ทางไวยากรณ์ ซึ่งคำไวยากรณ์จะเกิดขึ้นใหม่น้อยมาก คำไวยากรณ์จัดอยู่ในหมวดคำสรรพนาม (เช่น everyone, such, he) คำนำหน้านาม (เช่น a, an, the, such) คำบุพบท (เช่น between, off, whereas) คำสันธาน (เช่น and, or, whereas) คำแสดงตัวเลข (เช่น billion, eight, ten) (Nation, 2001; Murphy, 2010; ณัฐมา, 2554; ดิยูและคณะ, 2558)

ผลการศึกษาค้นคลังข้อมูลความถี่ของคำในบทความวิจัยทางวิชาการที่ผ่านมาพบว่า รายการคำศัพท์ที่มีความถี่สูงสุดนั้นมีชนิดของคำที่หลากหลายและให้ความหมายแตกต่างกัน (Schmitt, 2000) รายการคำศัพท์ที่มีความถี่สูงสุด คือ รายการคำศัพท์ที่ครอบคลุม 2,000 คำแรกหรือประมาณร้อยละ 80 ของบทความทั้งหมด (text coverage) และในจำนวนคำศัพท์ 2,000 คำแรกนี้ คำศัพท์ 1,000 คำแรกครอบคลุมบทความทั้งหมดประมาณร้อยละ 75 ขณะเดียวกันจำนวนคำศัพท์อีก 1,000 คำที่เหลือครอบคลุมบทความทั้งหมด

น้อยกว่าร้อยละ 10 (Paquot, 2010) คำที่มีความถี่สูงสุดคือ คำพื้นฐานที่ใช้บ่อยที่สุดในการใช้ภาษา (language use) (Paquot, 2010) รายการคำศัพท์พื้นฐานที่เป็นที่รู้จักกันดีคือรายการคำศัพท์ที่มีความถี่สูงสุด 2,000 คำแรกในรายการคำศัพท์พื้นฐานของ West (General Service List: GSL) (West, 1953) ซึ่งรายการคำศัพท์พื้นฐานของ West 2,000 คำแรกนี้คิดเป็นร้อยละ 80 ของคำศัพท์ทั้งหมด ส่วนผลการศึกษารายการคำศัพท์ที่มีความถี่สูงสุดของ Coxhead (2000) พบว่า รายการคำศัพท์พื้นฐานของ Coxhead 2,000 คำแรกนี้คิดเป็นร้อยละ 76 ของคำศัพท์ทั้งหมดซึ่งมีปริมาณต่ำกว่าของ West (1953) เล็กน้อย นอกจากนี้ นักวิจัยได้ศึกษารายการคำศัพท์ที่มีความถี่สูงสุดในหลาย ๆ สาขา อาทิ คลังข้อมูลคำศัพท์สาขาจิตวิทยา (Kaewphanngam, 2002) คลังข้อมูลคำศัพท์สาขาวิศวกรรมโยธา (Para, 2004) คลังข้อมูลคำศัพท์สาขาเกษตรศาสตร์ (Martínez, Beck, & Panza, 2009) คลังข้อมูลคำศัพท์สาขาการเงิน (Li & Qian, 2010) คลังข้อมูลคำศัพท์ในข้อสอบเข้ามหาวิทยาลัยในประเทศไทย (Cherngchawano & Jaturapitakkul, 2014) และคลังข้อมูลคำศัพท์สาขาวิทยาศาสตร์สิ่งแวดล้อม (Liu & Han, 2015)

แม้ว่าการศึกษาคำศัพท์ที่มีความถี่สูงสุดเป็นเรื่องที่น่าสนใจและมีความสำคัญต่อผู้เรียนภาษาอังกฤษในฐานะภาษาที่สองหรือภาษาที่ต่างประเทศ (Nation, 2001; Paquot, 2010) แต่ยังไม่มีการศึกษาคำศัพท์ทางการวิจัยสัตว์ทดลอง ทำให้ขาดข้อมูลคำศัพท์ทางการวิจัยสัตว์ทดลอง จากการค้นคว้าบทความวิจัยเกี่ยวกับคลังข้อมูลคำศัพท์ในหลายสาขาดังกล่าวข้างต้นเมื่อวันที่ 1 พฤษภาคม 2558 ในฐานะข้อมูล ‘ScienceDirect’ และ ‘Scopus’ ด้วยคำสำคัญ ‘คลังข้อมูลภาษา’ และ ‘สัตว์ทดลอง’ ไม่พบการศึกษาค้นคลังข้อมูลคำศัพท์ที่มีความถี่สูงสุดในสาขาสัตว์ทดลอง และในวันที่ 20 มีนาคม 2558 ผู้วิจัยได้สัมภาษณ์ผู้อำนวยการศูนย์วิจัยสัตว์ทดลองแห่งชาติ มหาวิทยาลัยมหิดล เกี่ยวกับปัญหาการใช้ภาษาอังกฤษของผู้วิจัยทางด้านวิทยาศาสตร์สาขาสัตว์ทดลองพบว่า นักวิจัยเหล่านี้เผชิญปัญหาการใช้ภาษาอังกฤษ โดยเฉพาะทักษะการอ่านและเขียนภาษาอังกฤษ เนื่องจากนักวิจัยเหล่านี้ต้องอ่านและเขียนบทความภาษาอังกฤษ จากการศึกษาค้นคลังข้อมูลคำศัพท์ในภาษาอื่นของ Nation (2001) พบว่า ความรู้ในด้านความหมายของคำศัพท์และการใช้คำศัพท์จะช่วยให้ผู้เรียนอ่านและเขียนภาษาอังกฤษได้ง่ายขึ้น การศึกษาคำศัพท์ที่มีความถี่สูงสุดในแต่ละสาขา จะเป็นการช่วยให้ผู้เรียนหรือนักวิจัยในสาขานั้น ๆ ทราบว่ามีความถี่จำนวนหนึ่งที่ผู้เรียนหรือนักวิจัยควรจะต้องศึกษาทำความเข้าใจในด้านความหมายและการใช้ เพราะคำศัพท์เหล่านี้จะพบบ่อยมากในการอ่านหรือใช้บ่อยมากในการเขียน เมื่อผู้เรียนหรือนักวิจัยมีความรู้ความเข้าใจในกลุ่มคำศัพท์ที่มีความถี่สูงสุดนี้แล้ว จะช่วยให้ผู้เรียนหรือนักวิจัยอ่านและการเขียนภาษาอังกฤษง่ายขึ้น สะดวกขึ้น นอกจากนี้ ยังไม่มีรายวิชาภาษาอังกฤษเพื่อวิทยาศาสตร์สาขาสัตว์ทดลองเปิดสอนในระดับอุดมศึกษา ดังนั้น ผลวิจัยครั้งนี้จะเป็นการเติมความรู้ทางด้านคำศัพท์ภาษาอังกฤษที่ใช้ในสาขาสัตว์ทดลอง

