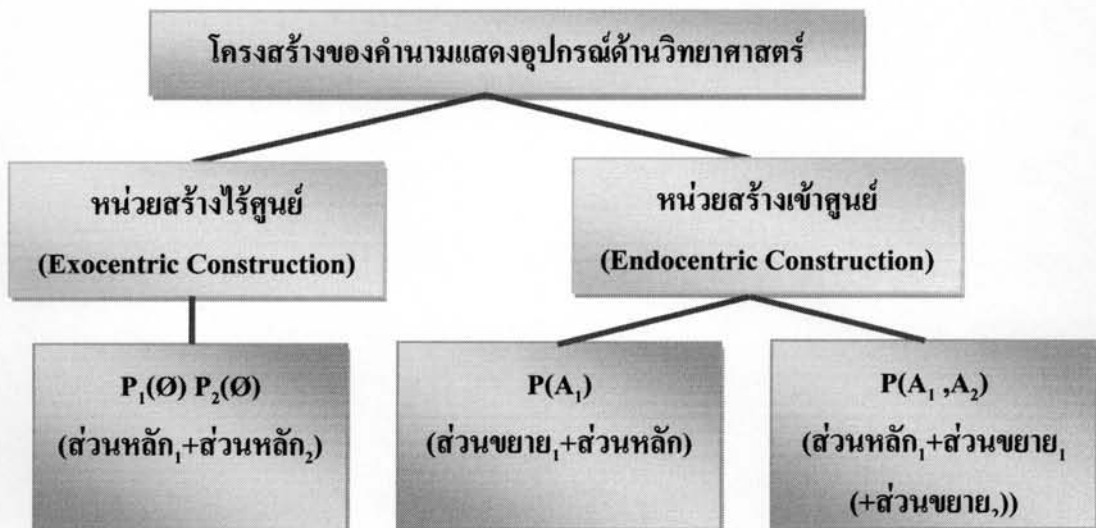


สรุป อภิปรายผล และข้อเสนอแนะ

6.1 สรุปและอภิปรายผล

ค่านามแสดงอุปกรณ์ด้านวิทยาศาสตร์เกิดจากการประกอบคำขึ้นเป็นหน่วยสร้างประเภทต่าง ๆ อาจรวมกันเป็นลำดับชั้น หรือเป็นเพียงคำ 2 - 3 คำเรียงกัน อย่างไรก็ตาม การพิจารณาเพียงโครงสร้างการประกอบคำที่เกิดจากคำชนิดใดนั้นไม่เพียงพอที่จะทำให้ทราบความสัมพันธ์ที่เกิดขึ้นระหว่างส่วนประกอบได้ การทำความเข้าใจความสัมพันธ์ที่เกิดขึ้นจึงต้องศึกษาโครงสร้างหลายระดับของค่านามแสดงอุปกรณ์ด้านวิทยาศาสตร์

ค่านามแสดงอุปกรณ์ด้านวิทยาศาสตร์ประกอบด้วยส่วนประกอบหลัก 2 ส่วน คือ ส่วนหลัก และส่วนขยาย การรวมส่วนประกอบดังกล่าวทำให้กลายเป็นหน่วยสร้างขึ้น ค่านามแสดงอุปกรณ์ด้านวิทยาศาสตร์จัดเป็นหน่วยสร้าง 2 ประเภท คือ หน่วยสร้างไร้ศูนย์กลางและหน่วยสร้างเข้าสู่ศูนย์กลางภาพด้านล่างนี้



ภาพที่ 41 ภาพแสดงโครงสร้างของค่านามแสดงอุปกรณ์ด้านวิทยาศาสตร์

ค่านามแสดงอุปกรณด้านวิทยาศาสตร์ประเภทแรกเป็นหน่วยสร้างไร้ศูนย์ หน่วยสร้างไร้ศูนย์ประกอบด้วยส่วนหลักที่ร่วมกันทำหน้าที่เป็นส่วนสำคัญของคำและร่วมกันแสดงความหมายของค่านามทั้งคำ ในที่นี้ ค่านามแสดงอุปกรณด้านวิทยาศาสตร์ที่เป็นหน่วยสร้างนี้ประกอบด้วยส่วนประกอบที่มีความเหนียวแน่นระหว่างคำสูง ส่วนประกอบแต่ละส่วนที่นำมารวมกันทำให้กลายเป็นคำที่มีความหมายใหม่ ไม่ตรงตามความหมายของส่วนประกอบแต่ละส่วน ต้องอาศัยการตีความจึงจะเข้าใจความหมายของส่วนประกอบแต่ละส่วนได้ จึงเรียกค่านามแสดงอุปกรณด้านวิทยาศาสตร์กลุ่มนี้ว่า คำประสมไร้ศูนย์ ค่านามกลุ่มนี้เกิดจากประสมคำ 2 ชนิด คือ ค่านามและคำกริยากรรม ส่วนประกอบทั้งสองร่วมกันทำหน้าที่เป็นส่วนหลักและเป็นคำแสดงที่ร่วมกันแสดงความหมายของค่านามทั้งคำ โครงสร้างหลายระดับของค่านามกลุ่มนี้เป็นดังตารางข้างล่างนี้

