

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

**หนังสืออ้างอิง
ห้ามนำออกนอกห้องสมุด**

การวิจัยและพัฒนาสื่อภาษากลางเพื่อการแปล
หลายภาษาด้วยคอมพิวเตอร์

A RESEARCH AND DEVELOPMENT ON INTERLINGUA
FOR MULTILINGUAL TRANSLATION WITH COMPUTER



นายวิเชียร อัสวธีระ
MR. VICHEAN ASSWATERA



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิศวกรรมไฟฟ้า

บัณฑิตวิทยาลัย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2537

ISBN 974-621-204-4

ลิขสิทธิ์ของบัณฑิตวิทยาลัย สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

เลขหมู่.....

เลขทะเบียน.....22016.....มิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วัน เดือน ปี.....15 S.A. 2537

A RESEARCH AND DEVELOPMENT ON INTERLINGUA
FOR MULTILINGUAL TRANSLATION WITH COMPUTER

MR. VICHEAN ASSWATERA

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE
MASTER OF ENGINEERING IN ELECTRICAL ENGINEERING
GRADUATE SCHOOL

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

1994

ISBN 974-621-204-4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อวิทยานิพนธ์	การวิจัยและพัฒนาสื่อภาษากลางเพื่อการแปลหลายภาษา ด้วยคอมพิวเตอร์
นักศึกษา	นายวิเชียร อัครธีระ
อาจารย์ผู้ควบคุมวิทยานิพนธ์	ผศ.ดร. รัตติกร วรากุลศิริพันธ์
ระดับการศึกษา	วิศวกรรมศาสตรมหาบัณฑิต สาขาวิศวกรรมไฟฟ้า
ภาควิชา	วิศวกรรมไฟฟ้า สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณ ทหารลาดกระบัง
ปีการศึกษา	2537

บทคัดย่อ

ระบบการแปลภาษาในปัจจุบันมีหลายวิธี มีการวิจัยและพัฒนาหลายลักษณะเพื่อให้ได้ภาษาเป้าหมายที่ได้ใจความสมบูรณ์ที่สุด โดยอาศัยหลักการทางด้านการประมวลผลภาษาธรรมชาติ และวิทยาการทางซอฟต์แวร์ วิธีการแปลโดยใช้ภาษากลางเป็นอีกวิธีหนึ่งที่ใช้ในการแปลภาษาด้วยคอมพิวเตอร์ โดยจะใช้ภาษากลางเป็นสื่อถ่ายทอดโครงสร้างทางไวยากรณ์และความหมายของภาษาต้นแบบไปสู่ภาษาเป้าหมาย ซึ่งเป็นวิธีหนึ่งที่ทำให้ความถูกต้อง และมีประสิทธิภาพในการแปลสูง นอกจากนี้ยังเหมาะสำหรับการแปลหลายๆภาษา ทั้งนี้เพราะจะลดจำนวนซอฟต์แวร์ จำนวนพจนานุกรม และขั้นตอนการถ่ายทอดไม่ว่าจะเป็นการวิเคราะห์หรือการสังเคราะห์ ให้ลดน้อยลง ซึ่งในวิทยานิพนธ์นี้ได้ออกแบบสื่อภาษากลางเพื่อให้ครอบคลุมหลายๆ ภาษา โดยใช้ฐานความรู้ทางด้านภาษาศาสตร์ และการประมวลผลภาษาธรรมชาติเป็นหลัก และคำนึงถึงความสามารถที่จะนำไปประยุกต์ทางซอฟต์แวร์ เพื่อให้ได้ความถูกต้องในการถ่ายทอดโครงสร้างประโยค ถ่ายทอดความหมาย และหน้าที่ของคำ ทั้งนี้เพื่ออำนวยความสะดวกให้กับการวิเคราะห์ภาษาต้นแบบ และการสร้างภาษาเป้าหมาย อันจะนำไปสู่การแปลภาษาที่มีประสิทธิภาพต่อไป

Thesis Title	A Research and Development on Interlingua for Multilingual Translation with Computer.
Student	Mr. Vichean Asswatera
Thesis Adviser	Asst.Prof.Dr. Ruttikorn Varakulsiripunth
Level of Study	Master of Engineering in Electrical Engineering
Department	Electrical Engineering King Mongkut's Institute of Technology Ladkrabang
Academic year	1994

ABSTRACT

In the present, there are many methods in the field of machine translation (MT) system that had been introduced so far. Aim of research and develop is the complete sentence. Most of methods are based on natural language processing and computer software techniques. The interlingua is another method used in MT system and used to be an intermediate language in order to transfer the syntactic and semantic information of source language into target language generation. This methodology has been successfully recognized in both accuracy and efficiency, besides, the interlingua approach is also suitable for multilingual MT system, because the amount software and dictionaries, and the transformation processes had largely been reduced. In this thesis. We have researched and designed our interlingua system to cover many languages, that based on linguistic knowledge base and the technique of natural language processing. This interlingua system is designed by considering to its simple software application. Moreover, we have designed it with emphasis on the most correctness of syntactic and semantic information transferring from source language to target language. This will fall into the achievement of high performance MT system.

กิตติกรรมประกาศ

ขอกราบขอบพระคุณคุณพ่อ และคุณแม่ของข้าพเจ้าที่ให้การสนับสนุนและให้โอกาสด้านการศึกษา ดูแลเอาใจใส่และให้กำลังใจต่อข้าพเจ้ามาโดยตลอด

ขอขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร. รัตติกร วรากุลศิริพันธ์ ที่ได้ให้ความอนุเคราะห์ ทั้งด้านความรู้ คำปรึกษา และแนะนำการจัดทำวิทยานิพนธ์ฉบับนี้จนสำเร็จลุล่วงไปด้วยดี

ขอขอบพระคุณคณาจารย์ทุกๆ ท่านของคณะวิศวกรรมศาสตร์ที่ให้ความรู้แก่ข้าพเจ้า

ขอขอบพระคุณ Center of the International Cooperation for Computerization ของรัฐบาลญี่ปุ่น (CICC) และห้องปฏิบัติการวิจัยภาษาและวิทยาการความรู้ (LINKS) ของกระทรวงวิทยาศาสตร์ เทคโนโลยีและสิ่งแวดล้อม และศูนย์วิจัยและพัฒนาปัญญาประดิษฐ์ของสถาบันเทคโนโลยีพระจอมเกล้าเจ้าอนบูร (สจธ) ที่กรุณาให้ข้อมูลเพื่อทำการวิจัยมาโดยตลอด

ขอขอบพระคุณคณาจารย์และเจ้าหน้าที่สำนักวิจัยและบริการคอมพิวเตอร์ที่กรุณาให้ความช่วยเหลือในทุกๆด้าน

ขอขอบพระคุณศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ มูลนิธิเพื่อการศึกษา คอมพิวเตอร์และการสื่อสาร และสมาคมประกันวินาศภัย ที่กรุณาให้ทุนในการวิจัย

ขอขอบคุณพี่ๆ เพื่อนๆ และน้องๆ ทุกคนที่ช่วยให้วิทยานิพนธ์นี้สำเร็จลุล่วงไปด้วยดี และสุดท้ายขอขอบพระคุณ

ผู้ช่วยศาสตราจารย์ ดร. รัตติกร วรากุลศิริพันธ์

รองศาสตราจารย์ ดร. ชม กัมปาน

ดร. บุญธีร์ เครือตราชู

ดร. เอื้อน ปิ่นเงิน

ผู้ช่วยศาสตราจารย์ สุขุมาลัย นิลรัตน์ (อาจารย์ประจำภาควิชาภาษาและสังคม
คณะครุศาสตร์)

ที่ได้ให้ความกรุณาสละเวลามาเป็นกรรมการสอบวิทยานิพนธ์ในครั้งนี้

สารบัญ

	หน้า
บทคัดย่อ	I
Abstract	II
กิตติกรรมประกาศ	III
สารบัญ	IV
สารบัญภาพประกอบ	VII
สารบัญตาราง	IX
บทที่ 1 บทนำ	1
1.1 ความเป็นมาของการแปลภาษาด้วยคอมพิวเตอร์	1
1.1.1 การแปลโดยตรง	3
1.1.2 การแปลแบบถ่ายทอด	3
1.1.3 การแปลแบบใช้ภาษากลางเป็นสื่อ	4
1.2 จุดประสงค์และขอบเขตของงานวิจัยในวิทยานิพนธ์นี้	7
1.2.1 การกำหนดรูปแบบของภาษากลาง	8
1.2.2 การออกแบบและสร้างพจนานุกรม	8
1.2.3 สร้างกฎทางความสัมพันธ์ของประโยค	9
1.2.4 การออกแบบและพัฒนาซอฟต์แวร์ของระบบ	9
1.3 โครงร่างของวิทยานิพนธ์	9
บทที่ 2 การออกแบบโครงสร้างภาษากลาง	11
2.1 บทนำ	11
2.2 โครงสร้างภาษากลางในเชิงภาษาศาสตร์คอมพิวเตอร์	13
2.3 สรุป	17

สารบัญ (ต่อ)

	หน้า
บทที่ 3 หน่วยความหมาย	18
3.1 บทนำ	18
3.2 รหัสแยกแยะหน่วยความหมาย : CPID	22
3.3 คำอธิบายหน่วยความหมาย : CPDC	32
3.4 สรุป	33
บทที่ 4 หน่วยความสัมพันธ์	34
4.1 บทนำ	34
4.2 แนวความคิดเกี่ยวกับความสัมพันธ์ทางอรรถศาสตร์ระหว่างหน่วยคำ	35
4.2.1 แนวความคิดของชาร์ลส์ เจ. ฟิลล์มอร์	35
4.2.2 แนวความคิดของกรูเบอร์	37
4.2.3 แนวความคิดของวอลเลซ แอล เซฟ	38
4.2.4 แนวความคิดของนิลเซน	40
4.2.5 แนวความคิดของอาร์ อี ลองเอเคอร์	42
4.2.6 แนวความคิดของจอห์น เอ็ม แอนเดอร์สัน	44
4.3 การศึกษาแนวความคิดเกี่ยวกับความสัมพันธ์ทางอรรถศาสตร์ของ บุคคลที่ได้กล่าวมา	46
4.4 กลุ่มความสัมพันธ์หลัก	47
4.5 กลุ่มความสัมพันธ์คุณสมบัติ	62
4.6 กลุ่มความสัมพันธ์เหตุการณ์	65
4.7 กลุ่มความสัมพันธ์คำเชื่อม	68
4.8 การกำหนดรหัสความสัมพันธ์	81

สารบัญ (ต่อ)

	หน้า
4.9 สรุป	86
บทที่ 5 พจนานุกรมเพื่อการแปลหลายภาษาด้วยคอมพิวเตอร์	87
5.1 บทนำ	87
5.2 พจนานุกรมความหมาย	88
5.3 พจนานุกรมภาษา	101
5.4 สรุป	107
บทที่ 6 ตัวอย่างการสังเคราะห์ภาษาเป้าหมายจากภาษากลาง	108
6.1 บทนำ	108
6.2 ขั้นตอนการสังเคราะห์ภาษาเป้าหมายจากภาษากลาง	109
6.3 ขั้นตอนการใช้โปรแกรมในการสังเคราะห์ภาษาเป้าหมายจากภาษากลาง	113
6.4 สรุป	140
บทที่ 7 บทสรุปและข้อเสนอแนะ	141
7.1 สรุป	141
7.2 ข้อเสนอแนะ	142
7.3 การนำผลงานวิจัยไปประยุกต์ใช้กับงานอื่น	143
เอกสารอ้างอิง	144
ภาคผนวก ก. พจนานุกรมความหมาย	146
ภาคผนวก ข. พจนานุกรมเฉพาะภาษาไทย	155
ภาคผนวก ค. พจนานุกรมเฉพาะภาษาอังกฤษ	163

สารบัญภาพประกอบ

	หน้า
รูปที่ 1.1 การสื่อความหมายด้วยภาษา	1
รูปที่ 1.2 การสื่อสารโดยการถ่ายทอด	2
รูปที่ 1.3 การแปลโดยตรง	3
รูปที่ 1.4 การแปลแบบถ่ายทอด	4
รูปที่ 1.5 การใช้ภาษากลางเป็นสื่อ	5
รูปที่ 1.6 โครงสร้างการแปลภาษาในลักษณะต่างๆ	5
รูปที่ 1.7 โครงสร้างของวิทยานิพนธ์	10
รูปที่ 2.1 โครงสร้างของระบบการแปลหลายภาษาโดยใช้ภาษากลางเป็นสื่อการแปล	12
รูปที่ 2.2 ตัวอย่างโครงข่ายของโครงสร้างภาษากลาง	15
รูปที่ 2.3 Interdependency	16
รูปที่ 2.4 Double dependency	17
รูปที่ 3.1 ตัวอย่างการแทนรหัสความหมายของความหมายของคำ	21
รูปที่ 3.2 แสดงการแบ่งกลุ่มรหัสความหมาย	30
รูปที่ 3.3 ตัวอย่างรหัสแยกแยะหน่วยความหมาย	31
รูปที่ 4.1 กลุ่มความสัมพันธ์ของหน่วยคำของนักภาษาศาสตร์	45
รูปที่ 4.2 การกำหนดรหัสความสัมพันธ์	82
รูปที่ 5.1 พจนานุกรมรวมภาษา	88
รูปที่ 5.2 Data Structure ของพจนานุกรมภาษาใน 1 record ของ 1 รหัส	101
รูปที่ 6.1 การสังเคราะห์ภาษาเป้าหมายจากภาษากลาง	108
รูปที่ 6.2 ตัวอย่างการกำหนดโนดของภาษากลางทางซอฟต์แวร์	110
รูปที่ 6.3 แสดงหน้าจอการเลือกคำสั่งในการทำงาน	114

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา VII ละต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญภาพประกอบ(ต่อ)

	หน้า
รูปที่ 6.4 แสดงหน้าจอการรับข้อมูลทางแป้นพิมพ์	115
รูปที่ 6.5 การแสดงโครงสร้างภาษากลางในรูปโครงสร้างต้นไม้ฟังก์ชัน	116
รูปที่ 6.6 จอภาพแสดงคำสั่งในการแปลภาษากลางไปเป็นภาษาเป้าหมาย	117
รูปที่ 6.7 จอภาพแสดงผลลัพธ์ของการสังเคราะห์ภาษาเป้าหมายจากภาษากลาง	118
รูปที่ 6.8 จอภาพแสดงกลุ่มพจนานุกรมของภาษาต่างๆ	119



สารบัญตาราง

	หน้า
ตารางที่ 1.1 การเปรียบเทียบวิธีการแปลภาษาด้วยคอมพิวเตอร์	6
ตารางที่ 3.1 ตัวอย่างคำที่มีหลายความหมายในภาษาต่างๆ	19
ตารางที่ 3.2 ตัวอย่างการกำหนดความหมายในระดับลึกของ CPID	33
ตารางที่ 5.1 พจนานุกรมความหมายของคำว่า "coach"	89
ตารางที่ 5.2 พจนานุกรมความหมายของคำว่า "lost"	90
ตารางที่ 5.3 พจนานุกรมความหมายของคำว่า "set"	94
ตารางที่ 5.4 สรุปข้อมูลคำศัพท์	99
ตารางที่ 5.5 ตัวอย่างพจนานุกรมความหมาย	100
ตารางที่ 5.6 ตัวอย่างพจนานุกรมเฉพาะสำหรับภาษาไทย	105
ตารางที่ 5.7 ตัวอย่างพจนานุกรมเฉพาะสำหรับภาษาอังกฤษ	106
ตารางที่ 6.1 โครงสร้างของการเก็บข้อมูลในซอฟต์แวร์	112

บทที่ 1

บทนำ

1.1 ความเป็นมาของการแปลภาษาด้วยคอมพิวเตอร์

นับตั้งแต่อดีตมากระบบการเรียนรู้ และการถ่ายทอดความรู้ของมนุษย์จะอยู่ในรูปของการบอกเล่า เพื่อให้ชนรุ่นหลังได้รู้ถึงที่มาของบรรพบุรุษ รวมทั้งวิทยาการต่างๆ ที่ได้มีการพัฒนา และสืบทอดต่อมา และการที่จะทำให้ต่างชนชาติ ต่างเผ่าพันธุ์ ต่างภาษาได้มีการเรียนรู้วิทยาการของเผ่าอื่นๆ จำเป็นต้องมีการถ่ายทอดกัน ด้วยรูปภาพ หรือตัวอักษร ซึ่งจะต้องมีการตีความหมายออกมา หรือเรียกอีกอย่างหนึ่งก็คือ การแปล ดังแสดงในรูปที่ 1.1

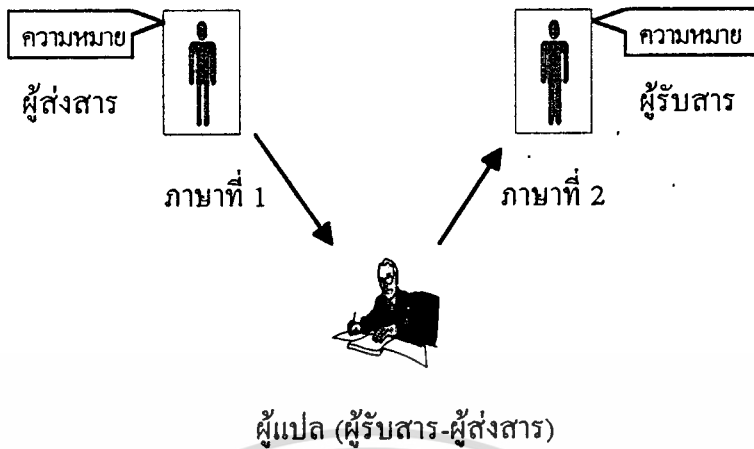


รูปที่ 1.1 การสื่อความหมายด้วยภาษา

จวบจนปัจจุบันมนุษยชาติได้มีการติดต่อสื่อสารกันไม่ว่าจะเป็นทางการเมือง ธุรกิจ หรือแม้กระทั่งการถ่ายทอดวิทยาการความเจริญต่างๆ การแปลภาษาก็ยังเป็นปัจจัยที่สำคัญที่จะถ่ายทอดความหมายไปสู่ชนชาติอื่นที่ไม่ทราบภาษาต่างชาติให้มีความรู้ ความเข้าใจ ในข่าวสารต่างๆ การแปลภาษาด้วยเครื่องคอมพิวเตอร์ (Machine Translation) นั่นก็เป็นการเลียนแบบการแปลภาษาของมนุษย์ซึ่งมีจุดประสงค์คือ เพื่อการถ่ายทอดความหมายจากภาษาหนึ่งไปเป็นอีกภาษาหนึ่ง และก็มีวิธีการที่คล้ายคลึงกันคือ การแปลภาษาของมนุษย์จะเริ่มจากการศึกษาภาษาต้นฉบับ แล้วทำการวิเคราะห์ต่อจากนั้นก็ตีความหมาย และจับใจความของภาษาต้นฉบับ แล้วถ่ายทอดออกมาในรูปของภาษาเป้าหมาย ส่วนการแปลภาษาด้วยเครื่องคอมพิวเตอร์นั้น ก็เริ่มจากการป้อนข้อมูลภาษาต้นฉบับ โดยที่ได้ป้อนความรู้ที่จะใช้ในการแปลแก่คอมพิวเตอร์แล้ว คอมพิวเตอร์จะใช้ความรู้นั้นทำการวิเคราะห์ภาษาต้นฉบับ ภาษาเป้าหมายที่ต้องการดังตัวอย่างที่แสดงในรูปที่ 1.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 1.2 การสื่อสารโดยการถ่ายทอด

การแปลภาษาด้วยคอมพิวเตอร์เป็นสาขาวิชาหนึ่งในการศึกษาด้านการประมวลผลภาษาธรรมชาติ (Natural Language Processing) หลักฐานชิ้นแรกๆ ที่ชี้ให้เห็นถึงความเป็นไปได้ที่จะนำเครื่องคอมพิวเตอร์มาใช้แปลภาษาได้คือ การยื่นขอจดทะเบียนลิขสิทธิ์ในกรุงมอสโคว์ เมื่อปี 1933¹¹ สำหรับระยะเริ่มแรกของวิวัฒนาการของระบบการแปลภาษาด้วยเครื่องคอมพิวเตอร์นั้น มีแนวความคิดในการสร้างระบบการแปลแบบคำต่อคำ โดยอาศัยพจนานุกรมคำศัพท์ (Dictionary) ขนาดใหญ่ ในชวงต่อมาเกิดการแตกแยกทางความคิดเป็น 2 แนวทาง แนวทางแรก เห็นว่าควรแปลเป็นคู่ภาษา แนวทางที่สอง เห็นว่าควรแปลเป็นภาษากลาง (Interlingua) ก่อน และต่อมาแนวความคิดในการใช้ภาษากลางได้ขยายตัวกว้างขึ้น และมีการพัฒนาเรื่อยมาเพื่อหาวิธีแก้ไขความกำกวมทางภาษา ซึ่งต้องใช้ความรู้ทางอรรถศาสตร์ (Semantic) มาประยุกต์ใช้และอาศัยความร่วมมือกันของนักวิจัยสาขาต่างๆ

ดังนั้นจากการศึกษางานวิจัยเกี่ยวกับการพัฒนาระบบการแปลภาษาด้วยคอมพิวเตอร์ จากอดีตถึงปัจจุบันสรุปได้เป็น 3 แนวทาง²¹ ได้แก่

