

รายการคำศัพท์ที่มีความถี่สูงสุดในบทความปริทัศน์สาขาสัตวทดลอง High Frequency Word List in Laboratory Animal Review Articles

ทรงศรี สรณสถาพร

สิริวิชญ์ 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” เกณฑ์การพิจารณารายการคำศัพท์ที่มีความถี่สูงสุดคือ 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)