ตารางที่ 19 โครงสร้างหลายระดับของหน่วยสร้างไร้ศูนย์

ระดับของโครงสร้าง				ค่านามฯ
ตรรกะ	หน้าที่	ชนิดของคำ	ความหมาย	
P ₁ (Ø) P ₂ (Ø)	ส่วนหลัก ₁ + ส่วนหลัก ₂	กส+กส	[วัตถุประสงค์-วิธีการ]	ไขควง
		น+กส	[วัตถุประสงค์-วิธีการ]	ตราชู
			[ส่วนประกอบ-วัตถุประสงค์]	ปากคียบ หัวเผา ขาหยั่ง หูฟัง
		น+น	[ส่วนประกอบ-หน้าที่]	หัวเทียน
			[ส่วนประกอบ-ทั้งส่วน]	ปากกา หัวนกดกระจอก หัวแร้ง หูช้าง
			[ความยิ่งใหญ่ ความสำคัญ- การกระทำ]	แม่ปั้ม แม่แรง
			[หน้าที่ - วัตถุประสงค์]	สะพานไฟ
			[รูปร่างลักษณะ-วัตถุ]	พวงมาลัย

ค่านามแสดงอุปกรณ์ด้านวิทยาศาสตร์ประเภทที่สองเป็นหน่วยสร้างเข้าสู่ศูนย์ หน่วยสร้างเข้าสู่ศูนย์ประกอบด้วยส่วนประกอบอย่างน้อยสองส่วน คือ ส่วนหลักและส่วนขยาย หากค่านามแสดงอุปกรณ์ด้านวิทยาศาสตร์ประกอบด้วยส่วนประกอบที่มีความเหนียวแน่นระหว่างคำสูง ส่วนประกอบแต่ละส่วนที่นำมารวมกันทำให้กลายเป็นคำที่มีความหมายใหม่ อาจคงเค้าความหมายเดิมของส่วนประกอบแต่ละส่วน จะเรียกค่านามแสดงอุปกรณ์ด้านวิทยาศาสตร์กลุ่มนี้ว่า คำประสมเข้าสู่ศูนย์

โครงสร้างหลายระดับของค่านามแสดงอุปกรณ์ด้านวิทยาศาสตร์กลุ่มนี้ มีดังต่อไปนี้

(1) โครงสร้างระดับชนิดของคำของค่านามแสดงอุปกรณ์ด้านวิทยาศาสตร์กลุ่มนี้แบ่งได้เป็น 2 โครงสร้างใหญ่ คือ โครงสร้าง “คำ + คำ” และ โครงสร้าง “คำ + หน่วยสร้างเข้าสู่ศูนย์อย่างวลีและอนุภาคย์” ทั้งสองโครงสร้างเกิดจากการประสมคำ 13 ชนิด ได้แก่

ค่านาม	คำกริยากรรม	คำกริยากรรม
คำกริยากรรมย่อย	คำบุพบท	คำลักษณนาม
คำจำนวนนับ	คำลำดับที่	คำหน้าจำนวน
คำหลังกริยา	คำวิเศษณ์	คำเชื่อม
คำปฏิเสธ		

(2) ค่านามแสดงอุปกรณ์ด้านวิทยาศาสตร์กลุ่มนี้ประกอบด้วยส่วนประกอบอย่างน้อย 2 ส่วน คือ ส่วนหลักและส่วนขยาย โดยส่วนหลักมักปรากฏเป็นสมาชิกทางซ้าย ตามด้วยส่วนขยายทางขวามือ ส่วนขยายสามารถปรากฏร่วมได้มากที่สุด 2 ส่วน โดยผ่านการรวมส่วนประกอบที่ประชิดกันเป็นลำดับขั้น ส่วนหลักทำหน้าที่เป็นส่วนหลักของคำ ส่วนขยายทำหน้าที่ขยายความส่วนหลักในแง่มุมต่าง ๆ

ในบางกรณี อาจปรากฏส่วนหลักเป็นสมาชิกทางขวาได้ ส่วนขยายจะเปลี่ยนตำแหน่งมาเป็นสมาชิกทางซ้ายมือแทน กรณีนี้ปรากฏกับค่านามแสดงอุปกรณ์ด้านวิทยาศาสตร์ที่ได้รับลักษณะการสร้างคำสมาสแบบบาลีสันสกฤตที่วางส่วนขยายไว้หน้าส่วนหลัก และอาจเกิดจากการเลียนแบบโครงสร้างของหน่วยจำนวนที่วางคำจำนวนนับซึ่งเป็นส่วนขยายไว้หน้าคำลักษณนามซึ่งเป็นส่วนหลัก จากข้อมูล พบเพียง 2 คำ คือ “กลอุปกรณ์” และ “สามขา”

(3) คำนามแสดงอุปกรณ์ด้านวิทยาศาสตร์กลุ่มนี้มีโครงสร้างระดับตรรกะที่ประกอบด้วย ส่วนประกอบ 2 ส่วน คือ คำแสดงและองค์ประกอบบังคับ คำแสดงทำหน้าที่เป็นส่วนสำคัญของ หน่วยสร้าง ปรากฏตำแหน่งเดียวกับส่วนหลัก ส่วนองค์ประกอบบังคับจะปรากฏร่วมกับคำแสดง เพื่อทำหน้าที่ขยายความคำแสดง องค์ประกอบบังคับจึงเป็นส่วนประกอบที่ทำหน้าที่เช่นเดียวกับ ส่วนขยาย และปรากฏตำแหน่งเดียวกับส่วนขยายด้วย องค์ประกอบบังคับในคำนามแสดงอุปกรณ์ ด้านวิทยาศาสตร์กลุ่มนี้เป็นองค์ประกอบบังคับภายใน สามารถปรากฏได้มากที่สุด 2 ส่วน คือ องค์ประกอบบังคับตัวที่ 1 (A_1) และองค์ประกอบบังคับตัวที่ 2 (A_2) ในที่นี้ จะไม่มีองค์ประกอบ บังคับภายนอก (A_0) เพราะองค์ประกอบบังคับภายนอกทำหน้าที่เป็นประธาน แต่องค์ประกอบ บังคับในคำนามแสดงอุปกรณ์ด้านวิทยาศาสตร์ล้วนทำหน้าที่เป็นส่วนขยาย