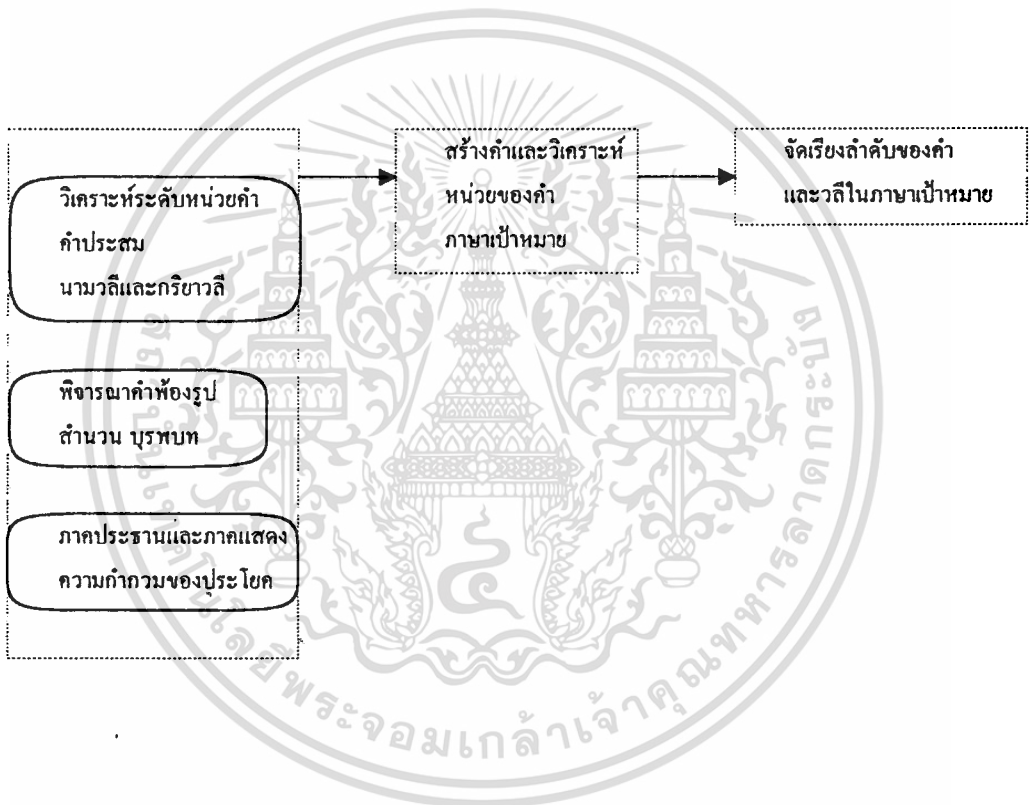
- 1) การแปลโดยตรง
- 2) การแปลแบบถ่ายทอด
- 3) การแปลแบบใช้ภาษากลางเป็นสื่อของการถ่ายทอด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 2. ละต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แต่ละวิธีการ มีขั้นตอนที่แตกต่างกันไป ดังต่อไปนี้

1.1.1 การแปลโดยตรง (Direct Machine Translation)

การแปลจากภาษาต้นฉบับไปสู่ภาษาเป้าหมาย ไม่มีการใช้ทฤษฎีภาษาศาสตร์ หรือหลักการทางวชิภาค (Parsing Principles) ใด ๆ ระบบขึ้นอยู่กับพจนานุกรมที่ใช้ การวิเคราะห์หน่วยคำ และโปรแกรมการวิเคราะห์ข้อความ เพื่อแปลภาษาต้นฉบับเป็นภาษาเป้าหมาย การแปลแบบนี้มีข้อจำกัดเฉพาะคู่ภาษา จะไม่คำนึงถึงภาษาอื่นๆ การแปลโดยตรงนี้มีขั้นตอนดังรูปที่ 1.3

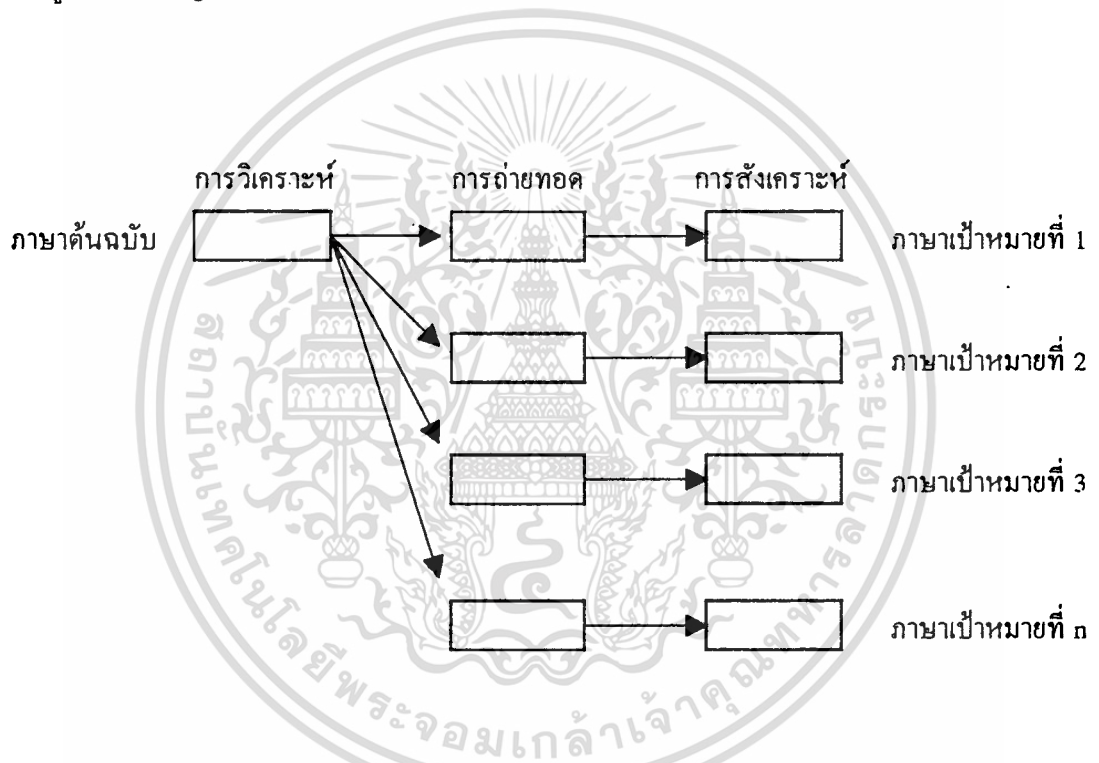


รูปที่ 1.3 การแปลโดยตรง

1.1.2 การแปลแบบถ่ายทอด (Transfer Machine Translation)

เป็นระบบการแปลที่มีการสร้างตัวแทนแสดงความหมายของหน่วยไวยากรณ์ระหว่างภาษาต้นฉบับ และภาษาเป้าหมาย จึงต้องมีช่วงเชื่อมต่อที่เทียบตัวแทนแสดงความหมายของภาษาหนึ่งแล้วเปลี่ยนไปเป็นอีกภาษาหนึ่ง ซึ่งขั้นตอนการเปลี่ยนจะประกอบด้วย การวิเคราะห์ (Analysis) การถ่ายทอด (Transfer) และการสังเคราะห์ (Generation) โดยเริ่มจากประโยคต้นฉบับถูกกระจายให้อยู่ในเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวแทนภายใน (Abstract Internal Representation) จากนั้นจะทำการเปลี่ยนในระดับโครงสร้าง และ ความหมายให้สอดคล้องกับโครงสร้าง และความหมายในภาษาเป้าหมาย และขั้นสุดท้ายก็จะแปล เป็นภาษาเป้าหมาย วิธีนี้จะใช้พจนานุกรม 3 กลุ่ม คือ พจนานุกรมภาษาต้นฉบับใช้ในขั้นวิเคราะห์ พจนานุกรมทวิภาษาใช้ในขั้นการถ่ายทอด และพจนานุกรมภาษาเป้าหมายใช้ในขั้นสังเคราะห์ การ แปลโดยวิธีการถ่ายทอดนี้สามารถประยุกต์ใช้กับระบบแปลหลากหลายภาษา (Multilingual) เพราะมี ลักษณะเป็นการแปลแบบคู่ภาษาที่มีระบบการวิเคราะห์ และการสังเคราะห์แยกเป็นอิสระต่อกัน ดัง แสดงในรูปที่ 1.4 ซึ่งระบบการวิเคราะห์ของการแปลโดยตรงจะทำไมถึงระดับความหมายแต่จะเป็น การเปรียบเทียบคำหลัก (Keyword Matching) ซึ่งเหมาะสำหรับงานเฉพาะด้าน แต่ไม่เหมาะสำหรับการ แปลข้อมูลขนาดใหญ่

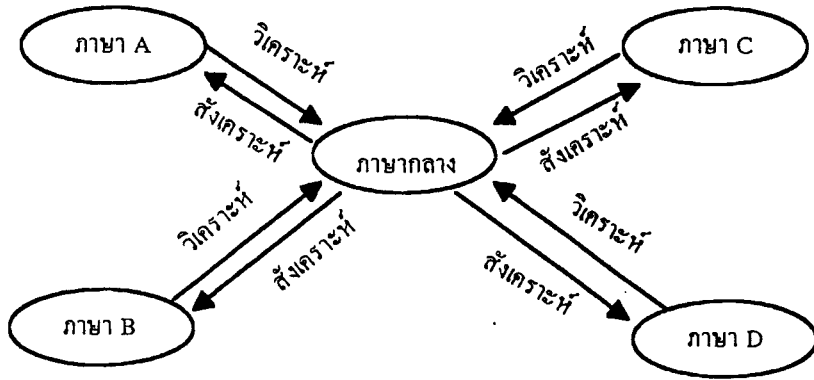


รูปที่ 1.4 การแปลแบบถ่ายทอด

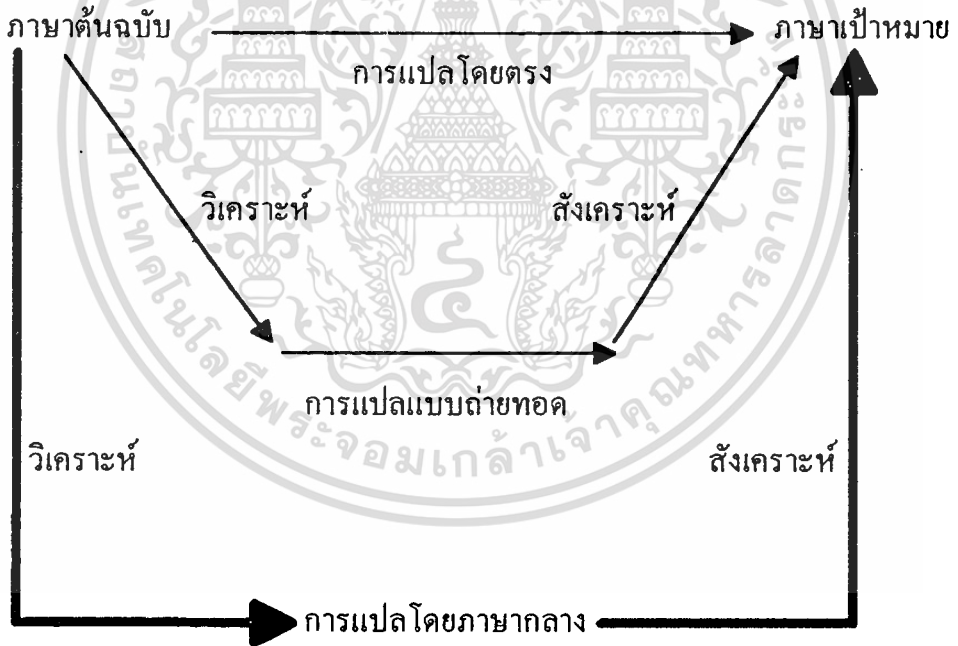
1.1.3 การแปลแบบใช้ภาษากลางเป็นสื่อ (Interlingua Machine Translation)

เป็นวิธีที่พยายามค้นหาความหมาย และหาตัวแทนความหมายมาแสดง โดยไม่ขึ้นอยู่กับ ภาษาใดภาษาหนึ่ง และลักษณะเป็นสากล ตัวแทนความหมายนี้เรียกว่า ภาษากลาง (Interlingua) วิธี การนี้จะลดขั้นตอนการถ่ายทอดทางภาษาดังวิธีที่กล่าวมาในข้อ 1.1.2 ดังนั้นวิธีนี้จึงประกอบด้วย การ วิเคราะห์ และการสังเคราะห์เท่านั้น ดังแสดงดังรูปที่ 1.5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 4 จะต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 1.5 การใช้ภาษากลางเป็นสื่อ



รูปที่ 1.6 โครงสร้างการแปลภาษาในลักษณะต่างๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 1.6 เป็นการแสดงระดับการวิเคราะห์และสังเคราะห์ของการแปลภาษาด้วยวิธีการต่างๆ กันจะเห็นว่าในระบบที่ใช้การถ่ายทอดจะเริ่มทำการถ่ายทอดที่จุดใดก็ได้ แล้วแต่ระดับความลึกของการวิเคราะห์ ถ้ามีการวิเคราะห์มากการถ่ายทอดจะซับซ้อนน้อยลง แต่เมื่อมีการวิเคราะห์ลึกถึงภาษากลางแล้วก็จะไม่มีการถ่ายทอดเกิดขึ้นเพราะจะสร้างประโยคภาษาเป้าหมายได้จากภาษากลางทันที นั่นคือระบบการใช้ภาษากลางนี้ การวิเคราะห์ และการสังเคราะห์ไม่ขึ้นต่อกัน จึงเป็นลักษณะหนึ่งของการแปลหลากหลายภาษาเช่นเดียวกับวิธีการถ่ายทอด วิธีการแปลโดยใช้ภาษากลางเป็นสื่อนี้จะเริ่มจากการวิเคราะห์ภาษาต้นฉบับ แล้วเปลี่ยนเป็นตัวแทนทางความคิด ซึ่งเป็นอิสระไม่ขึ้นกับภาษาใด เรียกว่า Language - Free Conceptual Representation ที่บรรจุด้วยความรอบรู้ (Contextual Knowledge) ทางภาษา จากตัวแทนทางความคิดหรือภาษากลางนี้ก็สามารถเปลี่ยนเป็นภาษาเป้าหมายต่อไป ซึ่งการใช้ภาษากลางจะแตกต่างจากการถ่ายทอดและการแปลโดยตรง เพราะภาษากลางจะลดขั้นตอนในการแปล เพราะภาษากลางจะสื่อความหมายได้ในระดับที่ลึกกว่าการถ่ายทอดและการแปลโดยตรง ทั้งภาษากลางสามารถลดความกำกวมของประโยคได้ เพราะภาษากลางเป็นการวิเคราะห์ทั้งระดับไวยากรณ์ (Syntax) และความหมาย (Semantic) ประสิทธิภาพของการแปลในระบบต่างๆ ที่กล่าวมามีจุดเด่นและข้อเสียเปรียบแตกต่างกันไปดังสรุปในตารางที่ 1.1

กลวิธีการแปลภาษา การทำงาน	การแปล โดยตรง	การแปลแบบ ถ่ายทอด	การแปลแบบใช้ ภาษากลางเป็นสื่อ
1.การวิเคราะห์โครงสร้างของภาษาต้นฉบับ	มี	มี	มี
2.การวิเคราะห์ความหมายของภาษาต้นฉบับ	ไม่มี	มี	มี
3.การแทนภาษาต้นฉบับด้วยภาษากลาง	ไม่มี	ไม่มี	มี
4.การสังเคราะห์โครงสร้างและความหมายของ ภาษาเป้าหมาย	ไม่มี	มี	มี
5.สร้างภาษาเป้าหมาย	มี	มี	มี
6.ชนิดของพจนานุกรมที่ใช้ในระบบ	ทิศทางเดียว	สองทิศทาง	หลายทิศทาง

ตารางที่ 1.1 การเปรียบเทียบวิธีการแปลภาษาด้วยคอมพิวเตอร์^[3]

1.2 จุดประสงค์และขอบเขตของงานวิจัยในวิทยานิพนธ์นี้

จากที่กล่าวมาข้างต้น จะเห็นว่าในการนำเครื่องคอมพิวเตอร์มาใช้กับการแปลหลายภาษา จะต้องมึระบบการจัดการที่ดี เพื่อลดความยุ่งยากซับซ้อน ผู้ที่ต้องการจะแปลภาษาไม่จำเป็นต้องรู้ภาษาเป้าหมายที่ต้องการจะแปล งานวิทยานิพนธ์นี้จึงได้ศึกษาวิจัย พบว่าการใช้ภาษากลางเป็นตัวแทนของภาษาต่างๆ หรือเป็นสื่อระหว่างภาษาต่างๆจะทำให้สามารถที่จะประยุกต์กับวิทยาการทางคอมพิวเตอร์เพื่อการวิเคราะห์ภาษาต้นแบบ (Source Language) ผ่านสื่อภาษากลางไปสู่การสร้างหรือสังเคราะห์ภาษาเป้าหมาย (Target Language) ที่ต้องการจะแปลได้อย่างสมบูรณ์ และถูกต้องมากกว่าวิธีการอื่นๆ

โดยหลักการแล้ว ภาษากลางควรจะเป็นตัวแทนทางโครงสร้างและความหมายให้กับทุกๆ ภาษาได้ ซึ่งจะต้องไม่อิงภาษาใดภาษาหนึ่ง แต่ในความเป็นจริงแล้วเป็นสิ่งที่ยุ่งยากและซับซ้อนมาก ในการที่จะกำหนดภาษากลางดังกล่าว ดังนั้นจึงจำเป็นต้องมีการจำกัด กลุ่มภาษาที่จะร่วมในระบบการแปลโดยใช้ภาษากลางเป็นสื่อ เพื่อเป็นขอบเขตของการกำหนด และออกแบบภาษากลาง ดังเช่น การพัฒนาในกลุ่มภาษายุโรปในระบบการแปลภาษาชื่อ EUROTRA ซึ่งเป็นโครงการวิจัยและพัฒนาของประชาคมเศรษฐกิจยุโรป (European Economic Community - EEC) โดยมีกลุ่มภาษาในระบบได้แก่ ภาษาเดนมาร์ก ดัตช์ อังกฤษ ฝรั่งเศส เยอรมัน กรีก และอิตาลี^[1] นอกจากนั้นยังมีโครงการของระบบแปลภาษาชื่อ TRANSLATOR ที่มหาวิทยาลัยคอลเกตในสหรัฐอเมริกา^[2] และระบบการแปลภาษาในกลุ่มภาษาเอเชียระหว่าง ไทย ญี่ปุ่น จีน อินโดนีเซีย มาเลเซีย ภายใต้โครงการความร่วมมือระหว่างรัฐบาลของประเทศดังกล่าว^[3] ตั้งแต่ปี พ.ศ. 2530 ถึงปี พ.ศ. 2537

[*] หมายเหตุ

รัฐบาลไทย ผ่านทาง ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์ (NECTEC) ซึ่งเป็นหน่วยงานหนึ่งของกระทรวงวิทยาศาสตร์เทคโนโลยีและสิ่งแวดล้อม
รัฐบาลญี่ปุ่น ผ่านทาง Center of International Cooperation for Computerization (CICC) ซึ่งเป็นหน่วยงานหนึ่งของ Ministry of International Trade and Industry (MITI)
รัฐบาลอินโดนีเซีย ผ่านทาง Agency for the Assessment and Application of Technology (BPPT)
รัฐบาลจีน ผ่านทาง China National Computer Software and Technology Service Corporation (CS&S)
รัฐบาลมาเลเซีย ผ่านทาง Ministry of Education (MOE)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 7 ละต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

งานวิทยานิพนธ์นี้ มีจุดเริ่มต้นมาจากการพัฒนาระบบการแปลของกลุ่มภาษาเอเชีย โดยใช้ภาษากลางดังกล่าวข้างต้น เพราะเล็งเห็นถึงความสำคัญของวิทยาการทางการแปลภาษาด้วยคอมพิวเตอร์โดยการใช้สื่อภาษากลาง และได้ทำการวิจัยขยายขอบเขตของภาษากลางให้สามารถถูกนำไปประยุกต์เป็นรูปแทนของภาษาต่างๆได้มากที่สุด โดยคำนึงถึงความถูกต้องทางหลักไวยากรณ์ทางภาษา ความสะดวกในการนำไปประยุกต์ทางซอฟต์แวร์คอมพิวเตอร์ของระบบการแปลหลายภาษา ความสะดวกในการทำความเข้าใจของนักวิเคราะห์ทางภาษาศาสตร์คอมพิวเตอร์ (Computational Linguistics) และการประยุกต์ใช้กับระบบทางภาษาอื่นๆต่อไป อันจะเป็นประโยชน์ต่องานวิจัยค้นคว้า และศึกษาด้านการประมวลผลภาษาธรรมชาติทั้งภายในและภายนอกประเทศ

หลักการทั่วไปของการวิจัยออกแบบและพัฒนาสื่อภาษากลางของวิทยานิพนธ์นี้ จะเน้นให้ภาษากลางเป็นตัวแทนของภาษาต่างๆ ในการเก็บความหมายความสัมพันธ์ของคำ หรือวลี โดยจะให้ความหมายเดียวกันนั้น ซึ่งจะทำให้สามารถที่จะสังเคราะห์ประโยคเป้าหมายได้ตรงตามความหมายของภาษาต้นฉบับ โดยอาศัยวิทยาการทางซอฟต์แวร์สร้างความรู้ทางภาษาให้ระบบคอมพิวเตอร์สามารถประมวลผลได้ และคำแต่ละคำของประโยคในภาษากลางจะต้องไม่คลุมเครือ เพื่อที่จะให้ได้ภาษากลางที่สามารถสื่อความหมายของประโยคได้อย่างถูกต้อง ซึ่งจะต้องรวมถึงการออกแบบสร้างพจนานุกรมและกฎทางภาษาศาสตร์ที่สอดคล้องกับภาษากลางนี้ด้วย รายละเอียดโดยพื้นฐานของการวิจัยพัฒนาภาษากลางมีดังต่อไปนี้