2. วัตถุประสงค์ในการวิจัย

2.1 เพื่อเปรียบเทียบความถี่ของ “คำเนื้อหา” และ “คำไวยากรณ์” ที่ปรากฏในรายการคำศัพท์ที่มีความถี่สูงสุดในคลังข้อมูลภาษาสาขาสัตว์ทดลอง

2.2 เพื่อค้นหาและรายงานคำเนื้อหา (content word) 100 คำแรกของรายการคำศัพท์ที่มีความถี่สูงสุด (high frequency word list) ในคลังข้อมูลภาษาสาขาสัตว์ทดลอง

3. วิธีการศึกษาและวัสดุอุปกรณ์

วิธีการวิจัยประกอบด้วยแหล่งข้อมูล การรวบรวมข้อมูล เครื่องมือวิจัยและการวิเคราะห์ข้อมูลซึ่งได้อธิบายไว้ ดังต่อไปนี้

3.1 แหล่งข้อมูล

ผู้วิจัยได้สืบค้นรายชื่อวารสารจากฐานข้อมูล Scopus ด้วยคำสำคัญ ‘laboratory animal’ เมื่อวันที่ 1 พฤษภาคม 2558 พบว่ามีจำนวนรายชื่อวารสารทั้งหมด 54,413 รายชื่อ ดังภาพที่ 1 มีบทความวิจัยจำนวนทั้งหมด 46,558 บทความและ บทความปริทัศน์จำนวนทั้งหมด 6,738 บทความ ดังภาพที่ 2

Source type	
<input checked="" type="checkbox"/> Journals	(54,413) >
<input type="checkbox"/> Book Series	(1,405) >
<input type="checkbox"/> Books	(1,119) >
<input type="checkbox"/> Conference Proceedings	(481) >
<input type="checkbox"/> Trade Publications	(50) >
<input type="checkbox"/> Undefined	(18) >

ภาพที่ 1 จำนวนรายชื่อวารสารที่สืบค้นจากฐานข้อมูล Scopus ด้วยคำหลัก “laboratory animal”

Document type	
<input checked="" type="checkbox"/> Article	(46,558) >
<input checked="" type="checkbox"/> Review	(6,738) >
<input type="checkbox"/> Conference Paper	(1,453) >
<input type="checkbox"/> Book Chapter	(983) >
<input type="checkbox"/> Note	(573) >

ภาพที่ 2 จำนวนบทความวิจัยและบทความปริทัศน์ที่สืบค้นจากฐานข้อมูล Scopus ด้วยคำหลัก “laboratory animal”

เมื่อสืบค้นจำนวนวารสารบทความวิจัยและบทความปริทัศน์จากฐานข้อมูล Scopus ด้วยคำสำคัญ ‘laboratory animal’ เมื่อวันที่ 9 มิถุนายน 2560 เพื่อเปรียบเทียบจำนวนวารสารและบทความทั้งหมด ในปี พ.ศ. 2558 และ ปี พ.ศ. 2560 ดังตารางที่ 1

รายการ	สืบค้นเมื่อ พ.ศ. 2558	สืบค้นเมื่อ พ.ศ. 2560
จำนวนวารสารทั้งหมด (Journals)	54,413	172,738
จำนวนบทความวิจัย (Research Articles)	46,558	147,388
จำนวนบทความปริทัศน์ (Review Articles)	6,738	18,379

ตารางที่ 1 เปรียบเทียบจำนวนวารสาร บทความวิจัย และบทความปริทัศน์ที่สืบค้นจากฐานข้อมูล Scopus ในปี พ.ศ. 2558 และ ปี พ.ศ. 2560

จากตารางข้างต้นแสดงให้เห็นว่า จำนวนวารสารและบทความทั้งสองประเภทในปี พ.ศ. 2560 มีจำนวนเพิ่มมากขึ้นจากปี พ.ศ. 2558 ซึ่งถือว่าเป็นประชากรที่มีขนาดใหญ่มาก ผู้วิจัยจึงทำการกำหนดตัวอย่างและสุ่มตัวอย่างแบบเฉพาะเจาะจงเพื่อนำวารสารที่มีคุณภาพตามเกณฑ์ที่กำหนดมาทำการศึกษา ตัวอย่างวารสารที่กำหนดคือ วารสารทางด้านสัตวทดลองเพื่อนำผลการวิจัยไปใช้ประโยชน์ในกลุ่มนักวิจัยที่ทำงานในศูนย์วิจัยสัตวทดลองแห่งชาติของมหาวิทยาลัยมหิดล เกณฑ์ที่ใช้ในการคัดเลือกวารสาร มี 4 เกณฑ์ คือ (1) ความถี่ของบทความวารสารโดยเฉลี่ย (impact factor) (2) นั้หน้าคุณภาพของวารสารตามค่าควอไทล์ (Q) (3) เป็นวารสารประเภทปริทัศน์ และ (4) เป็นวารสารที่ไม่ปรากฏในรายชื่อของวารสารที่ไม่ควรส่งไปตีพิมพ์ (Beall's list)

ผลจากการสุ่มพบว่า วารสาร Institute of Laboratory Animal Research Articles Journal (ILAR) เข้าข่าย คือ ตัววัดความถี่ของบทความวารสารโดยเฉลี่ย (impact factor) ของวารสาร ILAR ในปี 2557 มีค่าเฉลี่ยที่ 2.393 นั้หน้าคุณภาพของวารสารตามค่าควอไทล์ (Q) จากฐานข้อมูลการจัดอันดับวารสารของ SCImago journal rank (SJR) จากเว็บไซต์ SCImago Journal & Country Rank (<http://www.scimagojr.com>) วารสาร ILAR อยู่ในที่ควอไทล์ที่ 1 (Q1) ในสาขาวิชาสัตววิทยา จุลเคมี พันธุศาสตร์ และอนุชีววิทยา วารสาร ILAR เป็นวารสารประเภทปริทัศน์ และวารสาร ILAR ไม่ปรากฏในรายชื่อของวารสารที่ไม่ควรส่งไปตีพิมพ์ (Beall's list) ผู้วิจัยจึงใช้เป็นแหล่งข้อมูลในการวิเคราะห์คำศัพท์ จากนั้นผู้วิจัยได้รวบรวมบทความวิจัยจากวารสาร ILAR จำนวนทั้งหมด 160 บทความ ย้อนหลัง 5 ปี คือระหว่างปี 2553 ถึง 2557 เพื่อสร้างเป็นคลังข้อมูลภาษาในการศึกษาคั้งนี้