(4) ความสัมพันธ์ทางความหมายระหว่างส่วนประกอบของคำนามแสดงอุปกรณ์ด้าน วิทยาศาสตร์แบ่งได้เป็น 2 ประเภทใหญ่ คือ

(4.1) ความสัมพันธ์แบบไม่ซับซ้อน หมายความว่า ส่วนประกอบแต่ละส่วนไม่ได้ รวมเป็นหน่วยสร้างหนึ่งก่อนมาประสมกันเป็นคำนามแสดงอุปกรณ์ด้านวิทยาศาสตร์ ส่วนใหญ่เกิด จากคำ 2 – 3 คำมาเรียงต่อกัน ความสัมพันธ์ทางความหมายจึงไม่ซับซ้อน มีความสัมพันธ์เกิดขึ้น เพียงชั้นเดียว เช่น กระจกเงา เครื่องยนต์ สายดิน หม้อแปลง เป็นต้น

(4.2) ความสัมพันธ์แบบซับซ้อน หมายความว่า ส่วนประกอบแต่ละส่วนผ่าน กระบวนการรวมเป็นหน่วยสร้างหนึ่งก่อนมาประสมกันเป็นคำนามแสดงอุปกรณ์ด้านวิทยาศาสตร์ ส่วนใหญ่ หน่วยสร้างนั้นเป็นหน่วยสร้างเข้าสู่ศูนย์ที่มี โครงสร้างเลียนแบบวลีและประโยค (Kanyanut Kriengkiet, et al., 2007) หน่วยสร้างเข้าสู่ศูนย์นั้นจะเป็นนามวลี อนุพจน์ หรือเป็นหน่วย สร้างเข้าสู่ศูนย์ที่ขึ้นต้นด้วยคำกริยาแล้วตามด้วยคำนามหรือนามวลี คำนามแสดงอุปกรณ์ด้าน วิทยาศาสตร์ที่เกิดจากการประสมคำกับหน่วยสร้างเข้าสู่ศูนย์ที่เป็นนามวลีหรืออนุพจน์มักมี โครงสร้างเลียนแบบวลี ส่วนคำนามแสดงอุปกรณ์ด้านวิทยาศาสตร์ที่เกิดจากการประสมคำกับ หน่วยสร้างเข้าสู่ศูนย์ที่ขึ้นต้นด้วยคำกริยามักมีโครงสร้างเลียนแบบประโยค ลักษณะที่สองเป็น ลักษณะเด่นของคำนามแสดงอุปกรณ์ด้านวิทยาศาสตร์เพราะสามารถบอกความสัมพันธ์ทาง ความหมายได้ว่า อุปกรณ์ชิ้นนั้นใช้เพื่อวัตถุประสงค์ใดหรือมีวิธีการใช้เช่นใด มักเกิดจากคำตั้งแต่ 3 คำขึ้นไปมาเรียงต่อกัน ความสัมพันธ์ทางความหมายจึงซับซ้อนและมีความสัมพันธ์เป็นลำดับ เช่น

“กล้องจุลทรรศน์อิเล็กตรอน” (วรบ.) หมายถึง กล้องจุลทรรศน์ที่ใช้ อิเล็กตรอนเป็นส่วนสำคัญสำหรับดูสิ่งมีชีวิตขนาดเล็ก คำนี้เกิดจากส่วนประกอบ “กล้องจุลทรรศน์” ประสมกับคำว่า “อิเล็กตรอน” แสดงความสัมพันธ์ทางความหมายหลักเป็น [อุปกรณ์ - วิธีการ] โดย คำว่า “กล้องจุลทรรศน์” เกิดจากการประสมคำว่า “กล้อง” กับ “จุลทรรศน์” เป็นลำดับแรกเพื่อ แสดงถึงอุปกรณ์สำหรับดูสิ่งที่มีขนาดเล็ก แสดงความสัมพันธ์ทางความหมายได้เป็น [อุปกรณ์ - วัตถุประสงค์]

“ตัวกรองเชื้อเพลิง” (ยาน.) หมายถึง อุปกรณ์ที่ใช้สำหรับกรองเชื้อเพลิง คำนี้เกิดจากคำว่า “ตัว” ประสมกับส่วนประกอบ “กรองเชื้อเพลิง” แสดงความสัมพันธ์ทาง ความหมายหลักเป็น [อุปกรณ์ - หน้าที่] โดยส่วนประกอบ “กรองเชื้อเพลิง” เป็นหน่วยสร้างเข้าสู่ศูนย์ ที่ขึ้นต้นด้วยคำกริยา “กรอง” ตามด้วยคำนาม “เชื้อเพลิง” หน่วยสร้างนี้ได้รับความหมายเข้า ด้วยกันก่อนเพื่อแสดงวัตถุประสงค์ของอุปกรณ์ มีความสัมพันธ์ทางความหมายเป็น [วัตถุประสงค์- ผู้ถูก]