1.2.1 การกำหนดรูปแบบของภาษากลาง (Interlingua Structure)

รูปแบบภาษากลางประกอบด้วยหน่วยความหมายระดับแก่นของคำ (Conceptual Primitive unit:CP) หน่วยความสัมพันธ์ทางภาษา (Linguistic Relation unit:REL) และโครงข่ายแสดงความสัมพันธ์ (Relation Network) เพื่อเป็นตัวแทนของภาษาต่างๆ ซึ่งใช้หลักของ Hierarchy โดยแบ่งกลุ่มความหมายของคำ

1.2.2 การออกแบบและสร้างพจนานุกรม (Design and Create Dictionary)

พจนานุกรมนี้เพื่อใช้ในการวิเคราะห์และสังเคราะห์ประโยคภาษากลาง โดยกำหนดประเภทของคำ (Category) ประเภทย่อยของคำ (Subcategory) รูปแบบของกริยา (Verb Pattern) และกำหนดความหมายของคำเป็นรหัส (Identification) ซึ่งมีคำอธิบายความหมายของรหัสนั้น ให้มีความหมายเดียวเพื่อที่เวลาแปลไปเป็นภาษาอื่นๆ จะได้มีความหมายเดียวทำให้สามารถลดความกำกวมของคำได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา⁸ และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เพื่อใช้ในการสังเคราะห์ประโยคภาษาไทย และภาษาอังกฤษ ซึ่งแต่ละภาษาก็จะมีพจนานุกรมของแต่ละภาษาด้วย

1.2.3 การสร้างกฎทางความสัมพันธ์ของประโยค

กฎทางความหมายเป็นวิธีการสังเคราะห์ภาษาไทย และภาษาอังกฤษ เพื่อให้ได้ประโยคที่มีความหมายเดียวกับภาษาต้นฉบับ

1.2.4 การออกแบบและพัฒนาซอฟต์แวร์ของระบบ

เพื่อทำการทดสอบสังเคราะห์ประโยค และนำไปใช้กับการแปลข้อมูลต่างๆ ซึ่งต้องการความถูกต้องทั้งความหมายและไวยากรณ์

1.3 โครงร่างของวิทยานิพนธ์

วิทยานิพนธ์นี้ประกอบด้วยเนื้อหาทั้งหมด 7 บท โดยที่แต่ละบทจะมีเนื้อหาดังนี้

บทที่ 1 เป็นการกล่าวถึงความเป็นมาของระบบการแปลภาษาด้วยคอมพิวเตอร์ ซึ่งแบ่งออกเป็น 3 วิธีการคือ การแปลโดยตรง การแปลโดยการถ่ายถอด และการแปลโดยใช้ภาษากลางเป็นสื่อ นอกจากนี้ยังมีเนื้อหาเกี่ยวกับจุดประสงค์และขอบเขตของวิทยานิพนธ์นี้ด้วย

บทที่ 2 เป็นการออกแบบและกำหนดโครงสร้างของภาษากลาง

บทที่ 3 เป็นการกำหนดหน่วยความหมายของภาษากลางที่ใช้ในวิทยานิพนธ์นี้ ซึ่งแบ่งออกเป็นรหัสหน่วยความหมาย และคำอธิบายหน่วยความหมาย

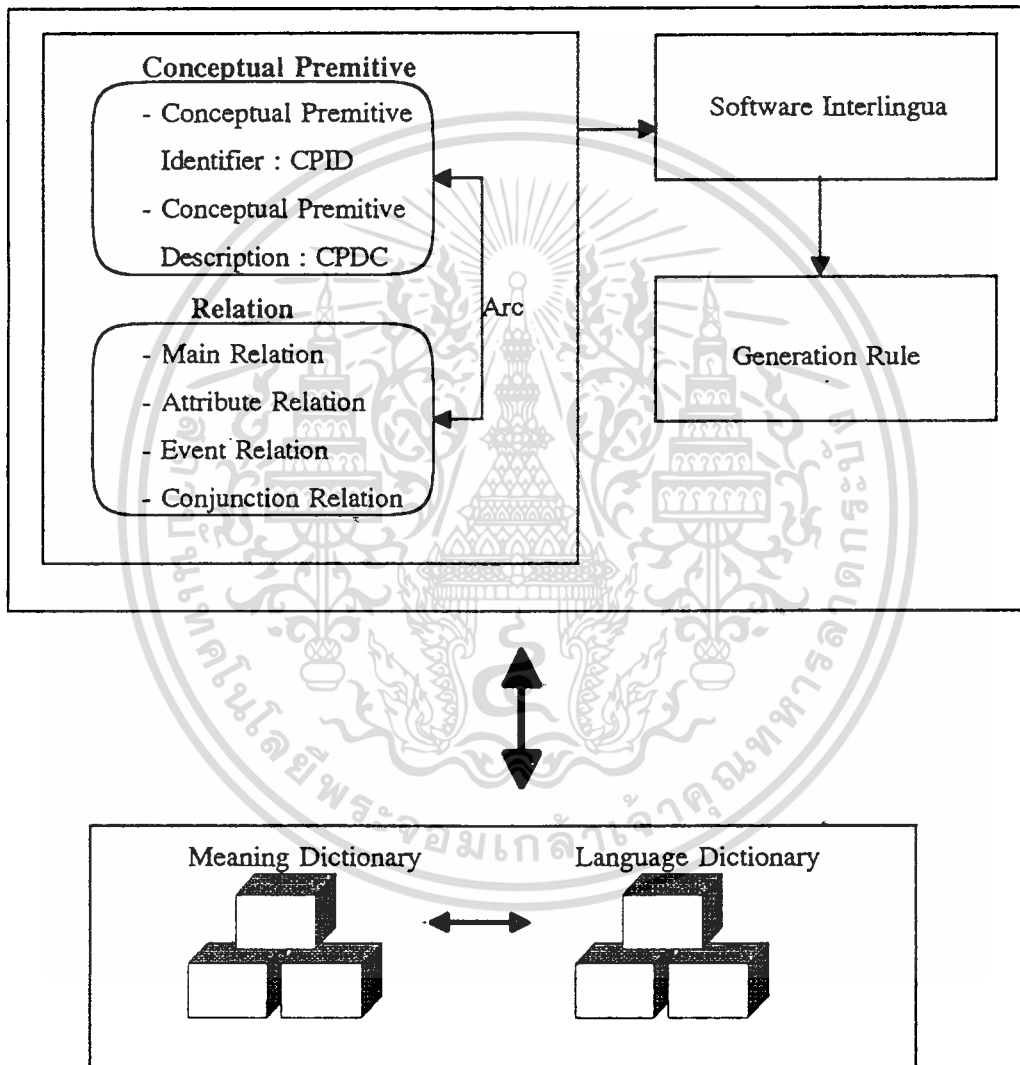
บทที่ 4 เป็นการกล่าวถึงแนวความคิดเกี่ยวกับความสัมพันธ์ทางอรรถศาสตร์ของนักภาษาศาสตร์ที่มีชื่อเสียง และเป็นการกำหนดความสัมพันธ์ของหน่วยความหมายของภาษากลางที่ใช้ในวิทยานิพนธ์นี้โดยแบ่งความสัมพันธ์ออกเป็น 4 ส่วนคือ ความสัมพันธ์หลัก ความสัมพันธ์คุณสมบัติ ความสัมพันธ์เหตุการณ์ และความสัมพันธ์คำเชื่อม

บทที่ 5 เป็นการออกแบบพจนานุกรมเพื่อการแปลภาษาด้วยคอมพิวเตอร์ โดยแบ่งออกเป็น 2 ส่วนคือ พจนานุกรมความหมาย และพจนานุกรมภาษา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 6 กล่าวถึงขั้นตอนการสังเคราะห์ภาษาเป้าหมายจากภาษากลาง และโครงสร้างภาษากลางทางซอฟต์แวร์

บทที่ 7 เป็นบทสรุปผลงานวิจัยของวิทยานิพนธ์นี้ และแนวทางในการพัฒนาต่อไป โครงสร้างของวิทยานิพนธ์นี้แสดงได้ดังรูปที่ 1.7



Multilingual Dictionary

รูปที่ 1.7 โครงสร้างของวิทยานิพนธ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา¹⁰ และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

การออกแบบโครงสร้างภาษากลาง

2.1 บทนำ

ในระบบการแปลหลายภาษาโดยใช้สื่อภาษากลางดังที่ได้กล่าวมาในบทที่ 1 แล้วนั้น โดยพื้นฐานจะมีกระบวนการประมวลผลทางภาษาอยู่ 2 ส่วนใหญ่ๆ^[1] ได้แก่

- 1) ส่วนการวิเคราะห์ (Analysis Process) ประโยคของภาษาต้นฉบับ (Source Language) สู่ภาษากลาง (Interlingua)
- 2) ส่วนการสังเคราะห์ (Generation Process) เพื่อสร้างประโยคของภาษาเป้าหมาย (Target Language) จากภาษากลาง

ทั้ง 2 ส่วนมีความสัมพันธ์กันดังแสดงในรูปที่ 2.1 และโดยพื้นฐานของส่วนกระบวนการวิเคราะห์จะแบ่งเป็น 3 ขั้นตอนใหญ่ๆ ที่มีกระบวนการประมวลผลทางภาษาตามลำดับต่อเนื่องกันเพื่อให้ได้ผลลัพธ์เป็นภาษากลาง^[4] ได้แก่

- การวิเคราะห์ประโยคในระดับหน่วยคำ (Morphological Analysis)
- การวิเคราะห์ประโยคในระดับโครงสร้างและความหมาย (Syntactic / Semantic Analysis)
- การวิเคราะห์ประโยคในระดับความหมายแก่น (Conceptual Analysis)

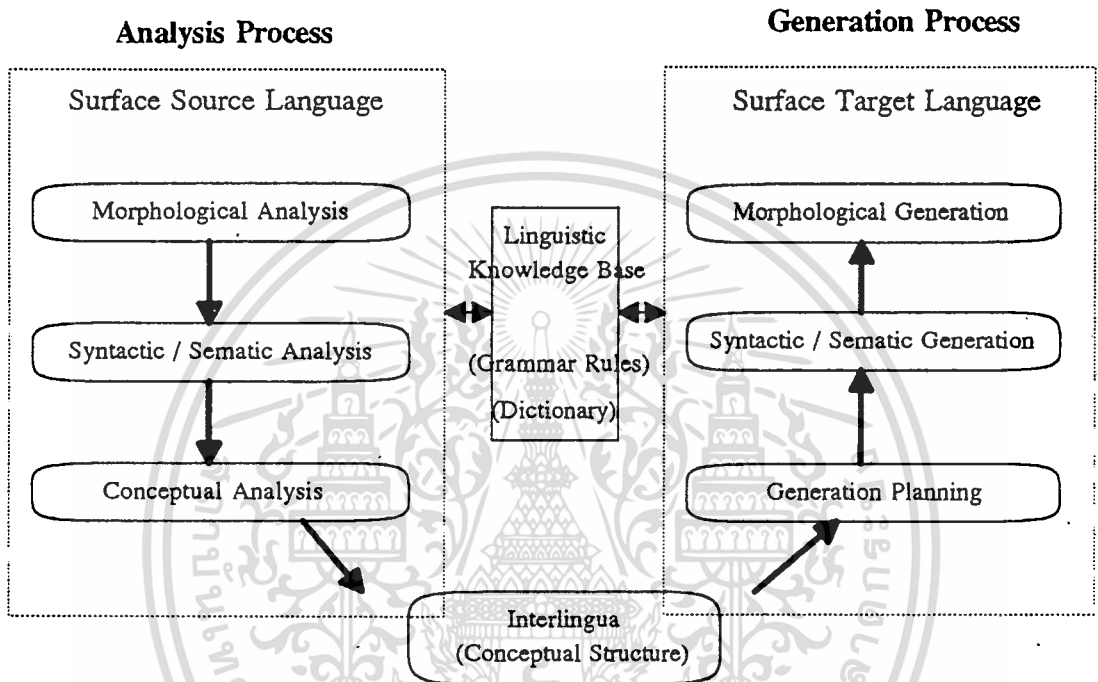
ผลของการวิเคราะห์ในระดับความหมายแก่น จะนำไปสู่ภาษากลางที่จะเป็นตัวแทนถ่ายทอดข้อมูลทางภาษาให้กับประโยคของภาษาต้นแบบไปสู่ภาษาเป้าหมายที่ต้องการต่อไป

ในทางตรงกันข้ามและทำนองเดียวกัน พื้นฐานของส่วนกระบวนการสังเคราะห์ได้แบ่งเป็น 3 ขั้นตอนใหญ่ๆ ที่มีการประมวลผลทางภาษาเป็นลำดับต่อเนื่องจากภาษากลางเพื่อนำไปสู่ผลลัพธ์ภาษาเป้าหมาย ได้แก่

- การสังเคราะห์ในระดับการเตรียมข้อมูล (Generation Planning) จากข้อมูลทางภาษาที่ได้จากภาษากลาง และข้อมูลทางภาษาจากฐานข้อมูลพจนานุกรม (Knowledge Base)
- การสังเคราะห์ในระดับโครงสร้างและความหมาย (Syntactic / Semantic Generation)

- การสังเคราะห์ในระดับหน่วยคำ (Morphological Generation)

ผลลัพธ์สุดท้ายจะได้ประโยคของภาษาเป้าหมายที่ต้องการจะแปล ในระหว่างการวิเคราะห์หรือสังเคราะห์ประโยคด้วยระบบคอมพิวเตอร์ จะต้องอาศัยฐานข้อมูลทางภาษา (Linguistic Knowledge Base) ได้แก่ ฐานข้อมูลทางภาษาของคำศัพท์ในรูปแบบของพจนานุกรม (Dictionary) และกฎไวยากรณ์ (Grammar Rules) ของภาษานั้นๆ




รูปที่ 2.1 โครงสร้างของระบบการแปลหลายภาษาโดยใช้ภาษากลางเป็นสื่อการแปล

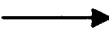
จากที่กล่าวมาทั้งหมดจะพบว่า ความสำคัญและความถูกต้องของระบบการแปลหลายภาษาด้วยคอมพิวเตอร์ โดยใช้สื่อภาษากลาง มิได้ขึ้นอยู่กับขบวนการวิเคราะห์ และขบวนการสังเคราะห์เพียงเท่านั้น แต่ยังขึ้นอยู่กับความสมบูรณ์และความถูกต้องของการถ่ายทอดข้อมูลทางภาษาของ "ภาษากลาง" และสิ่งสำคัญที่จะกำหนดความสมบูรณ์และความถูกต้องทางภาษาให้กับภาษากลางก็คือ โครงสร้างหรือรูปแบบของภาษากลางนั่นเอง ซึ่งเป็นจุดประสงค์หลักข้อหนึ่งของงานวิจัยในวิทยานิพนธ์นี้

งานวิทยานิพนธ์นี้ ได้วิจัยออกแบบโครงสร้าง ภาษากลางในเชิงภาษาศาสตร์คอมพิวเตอร์ ตลอดจนรูปแบบการเก็บบันทึกข้อมูลภาษากลางในระบบซอฟต์แวร์ ดังจะได้อธิบายในรายละเอียดต่อไป

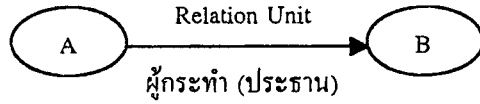
2.2 โครงสร้างภาษากลางในเชิงภาษาศาสตร์คอมพิวเตอร์ (Computational Linguistics Structure of Interlingua)

โดยพื้นฐานแล้วสื่อภาษากลางที่ใช้ในระบบการแปลภาษาด้วยเครื่องจะมีรูปแบบไม่เหมือนข้อความในภาษาธรรมชาติ (Natural Language) เพราะในภาษากลางจะต้องมีการกำหนดหน่วยแสดงความสัมพันธ์ของหน่วยคำ ในลักษณะภาษาทางปัญญาประดิษฐ์ที่ใช้แสดงโครงสร้าง และความหมายของประโยคของภาษาต้นแบบ การแสดงความสัมพันธ์จะต้องสามารถที่จะแสดงความสัมพันธ์ของหน่วยคำได้อย่างชัดเจนในลักษณะที่บอกให้ทราบว่าหน่วยคำนั้นมีความสัมพันธ์กับหน่วยคำอื่น และมีความสัมพันธ์กันอย่างไร ซึ่งลักษณะดังกล่าวมานี้ ตลอดจนรูปแบบของสื่อภาษากลาง ควรจะมีความเป็นอิสระไม่ขึ้นกับภาษาใดๆ จากเหตุผลดังกล่าวนี้ โครงสร้างต้นไม้ความสัมพันธ์แบบพึ่งพากัน (Dependency Relation Tree Structure) สามารถที่จะแสดงการเชื่อมต่อกันระหว่างหน่วยคำต่างๆ เป็นลำดับซึ่งสอดคล้องกับลักษณะทางวากยสัมพันธ์ (Syntactic Structure) ของภาษาต่างๆ ในวิทยานิพนธ์นี้จึงใช้โครงสร้างต้นไม้ความสัมพันธ์แบบพึ่งพากัน ซึ่งเป็นวิธีการกำหนดรูปแบบทางภาษาศาสตร์คอมพิวเตอร์แขนงหนึ่งในวิทยาการด้านการประมวลผลภาษาธรรมชาติ (Natural Language Processing) ที่มีส่วนประกอบหลักอยู่ 3 ส่วน ได้แก่

1) โนดของหน่วยคำ (Word's node) ในวิทยานิพนธ์นี้แบ่งเป็นโนดหลัก (Head node) และโนดพึ่งพา (Dependent node) ใช้สัญลักษณ์  แทนโนด

2) อาร์คแสดงความสัมพันธ์ (Relation Arc) ระหว่างโนดของหน่วยคำโดยทิศทางกำหนดความสัมพันธ์ ในวิทยานิพนธ์นี้จะใช้ เส้นตรงและหัวลูกศร  แทนอาร์ค

3) หน่วยแสดงความสัมพันธ์ (Relation Unit) ระหว่างโนดหลักที่อยู่ต้นอาร์คความสัมพันธ์กับโนดพึ่งพาที่อยู่ปลายลูกศรของอาร์คนั้น ความสัมพันธ์จะแสดงหน้าที่ของหน่วยคำที่มีต่อกันในทางความหมาย เช่น การเป็นผู้กระทำ (ประธาน) การเป็นผู้ถูกกระทำ (กรรม) เป็นต้น ตัวอย่างเช่น

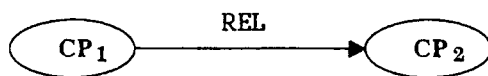


หมายความว่า A เป็นโน้ตหลัก และ B เป็นโน้ตพึ่งพา โดยที่โน้ต B เป็นผู้กระทำ (ประธาน) ของโน้ต A

เมื่อนำหลักการของโครงสร้างต้นไม้ความสัมพันธ์แบบพึ่งพาดังกล่าว มาประยุกต์ใช้กับการกำหนดโครงสร้างของภาษากลาง โดยคำนึงถึงหลักไวยากรณ์ทางภาษาแล้ว งานวิจัยนี้จึงได้กำหนดโครงสร้างภาษากลางโดยมีส่วนประกอบหลัก 3 ส่วน ที่สามารถแทนหรือแสดงได้ด้วยโครงสร้างต้นไม้ความสัมพันธ์แบบพึ่งพา ดังตารางเปรียบเทียบต่อไปนี้

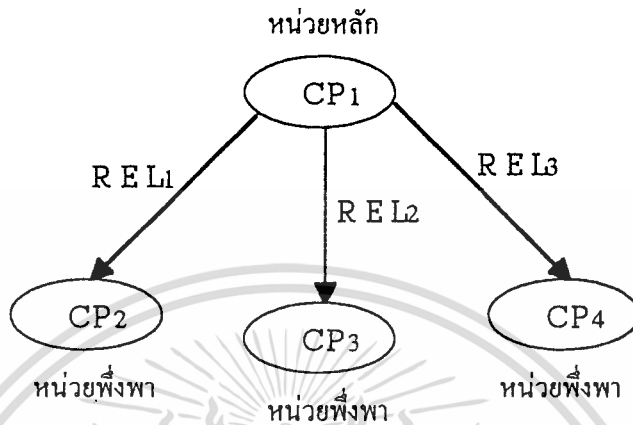
โครงสร้างภาษากลาง	โครงสร้างต้นไม้ความสัมพันธ์แบบพึ่งพา
1) หน่วยความหมายระดับแก่นของคำ	1) โน้ตของหน่วยคำ
2) หน่วยความสัมพันธ์ทางภาษา	2) หน่วยแสดงความสัมพันธ์
3) โครงข่ายแสดงความสัมพันธ์	3) อาร์คแสดงความสัมพันธ์

ในส่วนของโครงสร้างภาษากลางจะกำหนดสัญลักษณ์ CP แทนหน่วยความหมายระดับแก่นของคำ (Conceptual Primitive Unit) และกำหนดสัญลักษณ์ REL แทนหน่วยความสัมพันธ์ทางภาษา ตัวอย่างเช่น