3.2 การรวบรวมข้อมูล

ผู้วิจัยดาวน์โหลดบทความปริทัศน์สาขาสัตวทดลองจำนวนทั้งหมด 160 บทความจาก เว็บไซต์ <http://ilarjournal.oxfordjournals.org/content/by/year> ย้อนหลัง 5 ปี (2553 - 2557) เพื่อให้ได้ข้อมูลที่ทันสมัยและการตีพิมพ์บทความในแต่ละปี มีหัวข้อแตกต่างกันจำนวน 16 หัว ข้อดังตารางที่ 2

ปีที่พิมพ์ บทความ	จำนวน บทความ	หัวข้อเรื่อง	word token ¹	word type ²
	6	Regenerative Medicine	32,999	4,581
	8	Disaster Planning and Management	43,445	5,466
2553	10	The Intersection of Humans, Animals, and the Environment	42,103	5,239
	9	Birds as Animal Models in the Behavioral and Neural Sciences	49,760	5,771
2554	13	Animal Models of Aging	54,478	6,219
	8	Spineless Wonders	43,368	5,963
	11	Animal Models of Drug Addictions	50,835	4,940
2555	15	Neurobiology of Addiction-like Behaviors	69,854	6,026
	12	Zebrafish Health and Husbandry	62,458	5,953
	11	Epigenetics	63,066	5,574
2556	7	Ethical and IACUC Considerations for Field Biology Studies	36,508	4,483
	12	Progress in Genetics and Genomics of Nonhuman Primates	74,009	5,950
2557	9	Animal Models of Peripheral Neuropathy	35,619	4,449
	14	Naturally Occurring Diseases in Animals	89,556	7,669
	11	Behavioral Assessment in Animal Models	76,347	5,762
	4	Experimental Design and Statistics	16,368	2,150
Total	160	16	840,773	22,434

ตารางที่ 2 ข้อมูลเบื้องต้นของคลังข้อมูลภาษาบทความปริทัศน์สาขาสัตวทดลอง

หลังจากนั้นผู้วิจัยได้แปลงไฟล์ข้อมูล (file) จากนามสกุล .pdf เป็นนามสกุล .txt เนื่องจากโปรแกรม AntConc (Anthony, 2014) สามารถอ่านไฟล์นามสกุล .txt ได้เท่านั้น แล้วทำการตัดบรรณานุกรม เอกสารอ้างอิง ภาคผนวก เชิงอรรถ คำบรรยายใต้ภาพ และกิตติกรรมประกาศในแต่ละ

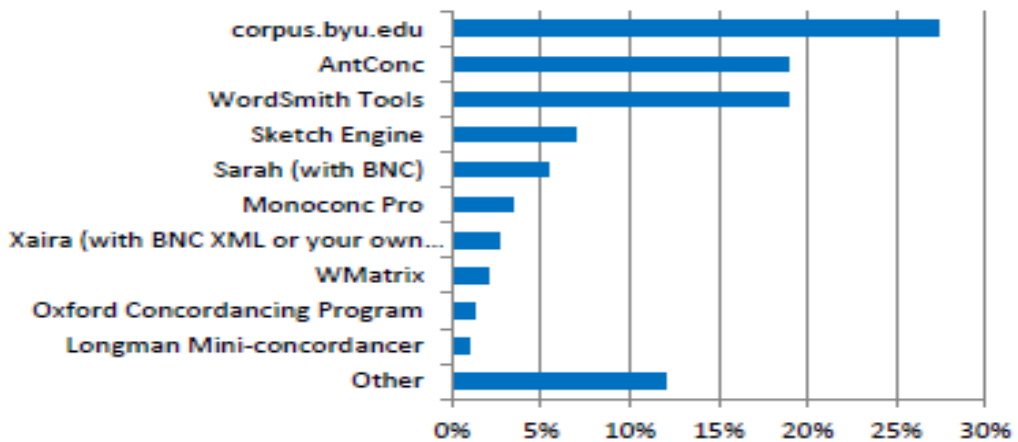
¹ word token คือ จำนวนคำทั้งหมดที่ปรากฏในคลังข้อมูลภาษาสัตวทดลอง (total number of words)

² word type คือ จำนวนคำศัพท์ทั้งหมดที่ปรากฏในคลังข้อมูลภาษาสัตวทดลอง (the number of distinct words)

บทความวิจัยออก หลังจากนั้นทำการจัดเก็บ (save) ไฟล์ข้อมูลเป็น 16 ไฟล์จาก 16 หัวข้อเพื่อเป็นแหล่งข้อมูลในการวิจัย ข้อมูลเบื้องต้นของคลังข้อมูลภาษาของบทความวิจัยสัตว์ทดลอง ประกอบด้วย word token ทั้งหมด 840,773 คำและ word type ทั้งหมด 22,434 คำ