ดังที่ได้กล่าวแล้วว่า คำนามแสดงอุปกรณ์ด้านวิทยาศาสตร์ที่เป็นหน่วยสร้างเข้าสู่ศูนย์จะ ประกอบด้วยส่วนประกอบสำคัญ คือ คำแสดงและองค์ประกอบบังคับ ส่วนใหญ่ คำแสดงมัก ปรากฏเป็นสมาชิกทางซ้าย ตามด้วยองค์ประกอบบังคับเป็นสมาชิกทางขวาตามลำดับ แต่จากข้อมูล คำนามแสดงอุปกรณ์ด้านวิทยาศาสตร์ พบว่า คำแสดงสามารถปรากฏได้ 2 ตำแหน่ง คือ สมาชิก ทางขวาและสมาชิกทางซ้าย ดังนั้น คำนามแสดงอุปกรณ์ด้านวิทยาศาสตร์ที่เป็นหน่วยสร้างเข้าสู่ศูนย์ จึงสามารถแบ่งได้เป็น 2 ประเภทใหญ่ ๆ ตามตำแหน่งของคำแสดง ดังนี้

(1) คำนามแสดงอุปกรณ์ด้านวิทยาศาสตร์ที่เป็นหน่วยสร้างเข้าสู่ศูนย์โดยคำแสดงเป็น สมาชิกทางขวา คำนามแสดงอุปกรณ์ด้านวิทยาศาสตร์ประเภทนี้ประกอบด้วยส่วนประกอบสำคัญ 2 ส่วน คือ คำแสดงและองค์ประกอบบังคับอย่างละ 1 ตัว คำแสดงทำหน้าที่เป็นส่วนหลักและแสดง ความหมายหลักของคำประสม ส่วนองค์ประกอบบังคับทำหน้าที่เป็นส่วนขยายและแสดง ความหมายเพิ่มเติมแก่ส่วนหลัก ชนิดของคำที่นำมาประกอบเป็นคำนามแสดงอุปกรณ์ด้าน วิทยาศาสตร์นั้น คำแสดงเป็นคำนามจะปรากฏร่วมกับองค์ประกอบบังคับที่เป็นคำนามหรือ

คำจำนวนนับก็ได้ ส่วนความสัมพันธ์ทางความหมายระหว่างส่วนประกอบมี 2 ลักษณะ คือ [ระบบ-อุปกรณ์] และ [ปริมาณ - ส่วนประกอบ]

คำนามแสดงอุปกรณ์ด้านวิทยาศาสตร์ประเภทนี้ปรากฏน้อยมากเพราะตำแหน่งคำแสดงที่เป็นส่วนหลักปรากฏตรงข้ามกับรูปแบบการขยายความในภาษาไทย โดยปกติ ส่วนหลักจะปรากฏเป็นสมาชิกทางซ้ายตามด้วยส่วนขยายเป็นสมาชิกทางขวาตามลำดับ ดังนั้น คำนามแสดงอุปกรณ์ด้านวิทยาศาสตร์ประเภทนี้จึงพบน้อยมากเพียง 2 คำ ได้แก่ “กลอุปกรณ์” และ “สามขา”

ในกรณีคำว่า “กลอุปกรณ์” ลักษณะการสร้างคำนี้ได้รับอิทธิพลมาจากภาษาบาลีสันสกฤต คือ ปรากฏส่วนขยายในตำแหน่งหน้าส่วนหลัก ส่วนคำว่า “สามขา” เป็นการสร้างคำโดยเลียนแบบโครงสร้างหน่วยจำนวน คือ คำจำนวนนับปรากฏหน้าคำลักษณนาม

โครงสร้างหลายระดับของคำนามกลุ่มนี้เป็นดังตารางข้างล่างนี้

ตารางที่ 20 แสดงโครงสร้างหลายระดับของหน่วยสร้างเข้าสู่ศูนย์ที่คำแสดงเป็นสมาชิกทางขวา

ระดับของโครงสร้าง				คำนามฯ
ตรรกะ	หน้าที่	ชนิดของคำ	ความหมาย	
P(A ₁)	ส่วนขยาย - ส่วนหลัก	น - น	[ระบบ - อุปกรณ์]	กลอุปกรณ์
		จน - لن	[จำนวน - ส่วนประกอบ]	สามขา

(2) คำนามแสดงอุปกรณ์ด้านวิทยาศาสตร์ที่มีหน่วยสร้างเข้าสู่ศูนย์โดยคำแสดงเป็นสมาชิกทางซ้าย คำนามแสดงอุปกรณ์ด้านวิทยาศาสตร์ประเภทนี้เกิดจากส่วนประกอบที่เป็นคำคำเดียวหรือเป็นหน่วยสร้างเข้าสู่ศูนย์ได้หลายรูปแบบ ทั้งนี้ ผู้วิจัยได้กล่าวโดยละเอียดแล้วในบทที่ 3 และ 4 ส่วนความสัมพันธ์ทางความหมายของส่วนประกอบของคำนามแสดงอุปกรณ์ด้านวิทยาศาสตร์สามารถปรากฏได้ 37 ลักษณะ โดยแบ่งตามจำนวนส่วนประกอบ ดังต่อไปนี้

(2.1) ส่วนหลัก 1 ส่วน และ ส่วนขยาย 1 ส่วน มีความสัมพันธ์ทางความหมายดังตารางข้างล่าง ดังนี้

ตารางที่ 21 ความสัมพันธ์ทางความหมายของส่วนประกอบ 2 ส่วน ของหน่วยสร้างเข้าสู่หน่วยที่มีส่วนหลักเป็นสมาชิกทางซ้าย