หมายความว่า หน่วยความหมาย CP1 จะเป็นหน่วยหลักและหน่วยความหมาย CP2 จะเป็นหน่วยพึ่งพา จากโครงสร้างพื้นฐานดังกล่าว เมื่อนำไปแสดงในลักษณะโครงข่ายความสัมพันธ์ทาง

ภาษา ทำให้สะดวกต่อการแสดงความสัมพันธ์ระหว่างหน่วยคำเป็นลำดับที่สอดคล้องกับลักษณะทางวากยสัมพันธ์ (Syntactic Structure) ของภาษาต่างๆ ดังโครงข่ายตัวอย่างในรูปที่ 2.2



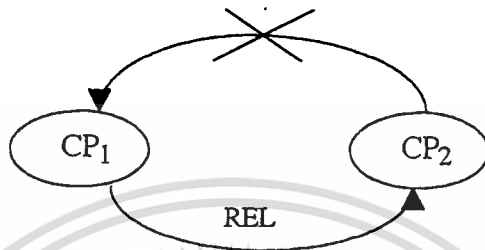
รูปที่ 2.2 ตัวอย่างโครงข่ายของโครงสร้างภาษากลาง

ในการสร้างภาษากลางในลักษณะโครงสร้างต้นไม้ความสัมพันธ์แบบพียงพานี้ จะต้องอยู่ภายใต้กฎพื้นฐาน 2 ข้อ ได้แก่

<กฎที่ 1> ความสัมพันธ์แบบพียงพาจะต้องไม่เป็นแบบที่ทั้งสองโหนด (node) ต่างก็เป็นหน่วยหลัก (head) และหน่วยพียงพา (dependor) ของกันและกัน (interdependency)

เช่น ในทางภาษาเมื่อวิเคราะห์ CP₁ เป็นหน่วยหลักและมี CP₂ เป็นหน่วยพียงพาแล้ว จะกำหนดให้ CP₂ กลับมาเป็นหน่วยหลักของ CP₁ ไม่ได้ ดังแสดงในรูปที่ 2.3 นั่นคือ อาร์คความสัมพันธ์จาก CP₂ มีทิศไปสู่ CP₁ ไม่ได้ เพราะคำในทางภาษาจะมีความสัมพันธ์กับคำอื่นได้ 1 ความหมายใน 1 ประโยค และคำที่ทำหน้าที่เป็นหน่วยหลักนั้นจะต้องมีคุณสมบัติทางภาษาจะเป็นคำกริยาหรือคำที่มีความหมายที่สามารถจะแสดงความหมายตามความสัมพันธ์ที่กำหนดไว้ได้ด้วย จึงไม่สามารถที่จะมีทั้ง 2 ความสัมพันธ์ได้ ดังประโยค "ฉันกินข้าว" ถ้ากำหนดให้ CP₁ เป็นคำว่า "กิน" เป็นหน่วยหลักและ CP₂ เป็นคำว่า "ฉัน" เป็นหน่วยพียงพา โดยมี REL แบบแสดงความเป็นแบบผู้กระทำ จะเห็นว่าจากความสัมพันธ์นี้ "ฉัน" ซึ่งเป็น CP₂ จะเป็นผู้กระทำอาการ "กิน" ซึ่งเป็น CP₁ แต่ถ้ามีความสัมพันธ์โดยให้ CP₂ เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 15 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

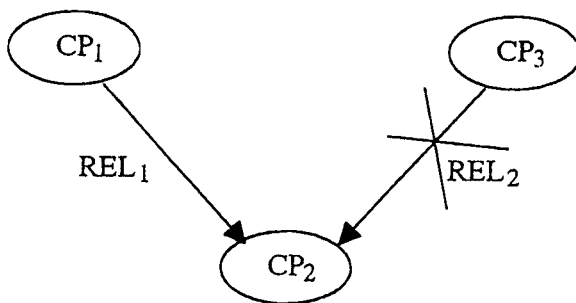
เป็นหน่วยหลัก และ CP1 เป็นหน่วยพึ่งพา โดยมีความสัมพันธ์แบบผู้กระทำอาการ ก็พบว่า "กิน" ไม่สามารถเป็นผู้กระทำได้เพราะ "กิน" มีประเภทของคำเป็นกริยาซึ่งผู้ที่สามารถจะเป็นผู้กระทำได้นั้นจะต้องเป็นคำนาม หรืออาจจะเป็นคำกริยาก็ได้ถ้าในส่วนของคำกริยานั้นเป็นส่วนหนึ่งของนามวลี เช่น ครูสั่งนักเรียนเขียนรายงาน คำว่า "สั่ง" จะเป็นหน่วยคำหลักของ "เขียน" โดยมีความสัมพันธ์แบบผู้ถูกกระทำ



รูปที่ 2.3 Interdependency

<กฎที่2> หน่วยใดๆจะเป็นหน่วยพึ่งพาของหน่วยหลักมากกว่าหนึ่งหน่วยไม่ได้ (double dependency)

เช่น ตัวอย่างในรูปที่ 2.4 ในกรณีนี้ CP2 จะเป็นหน่วยพึ่งพาของหน่วยหลัก CP1 และ CP3 พร้อมกันไม่ได้ ดังนั้นสมมติว่าในการวิเคราะห์ทางภาษา CP1 เป็นหน่วยหลักของ CP2 แล้วจะกำหนดให้ CP3 เป็นหน่วยหลักของ CP2 อีกไม่ได้ เพราะหน่วยพึ่งพานั้นจะมีความสัมพันธ์กับหน่วยหลักเพียงหน่วยเดียว แต่หน่วยหลักสามารถมีหน่วยพึ่งพาได้มากกว่า 1 หน่วย เช่นเดียวกับส่วนเสริมของภาษาจะเสริมในส่วนของคำหลักเพียงคำเดียวจะไม่มีกรไปเสริมคำหลักอื่น แต่คำหลักของภาษาสามารถที่จะมีคำมาเสริมมากกว่า 1 คำก็ได้เช่น ลิงกินแอปเปิลสีแดงห้าผล ในที่นี้ "แอปเปิล" จะเป็นหน่วยหลักของ "สีแดง" และ "ห้าผล"



รูปที่ 2.4 Double dependency

2.3 สรุป

ภาษากลางที่กำหนดขึ้นในวิทยานิพนธ์นี้ใช้ได้ทำการวิจัยและออกแบบโดยใช้หลักการของโครงสร้างต้นไม้พึ่งพา (Dependency Tree) เนื่องจากโครงสร้างต้นไม้นี้ สามารถที่จะกำหนดความสัมพันธ์ของแต่ละหน่วยคำที่ได้จากการวิเคราะห์ได้ชัดเจน ไม่ทำให้เกิดความกำกวมของภาษา ซึ่งต่างจากโครงสร้างของภาษาธรรมชาติ (Natural Language) ซึ่งจะเรียงต่อกันเป็นแนว (Linear) จะไม่ระบุความสัมพันธ์ของหน่วยคำไว้อย่างชัดเจน ซึ่งรายละเอียดในส่วนการวิจัยออกแบบและกำหนดหน่วยความหมายและหน่วยความสัมพันธ์ของภาษากลางในวิทยานิพนธ์นี้ จะได้กล่าวต่อไปในบทที่ 3 และบทที่ 4 ตามลำดับ

บทที่ 3

หน่วยความหมาย

3.1 บทนำ

การแทนความหมายของแต่ละคำ และของแต่ละภาษาจะต่างกันออกไป แต่จุดสำคัญของการแทนคำที่แต่ละภาษาใช้สื่อสารกันนั้นจะต้องมีความหมายของคำ (Meaning of Word) นั้น และแต่ละภาษาจะสื่อสารกันได้นั้นไม่ได้ขึ้นอยู่กับคำๆ นั้น แต่ขึ้นอยู่กับความหมายของคำนั้น ซึ่งมนุษย์สามารถที่จะเลือกความหมายของคำนั้นได้อย่างถูกต้อง แต่ถ้าเป็นการสื่อสารโดยใช้คอมพิวเตอร์การที่จะกำหนดให้คอมพิวเตอร์ทราบว่าคำๆ นั้นมีความหมายอย่างไรนั้นเป็นการยาก เพราะคำๆ หนึ่งในแต่ละภาษาอาจจะมีความหมายได้มากกว่าหนึ่งความหมาย ทั้งๆ ที่รูปแบบผิวของคำ (Surface form of Word) ที่เป็นตัวแสดงจะเหมือนกันก็ตามดังตัวอย่างในตารางที่ 3.1 เช่น

"ชั้น" ในภาษาไทยจะมีหลายความหมายและมีชนิดของคำหลายชนิดซึ่งอาจจะหมายถึง "ชั้นน้ำ" ซึ่งเป็นภาชนะที่ใช้สำหรับใสหรือตักน้ำโดยลักษณะนี้จะเป็นคำนาม หรืออาจจะหมายถึง "ไก่ชั้น" ซึ่งจะเป็นลักษณะการส่งเสียงร้องของไก่หรือนก หรืออาจจะหมายถึง "ชั้นสกูตอกร" จะเป็นการบิดหรือหมุนให้แน่น หรืออาจจะหมายถึง "ตลกขบขัน" จะเป็นการหัวเราะของบุคคลก็ได้

"Train" ในภาษาอังกฤษก็จะมีหลายความหมายและหลายชนิดของคำด้วย ถ้าเป็นกริยาจะหมายถึงการฝึกหัด การอบรม แต่ถ้าเป็นคำนามก็จะหมายถึงสิ่งที่เคลื่อนที่ไปเป็นทาง (รถไฟ) เป็นต้น

"KAERU" ในภาษาญี่ปุ่น ถ้าเป็นคำกริยานั้นภาษาญี่ปุ่นจะแบ่งคำกริยาออกเป็น 3 กลุ่ม¹คือกริยาที่ลงท้ายด้วยเสียงอิ (I) กริยาที่ไม่ได้ลงท้ายด้วยอิ (II) หรือ ชิมัดซี (SHIMASU) และกริยาที่ลงท้ายด้วยชิมัดซี (SHIMASU) แต่คำกริยาที่นำมาใช้นั้นจะอยู่ในรูปของคำในพจนานุกรมซึ่งมีการผันมาจากกลุ่มคำต่างๆ ทั้ง 3 กลุ่ม โดยถ้าเป็นกลุ่มที่ลงท้ายด้วยอิ (I) จะหมายถึงการเคลื่อนที่กลับไปสู่สถานที่เคยมีการกระทำ ถ้าเป็นกลุ่มกริยาที่สองจะหมายถึง แลกเปลี่ยนสิ่งของต่างๆ หรือเป็นการเปลี่ยนแปลง แต่ถ้าเป็นคำนามจะหมายถึงสิ่งมีชีวิตที่เป็นสัตว์สี่เท้าสะเทินน้ำสะเทินบกมีชื่อเรียกว่า "กบ"

[*] ภาษาญี่ปุ่น ชั้นต้นเล่ม 2 สยามเทคโนโลยี (ไทย-ญี่ปุ่น)

เอกสารนี้เป็นเอกสารสงวนลิขสิทธิ์ของเจ้าของเนื้อหา ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปเนยผิว (Surface)	ความหมาย (Meaning)
จัน ^[1] ในภาษาไทย	ภาชนะตักน้ำทำด้วยโลหะ dipper, metal bowl used for dipping water
	นกและไก่ส่งเสียงร้อง crow, caw
	บิดหรือหมุนให้แน่น wrench, twist, screw, tighten
	หมุนตามเกลียวให้แน่นหรือคลายออก wind, turn, move
	ขบขัน humorous, funny, jocular, comic, comical
Train ^[**] ในภาษาอังกฤษ	ฝึก, ฝึกปรี้อ, ฝึกหัด, จ่อหรือจ้อง (ปีน)
	สิ่งที่ไปเป็นทาง เช่น ชายกระโปรงที่ถูกลากไปขณะเดิน, คนที่ไปกันเป็นทาง เช่น คนติดตาม, รถไฟที่แล่นเป็นขบวน
Kaeru ^[***] ในภาษาญี่ปุ่น	กลับ(ไปที่), กลับคืนสู่ Return
	เปลี่ยน (แปลง), แลกเปลี่ยน change, transform exchange
	กบ a frog

ตารางที่ 3.1 ตัวอย่างคำที่มีหลายความหมายในภาษาต่างๆ

[1] เขียวชัย เอี่ยมวรเมธ, พจนานุกรม (ฉบับอธิบายสองภาษา), พิมพ์ที่นิยมวิทยา (แผนกการพิมพ์)

[**] สำนักพิมพ์ไทยวัฒนาพานิช, Modern English-Thai Dictionary

[***] กนิษฐา กาญจนจारी, พจนานุกรม ญี่ปุ่น-ไทย-อังกฤษ, แพร่พิทยาจัดจำหน่าย, โรงพิมพ์กล่องพัฒนา

จากตัวอย่างคำในภาษาต่างๆ ที่อธิบายข้างต้น ถ้าพิจารณาในทางตรงกันข้ามเกี่ยวกับรูปแบบผิวของคำในภาษาต่างๆ ที่ให้ความหมายเดียวกัน เช่น คำว่า "KAERU" ในภาษาญี่ปุ่น ถ้ามีความหมายคือ "สัตว์สี่เท้าสะเทินน้ำสะเทินบกและเป็นสัตว์ประเภท Anura" ก็จะมีรูปแบบผิวของคำตรงกับคำในภาษาไทยว่า "กบ" หรือตรงกับรูปแบบผิวของคำในภาษาอังกฤษว่า "Frog" เป็นต้น

เมื่อวิจัยพิจารณามาถึงจุดนี้จะพบว่า จากตัวอย่างคำที่มีความหมายตรงกันใน 3 ภาษาดังกล่าว ได้แก่ "KAERU" "กบ" และ "Frog" ถ้าจะกำหนดตัวแทนของคำเหล่านี้ในเชิงภาษาศาสตร์คอมพิวเตอร์ให้เป็นเพียงหน่วยคำเดียวในอดีต¹¹ จะใช้ภาษาใดภาษาหนึ่งซึ่งเป็นภาษาที่ยอมรับให้เป็นภาษาสากลในกลุ่มภาษาที่อยู่ในระบบที่กำลังพิจารณา กำหนดให้คำในภาษาสากลนั้นเป็นตัวแทนหน่วยคำในภาษาต่างๆที่มีความหมายเดียวกัน ดังเช่นตัวอย่าง 3 ภาษาข้างต้น ภาษาอังกฤษจะเป็นภาษาสากลที่ยอมรับ ดังนั้นคำว่า Frog จะถูกเลือกให้เป็นตัวแทนหน่วยความหมายของคำ "KAERU" ในภาษาญี่ปุ่น คำ "กบ" ในภาษาไทย และคำ "Frog" ในภาษาอังกฤษ เป็นต้น

ในการกำหนดตัวแทนของคำในภาษาต่างๆ ให้เป็นหน่วยความหมายของภาษากลางในทางทฤษฎีและนิยามแล้วไม่ควรจะอิงภาษาใด เพราะจะเกิดความสับสนและกำกวมได้ เนื่องจากรูปแบบผิวของคำในภาษาใดๆ มีโอกาสที่จะมีได้หลายความหมายดังเช่นได้อธิบายจากตัวอย่างในตารางที่ 3.1 ข้างต้นทั้งหมด หรือจากตัวอย่างที่เลือกใช้คำว่า "Frog" ในภาษาอังกฤษเป็นตัวแทนของคำ "KAERU" "กบ" และ "Frog" ดังกล่าว ถ้าวิจัยพิจารณาให้ลึกซึ้งลงไปในทางภาษา จะพบว่า "KAERU" ในภาษาญี่ปุ่นยังมีอีกหลายความหมาย (ดูตัวอย่างในตารางที่ 3.1) ทำนองเดียวกันคำว่า "กบ" ในภาษาไทย และคำว่า "Frog" ในภาษาอังกฤษยังมีอีกหลายความหมาย (ดูรูปที่ 3.1) ในภาษาของตนเองที่ต่างจากภาษาอื่น

ที่อธิบายข้างต้นเป็นเพียงตัวอย่างการกำหนดตัวแทนของคำในรูปหน่วยความหมายของภาษากลางให้กับภาษาญี่ปุ่น ไทย และอังกฤษ เพียง 3 ภาษา ถ้ามีจำนวนภาษามากมายความสับสนและความกำกวมก็จะเกิดมากขึ้น ถ้าหากเลือกตัวแทนจากภาษาใดภาษาหนึ่ง

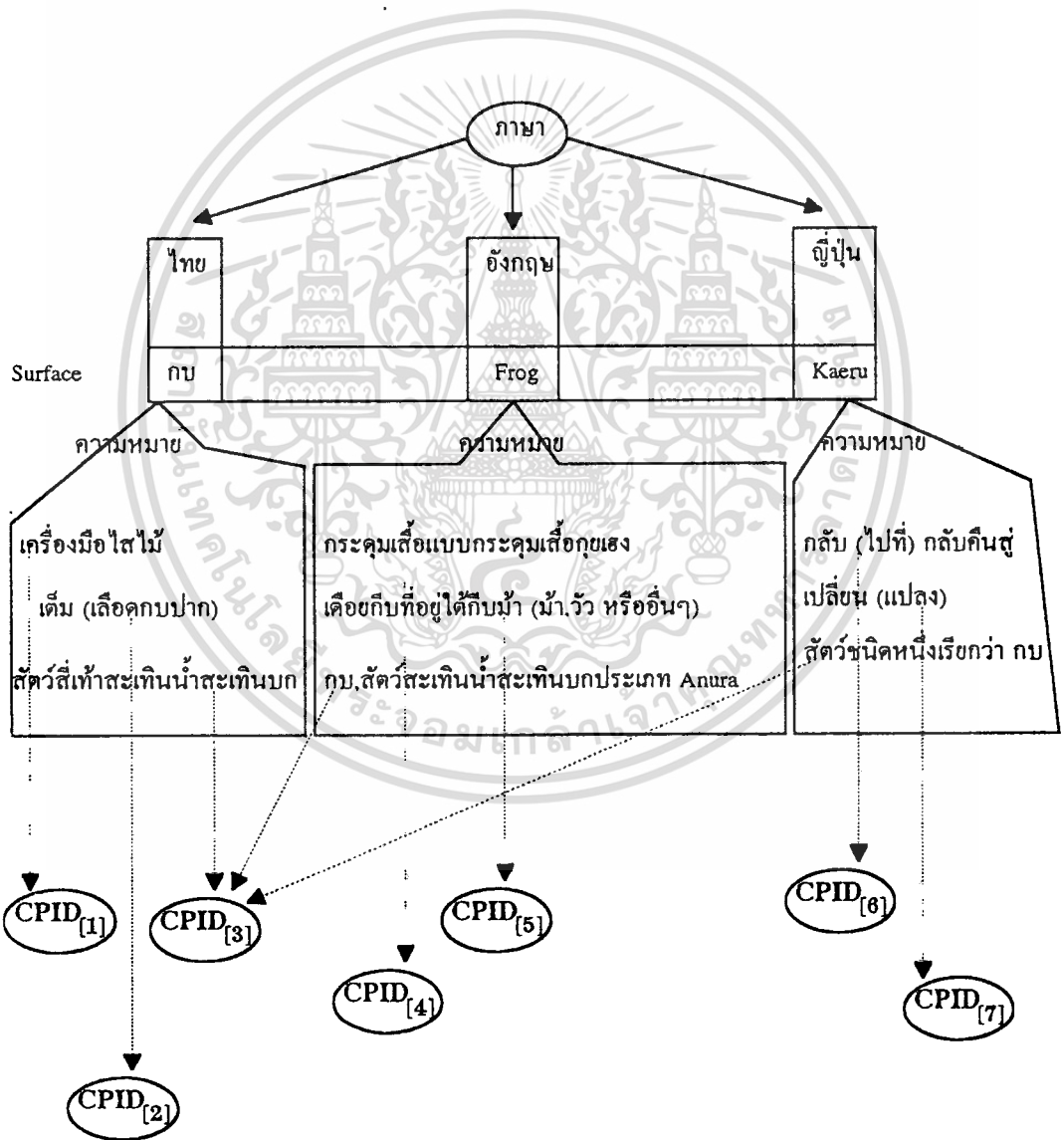
ดังนั้นงานวิจัยในวิทยานิพนธ์นี้ จึงใช้หลักการใหม่ในการกำหนดหน่วยความหมายของภาษากลางด้วยรหัสตัวเลขเรียกว่า "รหัสแยกแยะหน่วยความหมาย" (Conceptual Primitive Identifier) โดยการวิจัยพิจารณาหน่วยคำในระดับผิวของภาษาต่างๆ ที่มีความหมายลึก (Deep Meaning) หรือความหมายระดับแก่น (Conceptual Meaning) เหมือนกันหรือคล้ายคลึงกัน ตลอดจนมีคุณสมบัติทางวัฒนธรรมการใช้ภาษา (Cultural Characteristic Uses) เช่น หน้าที่ของคำ ประเภทของคำ หรือชนิดของคำที่คล้ายคลึงหรือเหมือนกัน งานวิจัยจะจัดให้กลุ่มคำในภาษาต่างๆ ดังกล่าวนี้ให้มี "รหัสแยกแยะหน่วยความหมาย" เดียวกัน และจะใช้รหัสตัวเลขนี้เป็นตัวแทนหน่วยความหมายของคำเหล่านั้นในภาษากลาง ดังเช่นตัวอย่างแสดงในรูปที่ 3.1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รหัสตัวเลขของการแยกแยะหน่วยความหมายนี้จะใช้สัญลักษณ์ย่อ "CPID" ตัวเลขแต่ละหลักในรหัสจะให้ข้อมูลทางภาษา ทั้งด้านความหมาย ประเภท ชนิด และหน้าที่ของคำ จึงเหมาะสมที่จะใช้เป็นตัวแทนของคำในภาษาต่างๆ ในรูปหน่วยความหมายของภาษากลาง ตลอดจนการประยุกต์ใช้ในเชิงภาษาศาสตร์คอมพิวเตอร์ โดยเฉพาะในระบบการแปลหลายภาษาด้วยคอมพิวเตอร์โดยภาษากลาง เพราะให้ข้อมูลทางภาษาได้อย่างสมบูรณ์ และมีความกำกวมในการเลือกรูปแบบผิวของคำในภาษาต่างๆ น้อย