3.3 เครื่องมือวิจัย

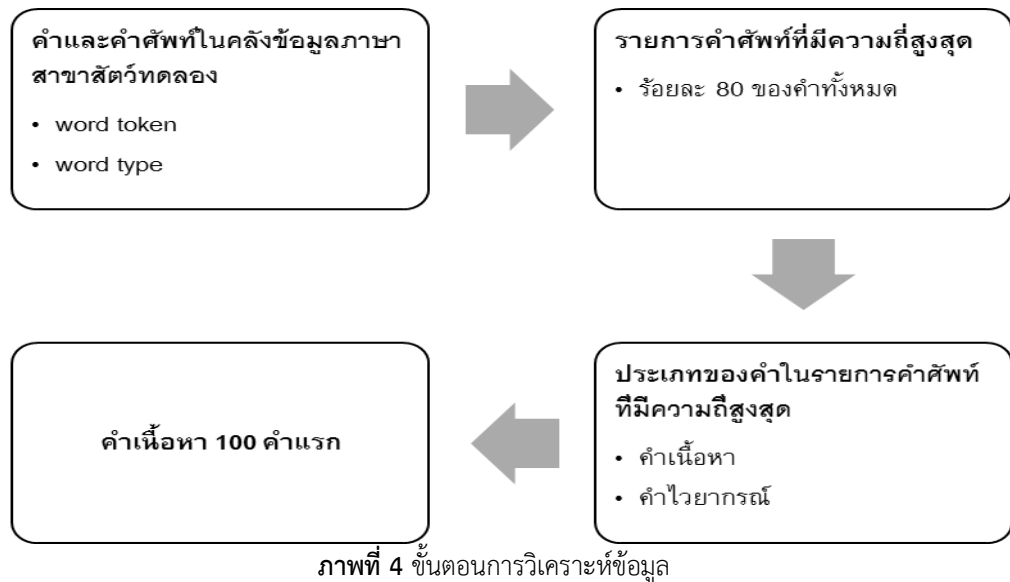
การวิจัยครั้งนี้ได้ใช้โปรแกรม AntConc Version 3.4.4 (Anthony, 2014) ในการวิเคราะห์ความถี่และจัดประเภทของคำในคลังข้อมูลภาษาสาขาสัตวทดลอง โปรแกรมนี้สร้างรายการคำศัพท์ทั้งหมดในคลังข้อมูล (Anthony, 2014) โปรแกรม AntConc ได้รับการพัฒนามาหลาย เวอร์ชันโดย Anthony และโปรแกรม AntConc เป็นโปรแกรมหนึ่งที่ใช้ในการวิเคราะห์คลังข้อมูลภาษามากที่สุดรองจากโปรแกรม corpus.byu.edu ดังภาพที่ 3 (Anthony, 2013) นอกจากนั้นโปรแกรม AntConc เป็นโปรแกรมฟรี ประมวลผลรวดเร็ว และใช้งานได้ง่าย (Anthony, 2013) ดังนั้นผู้วิจัยจึงเลือกโปรแกรม AntConc เพื่อหาความถี่ของคำศัพท์ในคลังข้อมูลภาษาสาขาสัตวทดลอง



ภาพที่ 3 โปรแกรมที่ใช้ในการวิเคราะห์คลังข้อมูลภาษา (Anthony, 2013)

3.4 การวิเคราะห์ข้อมูล

ผู้วิจัยทำการวิเคราะห์ข้อมูลโดยการเพิ่มไฟล์ 16 ไฟล์ลงไปในโปรแกรม AntConc เพื่อรวบรวมคำศัพท์และหาค่าความถี่ของคำในคลังข้อมูลภาษาสาขาสัตวทดลอง เมื่อได้คำศัพท์ทั้งหมดแล้วจึงจัดทำรายการคำศัพท์ที่มีความถี่สูงสุดพร้อมจำแนก “คำเนื้อหา” และ “คำไวยากรณ์” เพื่อเปรียบเทียบความถี่ของคำทั้ง 2 ประเภท หลังจากนั้น ผู้วิจัยได้ทำการรวบรวมคำเนื้อหา 100 คำแรก ของรายการคำศัพท์ที่มีความถี่สูงสุดในคลังข้อมูลภาษาสาขาสัตวทดลอง ขั้นตอนการดำเนินการแสดงในภาพที่ 4



ในการวิจัยครั้งนี้ มีข้อจำกัดดังนี้

- 1) คำกริยาที่เป็นได้ทั้ง “กริยาแท้” (main verb) และ “กริยาช่วย” (auxiliary) ได้แก่ ‘have’ ‘be’ และ ‘do’ ไม่นำมาวิเคราะห์เพื่อเลี่ยงความกำกวมในการจำแนกคำ
- 2) ไม่รวมจำนวนการปรากฏของคำหลัก (headword) และคำตั้ง (entry) เดียวกันไว้ด้วยกัน เนื่องจากงานวิจัยชิ้นนี้ต้องการศึกษาจำนวนคำศัพท์ทั้งหมด (word type) ที่ปรากฏจริงในคลังข้อมูลภาษา สาขาสัตว์ทดลอง

4. ผลการศึกษา

ผลการศึกษาพบว่าในคลังข้อมูลภาษาสาขาสัตว์ทดลอง สรุปได้ 2 ประเด็นหลัก ดังต่อไปนี้

(1) ผลการเปรียบเทียบความถี่ของ “คำเนื้อหา” และ “คำไวยากรณ์” ที่ปรากฏในรายการคำศัพท์ที่มีความถี่สูงสุดในคลังข้อมูลภาษาสาขาสัตว์ทดลอง

ผลการศึกษาพบว่า ในคลังข้อมูลภาษาสาขาสัตว์ทดลองมี word token จำนวน 840,773 คำ (ร้อยละ 100) และ word type พบน้อยกว่าจำนวน word tokens (ร้อยละ 2.67) นอกจากนี้ word token ประกอบด้วยคำเนื้อหาและคำไวยากรณ์ จำนวนคำเนื้อหา (ร้อยละ 55.85) พบมากกว่าจำนวนคำไวยากรณ์ (ร้อยละ 44.15) ดูตารางที่ 3

คำสถิติ	จำนวน (คำ)	ร้อยละ
จำนวนคำทั้งหมด (Word Token)	840,773	100
จำนวนคำศัพท์ทั้งหมดที่พบในคลังข้อมูล (Word Type)	22,434	2.67
จำนวนคำเนื้อหา (Content Word)	469,560	55.85
จำนวนคำไวยากรณ์ (Function Word)	371,213	44.15

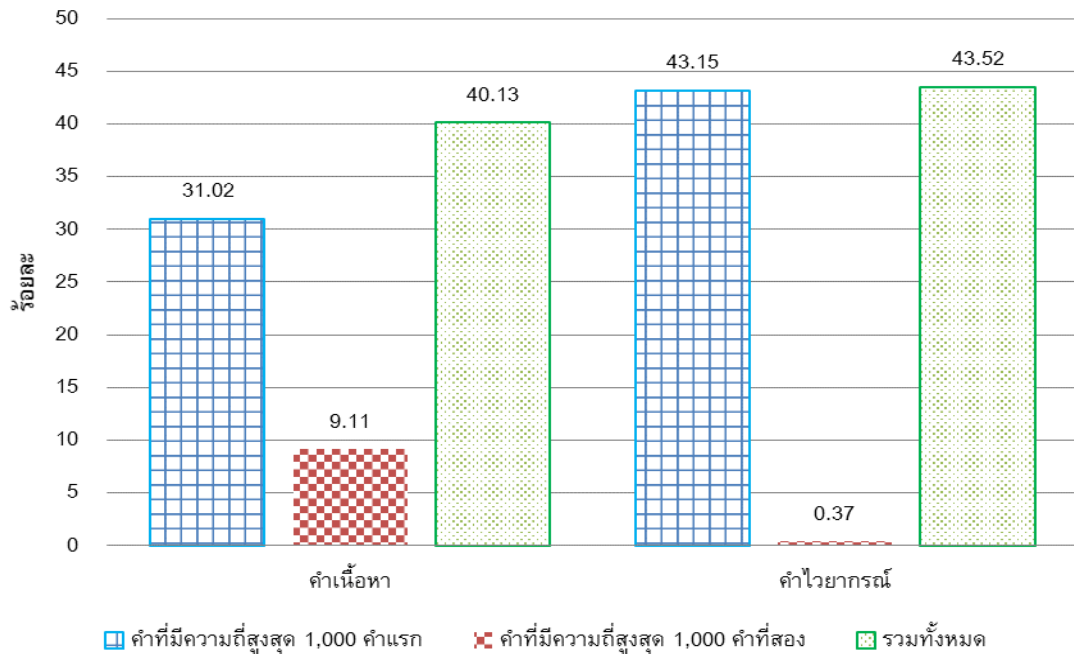
ตารางที่ 3 ข้อมูลทางสถิติในคลังข้อมูลภาษาสาขาสัตว์ทดลอง