คำแสดง	องค์ประกอบบังคับตัวที่ 1															
	[ไฉนอะไรไป๕๕]	[เอไป๕๕]	[ออกนอกอะไรไป]	[ไหน]	[อะไร๕๕ไป๕๕ไป]	[อะไร]	[ไหนไป๕๕]	[อะไร๕๕ไป๕๕]	[ไป๕๕]	[มา๕๕]	[ไป๕๕]	[ไหนไป๕๕]	[เอไป]	[นอกอะไรไป๕๕]	[ไหนไป๕๕]	[ไป๕๕ไป๕๕]
[อุปกรณ์]	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
[รูปร่างลักษณะ]					/											
[ทั้งส่วน]														/		
[ส่วนประกอบ]	/			/												
[วัตถุประสงค์]															/	/
[วัตถุ]										/						

จากข้อมูล ผู้วิจัยพบว่า คำเชื่อมเริ่มปรากฏในคานามแสดงอุปกรณ์ด้านวิทยาศาสตร์ที่มีหน่วยสร้างเข้าสู่ศูนย์โดยคำแสดงเป็นสมาชิกทางซ้ายนี้ ทำให้คานามแสดงอุปกรณ์ด้านวิทยาศาสตร์กลุ่มนี้มีลักษณะเป็นเพียงนามวลีสำหรับเรียกชื่ออุปกรณ์ กรณีนี้พบมากในคำเรียกอุปกรณ์ที่ผ่านการแปลและยังไม่ได้รับการสร้างคำเพื่อแสดงอุปกรณ์นั้น โดยมีปรากฏในพจนานุกรมไทย – อังกฤษที่รายการคำศัพท์ภายในถูกสร้างจากการแปลความจากคำหรือข้อความภาษาอังกฤษมาโดยตรง แต่ลักษณะดังกล่าวจะไม่ปรากฏในพจนานุกรมฉบับราชบัณฑิตยสถานเล่มใดเลย เพราะผู้บัญญัติศัพท์พยายามหลีกเลี่ยงคำเชื่อมต่าง ๆ เว้นเสียแต่ว่า การไม่ใส่คำเชื่อมจะทำให้เกิดความกำกวม จึงจะปรากฏคำเชื่อมในคานามนั้น ๆ

หลังจากวิเคราะห์คานามแสดงอุปกรณ์ด้านวิทยาศาสตร์ตามแนวภาษาศาสตร์แล้ว จึงนำผลการวิเคราะห์ทางภาษาศาสตร์มาประยุกต์กับเครื่องคอมพิวเตอร์ ด้วยโปรแกรม Unitex การประยุกต์ผลการวิเคราะห์ดังกล่าวทำได้โดยการสร้างกราฟแสดงความสัมพันธ์ ได้แก่ กราฟย่อยและกราฟหลัก ทั้งนี้ คลังข้อมูลที่นำมาใช้ทดสอบแบ่งเป็น 2 กลุ่ม คือ

(1) คลังข้อมูลที่เป็นรายการคำศัพท์ คลังข้อมูลนี้ได้มาจากการรวบรวมคานามแสดงอุปกรณ์ด้านวิทยาศาสตร์จากพจนานุกรม 4 เล่ม คือ

- พจนานุกรมฉบับราชบัณฑิตยสถาน พ.ศ. 2542 ซึ่งตีพิมพ์ในปี พ.ศ. 2546

- ศัพท์วิทยาศาสตร์ อังกฤษ – ไทย ไทย – อังกฤษ ฉบับราชบัณฑิตยสถาน ซึ่งตีพิมพ์ในปี พ.ศ. 2546

- พจนานุกรมศัพท์ยานยนต์ฉบับราชบัณฑิตยสถาน ซึ่งตีพิมพ์ในปี พ.ศ. 2542

- The Oxford – Duden Pictorial Thai – English Dictionary ซึ่งตีพิมพ์ในปี ค.ศ. 1999

(2) คลังข้อมูลที่เป็นบทความ เป็นคลังข้อมูลที่รวบรวมจากบทความนิยสาร คลังข้อมูลนี้จะใช้สำหรับทดสอบความสามารถของโปรแกรมในการตรวจจับโครงสร้างทั้งหมดที่วิเคราะห์ได้ว่าโปรแกรมสามารถตรวจจับคำที่มีโครงสร้างดังกล่าวได้ถูกต้องและแม่นยำเพียงใด แม้ว่าคำเหล่านั้น

จะปรากฏในประโยค แต่ไม่ได้ปรากฏเป็นรายการคำศัพท์ก็ตาม ทั้งนี้ ผู้วิจัยได้จัดทำรายการคำเหล่านั้นเพื่อตรวจสอบความถูกต้องของการตรวจจับคำด้วยโปรแกรม Unites คำดังกล่าวแยกได้เป็น 2 กลุ่ม คือ คำนามแสดงอุปกรณ์ด้านวิทยาศาสตร์และคำนามแสดงอุปกรณ์ที่อยู่นอกสาขาวิทยาศาสตร์

บทความที่นำมาเป็นคลังข้อมูลมาจากสิ่งตีพิมพ์และทางอินเทอร์เน็ต และแบ่งได้เป็น 4 สาขา ได้แก่

(2.1) บทความด้านวิทยาศาสตร์และเทคโนโลยีมาจากคอลัมน์ “สารคดีวิทยาศาสตร์” และ “มองไทย” นิตยสารอแพค

(2.2) บทความด้านวิทยาการทางการแพทย์และสุขภาพมาจากคอลัมน์ “โลกกว้างทางการแพทย์” “ตามประสาหมอ” และ “Drug & Supplement” นิตยสารใกล้หมอ