งานวิทยานิพนธ์นี้ ได้ทำการวิจัยออกแบบการกำหนดรหัสตัวเลขในการแยกแยะหน่วยความหมาย บนพื้นฐานข้อมูลทางภาษาศาสตร์ ดังจะได้กล่าวโดยละเอียดในข้อ 3.2 ต่อไป



เอกสารนี้เป็นเอกสารที่รูปที่ 3.1 ตัวอย่างการแทนรหัสความหมายของความหมายของคำระโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2 รหัสแยกแยะหน่วยความหมาย :CPID (Conceptual Primitive Identifier)

งานวิทยานิพนธ์ได้วิจัยออกแบบรหัสแยกแยะหน่วยความหมายเพื่อใช้เป็นสัญลักษณ์แทนความหมายของคำในภาษาต่างๆ ที่ใช้กันอยู่ โดยการกำหนดนี้จะมีเพียงหนึ่งความหมายต่อหนึ่งรหัสแยกแยะหน่วยความหมาย โดยกำหนดเป็นสัญลักษณ์เป็นตัวเลขรวมผสมตัวอักษรจำนวน 9 หลักภายใต้โครงสร้างการแบ่งประเภทของคำตามลำดับชั้นความหมาย (of Word Meaning)^[5] โดยแบ่งเป็นชนิดของคำแต่ละคำในประโยค เช่น คำนาม (noun) คำกริยา (verb) คำกริยาวิเศษณ์ (adverb) ประเภทของคำ (A kind of) โดยสามารถแบ่งตามลักษณะรูปธรรม (Concrete) และนามธรรม (Abstract) และยังสามารถที่จะบอกถึงประเภทของคำๆ นั้นว่า เป็นบุคคล, สัตว์, สิ่งของ, สถานที่ หรือแม้แต่ลักษณะของการกระทำที่เกิดขึ้นเองตามธรรมชาติ หรือเกิดจากการกระทำของมนุษย์ ซึ่งสิ่งเหล่านี้มีความสำคัญต่อระบบการแปลหลายภาษาด้วยคอมพิวเตอร์โดยใช้ภาษากลางเป็นสื่อ (Multilingual Machine Translation System used by Interlingua) อย่างมาก โดยการจัดกลุ่มความหมายให้เป็นหมวดหมู่นี้จะทำให้การวิเคราะห์ประโยคเพื่อที่จะให้ได้รูปแทนความหมายในรูปของภาษากลางได้อย่างถูกต้อง และเมื่อทำการสังเคราะห์ประโยคภาษาเป้าหมายก็สามารถที่จะทราบความหมายที่แท้จริงของประโยค ซึ่งเป็นการลดความกำกวม (ambiguous) ของการสังเคราะห์ประโยคภาษาเป้าหมายที่จะเกิดขึ้น โดยการกำหนดหมายเลขใน CPID เพื่อเป็นตัวแทนของลักษณะของคำและตัวอักษรจะเป็นการกำหนดความหมายเฉพาะของคำในรหัส CPID นั้น ซึ่งรหัสความหมายกำหนดไว้ 9 รหัสตัวอักษร โดยรหัส 5 ตัวแรกจะเป็นตัวเลขซึ่งแทนลักษณะของคำ และรหัส 4 ตัวท้ายจะเป็นตัวอักษรในการกำหนดความหมายเฉพาะของคำ ดังแสดงในรูปที่ 3.2 ซึ่งการกำหนดโครงสร้างตามลักษณะนี้จะกำหนดแต่ละหลักของรหัสแยกแยะหน่วยความหมาย ดังต่อไปนี้

รหัสที่ 1 จะเป็นการแบ่งกลุ่มความหมายเป็น 2 กลุ่มใหญ่ๆ คือ

กลุ่มรูปธรรม (Concrete)	คือกลุ่มความหมายที่มีความหมายเกี่ยวกับสิ่งที่มีตัวตนหรืออยู่ในรูปของวัตถุ เช่น คน สัตว์ สิ่งของ สถานที่ เป็นต้น
กลุ่มนามธรรม (Abstract)	คือกลุ่มคำที่มีความหมายเกี่ยวกับสภาพของวัตถุ เหตุการณ์ หรือการกระทำ เช่น กิน วิ่ง เดิน เป็นต้น

รหัสที่ 2 จะเป็นการแบ่งกลุ่มความหมายลดหลั่นลงไปจากรหัสที่ 1 เพื่อให้ได้ขอบเขตของความหมายที่แคบลงไปอีก ดังในกลุ่มของรูปธรรมจะสามารถแบ่งออกได้เป็น 3 กลุ่มคือ

กลุ่มของรูปธรรมที่เป็นกลุ่มบุคคลและกลุ่มที่ประกอบด้วยคน

กลุ่มของรูปธรรมที่เป็นสถานที่

กลุ่มของรูปธรรมที่เป็นสิ่งของเป็นสิ่งมีชีวิตหรือสิ่งไม่มีชีวิตที่เกิดขึ้นเองตามธรรมชาติหรือมนุษย์สร้างขึ้นก็ได้

ในกลุ่มของนามธรรมสามารถแบ่งออกได้เป็น 3 กลุ่มเช่นเดียวกับในกลุ่มของรูปธรรม คือ

กลุ่มของนามธรรมที่อยู่ในรูปของเหตุการณ์และกิริยาอาการต่างๆ

กลุ่มของนามธรรมที่มีลักษณะคล้ายวงที่สามารถจับต้องได้

กลุ่มของนามธรรมที่อยู่ในรูปของความสัมพันธ์ในเชิงนามธรรม

รหัสที่ 3 จะเป็นการแบ่งกลุ่มลดหลั่นลงไปจากรหัสที่ 2 เพื่อกำหนดขอบเขตความหมายให้ลึกลงไป เช่น ในกลุ่มของรูปธรรมที่เป็นกลุ่มบุคคลและกลุ่มที่ประกอบด้วยคนสามารถกำหนดความหมายลดหลั่นของกลุ่มนี้ได้ 2 กลุ่มคือ

กลุ่มบุคคลที่เป็นสิ่งมีชีวิตสามารถกระทำอาการได้

กลุ่มบุคคลที่ประกอบกันเป็นองค์กร หรือสถาบันต่างๆ เป็นสิ่งไม่มีชีวิต และจะอยู่ในรูปของนิติบุคคล

และในกลุ่มของรูปธรรมที่เป็นสิ่งของเป็นสิ่งมีชีวิตหรือสิ่งไม่มีชีวิตที่เกิดขึ้นเองตามธรรมชาติหรือมนุษย์สร้างขึ้นก็ได้ ก็สามารถกำหนดความหมายลดหลั่นของกลุ่มความหมายนี้ได้เป็น 2 กลุ่มคือ

กลุ่มสิ่งของที่เกิดขึ้นเองตามธรรมชาติไม่มีการนำมาแปรสภาพ

กลุ่มสิ่งของที่มนุษย์สร้างขึ้นมาโดยการแปรสภาพจากสิ่งที่เกิดขึ้นเองตามธรรมชาติ

รหัสที่ 4 จะเป็นการแบ่งกลุ่มลดหลั่นลงไปจากรหัสที่ 3 เพื่อกำหนดขอบเขตความหมายให้ลึกลงไป เช่น กลุ่มสิ่งของที่เกิดขึ้นเองตามธรรมชาติไม่มีการนำมาแปรสภาพ ก็สามารถกำหนดความหมายลดหลั่นของกลุ่มความหมายนี้ได้เป็น 2 กลุ่มคือ

กลุ่มของสิ่งมีชีวิตซึ่งกลุ่มนี้จะแบ่งแยกกลุ่มของบุคคลออกจากกลุ่มนี้ โดยกลุ่มนี้จะ
เป็นพวกพืช และสัตว์

กลุ่มของสสาร แร่ธาตุต่างๆ โดยสิ่งเหล่านี้จะต้องเกิดจากธรรมชาติ

รหัสที่ 5 จะเป็นการแบ่งกลุ่มลดหลั่นลงไปจากรหัสที่ 4 เพื่อกำหนดขอบเขตความหมายให้ลึก
ลงไป เช่น กลุ่มของสิ่งมีชีวิตซึ่งกลุ่มนี้จะแบ่งแยกกลุ่มของบุคคลออกจากกลุ่มนี้ โดยกลุ่มนี้จะ
เป็นพวกพืช และสัตว์ ก็สามารถกำหนดความหมายลดหลั่นของกลุ่มความหมายนี้ได้เป็น 8 กลุ่มคือ

กลุ่มของสัตว์ที่อาศัยอยู่บนบก

กลุ่มของสัตว์ที่อาศัยอยู่ในน้ำ

กลุ่มของสัตว์ที่มีปีกซึ่งอาจจะบินได้หรือไม่ก็ได้

กลุ่มของสัตว์ประเภทเลี้ยงลูก

กลุ่มของพืชที่มีต้นกำเนิดบนพื้นดิน

กลุ่มของพืชที่มีต้นกำเนิดในน้ำ

กลุ่มของส่วนประกอบของสิ่งมีชีวิตที่เป็นบุคคล หรือสัตว์ หรือสิ่งที่ได้มาจากบุคคล
หรือสัตว์

กลุ่มของส่วนประกอบของพืช หรือสิ่งที่ได้มาจากพืช

โดยที่ในกลุ่มรหัสที่ 1-5 นั้นจะกำหนดด้วยตัวเลข 1-9 และในกลุ่มใดไม่มีการกำหนดรหัสให้
แทนตำแหน่งนั้นด้วยตัวเลขศูนย์ ดังมีรายละเอียดดังต่อไปนี้

3.2.1 กลุ่มรูปธรรม⁽⁵⁾ (Concrete) แบ่งตามรหัสที่กำหนดไว้เป็น

กลุ่มรหัส 11000 Subject : บุคคลและกลุ่มที่ประกอบด้วยคน แบ่งเป็น

กลุ่มรหัส 11100 Person : บุคคล เช่น ผู้ปกครอง เขา คู่รัก ครู เพื่อน

กลุ่มรหัส 11110 Person Name : ชื่อเฉพาะของบุคคล เช่น สมศักดิ์

ศักดิ์ชัย จอห์น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 24 จะต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กลุ่มรหัส 11200 Organization : กลุ่มหรือองค์กรต่าง ๆ เช่น บริษัท กรรมการ
กลุ่ม คณะ

กลุ่มรหัส 11210 Organization Name: ชื่อกลุ่ม องค์กร หรือสถาบันต่างๆ
เช่น NECTEC KMITL

กลุ่มรหัส 12000 Concrete Place : สถานที่ แบ่งเป็น

กลุ่มรหัส 12100 Region.: สถานที่ที่เป็นดินแดนหรือมีขอบเขตที่แน่นอน เช่น
จังหวัด อำเภอ

กลุ่มรหัส 12110 Region Name : ชื่อเฉพาะของสถานที่ เช่น กรุงเทพฯ
ประเทศไทย

กลุ่มรหัส 12200 Nature : สถานที่ที่เป็นธรรมชาติ

กลุ่มรหัส 12210 Topography : สถานที่ที่เป็นภูมิประเทศ เช่น ชายหาด
เทือกเขา ภูเขา

กลุ่มรหัส 12211 Topography Name : สถานที่ที่เป็นชื่อของ
ภูมิประเทศ เช่น ชายหาดป่าตอง ภูเขาบรรทัด

กลุ่มรหัส 12220 Astronomy : สถานที่ที่เกี่ยวข้องกับดาราศาสตร์ เช่น
โลก ดวงอาทิตย์

กลุ่มรหัส 12300 Man-Made-Place : สถานที่ที่มนุษย์สร้างขึ้น เช่น ดึก บ้าน
โรงเรียน

กลุ่มรหัส 13000 Concrete Thing สิ่งของ แบ่งเป็น

กลุ่มรหัส 13100 Nature Thing : สิ่งของที่เกิดตามธรรมชาติ แบ่งเป็น

กลุ่มรหัส 13110 Living Thing : สิ่งมีชีวิต แบ่งเป็น

กลุ่มรหัส 13111 Land Animal : สัตว์ เช่น สุนัข แมว

กลุ่มรหัส 13112 Aquatic Animal : สัตว์น้ำ เช่น ปลา ปู

กลุ่มรหัส 13113 Fowl : สัตว์ปีก เช่น นก ไก่

กลุ่มรหัส 13114 Reptile : สัตว์เลื้อยคลาน เช่น จระเข้

กลุ่มรหัส 13115 Plant : พืชบก เช่น ต้นข้าว ต้นล็ก

กลุ่มรหัส 13116 Plant in water : พืชน้ำ เช่น สาหร่าย

กลุ่มรหัส 13117 A Part of Animal : ส่วนของสิ่งมีชีวิตหรือสิ่งที่ได้จากสิ่งมีชีวิตเช่น ขา แขน

กลุ่มรหัส 13118 A Part of Plant / Derivative : ส่วนของพืชหรือสิ่งที่ได้จากพืช เช่น กิ่ง ดอก ใบ

กลุ่มรหัส 13120 Substance : สสาร แบ่งเป็น

กลุ่มรหัส 13121 Solid : ของแข็ง เช่น น้ำแข็ง ทอง เงิน

กลุ่มรหัส 13122 Liquid : ของเหลว เช่น น้ำ น้ำมัน

กลุ่มรหัส 13123 Gas : ก๊าซ เช่น ไอน้ำ อากาศ ก๊าซธรรมชาติ

กลุ่มรหัส 13200 Man-Made-Thing : สิ่งของที่มนุษย์สร้างขึ้น แบ่งเป็น

กลุ่มรหัส 13210 Material : วัตถุดิบ เช่น ผ้า พลาสติก

กลุ่มรหัส 13220 Finished Product : ผลิตภัณฑ์ เช่น โต๊ะ รถยนต์ คอมพิวเตอร์

กลุ่มรหัส 13230 A Part of Product or it's Derivative : ส่วนของผลิตภัณฑ์หรือผลที่ได้จากผลิตภัณฑ์ เช่น ขาโต๊ะ ใบมีด

3.2.2. กลุ่มนามธรรม^[5] (Abstract) แบ่งเป็นกลุ่มลำดับย่อยได้เป็น

กลุ่มรหัส 21000 Abstract Matter : นามธรรมที่ครอบคลุมเหตุการณ์และกริยาอาการต่างๆ แบ่งเป็น

กลุ่มรหัส 21100 Activity : กริยาและอาการต่างๆ แบ่งเป็น

กลุ่มรหัส 21110 Action : การกระทำ เช่น ปฏิกริยา เดินทาง ใช้ กิน

กลุ่มรหัส 21120 Mental Action : การกระทำที่เกี่ยวกับความรู้สึกนึกคิด แบ่งเป็น

กลุ่มรหัส 21121 Rational Action : การกระทำที่ต้องใช้เหตุผล
และความคิด เช่น คิด ควบคุม

กลุ่มรหัส 21122 Emotional Action : การกระทำที่เกี่ยวข้องกับ
ความรู้สึก เช่น รัก หวัง โกรธ

กลุ่มรหัส 21200 Phenomenon : ปรากฏการณ์ แบ่งเป็น

กลุ่มรหัส 21210 Event : เหตุการณ์ เช่น ดวง โชค ความสำเร็จ

กลุ่มรหัส 21220 National Phenomenon ; ปรากฏการณ์ธรรมชาติ เช่น
พายุฝน สุริยุปราคา

กลุ่มรหัส 21230 Social Phenomenon : ปรากฏการณ์ทางสังคม เช่น
สงคราม การจลาจล

กลุ่มรหัส 21240 Physiological Phenomenon : ปรากฏการณ์ทางสรีระ
หรือทางร่างกาย เช่น อาการป่วย การเกิด การตาย

กลุ่มรหัส 22000 Abstract Thing : นามธรรมที่มีลักษณะคล้ายสิ่งที่สามารถจับต้องได้ แบ่ง
เป็น

กลุ่มรหัส 22100 Product of Activity : ผลของกริยาและอาการต่าง ๆ แบ่งเป็น

กลุ่มรหัส 22110 Result of Action : ผลของการกระทำ เช่น คำอธิบาย
แผนการณื ข้อเสนอ

กลุ่มรหัส 22120 Result of Mental Action : ผลของการกระทำที่เกี่ยวข้อง
กับความรู้สึกนึกคิด เช่น ความจำ ความรัก

กลุ่มรหัส 22130 Abstract Thing Fixed as Socially : สิ่งที่เป็นนาม
ธรรมที่เกี่ยวข้องกับสังคม เช่น งานสงกรานต์ งานปีใหม่

กลุ่มรหัส 22140 Abstract Thing Fixed as Culture : สิ่งที่เป็นนามธรรม
ที่เกี่ยวข้องกับวัฒนธรรม เช่น ภาษาไทย รำวง มวยไทย

กลุ่มรหัส 22150 Inmaterial Man-Made-Thing : สิ่งที่ไม่ใช่วัตถุที่มนุษย์
สร้างขึ้น เช่น โปรแกรมคอมพิวเตอร์ บทความ สุภาษิต

กลุ่มรหัส 23000 Abstract Relation : ความสัมพันธ์เชิงนามธรรม แบ่งเป็น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา ²⁷ และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กลุ่มรหัส 23100 Place : สถานที่ เช่น บน ล่าง ซ้าย ขวา ที่นี้

กลุ่มรหัส 23200 Time : เวลา เช่น เมื่อวานนี้ พรุ่งนี้ วันจันทร์

กลุ่มรหัส 23300 Existence : สถานะ เช่น มี คงอยู่

กลุ่มรหัส 23400 Set : กลุ่ม เช่น กลุ่ม แถว

กลุ่มรหัส 23500 Attribute : คุณสมบัติของวัตถุหรือสิ่งของ แบ่งเป็น

กลุ่มรหัส 23510 Quality : คุณลักษณะ เช่น ลักษณะ อุบัติภัย

กลุ่มรหัส 23520 Condition : สภาวะ เช่น สถานะการณ์ สภาพร่างกาย

กลุ่มรหัส 23530 Appearance : ลักษณะที่ปรากฏให้เห็น เช่น สี รูปร่าง ลาย

กลุ่มรหัส 23540 Sort/Type : ประเภทหรือชนิด เช่น ชนิด แบบ ประเภท

กลุ่มรหัส 23550 Number/Quantity : จำนวน เช่น ปริมาตร ความสูง

กลุ่มรหัส 23560 Relation : ความสัมพันธ์ เช่น ความเกี่ยวข้อง

กลุ่มรหัส 23570 Function : หน้าที่และประสิทธิภาพ เช่น ความสามารถ ประสิทธิภาพ คุณภาพ

กลุ่มรหัส 23600 Value : ค่าของคุณสมบัติหรือสิ่งของ แบ่งเป็น

กลุ่มรหัส 23610 Quality : ค่าที่แสดงลักษณะ เช่น ใจดี สูง ง่าย

กลุ่มรหัส 23620 Condition : ค่าที่แสดงสภาวะ เช่น เป็นอันตราย ปลอดภัย

กลุ่มรหัส 23630 Appearance : ค่าที่แสดงลักษณะที่ปรากฏ เช่น กลุ่ม สี เหลี่ยม สว่าง แดง

กลุ่มรหัส 23640 Sort/Type : ค่าที่แสดงชนิดหรือประเภท เช่น รุ่น 2503 แบบดิจิทัล

กลุ่มรหัส 23650 Number/Quantity : ค่าที่แสดงจำนวนปริมาณ เช่น 10 วัน มาก นาน

กลุ่มรหัส 23660 Relation : ค่าที่แสดงความสัมพันธ์ เช่น ใกล้เคียง คล้าย ติดกัน

กลุ่มรหัส 23670 Function : ค่าที่แสดงหน้าที่และประสิทธิภาพ เช่น (ประสิทธิภาพ)สูง (คุณภาพ)เยี่ยม

รหัสที่ 6 รหัสที่ 7 รหัสที่ 8 และรหัสที่ 9 นั้นเป็นการกำหนดขอบเขตความหมายให้ลึกลงไปจนสามารถระบุคำศัพท์ได้ เพื่อสามารถที่จะกำหนดรูปแบบผิวของคำในภาษาต่างๆ โดยที่ในกลุ่มรหัสที่ 6-9 นั้นจะกำหนดด้วยตัวอักษร a-z และในกลุ่มใดที่ไม่มีการกำหนดรหัสไว้จะแทนตำแหน่งนั้นด้วยเลขศูนย์ ดังแสดงตัวอย่างในตารางที่ 3.2 การแสดงรหัสแยกแยะความหมายในวิทยานิพนธ์นี้ได้คำนึงถึงจำนวนความหมายที่จะมารองรับภาษาต่างๆ โดยใน 1 หมวดของกลุ่มความหมายจะสามารถกำหนดความหมายเฉพาะของคำได้ถึง 475254 ความหมายดังนี้

$$\boxed{1} \times \boxed{1} \times \boxed{1} \times \boxed{1} \times \boxed{1} \times \boxed{1} \times \boxed{1} \times \boxed{1} \times \boxed{1} \times \boxed{1} \times \boxed{26} = 26 \text{ ความหมาย} \quad \text{---(1)}$$

$$\boxed{1} \times \boxed{1} \times \boxed{1} \times \boxed{1} \times \boxed{1} \times \boxed{1} \times \boxed{1} \times \boxed{1} \times \boxed{26} \times \boxed{26} = 676 \text{ ความหมาย} \quad \text{---(2)}$$

$$\boxed{1} \times \boxed{1} \times \boxed{1} \times \boxed{1} \times \boxed{1} \times \boxed{1} \times \boxed{26} \times \boxed{26} \times \boxed{26} = 17576 \text{ ความหมาย} \quad \text{---(3)}$$