ตารางที่ 4 แสดงค่าสถิติของคำที่มีความถี่สูงสุดในคลังข้อมูลภาษาสาขาสัตว์ทดลองพบว่า รายการคำศัพท์ 2,000 คำแรกที่มีความถี่สูงสุดครอบคลุมร้อยละ 83.65 ของจำนวนทั้งหมด นอกจากนี้ คำที่มีความถี่สูงสุด 1,000 คำแรก (ร้อยละ 74.17) เกิดขึ้นมากกว่าคำที่มีความถี่สูงสุด 1,000 คำที่สอง (ร้อยละ 9.48) ประมาณ 8 เท่า

คำที่มีความถี่สูงสุด	จำนวนคำทั้งหมด	ร้อยละ
1,000 คำแรก (The first 1,000 words)	623,583	74.17
1,000 คำที่สอง (The second 1,000 words)	79,713	9.48
รวมทั้งหมด	703,296	83.65

ตารางที่ 4 ข้อมูลทางสถิติของคำที่มีความถี่สูงสุด 2,000 คำแรกในคลังข้อมูลภาษาสาขาสัตว์ทดลอง

เมื่อพิจารณาสัดส่วนของคำเนื้อหาและคำไวยากรณ์ในคำที่มีความถี่สูงสุดดังภาพที่ 5 พบว่า คำเนื้อหา (ร้อยละ 40.13) มีความถี่ต่ำกว่าคำไวยากรณ์ (ร้อยละ 43.52) กล่าวคือ คำเนื้อหาในคำที่มีความถี่สูงสุด 1,000 คำแรก (ร้อยละ 31.02) มีความถี่น้อยกว่าคำไวยากรณ์ (ร้อยละ 43.15) เนื่องจากคำที่มีความถี่สูงสุด 22 คำแรกคือ “คำไวยากรณ์” ได้แก่ the, of, and, in, to, a, that, is, for, with, are, as, or, be, by, on, have, this, from, an, not, และ these ในขณะที่คำเนื้อหาเกิดขึ้นลำดับที่ 23 (studies) อย่างไรก็ตาม คำเนื้อหาในคำที่มีความถี่สูงสุด 1,000 คำที่สอง (ร้อยละ 9.11) มีความถี่มากกว่าคำไวยากรณ์ (ร้อยละ 0.37) ประมาณ 25 เท่า



ภาพที่ 5 สัดส่วนของคำเนื้อหาและคำไวยากรณ์ในรายการคำศัพท์ที่มีความถี่สูงสุด

(2) ผลการค้นหาและรายงานคำเนื้อหา (content word) 100 คำแรกที่มีความถี่สูงสุดในบทความวิจัยสัตว์ทดลอง

ผลการค้นหาและรายงานคำเนื้อหา (content word) 100 คำแรกที่มีความถี่สูงสุดในบทความวิจัยสัตว์ทดลอง แสดงในตารางที่ 5

ที่	คำเนื้อหา	f	ร้อยละ
1	studies	3031	0.36
2	animal	2942	0.35
3	animals	2868	0.34
4	research	2703	0.32
5	human	2252	0.27
6	species	1954	0.23
7	used	1646	0.20
8	models	1596	0.19
9	study	1571	0.19
10	Use	1536	0.18
11	humans	1529	0.18

ที่	คำเนื้อหา	f	ร้อยละ
51	brain	761	0.09
52	behavioral	754	0.09
53	similar	754	0.09
54	drug	748	0.09
55	related	743	0.09
56	data	742	0.09
57	new	735	0.09
58	control	729	0.09
59	based	711	0.08
60	risk	707	0.08
61	found	702	0.08

ที่	คำเนื้อหา	f	ร้อยละ
12	model	1507	0.18
13	disease	1488	0.18
14	cells	1355	0.16
15	effects	1318	0.16
16	mice	1195	0.14
17	development	1177	0.14
18	Cell	1139	0.14
19	associated	1117	0.13
20	genetic	1054	0.13
21	well	1041	0.12
22	gene	1038	0.12
23	specific	1030	0.12
24	behavior	994	0.12
25	levels	989	0.12
26	expression	978	0.12
27	response	942	0.11
28	stress	934	0.11
29	time	923	0.11
30	important	910	0.11
31	using	907	0.11
32	Rats	906	0.11
33	exposure	902	0.11
34	dogs	894	0.11
35	including	879	0.10
36	genes	867	0.10
37	water	865	0.10
38	zebrafish	860	0.10
39	different	857	0.10
40	health	855	0.10
41	increased	840	0.10
42	results	825	0.10
43	changes	824	0.10

ที่	คำเนื้อหา	f	ร้อยละ
62	social	696	0.08
63	significant	694	0.08
64	early	693	0.08
65	factors	685	0.08
66	differences	665	0.08
67	group	662	0.08
68	often	658	0.08
69	reported	657	0.08
70	potential	646	0.08
71	cancer	635	0.08
72	long	613	0.07
73	years	602	0.07
74	age	598	0.07
75	infection	595	0.07
76	fish	594	0.07
77	food	591	0.07
78	effect	590	0.07
79	include	578	0.07
80	large	575	0.07
81	role	574	0.07
82	likely	571	0.07
83	same	569	0.07
84	common	561	0.07
85	monkeys	561	0.07
86	macaques	558	0.07
87	addition	557	0.07
88	activity	556	0.07
89	diseases	555	0.07
90	provide	554	0.07
91	responses	552	0.07
92	experimental	550	0.07
93	methylation	547	0.07

ที่	คำเนื้อหา	f	ร้อยละ
44	High	802	0.10
45	laboratory	802	0.10
46	system	802	0.10
47	clinical	798	0.09
48	number	791	0.09
49	example	782	0.09
50	treatment	781	0.09