(2.3) บทความด้านบ้านและการตกแต่งมาจาก 2 แหล่ง คือ

(2.3.1) จากสิ่งตีพิมพ์ คือ คอลัมน์ “ผลิตภัณฑ์” นิตยสารบ้านและสวน

(2.3.2) จากอินเทอร์เน็ต คือ คอลัมน์ “ผลิตภัณฑ์” นิตยสารบ้านและสวน ทางเว็บไซต์ <http://www.baanlaesuan.com>

(2.4) บทความด้านยานยนต์ นิตยสารฟอร์มูลาซึ่งเผยแพร่ผ่านทางเว็บไซต์ <http://autoinfo.co.th/formula>

จากการทดสอบ พบว่า เมื่อทดสอบกับรายการคำศัพท์ ทั้งค่าความครบถ้วน (Recall) และค่าความแม่นยำ (Precision) สูงมาก คือ ประมาณร้อยละ 99 เพราะเป็นรายการคำศัพท์ที่ปรากฏ 1 คำต่อ 1 บรรทัด ทำให้ไม่ต้องพึ่งพาบริบทใด ๆ แต่เมื่อทดสอบกับคลังข้อมูลที่เป็นบทความค่าครบถ้วน (Recall) ต่ำ คือ ประมาณร้อยละ 18 สำหรับคำนามแสดงอุปกรณ์ด้านวิทยาศาสตร์ และประมาณร้อยละ 36 สำหรับคำนามที่อยู่นอกสาขา และค่าความแม่นยำ (Precision) ต่ำมากเช่นกัน คือ โดยเฉลี่ยเพียงร้อยละ 1 เนื่องจากพบข้อผิดพลาด (error) จำนวนมาก ข้อผิดพลาดดังกล่าว ได้แก่

(1) โปรแกรมตรวจจับคำที่ไม่ต้องการจำนวนมากเพราะไม่ได้ระบุคุณลักษณะ (feature) ของคำแต่ละคำในส่วนประกอบ คำที่ปรากฏมาจากชนิดของคำที่ตรงกันเท่านั้น กรณีนี้ ผู้วิจัยแก้ปัญหาด้วยการกำกับหมวดหมู่ทางความหมาย (คำ AKO) เพิ่มลงไป

(2) โปรแกรมตรวจจับคำทุกรูปแบบที่มีโครงสร้างตรงกับกราฟที่สร้างไว้ โดยไม่คำนึงถึงการเว้นวรรค ทำให้ปรากฏโครงสร้างหน่วยจำนวนเป็นจำนวนมาก กรณีนี้ ผู้วิจัยแก้ไขโดยเพิ่มเงื่อนไขทางบริบทว่า ไม่ให้ตรวจจับคำที่มีการเว้นวรรคภายในคำด้วยเครื่องหมาย #

(3) ปัญหาทางภาษาเรื่องคำหลายหน้าที่

(4) โปรแกรมตัดคำผิด

ผู้วิจัยได้แก้ปัญหาด้วยวิธีการกำหนดเงื่อนไขทางบริบท การกำกับหมวดหมู่ทางความหมาย (คำ AKO) ให้แก่คำ และการสร้างกราฟสำหรับแก้ไขการตัดคำผิดในขั้นตอนก่อนการประมวลผล (Preprocessing) กราฟนี้จะกำหนดคำที่ถูกต้องซึ่งโปรแกรมเคยตัดผิด ทำให้โปรแกรมระบุลงไปในข้อมูลอย่างชัดเจนว่าคำคำนั้นเป็นคำใด เป็นคำชนิดใด และมีหมวดหมู่ทางความหมาย (คำ AKO) ใด โปรแกรมจะกำหนดลักษณะที่ถูกต้องทันทีโดยไม่ตรวจสอบในพจนานุกรมอีก ดังนั้น เมื่อผู้วิจัยแก้ปัญหาด้วยวิธีการต่าง ๆ ข้างต้นแล้ว ทำให้สามารถเพิ่มค่าความครบถ้วน (Recall) เป็นร้อยละ 71.41 สำหรับค่านามแสดงอุปกรณ์ด้านวิทยาศาสตร์ และเป็นร้อยละ 81.47 สำหรับค่านามที่อยู่นอกสาขา และเพิ่มค่าความแม่นยำ (Precision) เป็นร้อยละ 30.62 สำหรับค่านามแสดงอุปกรณ์ด้านวิทยาศาสตร์ และเป็นร้อยละ 51.81 สำหรับค่านามที่อยู่นอกสาขา

ทั้งนี้ การที่ค่าความแม่นยำสำหรับค่านามที่อยู่นอกสาขามากกว่าค่าความแม่นยำสำหรับค่านามแสดงอุปกรณ์ด้านวิทยาศาสตร์ เนื่องด้วยเหตุผล ดังนี้

(1) ค่านามแสดงอุปกรณ์ทั่วไปที่อยู่นอกสาขาส่วนใหญ่มีโครงสร้างไม่ซับซ้อน โครงสร้างนั้นคล้ายกับโครงสร้างของค่านามแสดงอุปกรณ์ด้านวิทยาศาสตร์ที่มีคำปรากฏใน โครงสร้างจำนวนมาก ตัวอย่างเช่น

โครงสร้าง [อุปกรณ์ - วัตถุประสงค์] มีค่านามแสดงอุปกรณ์ด้านวิทยาศาสตร์ในรายการคำศัพท์จำนวน 536 คำ

โครงสร้าง [อุปกรณ์ - วิธีการ] มีค่านามแสดงอุปกรณ์ด้านวิทยาศาสตร์ในรายการคำศัพท์จำนวน 95 คำ