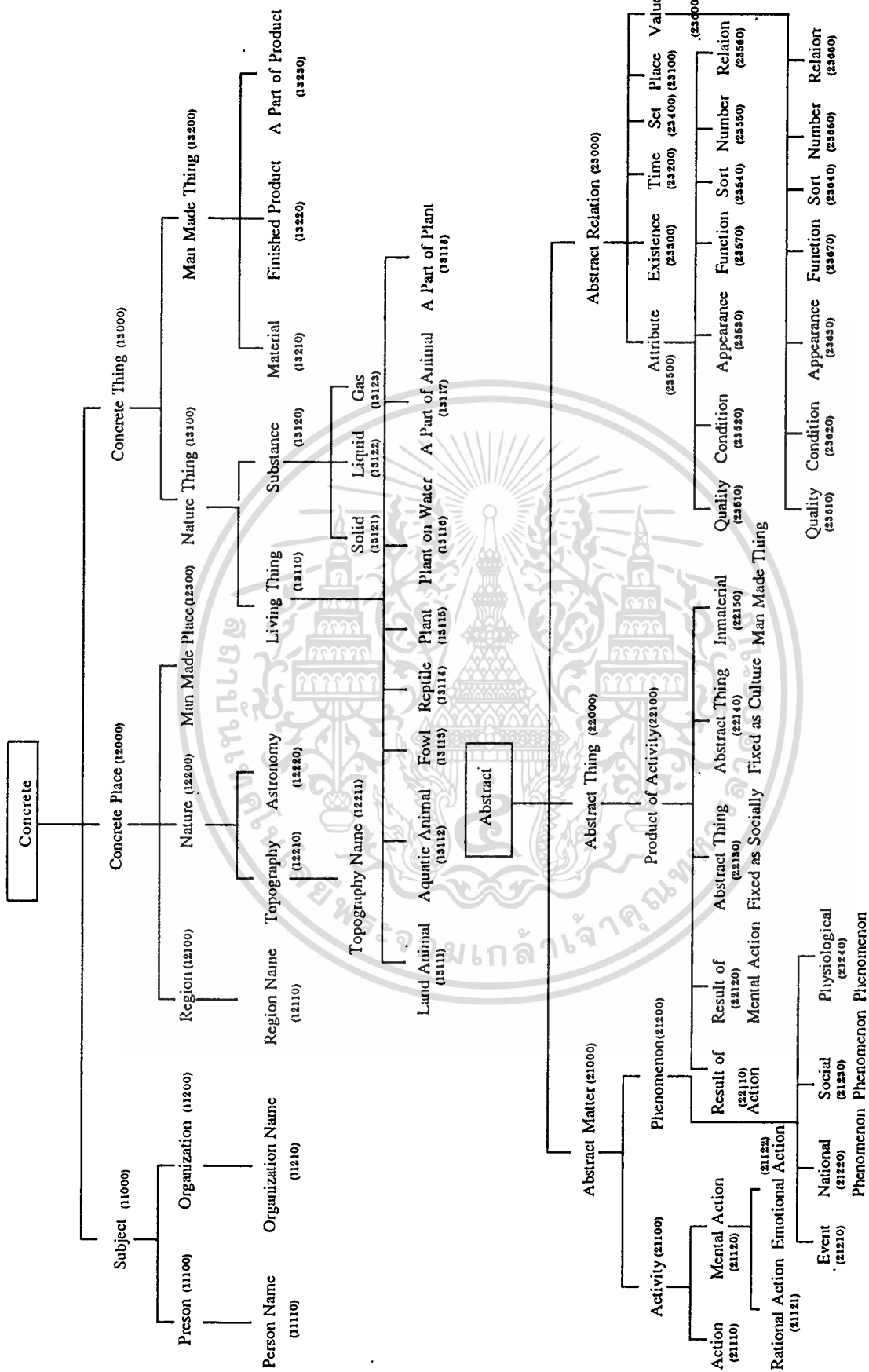
$$\boxed{1} \times \boxed{1} \times \boxed{1} \times \boxed{1} \times \boxed{1} \times \boxed{26} \times \boxed{26} \times \boxed{26} \times \boxed{26} = 456976 \text{ ความหมาย} \quad \text{---(4)}$$

$$(1) + (2) + (3) + (4) = 475254 \text{ ความหมาย}$$

จากการแบ่งความหมายลดหลั่นนี้ จะเห็นว่าการจำกัดขอบเขตของความหมายของกลุ่มต่างๆ คล้ายเป็นการ Focus ข้อมูลให้ละเอียดเพิ่มขึ้นเพื่อที่จะสามารถกำหนดรูปแบบผิวของคำ การกำหนดรหัสแยกแยะหน่วยความหมายนี้มีความสำคัญต่อระบบการแปลหลายภาษาด้วยคอมพิวเตอร์โดยใช้ภาษากลางเป็นสื่ออย่างมาก ทำให้การวิเคราะห์ประโยคเพื่อที่จะให้ได้รูปแบบความหมายในรูปของภาษากลางได้อย่างถูกต้อง และเมื่อทำการสังเคราะห์ประโยคภาษาเป้าหมายก็สามารถที่จะทราบความหมายที่แท้จริงของประโยค ซึ่งเป็นการลดความกำกวม (ambiguous) ของการสังเคราะห์ประโยคภาษาเป้าหมายที่จะเกิดขึ้น

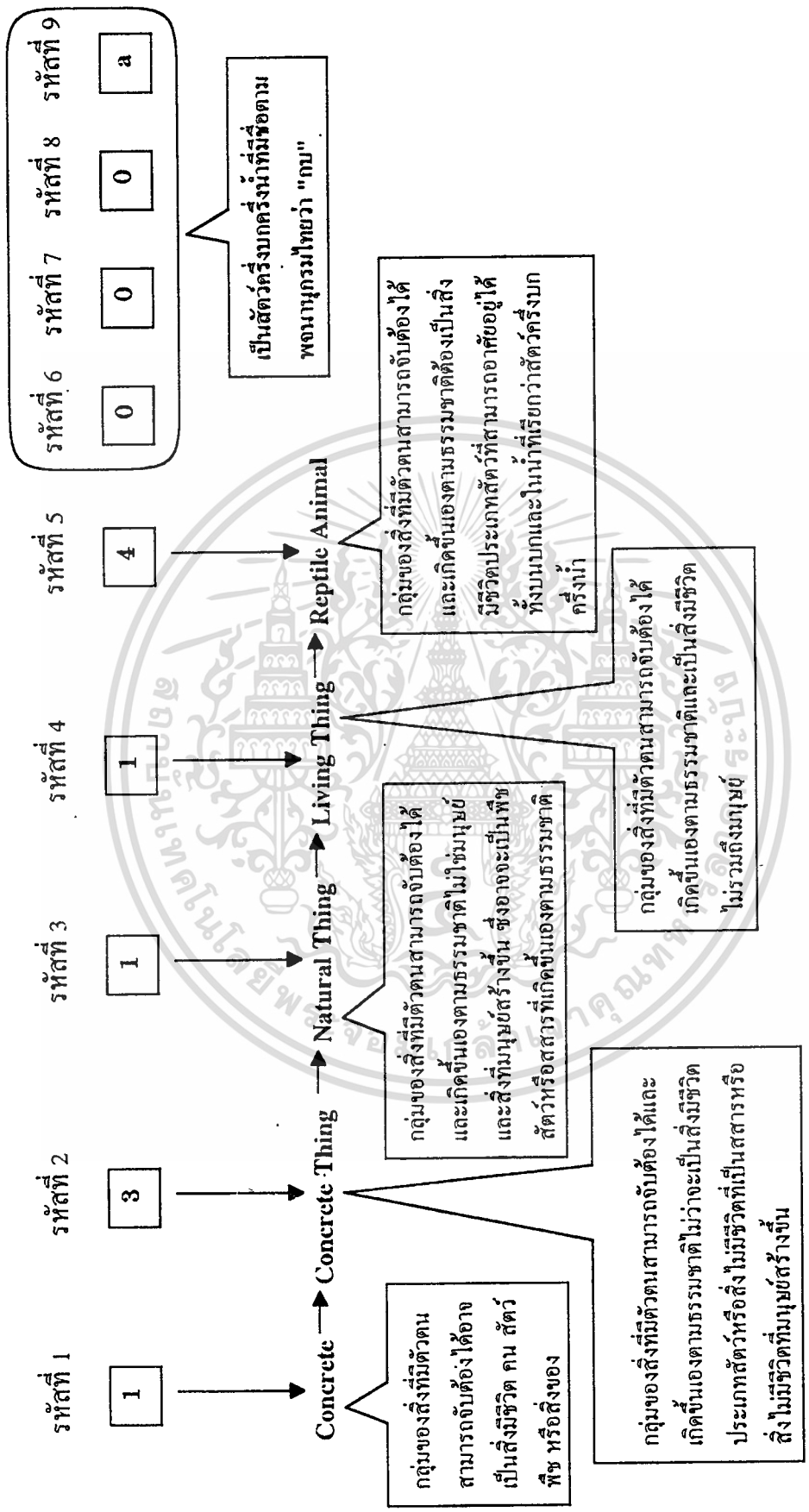
เอกสารนี้เป็นทรัพย์สินทางปัญญาสำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 29 จะต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.2 แสดงการแบ่งกลุ่มรหัสความหมาย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.3 ตัวอย่างรหัสแยกแยะหน่วยความหมาย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3 คำอธิบายหน่วยความหมาย:CPDC (Conceptual Premitive Description)

เป็นข้อมูลระดับความหมายของ CPID ในที่นี้จะใช้รหัสย่อว่า CPDC ซึ่งเป็นการให้ข้อมูลทางด้านกรอบทางความหมายของการใช้คำศัพท์ในภาษาต่างๆ โดยเฉพาะคำศัพท์ที่มีหลายความหมาย CPDC นี้จะอธิบายความหมายของ CPID เพื่อให้เกิดความเข้าใจในความหมายของ CPID ที่ตรงกัน โดยจะอธิบายความหมายของ CPID นั้น เพื่อให้กำหนดขอบเขตของความหมายของ CPID แต่ละตัว เพื่อสามารถกำหนดคำในภาษานั้นๆ และในกรณีที่ยังไม่ได้กำหนดไว้ในภาษานั้น หรือกรณีที่ไม่สามารถกำหนดเป็นคำในภาษานั้น ก็สามารถจะเข้าใจความหมายของ CPID ในรูปของคำอธิบาย ก็สามารถที่จะอธิบายให้เข้าใจถึงความหมายของ CPID นั้นได้ เป็นส่วนที่ใช้แทนคำในภาษาต่างๆ ในส่วนที่มีหลายความหมายในลักษณะของการอธิบายความหมายของ CPID นั้นเพื่อให้เกิดความเข้าใจใน CPID ที่ตรงกัน ซึ่งในหนึ่งโครงสร้างแบบผิว อาจจะแสดงได้หลายคนสมบัติก็ได้ เช่น คำว่า "กบ" ในภาษาไทย ถ้าอยู่ในประโยคว่า สมศักดิ์กินกบผัดเผ็ด คำว่า "กบ" ในที่นี้จะหมายถึงสิ่งมีชีวิตเป็นสัตว์สะเทินน้ำสะเทินบกสามารถที่จะเป็นอาหารของมนุษย์ได้ แต่ถ้าเป็นประโยค เขาใช้กบไล่ไม้ คำว่า "กบ" ในประโยคนี้อาจหมายถึงเครื่องมือที่ใช้ในการทำให้ไม้บางลง และทำให้ไม้เรียบซึ่งจะมีลักษณะเป็นค่านามเหมือนกัน ดังตัวอย่างที่แสดงในตารางที่ 3.2

CPID										CPDC
1	2	3	4	5	6	7	8	9		
1	1	1	0	0	0	0	0	a		a person who has the qualification to teach someone something in school ⁽¹⁾
(ครู)										
1	3	1	1	2	0	0	0	a		cold-blooded animal living wholly in water and breathing through gills, with fins for swimming ⁽¹⁾
(ปลา)										
1	3	1	1	4	0	0	0	a		cold-blooded ,tailless jumping animal living in water and on land ⁽¹⁾
(กบ)										
1	3	2	2	0	0	0	0	s		a tool with a blade that takes very thin pieces of wooden surfaces to make them smooth ⁽¹⁾
(กบ)										
2	1	1	1	0	0	0	0	a		take (food etc.) into the mouth and swallow it ⁽¹⁾
(กิน)										

⁽¹⁾ Oxford Advanced Learner's Dictionary of Current English Third Edition ของ A S Hornby

ตารางที่ 3.2 ตัวอย่างการกำหนดความหมายระดับลึกของ CPID

() คือตัวอย่างรูปแบบผิวของคำในภาษาไทย

3.4 สรุป

หน่วยความหมายของภาษากลางในวิทยานิพนธ์นี้ได้กำหนดกลุ่มคำที่มีความหมายในกลุ่มเดียวกัน โดยการแบ่งกลุ่มในลักษณะตามลำดับขั้นความหมาย (of Word Meaning) ได้ 2 กลุ่มใหญ่ๆ คือ รูปธรรม (Concrete) และนามธรรม (Abstract) แต่ละกลุ่มก็แบ่งย่อยทางความหมายลึกลงไปโดยกำหนดในรูปของรหัสแยกแยะความหมาย (Conceptual Primitve Identifier) ซึ่งรหัสแยกแยะความหมาย 1 รหัสจะมีความหมายเดียวแต่อาจจะมีหลายคำที่สามารถนำมาแทนความหมายนี้ ในภาษากลางจะต้องมีทั้งหน่วยความหมายและหน่วยความสัมพันธ์ถึงสามารถที่จะสื่อความหมายของภาษาธรรมชาติได้ ซึ่งจะกล่าวต่อไปในบทที่ 4 หน่วยความสัมพันธ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 33 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

หน่วยความสัมพันธ์

4.1 บทนำ

ในการแปลภาษาด้วยระบบคอมพิวเตอร์โดยใช้ภาษากลางเป็นสื่อในการถ่ายทอดความรู้ทางภาษาจากการวิเคราะห์ภาษาต้นแบบไปสู่การสร้างภาษาเป้าหมายดังที่ได้กล่าวอธิบายมาในบทต้นๆ ของวิทยานิพนธ์แล้วว่า ส่วนสำคัญอยู่ที่โครงสร้างภาษากลางซึ่งประกอบไปด้วย หน่วยความหมาย ระดับแก่นของคำ โครงข่ายแสดงความสัมพันธ์ และหน่วยความสัมพันธ์ทางภาษา โดยแต่ละหน่วยจะต้องถูกวิจัยออกแบบให้เป็นกลาง นั่นคือไม่อิงภาษาใดภาษาหนึ่งโดยเฉพาะ และจะต้องมีโครงสร้างภายในหน่วยที่สามารถบรรจุหรือแสดงข้อมูลทางภาษาเพียงพอที่จะถ่ายทอดการแปลจากภาษาต้นแบบไปสู่การสร้างภาษาเป้าหมายได้อย่างถูกต้อง ความถูกต้องของการแปลในที่นี้หมายถึงว่า ภาษาเป้าหมายที่สร้างขึ้นมาได้ด้วยระบบซอฟต์แวร์ของการแปล นอกจากจะต้องมีโครงสร้างที่ถูกหลักไวยากรณ์ของภาษานั้นๆ แล้วยังจะต้องมีความหมายที่ได้ใจความตรงกับภาษาต้นแบบ

ในบทนี้จะกล่าวถึง การวิจัยออกแบบหน่วยความสัมพันธ์ทางภาษาของภาษากลางของงานวิทยานิพนธ์นี้ โดยคำนึงถึงการให้ข้อมูลทางภาษาในด้านความสัมพันธ์ระหว่างหน่วยคำ (หรือหน่วยความหมายที่กล่าวโดยละเอียดมาแล้วในบทที่ 3) ในระดับแก่นเกี่ยวกับการแสดงความสัมพันธ์ของการกระทำ ขอบเขตของการกระทำ ผู้หรือสิ่งแสดงการกระทำ จุดประสงค์ของการกระทำ เป้าหมายของการกระทำหรือผู้ถูกกระทำ เวลาที่เกิดเหตุการณ์เหล่านั้นของการกระทำ สถานที่ที่เกิดของการกระทำ รายละเอียดอื่นๆ ของการกระทำ รายละเอียดของผู้กระทำหรือผู้ถูกกระทำหรือสิ่งนั้น ปริมาณของสิ่งเหล่านั้น ส่วนประกอบของสิ่งเหล่านั้น และเหตุผลของการกระทำ โดยที่ได้กล่าวมาแล้วว่า รูปแบบภาษากลางของงานวิทยานิพนธ์นี้ได้วิจัยกำหนดให้การกระทำหรือหน่วยคำกริยาหลักของประโยคเป็นศูนย์กลางของการกระจายความสัมพันธ์กับหน่วยคำอื่นในรูปโครงสร้างความพึงพาแบบต้นไม้ เพราะเป็นที่ยอมรับกันโดยพื้นฐานของนักภาษาศาสตร์แล้วว่า ศูนย์กลางของความหมายของประโยคนั้นอยู่ที่กริยาหลักของประโยค²¹ ซึ่งชี้ให้เห็นวัตถุประสงค์ของการกระทำหรือการเกิดเหตุการณ์ในประโยคนั้น ดังนั้นการประยุกต์หลักการนี้กับภาษากลางจึงเหมาะสม เพราะจะสามารถแสดงความสัมพันธ์ทางหน้าที่ระหว่างหน่วยคำได้สมบูรณ์ และสะดวกต่อการถ่ายทอดข้อมูลทางหลักไวยากรณ์ของภาษาต่างๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

งานวิทยานิพนธ์นี้ได้วิจัยออกแบบหน่วยความสัมพันธ์โดยได้ศึกษาทฤษฎีความสัมพันธ์ (Relation Theory) ทางการประมวลผลภาษาธรรมชาติของ ชาร์ลส์ เจ. ฟิลล์มอร์ (Charles J. Fillmore) กรูเบอร์ (Gruber) วอลเลซ แอล เชฟ (Wallace L. Chafe) นิลเซน (Nilsen) อาร์ อี ลองเอเคอร์ (R.E Longacre) และจอห์น เอ็ม แอนเดอร์สัน (John M. Anderson) และได้วิจัยหาจุดรวมของทฤษฎีของนักวิจัยทั้ง 6 ท่านนี้เพื่อลดรูปแบบของหน่วยความสัมพันธ์ไม่ให้เกิดความซ้ำซ้อน และวิจัยเพิ่มหลักการของวิทยานิพนธ์เพื่อให้หน่วยความสัมพันธ์มีความหมายที่ชัดเจนขึ้น นอกจากนี้ยังได้จัดแบ่งหน่วยความสัมพันธ์ที่มีความหมายใกล้เคียงไว้ในกลุ่มเดียวกันเพื่อสะดวกต่อการวิเคราะห์ในเชิงภาษาศาสตร์คอมพิวเตอร์ต่อไป

4.2 แนวความคิดเกี่ยวกับความสัมพันธ์ทางอรรถศาสตร์ระหว่างหน่วยคำ⁶⁾

ในหัวข้อนี้จะเป็นแนวความคิดของนักวิจัยจำนวน 6 ท่าน โดยได้ทำการวิจัยเกี่ยวกับความสัมพันธ์ของหน่วยคำ ซึ่งทั้งแนวความคิดและการกำหนดความหมายของความสัมพันธ์ของแต่ละท่านก็จะแตกต่างกันไป โดยอาศัยหลักพื้นฐานทางภาษาศาสตร์ในการเสนอแนวความคิด ดังต่อไปนี้

4.2.1 แนวความคิดของ ชาร์ลส์ เจ. ฟิลล์มอร์ (Charles J. Fillmore)

ได้เสนอแนวความคิดเกี่ยวกับความสัมพันธ์ทางอรรถศาสตร์ และแสดงแนวคิดเกี่ยวกับความสัมพันธ์ระหว่างคำนามกับกริยาในระดับโครงสร้างลึก (deep structure) ที่ไม่ใช่ความสัมพันธ์ในระดับโครงสร้างผิว (surface structure) ซึ่งเป็นความสัมพันธ์ระหว่างคำนามกับกริยาในลักษณะที่เรียกว่าเป็นประธาน (subject) และกรรม (object) ของประโยค ลักษณะของความสัมพันธ์ในโครงสร้างลึกนี้จะแสดงถึงเหตุการณ์ต่างๆ รอบๆ ตัวว่าใครทำอะไร ใครมีประสบการณ์เรื่องใด ใครได้รับผลประโยชน์จากอะไร สิ่งนั้นเกิดที่ไหน อะไรเปลี่ยนแปลงไป อะไรเคลื่อนที่ สิ่งนั้นเริ่มต้นจากอะไร ที่ไหน และสิ้นสุดเมื่อไร ที่ไหน ฟิลล์มอร์เรียกลักษณะความสัมพันธ์เหล่านี้ว่า การกสัมพันธ์ (Case Relationships) และเสนอแนวความคิดว่า ความสัมพันธ์ทางอรรถศาสตร์ระหว่างคำนามกับคำกริยาจะอยู่ที่ส่วนพื้นฐานของไวยากรณ์ของภาษาทุกๆ ภาษา และควรจะมีลักษณะดังนี้