ที่	คำเนื้อหา	f	ร้อยละ
94	information	544	0.06
95	induced	538	0.06
96	pain	536	0.06
97	environmental	535	0.06
98	compared	534	0.06
99	analysis	524	0.06
100	systems	522	0.06

ตารางที่ 5 คำเนื้อหา (content word) 100 คำแรกที่มีความถี่สูงสุดในบทความวิจัยสัตว์ทดลอง

คำเนื้อหาที่เกิดขึ้นสูงสุด 100 คำแรกในคลังข้อมูลภาษาสัตวศาสตร์ทดลองคือ คำว่า ‘studies’ ปรากฏในลำดับที่ 23 ของคลังข้อมูลทั้งหมด ซึ่งเป็นคำที่เกิดขึ้นสูงสุด (ร้อยละ 0.36) และคำที่เกิดขึ้นเป็นลำดับที่ 100 คือคำว่า ‘systems’ (ร้อยละ 0.06) คำที่มีความถี่สูงสุด 5 คำแรกได้แก่ ‘studies’ (ร้อยละ 0.36), ‘animal’ (ร้อยละ 0.35), ‘animals’ (ร้อยละ 0.34), ‘research’ (ร้อยละ 0.32) และ ‘human’ (ร้อยละ 0.27)

สรุป ผลการศึกษาพบว่า word token มีจำนวน 840,773 คำและ word type มีจำนวน 22,434 คำ และรายการคำศัพท์ที่มีความถี่สูงสุดในคลังข้อมูลภาษาสัตวศาสตร์ทดลองเกิดขึ้นร้อยละ 83.65 และมีคำเนื้อหาเกิดขึ้นร้อยละ 40.13 ของรายการคำศัพท์ที่มีความถี่สูงสุดคำเนื้อหาที่เกิดขึ้นสูงสุดคือ คำว่า ‘studies’ และลำดับที่ 100 คือ คำว่า ‘systems’

5. อภิปรายผลการศึกษา

จากผลการศึกษาพบว่า จำนวนคำทั้งหมดมี 840,773 คำและจำนวนคำศัพท์ทั้งหมดที่พบในคลังข้อมูลมี 22,434 คำ ในรายการคำศัพท์ที่มีความถี่สูงสุดในคลังข้อมูลภาษาสัตวศาสตร์ทดลองมีคำไวยากรณ์ จำนวน 371,213 คำ และคำเนื้อหา จำนวน 469,560 คำ และรายการคำศัพท์ที่มีความถี่สูงสุดทั้งหมดในคลังข้อมูลภาษาสัตวศาสตร์ทดลองเกิดขึ้นร้อยละ 83.65 คำเนื้อหาเกิดขึ้นร้อยละ 40.13 ในขณะที่คำไวยากรณ์เกิดขึ้นร้อยละ 43.52

ค่าความถี่ของคำในรายการคำศัพท์ที่มีความถี่สูงสุด (high frequency word list) ในคลังข้อมูลภาษาสัตวศาสตร์ทดลอง (ร้อยละ 83.65) สูงกว่าค่าความถี่ของคำในรายการคำศัพท์ที่มีความถี่สูงสุดของ West (1953) (ร้อยละ 80) และรายการคำศัพท์ที่มีความถี่สูงสุดของ Coxhead (2000) (ร้อยละ 76) นอกจากนี้ ค่าความถี่ของคำในรายการคำศัพท์ของการศึกษารังนี้สูงกว่าความถี่ของคำในคลังข้อมูลภาษาสัตวศาสตร์สาขาอื่น ๆ อีกด้วย อาทิ สาขาจิตวิทยา (ร้อยละ 51.48) (Kaewphanngam, 2002)

วิศวกรรมโยธา (ร้อยละ 72.54) (Para, 2004) สาขาเกษตรศาสตร์ (ร้อยละ 67.53) (Martínez, Beck, & Panza, 2009) และสาขาวิทยาศาสตร์สิ่งแวดล้อม (ร้อยละ 70.61) (Liu & Han, 2015) แต่รายการคำศัพท์ที่มีความถี่สูงสุดในการศึกษาครั้งนี้มีความถี่ต่ำกว่าคลังข้อมูลภาษาในข้อสอบเข้ามหาวิทยาลัยของประเทศไทย (ร้อยละ 85.05) (Cherngchawano & Jaturapitakkul, 2014) จะเห็นได้ว่า ความถี่ของคำในรายการคำศัพท์ที่มีความถี่สูงสุดในคลังข้อมูลภาษาที่ต่างกันจะมีค่าความถี่ต่างกัน เนื่องจากความแตกต่างทางภาษาและวัฒนธรรม (Paquot, 2010) เช่น รายการคำศัพท์ที่มีความถี่สูงสุดในคลังข้อมูลภาษาข้อสอบเข้ามหาวิทยาลัยของประเทศไทยมีค่าความถี่สูงกว่าคลังข้อมูลภาษาอื่น เพราะคำศัพท์ในข้อสอบเข้ามหาวิทยาลัยเป็นคำศัพท์พื้นฐานสำหรับนักเรียนในระดับมัธยมศึกษาส่วนคำในรายการคำศัพท์ที่มีความถี่สูงสุดในคลังข้อมูลภาษาสาขาสัตวทดลองสูงกว่าคลังข้อมูลภาษาสาขาอื่น ๆ อาจเป็นเพราะกลุ่มตัวอย่างในการศึกษาครั้งนี้เป็นบทความปริทัศน์ ซึ่งแตกต่างจากงานวิจัยอื่นที่ใช้บทความวิจัยเป็นกลุ่มตัวอย่าง

นอกจากนี้ รายการคำศัพท์ที่มีความถี่สูงสุด 1,000 คำแรก (ร้อยละ 74.17) เกิดขึ้นมากกว่ารายการคำที่มีความถี่สูงสุด 1,000 คำที่สอง (ร้อยละ 9.48) เป็นจำนวนประมาณ 8 เท่า อาจเนื่องจากว่า คำที่มีความถี่สูงสุดปรากฏในรายการคำศัพท์ที่มีความถี่สูงสุด 1,000 คำแรก เช่น คำว่า *'the'* ครอบคลุมเนื้อหาทั้งหมดร้อยละ 5.41 แต่รายการคำศัพท์ที่มีความถี่สูงสุด 1,000 คำที่สองครอบคลุมเนื้อหาทั้งหมดน้อยกว่า เพราะคำแรกในรายการคำศัพท์ที่มีความถี่สูงสุด 1,000 คำที่สองคือคำว่า *'healthy'* ซึ่งครอบคลุมเนื้อหาทั้งหมดร้อยละ 0.014 และเมื่อเปรียบเทียบกับการศึกษาของ West (1953) และ Coxhead (2000) พบว่า คำที่มีความถี่สูงสุด 1,000 คำแรกครอบคลุมเนื้อหาทั้งหมดน้อยกว่าการศึกษาของ West (1953) (ร้อยละ 77) และมากกว่าการศึกษาของ Coxhead (2000) (ร้อยละ 71.4) เมื่อพิจารณาคำที่มีความถี่สูงสุด 1,000 คำที่สองพบว่า ผลของการศึกษาค้นคว้าครอบคลุมเนื้อหาทั้งหมดมากกว่าการศึกษาของ West (1953) และ Coxhead (2000)