(2) คำนามแสดงอุปกรณืทั่วไปที่อยู่นอกสาขามีจำนวนคำในคลังข้อมูลมากกว่าคำนามแสดงอุปกรณืด้านวิทยาศาสตร์ คือ คำนามแสดงอุปกรณืด้านวิทยาศาสตร์มีจำนวน 815 คำ ในขณะที่คำนามแสดงอุปกรณืที่อยู่นอกสาขาวิทยาศาสตร์มีจำนวน 1,209 คำซึ่งมีจำนวนคำมากกว่า 394 คำ

(3) ผู้วิจัยได้คัดเลือกคำนามแสดงอุปกรณืทั้งด้านวิทยาศาสตร์และนอกสาขาวิทยาศาสตร์จากการอ่านบทความทั้งหมด โดยไม่ได้วิเคราะห์โครงสร้างภายในของคำนามแสดงอุปกรณืเหล่านั้น จึงมีความเป็นไปได้ที่ว่ อาจมีคำนามแสดงอุปกรณืด้านวิทยาศาสตร์บางคำไม่ถูกตรวจจับได้

ด้วยสาเหตุดังกล่าว ค่าความแม่นยำสำหรับคำนามที่อยู่นอกสาขาจึงมากกว่าค่าความแม่นยำสำหรับคำนามแสดงอุปกรณืด้านวิทยาศาสตร์

ผู้วิจัยยังพบว่า การกำกับหมวดหมู่ทางความหมาย (ค่า AKO) ให้แก่คำเป็นส่วนสำคัญอย่างมากในการแก้ไขปัญหาการตรวจจับคำได้มากเกินไป เพราะการกำกับหมวดหมู่ทางความหมาย (ค่า AKO) ให้แก่คำช่วยให้สามารถคัดกรองคำที่ไม่เกี่ยวข้องได้สูงกว่าการระบุเพียงชนิดของคำเท่านั้น

ทั้งนี้ การได้ค่าความครบถ้วน (Recall) สูงจะเป็นประโยชน์ต่อการนำโครงสร้างที่วิเคราะห์ไปประยุกต์กับการประมวลผลภาษาธรรมชาติในด้านต่าง ๆ ต่อไปได้เพราะค่าความครบถ้วนสูงแสดงให้เห็นถึงความสามารถในการรู้จำด้วยเครื่องว่า เครื่องสามารถเรียนรู้โครงสร้างและสามารถตรวจจับคำที่มีโครงสร้างที่ต้องการได้ครบถ้วนอย่างมาก

นอกจากนี้ จากข้อมูลคำนามแสดงอุปกรณืด้านวิทยาศาสตร์ ผู้วิจัยยังพบว่า ปัจจุบัน ได้เกิดการสร้างคำวิธีใหม่ขึ้นเพื่อเรียกสิ่งต่าง ๆ กล่าวคือ แต่เดิม เมื่อต้องการสร้างคำใหม่ในภาษาไทย มักใช้วิธีการสร้างคำด้วยวิธีประสมคำให้เป็นคำประสม วิธีซ้อนคำให้เป็นคำซ้อน และวิธีซ้ำคำให้กลายเป็นคำซ้ำ การสร้างคำด้วยวิธีเหล่านี้มักประกอบด้วยคำเพียง 2-3 คำ ส่วนการสร้างคำเพื่อใช้เรียกวิชาการความรู้ หรือสิ่งประดิษฐ์ใหม่ ๆ ที่เข้ามาในประเทศไทยและสิ่งเหล่านั้นยังไม่มีคำไทยใช้ มักสร้างคำเพื่อใช้เรียกสิ่งเหล่านั้นด้วยวิธีการบัญญัติศัพท์ การบัญญัติศัพท์จะประกอบด้วยคำ

จำนวนน้อย ส่วนมากเพียง 2 – 3 คำ ทำให้คำใหม่ที่เกิดขึ้นมักเป็นคำสั้น ๆ ไม่ยืดยาว ฟังดูกะทัดรัด แต่ปัจจุบัน การสร้างคำเพื่อเรียกสิ่งต่าง ๆ ได้มีการพัฒนาและเปลี่ยนแปลงรูปแบบไปจากเดิม คือ คำที่สร้างขึ้นมักเป็นคำที่ยาวและมีโครงสร้างเลียนแบบวลีหรือประโยค ผู้สร้างคำไม่นิยมการประสมคำเพียง 2 – 3 คำเป็นคำสั้น ๆ ดังแต่ก่อน แต่นิยมขยายความจากส่วนหลักมากกว่า ดังที่ อมรา ประสิทธิ์รัฐสินธุ์ (2545) กล่าวถึงศัพท์บัญญัติว่า ศัพท์บัญญัติในปัจจุบันก็ยังมีลักษณะยาวจนเกือบเหมือนคำอธิบาย ทั้งยังเริ่มปรากฏการใช้อนุภาคเป็นส่วนขยายด้วย