- 1) ผู้ทำ (agentive) ความสัมพันธ์ระหว่างคำนามกับคำกริยาในลักษณะผู้ทำให้เกิดเหตุการณ์ขึ้น โดยปกติผู้ทำให้เกิดเหตุการณ์ขึ้นจะเป็นสิ่งมีชีวิต เช่น จอห์นเปิดประตู, ประตูถูกจอห์นเปิด
- 2) เครื่องมือ (instrumental) สัมพันธ์ระหว่างคำนามกับคำกริยาในลักษณะบ่งบอกเครื่องมือ หรือ เครื่องที่ทำให้เกิดเหตุการณ์ หรือสภาพ ซึ่งแสดงโดยคำกริยา เช่น จอห์นเปิดประตูด้วยกุญแจ, จอห์นใช้กุญแจเปิดประตู
- 3) ผู้ถูกกระทำ (object) ความสัมพันธ์ระหว่างคำนามกับคำกริยาในลักษณะผู้ทรงสภาพ หรือ ผู้ที่รับสภาพ หรือได้รับความกระทบกระเทือนจากการกระทำที่แสดง โดยคำกริยาความสัมพันธ์นี้เป็นความสัมพันธ์ที่มีความหมายเป็น กลาง (neutral case) ที่สุด เช่น จอห์นเปิดประตู, ประตูเปิด
- 4) สถานที่ (location) ความสัมพันธ์ระหว่างคำนามกับคำกริยาในลักษณะบ่งบอกบริเวณที่เกิดเหตุการณ์ หรือการกระทำซึ่งแสดงโดยคำกริยา เช่น จอห์นเก็บรถไว้ในโรงรถ, เมืองซิดนีย์ไกลมแรง
- 5) ผู้มีประสบการณ์ (experiencer) สัมพันธ์ระหว่างคำนามกับคำกริยาในลักษณะผู้มีความรู้สึก มีอารมณ์ มีการเรียนรู้เหตุการณ์ต่างๆ ที่แสดงโดยคำกริยา เช่น ซูปร็อนแล้ว
- 6) ผู้ร่วมกระทำ (comitative) สัมพันธ์ระหว่างคำนามกับคำกริยาในลักษณะทำอาการหรือทำให้เกิดเหตุการณ์ที่แสดงโดยคำกริยา ร่วมกันกับผู้อื่น เช่น จอห์นทะเลาะกับแม่
- 7) ผู้ได้รับประโยชน์ (benefactive) สัมพันธ์ระหว่างคำนามกับคำกริยา ในลักษณะผู้ได้รับผลแห่งการกระทำซึ่งแสดงโดยกริยา เช่น จอห์นให้หนังสือแก่น้องชายของฉัน
- 8) เวลา (time) สัมพันธ์ระหว่างคำนามกับคำกริยา ในลักษณะบ่งบอกเวลาของเหตุการณ์ หรือสภาพที่แสดงโดยกริยา เช่น จอห์นตื่นแต่เช้า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 36 จะต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- 9) จุดเริ่มต้น (source) สัมพันธ์ระหว่างคำนามกับคำกริยาในลักษณะบอกจุดตั้งต้นของการกระทำ ซึ่งแสดงโดยคำกริยา เช่น จอห์นมาจากบ้าน
- 10) จุดหมาย (goal) ความสัมพันธ์ระหว่างคำนามกับคำกริยาในลักษณะจุดสิ้นสุดของการกระทำ ซึ่งแสดงโดยคำกริยา เช่น จอห์นวาดรูป

4.2.2 แนวความคิดของ กรูเบอร์ (Gruber)

ได้เสนอความคิดของเขาในวิทยานิพนธ์ชื่อ "Studies in Lexical Relations" โดยคิดว่าความสัมพันธ์ทางไวยากรณ์ (Grammatical Relations) ไม่สอดคล้องกับความสัมพันธ์ทางความหมาย (Semantic Relations) และได้ยกตัวอย่างคือ

ก ประตุเปิด

ข ซาลิเปิดประตุ

กรูเบอร์ชี้ให้เห็นว่า คำว่า "ประตุ" ในประโยค ก และประโยค ข นั้นมีหน้าที่ทางความหมาย (Semantic Function) เหมือนกัน แม้ว่า "ประตุ" จะมีความสัมพันธ์ทางไวยากรณ์เป็นประธานในประโยค ก และเป็นกรรมในประโยค ข เมื่อเป็นเช่นนี้ กรูเบอร์จึงได้เสนอความสัมพันธ์แบบหนึ่งเรียกว่า ความสัมพันธ์แบบผู้ถูกกระทบ (theme) ในทุกๆ ประโยคจะมีนามวลี ซึ่งทำหน้าที่เป็น ผู้ถูกกระทบ

1) ผู้ถูกกระทบ

โดยมีกรณีดังต่อไปนี้

(theme)

1.1 ในประโยคที่มีคำกริยาที่เกี่ยวกับการเคลื่อนที่ คำกริยาบางคำที่แสดงการเคลื่อนที่ที่มองไม่เห็นการกระทำ รวมถึงการถูกเปลี่ยนเจ้าของ หรือถูกถ่ายทอด นามวลีที่ถูกเคลื่อนที่ไปในประโยคดังกล่าวจะเป็นผู้ถูกกระทบ เช่น ก้อนหินเคลื่อนที่ไป, จอห์นกลิ้งก้อนหินจากกองหินในบ้าน, บิลผลักก้อนหินลงหลุม

1.2 ในประโยคที่มีคำกริยาที่บ่งบอกสถานที่ นามวลีที่ถูกบ่งสถานที่ในประโยคดังกล่าวจะเป็นผู้ถูกกระทบ เช่น เฮอร์แมนเก็บหนังสือไว้บนชั้น, แมกซ์รู้คำตอบแล้ว, แมกซ์เป็นเจ้าของหนังสือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 37 จะต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- 2) สถานที่ (location) สัมพันธระหว่างคำนามวลีกับกริยาที่บ่งสถานที่ มักมีคำบุรพบทปรากฏอยู่ในประโยคด้วย เช่น จอห์นอยู่ในห้อง, จอห์นอยู่ในอารมณโกรธ
- 3) แหล่งเดิม (source) สัมพันธระหว่างนามวลีที่แสดงจุดตั้งต้นของการเคลื่อนที่ หรือการเปลี่ยนเจ้าของ กับคำกริยาที่เกี่ยวข้องกับการเคลื่อนที่ มักมีคำบุรพบทปรากฏอยู่ในประโยคดังกล่าว เช่น จอห์นกลิ้งก้อนหินจากกองหินในบ้าน, จอห์นซื้อตะเกียงจากแมกซ์
- 4) จุดหมาย (goal) สัมพันธระหว่างนามวลีที่แสดงจุดปลายของการเคลื่อนที่ หรือการเปลี่ยนเจ้าของ กับคำกริยาที่เกี่ยวข้องกับการเคลื่อนที่ มักมีคำบุรพบทปรากฏอยู่ในประโยคดังกล่าว เช่น บิลผลั๊กก้อนหินลงหลุม, เดฟอธิบายข้อพิสูจน์แก่นักเรียน
- 5) ผู้ทำ (agent) ความสัมพันธ์ระหว่างนามวลีที่มีความตั้งใจที่จะกระทำสิ่งหนึ่งสิ่งใด กับคำกริยา นามวลีดังกล่าวจะต้องเป็นสิ่งมีชีวิตเท่านั้น เช่น จอห์นกลิ้งก้อนหินจากกองหินในบ้าน

4.2.3 แนวความคิดของ วอลเลซ แอล เชฟ (Wallace L. Chafe)

กล่าวถึงความสัมพันธ์ทางอรรถศาสตร์ระหว่างคำกริยา โดยแบ่งกริยาออกเป็นประเภทต่างๆ เป็น 5 ประเภทคือ กริยาแสดงสภาพ (State) กริยาแสดงการเปลี่ยนแปลง (Process) กริยาแสดงอาการ (Action) กริยาแสดงทั้งอาการและการเปลี่ยนแปลง (Action-Process) กริยาแสดงสภาพแวดล้อม (Ambient) ตามลักษณะเฉพาะทางความหมาย (Semantic feature) แล้วพิจารณาว่ากริยาแต่ละประเภทจะปรากฏกับคำนามที่มีความสัมพันธ์กับกริยานั้นๆ ต่างๆ กัน ได้แก่

- 1) ผู้ทำ (agent) ความสัมพันธ์ระหว่างคำนามที่หมายถึงสิ่งมีชีวิตที่สามารถกระทำ หรือแสดงอาการ กับคำกริยาที่บ่งบอกอาการ หรือคำกริยาที่บ่งบอกทั้งอาการและการเปลี่ยนแปลง เช่น แสบเรียตร้องเพลง, แสบเรียตทำจานแตก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 38 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- 2) ผู้ก่อให้เกิดกริยา (potent) ความสัมพันธ์ระหว่างค่านามซึ่งเป็นสิ่งไม่มีชีวิต แต่มีพลังที่จะทำให้เกิดอาการได้ กับคำกริยาแสดงอาการและการเปลี่ยนแปลง เช่น ความร้อนละลายน้ำตา
- 3) ผู้ทรง-รับ (patient) ความสัมพันธ์ระหว่างค่านามซึ่งเป็นสิ่งมีชีวิต หรือไม่มีก็ได้ และได้รับผลหรือเปลี่ยนแปลงสภาพ กับคำกริยาแสดงสภาพ แสดงการเปลี่ยนแปลงหรือแสดงทั้งอาการและการเปลี่ยนแปลง เช่น พินแห้งแล้ว, จานแตก, ความร้อนละลายน้ำตา, ทอมเห็นงู
- 4) ผู้ประสบ (experiencer) ความสัมพันธ์ระหว่างค่านามซึ่งเป็นสิ่งมีชีวิต และมีความรู้สึก กับคำกริยาที่แสดงสภาพ หรือแสดงการเปลี่ยนแปลง เช่น ทอมต้องการเครื่องดื่ม, ทอมรู้คำตอบ, ทอมเห็นงู
- 5) ผู้เสริมให้สมบูรณ์ (complement) ความสัมพันธ์ระหว่างค่านามกับคำกริยา โดยที่ค่านามนั้นช่วยทำให้ความหมายของคำกริยาสมบูรณ์ หรือช่วยทำให้ความหมายของคำกริยาแคบเข้า แต่ไม่ใช่ค่านามที่เป็นผู้ทรง-รับ เช่น แมรี่ร้องเพลง, ลูกกวาดราคา 10 เซนต์
- 6) ผู้ได้รับประโยชน์ (beneficiary) ความสัมพันธ์ระหว่างค่านามซึ่งเกี่ยวเนื่องกับประโยชน์ กับคำกริยาแสดงการเปลี่ยนแปลง หรือคำกริยาที่แสดงทั้งอาการและการเปลี่ยนแปลง หรือคำกริยาที่แสดงสภาพ เช่น ทอมมีตัว, ทอมได้รางวัลเครื่องรับโทรทัศน์, แมรี่ให้ตัวแก่ทอม
- 7) เครื่องมือ (instrumental) ความสัมพันธ์ระหว่างค่านามซึ่งเป็นสิ่งไม่มีชีวิต ที่เป็นเครื่องมือในการทำกริยากับคำกริยาแสดงอาการ หรือคำกริยาที่แสดงทั้งอาการและการเปลี่ยนแปลง เช่น ทอมเปิดประตูด้วยกุญแจ, ตัดเชือกด้วยมีด

- 8) สถานที่ (location) ความสัมพันธ์ระหว่างคำนามซึ่งเป็นสิ่งไม่มีชีวิต และบ่งสถานที่ของกริยากับคำกริยาแสดงอาการ กริยาแสดงสภาพ หรือกริยาที่แสดงทั้งอาการและการเปลี่ยนแปลง เช่น มีดอยู่ในกล่อง, แมวคลานบนหลังคา, เด็กขว้างก้อนหินไปที่ประตู

4.2.4 แนวความคิดของ นิลเซน (Nilsen)

เสนอความคิดเกี่ยวกับความสัมพันธ์ทางอรรถศาสตร์ระหว่างคำนามกับคำกริยาโดยเสนอความสัมพันธ์แบบต่างๆ ดังนี้

- 1) ผู้ถูก (object) ความสัมพันธ์ระหว่างคำนามซึ่งหมายถึงสิ่งไม่มีชีวิตที่ได้รับการกระทบกระเทือนจากกริยากับคำกริยา โดยอาจมีคำนามซึ่งหมายถึงเครื่องมือในการทำกริยาอยู่ในโครงสร้างลึก (deep case) เช่น แม่รีมอบหัวใจของเธอแก่จอห์น, จอห์นทุบหน้าต่างแตก (ด้วยค้อน)
- 2) ผู้ทำ (agent) ความสัมพันธ์ระหว่างคำนามซึ่งหมายถึงสิ่งมีชีวิตที่เป็นเหตุของกริยาหรือก่อให้เกิดกริยากับคำกริยา ความสัมพันธ์แบบผู้ทำสามารถปรากฏลำพังได้ไม่ต้องมีความสัมพันธ์ระหว่างคำนามกับคำกริยาแบบอื่นๆ ปรากฏร่วมด้วยแม้ในโครงสร้างลึก เช่น จอห์นทุบ
- 3) พลัง (force) ความสัมพันธ์ระหว่างคำนามซึ่งหมายถึงสิ่งไม่มีชีวิตแต่เป็นพลังที่เป็นเหตุให้เกิดกริยาได้ กับคำกริยา และเป็นความสัมพันธ์อีกแบบหนึ่งที่ปรากฏลำพังได้ในประโยค เช่นเดียวกับผู้ทำ เช่น ลมพัดหิว
- 4) ผู้ประสบ (experiencer) ความสัมพันธ์ระหว่างคำนามที่หมายถึงสิ่งมีชีวิตที่มีประสบการณ์อย่างใดอย่างหนึ่งกับคำกริยา เช่น แม่รีมอบหัวใจของเธอแก่จอห์น ซึ่งโดยทั่วไป ถ้าในประโยคมีความสัมพันธ์แบบผู้ประสบก็จะต้องมีความสัมพันธ์แบบผู้ถูกด้วย เช่น ความมืดทำให้จอห์นตกใจ

- 5) เครื่องมือ (instrument) ความสัมพันธ์ระหว่างคำนามซึ่งหมายถึงสิ่งไม่มีชีวิตที่เป็นเหตุให้เกิดกริยากับคำกริยา เช่น จอห์นทุบกระจกหน้าต่างด้วยก้อน
- 6) แหล่งเดิม (source) ความสัมพันธ์ระหว่างคำนามซึ่งหมายถึงสิ่งไม่มีชีวิต ซึ่งบ่งบอกแหล่งเดิมของกริยา กับคำกริยา ซึ่งจะปรากฏเฉพาะกับคำกริยาที่แสดงอาการและต้องเป็นผู้ต้องการทำ โดยจะปรากฏร่วมกับความสัมพันธ์แบบผู้ทำเสมอ เช่น จอห์นเดินทางจากซานฟรานซิสโก
- 7) จุดผ่าน (path) ความสัมพันธ์ระหว่างคำนามซึ่งหมายถึงสิ่งไม่มีชีวิต ซึ่งบ่งบอกจุดผ่านของกริยา กับคำกริยา ซึ่งจะปรากฏเฉพาะกับคำกริยาแสดงอาการและต้องการผู้ทำ โดยจะปรากฏร่วมกับความสัมพันธ์แบบผู้ทำเสมอ เช่น จอห์นเดินทางจากซานฟรานซิสโกโดยผ่านสะพานโกลเดนเกต
- 8) จุดหมาย (goal) ความสัมพันธ์ระหว่างคำนามซึ่งหมายถึงสิ่งไม่มีชีวิต ซึ่งบ่งบอกจุดหมายของกริยากับคำกริยา ซึ่งจะปรากฏเฉพาะกับคำกริยาแสดงอาการ และต้องการผู้ทำ โดยจะปรากฏร่วมกับความสัมพันธ์แบบผู้ทำเสมอ เช่น จอห์นเดินทางจากซานฟรานซิสโกสู่บอร์กลีย์โดยผ่านสะพานโกลเดนเกต
- 9) เวลา (time) ความสัมพันธ์ระหว่างคำนาม ซึ่งบ่งบอกเวลา กับคำกริยา เช่น จอห์นจะไม่ทำงานหนักหลังจากเทศกาลคริสต์มาส
- 10) สถานที่ (place หรือ location) ความสัมพันธ์ระหว่างคำนามซึ่งบ่งบอกสถานที่กับคำกริยา เช่น ฉันใช้เวลาฤดูร้อนปีที่แล้วในบลูมมิงตัน
- 11) ลักษณะอาการ (manner) ความสัมพันธ์ระหว่างคำนามกับคำกริยา โดยที่คำนามจะช่วยบ่งบอกอาการของการกระทำ เช่น เธอทำให้ความเข้าใจผิดลดลงได้ด้วยมารยาทอันงดงาม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา ⁴¹ และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.5 แนวความคิดของ อาร์ อี ลองเอเคอร์ (R.E Longacre)

เสนอความคิดเกี่ยวกับความสัมพันธ์ทางอรรถศาสตร์ระหว่างคำนามกับคำกริยา ดังนี้

- 1) ผู้ประสบ (experiencer)
 - 1.1 ผู้มีปฏิกริยาต่อสิ่งแวดล้อม เช่น ฉันร้อน
 - 1.2 ผู้มีสภาพอารมณ์ เช่น เขากำลังสับสน
 - 1.3 ผู้ที่ตกอยู่ภายใต้กระบวนการทางจิตวิทยา เช่น เขากำลังตกอยู่ในความกลัว
 - 1.4 ผู้ที่ได้รับผลจากกิจกรรมของผู้อื่น เช่น ฉันให้กำลังใจเธอ
 - 1.5 ผู้ที่ได้รับการกระทำ แต่สภาพทางร่างกายไม่เปลี่ยนแปลง หรือไม่ได้เคลื่อนที่ไป เช่น จอห์นตีบิล, จอห์นจูบภรรยาของเขา
 - 1.6 ผู้ที่ถูกสอน ในประโยคที่มีคำกริยาเกี่ยวกับการเรียนการสอน เช่น ครูกำลังสอนนักเรียน
- 2) ผู้ทรง-รับ (patient)

คำนามที่หมายถึงสิ่งไม่มีชีวิตซึ่งได้รับการบ่งบอกสภาพ หรือสถานที่ โดยภาคแสดง เช่น ฤดูหนาวอยู่ในลินซัค หรือคำนามที่หมายถึงสิ่งมีชีวิต ซึ่งเปลี่ยนสภาพหรือสถานที่ เช่น ดอนุตตกจากเก้าอี้ หรือคำนามที่หมายถึงคำนามที่มีเจ้าของ มีผู้ต้องการ หรือถูกเปลี่ยนมือไป เช่น ดิคมี่หนังสือใหม่, ทอมให้หนังสือแกedik
- 3) ผู้ทำ (agent)

คำนามที่หมายถึงสิ่งมีชีวิต หรือสิ่งไม่มีชีวิตที่แสดงอาการ เช่น มิสเตอร์สมิทสอนวิชาพีชคณิตแก่ซูซาน บางครั้งคำนามคำหนึ่งเป็นได้ทั้งผู้ทำและผู้ประสบในขณะเดียวกัน เช่น ทอมฟังเสียงนกฮูก

- 4) ผู้เสริม (range) คำนามที่ปรากฏในโครงสร้างผิว และช่วยให้ภาคแสดงสมบูรณ์ โดยแบ่งผู้เสริมออกเป็นลักษณะย่อยๆ ได้ดังนี้
- 4.1 ผู้เสริมซึ่งต้องเกิดคู่กับภาคแสดง เช่น คารุโฆร้องเพลง
- 4.2 ผู้เสริมซึ่งเป็นผลของภาคแสดง เช่น คิตกวีแต่งบทเพลง
- 5) ปริมาณ (measure) คำนามที่ทำให้ภาคแสดงสมบูรณ์ในแง่ของจำนวน หรือปริมาณ เช่น อุปกรณ์ชิ้นนี้ราคา 500 ดอลลาร์
- 6) เครื่องมือ (instrument) คำนามที่หมายถึง สิ่งมีชีวิตหรือสิ่งไม่มีชีวิตที่ถูกใช้ให้ทำให้สำเร็จลงด้วยความตั้งใจ ส่วนใหญ่จะปรากฏกับคำกริยาที่แสดงการกระทำ เช่น จอห์นตัดเชือกด้วยมีด คำกริยาบางคำจะใช้เครื่องมือเฉพาะอย่างจึงไม่จำเป็นต้องปรากฏเครื่องมือในโครงสร้างผิว เว้นแต่กรณีที่มีอะไรต่างไปจากปกติ เช่น เอ็ดเวิร์ดคตปลา (ด้วยเบ็ด), เอ็ดเวิร์ดคตปลาด้วยเบ็ดคันที่เขาทำขึ้นเอง
- 7) สถานที่ (location) คำนามซึ่งบ่งบอกสถานที่ของภาคแสดง และไม่มีนัยความหมายของการเคลื่อนที่ เช่น เรือจมในทะเล, บ้านตั้งอยู่ในสวนสาธารณะ
- 8) แหล่งเดิม (source) คำนามซึ่งบ่งสถานที่ที่บอกจุดเริ่มที่มา หรือกำเนิด หรือได้แก่คำนามที่หมายถึงสิ่งมีชีวิตผู้เป็นที่มา หรือเป็นเจ้าของในการมอบ การให้ หรือการแลกเปลี่ยน แหล่งเดิมจะปรากฏกับกริยาที่มีการเคลื่อนที่ เช่น หอมตกลูกจากแก้ว บางครั้งคำนามก็อาจมีความสัมพันธ์กับคำกริยา ทั้งแบบผู้ทำและแหล่งเดิมในขณะเดียวกัน เช่น หอมขว้างมีดไปยังกล่อง, แม่เล่านิทานให้ลูกฟัง