เมื่อเปรียบเทียบความถี่ของคำเนื้อหาและคำไวยากรณ์ ในการศึกษาครั้งนี้พบว่า คำเนื้อหา (ร้อยละ 40.13) มีความถี่ต่ำกว่าคำไวยากรณ์ (ร้อยละ 43.52) เพราะ 22 คำแรกของรายการคำที่มีความถี่สูงสุดใน 1,000 คำแรกเป็นคำไวยากรณ์และมีค่าความถี่สูงสุดซึ่งค่าความถี่ของคำเนื้อหาและคำไวยากรณ์สอดคล้องกับงานวิจัยของ Nation (2001) ที่พบว่าคำไวยากรณ์ครอบคลุมเนื้อหาทั้งหมดร้อยละ 43-44

ในคำเนื้อหา 100 คำแรกประกอบด้วยคำนาม คำกริยา และคำคุณศัพท์ ตามลำดับ คำเนื้อหาเหล่านี้มีการสร้างคำใหม่ 3 แบบ ได้แก่ การเติมหน่วยคำวิภัติปัจจัยเติมหลัง การเติมหน่วยคำอนุพันธ์ และคำประสม ตามลำดับ เช่น คำนามหลายคำในรายการคำเนื้อหา 100 คำแรกมี “การผันรูปคำ” (inflection) ตามหลักไวยากรณ์ คือ มีการเติมหน่วยคำเติมหลังหรือปัจจัย (suffix) “-s” หรือ “-es” หลังคำนามเพื่อให้คำนามเอกพจน์เป็นคำนามพหูพจน์ เช่น *animals, humans, levels, rats* และ *dogs*

เป็นต้น นอกจากนี้ คำกริยาก็มีการผันรูปคำเช่นกัน เช่น เติมปัจจัย (suffix) “-s” หรือ “-es” หลังคำกริยาเพื่อแสดงรูปกริยากาลปัจจุบันที่ใช้กับประธานบุรุษที่ 3 เอกพจน์ เช่น *studies*, *uses* และ *responses* เป็นต้น หรือการเติมปัจจัย (suffix) “-ed” หรือ “-d” หลังคำกริยาเพื่อแสดงรูปกริยากาลอดีต เช่น *related*, *based* และ *reported* เป็นต้น และมีการสร้างคำโดยการเติม “หน่วยคำอนุพันธ์” (derivational morphemes) เช่น คำคุณศัพท์ในรายการคำเนื้อหา 100 คำแรก มีการเติมปัจจัย (suffix) “-al” หลังคำนาม เช่น *behavioral*, *clinical* และ *environmental* นอกจากนี้ ยังพบการสร้างคำใหม่ด้วยการประสมคำเช่นกัน ได้แก่ *zebrafish* เกิดจากการประสมคำนาม *zebra* และคำนาม *fish* เข้าด้วยกันและทำให้เกิดคำนามคำใหม่ขึ้น

6. สรุปผลการศึกษา

เนื่องจากยังไม่มีการศึกษาเกี่ยวกับคำศัพท์ในบทความวิจัยสาขาสัตวทดลอง ผู้วิจัยได้รวบรวมคำศัพท์ทั้งหมดในคลังข้อมูลภาษาสาขาสัตวทดลอง และเปรียบเทียบความถี่ของ “คำเนื้อหา” และ “คำไวยากรณ์” ที่ปรากฏในรายการคำศัพท์ที่มีความถี่สูงสุด รวมทั้งค้นหาและรายงานคำเนื้อหา (content word) 100 คำแรกของรายการคำศัพท์ที่มีความถี่สูงสุด (high frequency word list) ในคลังข้อมูลภาษาสาขาสัตวทดลองโดยใช้เกณฑ์การพิจารณาความถี่ของรายการคำศัพท์ที่มีความถี่สูงสุดของ West (1953) ผลของการศึกษาพบว่า ในคลังข้อมูลภาษาสาขาสัตวทดลองมี word token จำนวน 840,773 คำ และ word type จำนวน 22,434 คำซึ่งประกอบด้วยคำเนื้อหา (ร้อยละ 55.85) และคำไวยากรณ์ (ร้อยละ 44.15) ในรายการคำศัพท์ที่มีความถี่สูงสุดมีคำเนื้อหาร้อยละ 40.13 และคำไวยากรณ์ ร้อยละ 43.52 คำเนื้อหา 100 คำแรกของรายการคำศัพท์ที่มีความถี่สูงสุดประกอบด้วยคำนาม คำกริยา และคำคุณศัพท์ ตามลำดับ

ฉะนั้น งานวิจัยชิ้นนี้จึงเป็นองค์ความรู้ใหม่สำหรับผู้เรียน ผู้วิจัย และผู้สอนในสาขาดังกล่าว นอกจากนี้ รายการคำศัพท์ที่มีความถี่สูงสุดในคลังข้อมูลภาษาสาขาสัตวทดลองจะเป็นประโยชน์ต่อทั้งผู้เรียนและผู้สอน กล่าวคือ ผู้เรียนสามารถใช้รายการคำศัพท์ดังกล่าวเพื่อเป็นประโยชน์ต่อการอ่านและเขียนบทความวิจัยสาขาสัตวทดลองที่ตีพิมพ์เป็นภาษาอังกฤษ และครูผู้สอนก็สามารถใช้รายการคำศัพท์ดังกล่าวในการออกแบบบทเรียนในการสอนรายวิชาภาษาอังกฤษเพื่อการเขียนทางวิชาการ

7. ข้อเสนอแนะในการศึกษาครั้งต่อไป

7.1 ศึกษาคำศัพท์เฉพาะทางที่ปรากฏในสาขาสัตวทดลองเพื่อเป็นประโยชน์ต่อการจัดทำพจนานุกรมศัพท์ในสาขานี้ต่อไป