นอกเหนือจากประเด็นข้างต้น คำนามแสดงอุปกรณ์ด้านวิทยาศาสตร์ที่ผู้วิจัยเลือกใช้เป็นแหล่งข้อมูลนั้นเป็นคำที่ได้มาจากการแปลคำเรียกอุปกรณ์ด้านวิทยาศาสตร์ที่เป็นภาษาอังกฤษ คำนามแสดงอุปกรณ์ด้านวิทยาศาสตร์ที่เกิดจากการแปลนี้แสดงถึงมโนทัศน์ทางวิทยาศาสตร์จากภาษาอังกฤษค่อนข้างมาก เนื่องจากอุปกรณ์ด้านวิทยาศาสตร์ที่ปรากฏในแหล่งข้อมูลเป็นอุปกรณ์ที่ยังไม่มีคำเรียกเป็นภาษาไทย ไม่มีคำอธิบายเป็นภาษาไทยที่สมบูรณ์ด้วยคำสั้น ๆ และผู้ที่อยู่ในสาขาวิทยาศาสตร์นั้น ๆ เองก็นิยมเรียกชื่ออุปกรณ์ด้วยวิหัตศัพท์ ดังนั้น เมื่อต้องการสร้างคำเรียกอุปกรณ์เหล่านั้น จึงเกิดการสร้างคำด้วยวิธีการแปลเพื่อให้สามารถอธิบายลักษณะของอุปกรณ์ได้ชัดเจน คำนามแสดงอุปกรณ์ด้านวิทยาศาสตร์ในปัจจุบันจึงมีลักษณะยืดยาว ชับซ้อน และคล้ายกับประโยค ทั้งนี้ คำนามแสดงอุปกรณ์ด้านวิทยาศาสตร์ที่ใช้วิธีการแปลนี้ได้แสดงมโนทัศน์ทางวิทยาศาสตร์ที่ว่า อุปกรณ์ด้านวิทยาศาสตร์มักมีชื่อที่บ่งบอกว่าเป็นอุปกรณ์ เช่น กล้อง เครื่องอุปกรณ์ เป็นต้น หรือบ่งบอกรูปร่างของอุปกรณ์ ชื่อของอุปกรณ์ด้านวิทยาศาสตร์มักปรากฏชื่อผู้ประดิษฐ์หรือผู้ค้นพบอุปกรณ์นั้น ๆ เพื่อเป็นการให้เกิดเกียรติ มักบอกรหัสประจำตัวประจำตัวการใช้งาน ส่วนประกอบเด่น วัตถุประสงค์ และวิธีการทำงานของอุปกรณ์ และมักปรากฏคำศัพท์เฉพาะทางด้านวิทยาศาสตร์เป็นส่วนขยายในคำเรียกอุปกรณ์ เมื่อแปลคำเรียกอุปกรณ์ด้านวิทยาศาสตร์เป็นภาษาไทยแล้ว ก็ยังคงปรากฏคำศัพท์เฉพาะทางด้านวิทยาศาสตร์อยู่ในลักษณะการทับศัพท์ภาษาอังกฤษ เช่น ปรากฏชื่อผู้ประดิษฐ์คิดค้นในคำว่า “เครื่องยนต์วังเกล” “อุปกรณ์แบบ Kipp” ปรากฏคำศัพท์เฉพาะทางด้านวิทยาศาสตร์ในคำว่า “กล้องจุลทรรศน์อิเล็กทรอนิกส์” “บารอมิเตอร์แบบแอนิรอยด์” “คาร์บูเรเตอร์แบบคูลลง” “เครื่องเอกซเรย์ฟลิน” เป็นต้น

6.2 ข้อเสนอแนะ

จากการศึกษาคำนามแสดงอุปกรณ์ด้านวิทยาศาสตร์ ผู้วิจัยเห็นว่า คำนามแสดงอุปกรณ์ด้านวิทยาศาสตร์ส่วนใหญ่เป็นคำประสม หากไม่ใช้ก็เนื่องจากผ่านกระบวนการแปลโดยยังไม่ผ่านการเลือกใช้คำเพื่อแสดงความหมายเป็นคำคำเดียวโดยไม่มีคำเชื่อม แต่โดยรวมแล้ว คำนามแสดงอุปกรณ์ด้านวิทยาศาสตร์มีลักษณะเป็นคำประสม โครงสร้างของคำมีความเป็นสากลเพราะสามารถตรวจจับคำประสมที่อยู่นอกสาขาได้ด้วย จึงควรขยายการศึกษาโครงสร้างคำประสมในกลุ่มอื่นๆ ด้วยเพื่อให้ทราบโครงสร้างคำประสมทั้งหมดในภาษาไทย เช่น ศิลปะ งานช่าง อาชีพ เป็นต้น

การทราบโครงสร้างคำประสมในภาษาไทยจะเป็นประโยชน์มากต่อการนำโครงสร้างไปประยุกต์ใช้ทางการประมวลผลภาษาธรรมชาติ โดยเฉพาะการสร้างพจนานุกรมอิเล็กทรอนิกส์ที่ทันสมัย การสร้างเครือข่ายคำ การแปลภาษา และการตัดคำแบบเบ็ดเสร็จด้วยคอมพิวเตอร์

นอกจากนี้ ยังควรศึกษาโครงสร้างของนามวลีในภาษาไทยเพื่อหาความแตกต่างในการปรากฏระหว่างคำประสม และนามวลี สิ่งนี้จะช่วยให้สามารถแยกแยะคำประสมเพื่อเก็บเป็นคำศัพท์ในพจนานุกรมได้ง่ายยิ่งขึ้นด้วย

สำหรับการประยุกต์ผลการวิเคราะห์ทางภาษาศาสตร์กับเครื่องคอมพิวเตอร์ ควรมีการศึกษาเพื่อเติมเพื่อทำการทดลองด้วยวิธีการอื่น ๆ ที่หลากหลายมากขึ้น อันจะส่งผลให้สามารถพบแนวทางที่เหมาะสมและใช้ได้ผลที่สุดในการประยุกต์ผลการวิเคราะห์กับเครื่องคอมพิวเตอร์