7.2 ในการศึกษาครั้งต่อไป ผู้วิจัยที่สนใจอาจจะศึกษาการสร้างคำและตระกูลคำศัพท์ที่ปรากฏในรายการคำศัพท์ของคลังข้อมูลภาษาสาขาสัตว์ทดลอง

7.3 ในการศึกษาครั้งต่อไป ผู้วิจัยที่สนใจอาจศึกษารูปแบบและหน้าที่ของภาษา (form and function of language) ในงานวิจัยสาขาสัตว์ทดลอง เช่น รูปแบบและหน้าที่ภาษาในการแสดงวัตถุประสงค์ แสดงความสำคัญ แสดงปัญหาหรือช่องว่างในการวิจัย เพื่อช่วยให้ผู้เรียนและนักวิจัยได้ทราบรูปแบบและหน้าที่ของภาษาทำให้เกิดความเข้าใจ มีกำลังใจที่จะอ่านหรือเขียนบทความวิจัยเป็นภาษาอังกฤษต่อไปเนื่องจากอ่านเข้าใจและเขียนได้

7.4 ในการศึกษาครั้งนี้ ผู้วิจัยศึกษาเฉพาะวารสาร Institute of Laboratory Animal Research Articles Journal จำนวนทั้งหมด 160 บทความ ย้อนหลัง 5 ปี (2553-2557) ตัวอย่างที่ศึกษาไม่ได้เก็บจากวารสารอื่น ๆ จึงอาจไม่มีความหลากหลาย อีกทั้งมีข้อจำกัดด้านเวลาและงบประมาณ ดังนั้น ในการศึกษาครั้งต่อไป ผู้วิจัยอาจสุ่มวารสารและบทความจำนวนมากกว่านี้

เอกสารอ้างอิง

- ณัฐมา พงศ์ไพโรจน์. (2554). *ระบบหน่วยคำในภาษาอังกฤษ English Morphology*. กรุงเทพฯ: โครงการเผยแพร่ผลงานวิชาการ คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- ติดยุ ศรินราววัฒน์ และชลธิชา บำรุงรักษ์. (บรรณาธิการ). (2558). *ภาษาและภาษาศาสตร์*. กรุงเทพฯ: มหาวิทยาลัยธรรมศาสตร์.
- นันทนา รณเกียรติ. (2555). *สัตวศาสตร์เพื่อการสอนการออกเสียงภาษาอังกฤษ*. กรุงเทพฯ: มหาวิทยาลัยธรรมศาสตร์.
- น้ำทิพย์ ภัทธรวิวัฒน์. (2551). *การเปลี่ยนแปลงของภาษา: ภาษาอังกฤษผ่านกาลเวลา*. กรุงเทพฯ: โครงการเผยแพร่ผลงานวิชาการ คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- ราชบัณฑิตยสถาน. (2546). *ศัพท์ภาษาศาสตร์ ฉบับราชบัณฑิตยสถาน*. กรุงเทพฯ: สหมิตรพรีนติ้ง.
- ราชบัณฑิตยสถาน. (2557). *พจนานุกรมศัพท์ภาษาศาสตร์ (ภาษาศาสตร์ประยุกต์) ฉบับราชบัณฑิตยสถาน*. กรุงเทพฯ: รุ่งศิลป์การพิมพ์.

References

- Anthony, L. (2013). *Developing AntConc for a new generation of corpus linguists* [PowerPoint Slides]. Retrieved from http://cass.lancs.ac.uk/wp-content/uploads/2014/10/aacl_2014_corpus_tools_brainstorming_printing.pdf

- Anthony, L. (2014). *AntConc (Version 3.4.4)* [Software]. Available from <http://www.laurenceanthony.net/software/antconc/>
- Cherngchawano, W. & Jaturapitakkul, N. (2014). Lexical profiles of Thailand university admission tests. *PASAA*, 48, 1-27.
- Kenneth, I. A. (2000). A Buddhist response to the nature of human rights. *Journal of Buddhist Ethics*, 8. Retrieved from <http://www.cac.psu.edu/jbe/twocont.html>
- Coxhead, A. (2000). A new academic word list. *TESOL quarterly*, 34(2), 213-238. (Retrieved from <http://www.vuw.ac.nz/lals/research/ael/aelinfo.html>)
- Gass, S. M. & Selinker, L. (2008). *Second language acquisition: An introductory course* (3rd ed.). New York: Routledge.
- Hamel, R. E. (2007). *The dominance of English in the international scientific periodical literature and the future of language use in science*. Philadelphia: John Benjamins.
- Hutchinson, T., & Waters, A. (1987). *English for specific purposes: A learning centered approach*. Cambridge: Cambridge University Press.
- Katamba, F. & Stonham, J. (2006). *Morphology*. Basingstok: Palgrave Macmillan.
- Kaewphanngam, C. (2002). *A corpus analysis of psychology texts as a basic for the development of teaching materials in English for academic purposes* (Unpublished master's thesis). Mahidol University, Thailand.
- Li, Y. & Qian, D. (2010). Profiling the Academic Word List (AWL) in a financial corpus. *SYSTEM*, 38, 402-411.
- Liu, J., & Han, L. (2015). A corpus-based environmental academic word list building and its validity test. *English for Specific Purposes*, 39, 1-11.
- Martínez, I. A., Silvia, C. B., & Carolina B. P. (2009). Academic Vocabulary in agriculture research articles: A corpus-based study. *English for Specific Purposes*, 28(3), 183-198.
- Murphy, M. L. (2010). *Lexicon meaning*. Cambridge: Cambridge University Press.

- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Para, C. (2004). *A corpus study of high-frequency words in civil engineering research article: Sub-disciplinary difference between structure and transportation (Unpublished master's thesis)*. Mahidol University. Thailand.
- Paltridge, B., & Starfield, S. (2013). *The handbook of English for specific purposes*. Malaysia: John Wiley & Sons.
- Paquot, M. (2010). *Academic vocabulary in learner writing from extraction to analysis*. New York: Continuum.
- Schmitt, N. (2000). *Vocabulary in language teaching*. United Kingdom: Cambridge University Press.
- Trimble, L. (1985). *English for science and technology: A discourse approach*. Cambridge: Cambridge University Press.
- Valipouri, L., & Nassaji, H. (2013). A corpus-based study academic vocabulary in chemistry research articles. *Journal of English for Academic Purpose*, 12, 248-263.
- Van Weijen, D. (2012). The language of (future) scientific communication. *Research Trends*, 31, 7-8.
- West, M. (1953). *A general service list of English words*. London: Longman, Green & Co.