9) จุดหมาย (goal) คำนามที่บ่งจุดหมายของภาคแสดง หรือคำนามที่หมายถึงสิ่งที่ภาคแสดงมุ่งตรงไป โดยปราศจากการเปลี่ยนแปลงสภาพ เช่น แซมว้าย่น้ำตรงไปยังแพ ถ้าเป็นคำกริยาที่แสดงความปรารถนา ความระลึกได้ หรือการประเมินค่าแล้ว จุดหมายจะเป็นสิ่งที่ภาคแสดงบ่งถึง โดยที่วິงนั้นมิได้มีการเปลี่ยนแปลงสภาพแต่อย่างใด เช่น แมรี่รักทอม ถ้าเป็นคำกริยาที่แสดงความเป็นเจ้าของ ผู้เป็นเจ้าของคนสุดท้ายจะเป็นจุดหมาย เช่น ทอมให้หนังสือแก่บิล และคำนามก็อาจมีความสัมพันธ์กับคำกริยาทั้งแบบผู้ทำ และจุดหมายในขณะเดียวกัน เช่น จอห์นรับหนังสือจากแมรี่

10) จุดผ่าน (path) คำนามที่บ่งสถานที่ขณะเกิดกริยาที่แสดงอาการเคลื่อนไหว เช่น ทอมขว้างมีดข้ามห้องไปยังกล่อง ถ้าเป็นคำกริยาที่เกี่ยวกับการส่งผ่าน จุดผ่านจะได้แก่กรรมรองที่ปรากฏในโครงสร้างผิว เช่น สมิธขายรถเปิดประทุนให้ทอมซึ่งซื้อให้ภรรยาของเขา

4.2.6 แนวความคิดของจอห์น เอ็ม แอนเดอร์สัน (John M. Anderson)

เสนอความคิดเกี่ยวกับความสัมพันธ์ทางอรรถศาสตร์ระหว่างคำนามกับคำกริยา 4 แบบดังนี้

1) ผู้มีสภาพ (nominative) คำนามที่หมายถึงผู้มีสภาพที่แสดงโดยคำกริยา หรือคำคุณศัพท์ เช่น จอห์นตาย คำกริยาที่จะปรากฏกับผู้มีสภาพจะเป็นคำกริยาที่แสดงสภาพ (stative) เช่น เอ็ดเบอร์ดรู้ความจริง

2) ผู้เริ่ม (ergative) คำนามที่หมายถึงผู้นำมาซึ่งการกระทำซึ่งจะเป็นสิ่งมีชีวิตหรือสิ่งไม่มีชีวิตก็ได้ ซึ่งแสดงโดยกริยา เช่น จอห์นอ่านหนังสือ, ลูกบอลกลิ้งจากเนินไปยังแมรี่

3) สถานที่ (location) คำนามที่หมายถึงสถานที่ของผู้มีสภาพมักจะมีคำบุพบท ใน และ บน ปรากฏอยู่ด้วย เช่น รูปปั้นตั้งอยู่บนฐาน สถานที่ที่ยังอาจหมายถึงคำนามที่เป็นผู้รับสภาพ เช่น ละครทำความพึงพอใจแก่ผู้ชม

4) จุดเริ่ม

(ablative)

คำนามที่หมายถึงสถานที่ที่เป็นจุดเริ่มของกริยา เช่น ลูกบอลกลิ้งจาก เนินไปยังแมรี คำกริยาที่ปรากฏกับจุดเริ่มต้นจะไม่ใช่คำกริยาที่แสดงสภาพ (non-stative) ดังตัวอย่าง ซึ่งต่างจากคำกริยาที่ปรากฏกับสถานที่ที่เป็นกริยาแสดงสภาพ เช่น ลูกบอลอยู่บนพื้น จุดเริ่มนอกจากจะเป็นคำนามที่เป็นรูปธรรมดังตัวอย่าง ยังอาจเป็นนามธรรมได้ด้วย เช่น อารมณ์ของเขาเปลี่ยนจากไม่ยินดีเป็นยินดีเป็นอนุกรร

ชาร์ลส์ เจ ฟิลล์มอร์	กรูเบอร์	วอลเลซ แอล เซฟ
ผู้ทำ (agentive)	ผู้ถูกกระทบ (theme)	ผู้ทำ (agent)
เครื่องมือ (instrumental)	สถานที่ (location)	ผู้ก่อให้เกิดกริยา (potent)
ผู้ถูก (objective)	แหล่งเดิม (source)	ผู้ทรง-รับ (patient)
สถานที่ (location)	จุดหมาย (goal)	ผู้ประสบ (experiencer)
ผู้มีประสบการณ์ (experiencer)	ผู้ทำ (agent)	ผู้เสริมให้สมบูรณ์ (complement)
ผู้ร่วมกระทำ (comitative)		ผู้ได้รับประโยชน์ (beneficiary)
ผู้ได้รับประโยชน์ (benefactive)		เครื่องมือ (instrumental)
เวลา (time)		สถานที่ (location)
จุดเริ่มต้น (source)		
จุดหมาย (goal)		
นิสเซน	อาร์ ฮี ลองเอเคอร์	จอห์น เอ็ม แอนเดอร์สัน
ผู้ถูก (object)	ผู้ประสบ (experiencer)	ผู้มีสภาพ (nominative)
ผู้ทำ (agent)	ผู้ทรง-รับ (patient)	ผู้เริ่ม (ergative)
พลัง (force)	ผู้ทำ (agent)	สถานที่ (location)
ผู้ประสบ (experiencer)	ผู้เสริม (range)	จุดเริ่ม (ablative)
เครื่องมือ (instrument)	ปริมาณ (measure)	
แหล่งเดิม (source)	เครื่องมือ (instrument)	
จุดผ่าน (path)	สถานที่ (location)	
จุดหมาย (goal)	แหล่งเดิม (source)	
เวลา (time)	จุดหมาย (goal)	
สถานที่ (location)	จุดผ่าน (path)	
ลักษณะอาการ (manner)		

รูปที่ 4.1 กลุ่มความสัมพันธ์ของหน่วยคำของนักภาษาศาสตร์

4.3 การศึกษาแนวความคิดเกี่ยวกับความสัมพันธ์ทางอรรถศาสตร์ของบุคคลที่ได้กล่าวมา

เมื่อได้ทำการศึกษางานวิจัยของบุคคลต่างๆแล้ว ผู้วิจัยพบว่ามีความแตกต่างกันหลายด้าน สามารถสรุปได้ดังนี้

1. จำนวนของลักษณะความสัมพันธ์ทางอรรถศาสตร์ที่เสนอไว้นั้นไม่ตรงกัน
2. ชื่อที่ใช้เรียกลักษณะความสัมพันธ์ไม่ตรงกัน บางชื่อตรงกันแต่คำจำกัดความ หรือความหมายไม่ตรงกัน เช่น agent ของฟิลล์มอร์ หมายถึงผู้กระทำ ซึ่งต้องเป็นคำนามที่หมายถึงสิ่งมีชีวิตเท่านั้น แต่ agent ของอาร์ อี ลองเอเคอร์ หมายถึงผู้กระทำซึ่งอาจเป็นคำนามที่หมายถึงสิ่งมีชีวิต หรือไม่มีชีวิตก็ได้
3. หลักเกณฑ์ในการพิจารณาลักษณะความสัมพันธ์ระหว่างคำนามกับคำกริยาแตกต่างกัน ทำให้การตัดสินใจว่าคำนามกับคำกริยาใดมีลักษณะความสัมพันธ์อย่างไรจึงมักไม่ตรงกัน ทั้งที่เป็นลักษณะความสัมพันธ์ระหว่างคำนามกับคำกริยาคู่เดียวกัน เช่น คีตกวีแต่งเพลง อาร์ อี ลองเอเคอร์ กำหนดให้ เพลง เป็นผู้เสริมของภาคแสดง แต่ฟิลล์มอร์ และนิสเซน กำหนดให้ เพลง เป็นผู้ถูก
4. หลักเกณฑ์ในการพิจารณาระหว่างคำนามกับคำนาม หรือคำนามกับคำคุณศัพท์ ยังไม่เด่นชัด ทำให้ยากต่อการกำหนดความสัมพันธ์ทางอรรถศาสตร์ได้
5. การกำหนดความสัมพันธ์มีความซ้ำซ้อนกันทำให้ไม่สามารถตัดสินใจว่าเป็นความสัมพันธ์ประเภทไหน เช่น จากแนวความคิดของ อาร์ อี ลองเอเคอร์ กำหนดความสัมพันธ์แบบผู้ทำ กับความสัมพันธ์แบบแหล่งเดิม มีความซ้ำซ้อนกัน ในประโยค ทอมขว้างมีดไปยังกล่อง ซึ่งทอมจะมีความสัมพันธ์กับอาการขว้างในรูปของผู้กระทำก็ได้ และในขณะเดียวกัน ทอมก็สามารถที่จะแหล่งเดิมก็ได้

ในงานวิจัยนี้จึงได้รวบรวมหลักการ แนวความคิดและความสัมพันธ์ของคำจากบุคคลที่ได้กล่าวมาแล้ว และเพิ่มเติมความสัมพันธ์ที่ยังกำหนดไม่ชัดเจนและที่ยังไม่ได้กำหนดไว้ เป็นความสัมพันธ์ขึ้นมา 4 ประเภทคือ ^{(3), (6), (7), (9), (10), (11)}

- กลุ่มความสัมพันธ์หลัก (Main Relation)
- กลุ่มความสัมพันธ์คุณสมบัติ (Attribute Relation)
- กลุ่มความสัมพันธ์สถานการณ์ (Event Relation)
- กลุ่มความสัมพันธ์คำเชื่อม (Conjunction Relation)

ซึ่งทำให้การจำกัดความหมายของแต่ละความสัมพันธ์ชัดเจนยิ่งขึ้นและไม่กำกวม ความสัมพันธ์ที่กำหนดนี้ไม่ซ้ำซ้อน โดยการแบ่งออกเป็นกลุ่มต่างๆเพื่อให้สามารถใช้กับคอมพิวเตอร์ได้อย่างมีประสิทธิภาพ และสะดวกต่อการวิเคราะห์หรือสังเคราะห์ทางภาษาศาสตร์ต่อไป ดังมีรายละเอียดของนิยามของแต่ละหน่วยความสัมพันธ์ดังต่อไปนี้

4.4 กลุ่มความสัมพันธ์หลัก (Main Relation)

เป็นกลุ่มความสัมพันธ์หลักของประโยค เพื่อที่จะสามารถอธิบายความสัมพันธ์ของหน่วยความหมายระดับแก่น (Conceptual Primitive) ที่มีอยู่ในประโยค ได้วิจัยออกแบบไว้ทั้งหมด 15 หน่วยความสัมพันธ์ ดังต่อไปนี้

1) ผู้กระทำ (Agentive) : จะใช้รหัส r1001

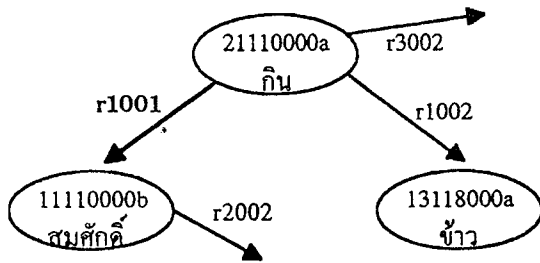
ความสัมพันธ์ระหว่างคำนามกับกริยาในลักษณะผู้ทำให้เกิดเหตุการณ์ขึ้น จะรวมทั้งสิ่งมีชีวิตและสิ่งไม่มีชีวิต แต่เป็นลักษณะทำให้เกิดเหตุการณ์นั้นขึ้น เหตุการณ์นั้นเป็นเหตุการณ์ที่ผู้ทำสามารถปรากฏลำพังได้ไม่ต้องมีความสัมพันธ์ระหว่างคำนามกับคำกริยาแบบอื่นๆ ปรากฏรวมด้วยแม้ในโครงสร้างลึก เป็นประธานของประโยคโดยความตั้งใจ การกระทำต่างๆ และเหตุการณ์ที่เกิดจากความตั้งใจทั้งทางตรง และทางอ้อมมีทั้งคน สัตว์ องค์กร หรือแม้แต่สิ่งไม่มีชีวิตก็สามารถเป็นประธานของประโยค รวมทั้งเหตุการณ์ที่เกิดขึ้นจากธรรมชาติ

ผู้กระทำแสดงอาการนั้นโดยสามารถเห็นการเปลี่ยนแปลง เช่น ประโยคภาษาไทย

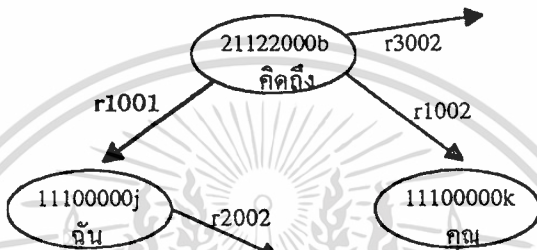
"สมศักดิ์กินข้าว"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

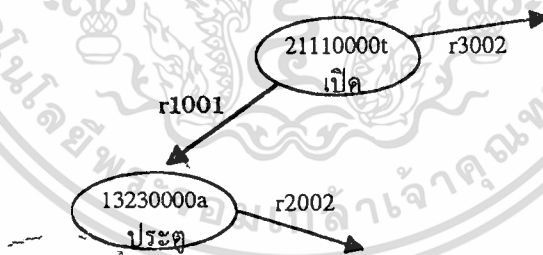
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 47 จะต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



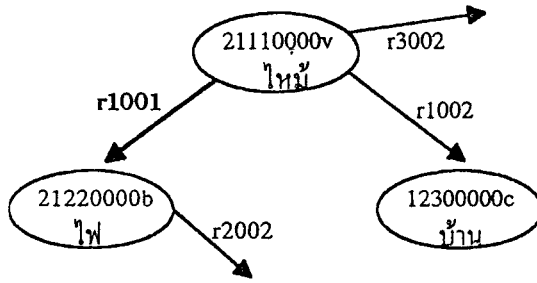
ผู้กระทำแสดงความรู้สึก ความเข้าใจ ฯลฯ เช่น ประโยคภาษาไทย "ฉันคิดถึงคุณ"



สิ่งของซึ่งสามารถแสดงอาการเคลื่อนไหว รวมทั้งการทำงานของเครื่องจักร เช่น ประโยคภาษาไทย "ประตูเปิด"



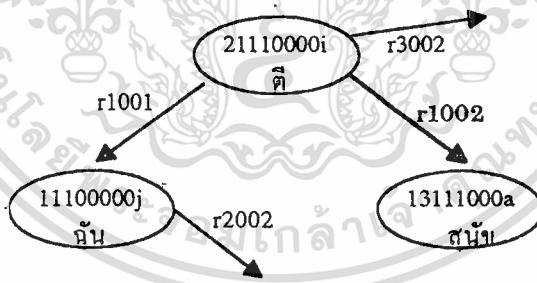
เป็นปรากฏการณ์ที่เกิดจากธรรมชาติ เช่น ประโยคภาษาไทย "ไฟไหม้บ้าน"



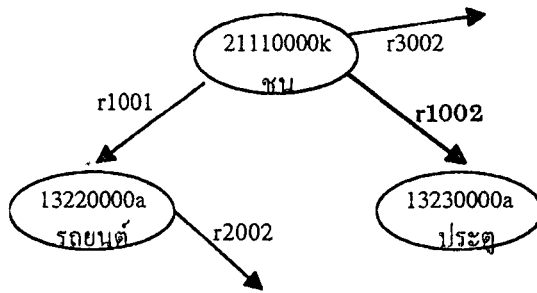
2) ผู้ถูกกระทำ (Object) : จะใช้รหัส r1002

ความสัมพันธ์ระหว่างคำนามกับคำกริยาในลักษณะผู้รับสภาพหรือได้รับผลกระทบกระเทือนจากการกระทำรวมทั้งสิ่งไม่มีชีวิตที่แสดงโดยกริยา ไม่รวมถึงเป้าหมาย จุดประสงค์ หรือข้อเสนอ ลักษณะนี้จะเป็นจุดหมายของการกระทำ มีอิทธิพลโดยตรงในท่ากรกระทำนั้น ผู้ถูกกระทำต้องเป็นสิ่งจำเป็น (กริยาที่ต้องมีกรรม) ผู้ถูกกระทำอาจเป็นผลของการกระทำ สามารถ หรือคาดหวัง การกระทำที่ทำให้เกิดความหมายก็เป็นกรรม

สิ่งมีชีวิตนั้นรับผลจากการกระทำของผู้กระทำ เช่น ประโยคภาษาไทย "ฉันทีสุนัข"



วัตถุนั้นรับผลของการกระทำของอาการนั้น เช่น ประโยคภาษาไทย "รถยนต์ชนประตู"

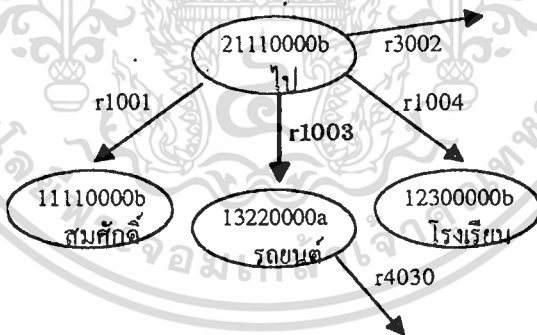


3) เครื่องมือ (Instrumental) : จะใช้รหัส r1003

ความสัมพันธ์ระหว่างคำนามกับคำกริยาในลักษณะบ่งบอกเครื่องมือที่ใช้กับกริยาอย่างตัว ทฤษฎีหรือวิธีการของการกระทำที่ทำให้เกิดเหตุการณ์ และสิ่งอื่นๆ ซึ่งทำให้การกระทำนั้นสมบูรณ์ไม่ว่าการกระทำ หรือการเปลี่ยนนั้นจะใช้วิธีการใดๆ

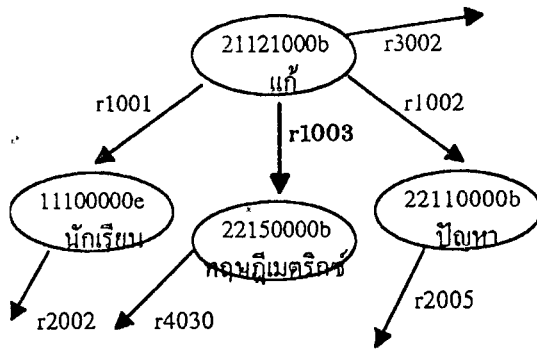
วัตถุนั้นถูกกระทำโดยวัตถุซึ่งเป็นเครื่องมือ เช่น ประโยคภาษาไทย

"สมศักดิ์ไปโรงเรียนโดยรถยนต์"



ทฤษฎี หรือวิธีการที่ทำให้เหตุการณ์เกิดขึ้น เช่น ประโยคภาษาไทย

"นักเรียนแก้ปัญหานี้ด้วยทฤษฎีเมตริกซ์"

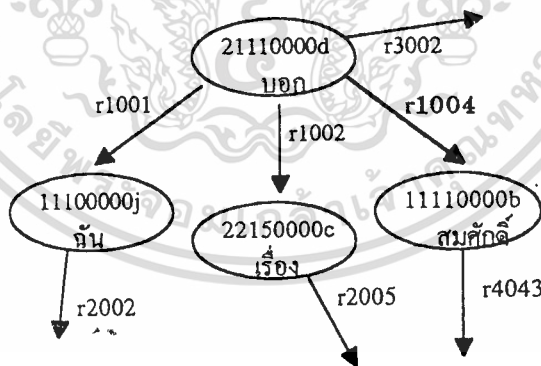


4) จุดหมาย (Target) : จะใช้รหัส r1004

ความสัมพันธ์ระหว่างคำนามกับคำกริยาในลักษณะบ่งบอกถึงจุดมุ่งหมาย จุดปลายทาง จุดสิ้นสุดของการกระทำ หรือเป้าหมายของการกระทำที่ได้เป็นกรรมแท้ของกริยา หรือสถานที่ อธิบายเกี่ยวกับการเคลื่อนที่ของสิ่งนั้นในตำแหน่งสุดท้าย

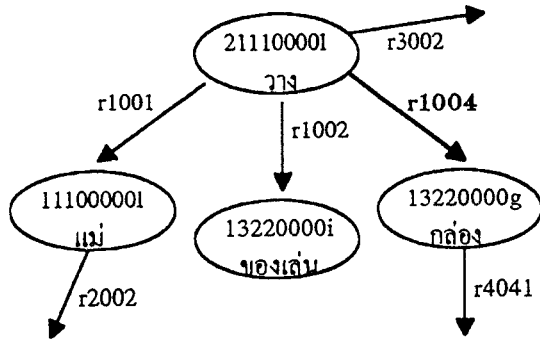
จุดมุ่งหมายหรือจุดหมายปลายทางของการกระทำ เช่น ประโยคภาษาไทย

"ฉันบอกเรื่องนี้แก่สมศักดิ์"



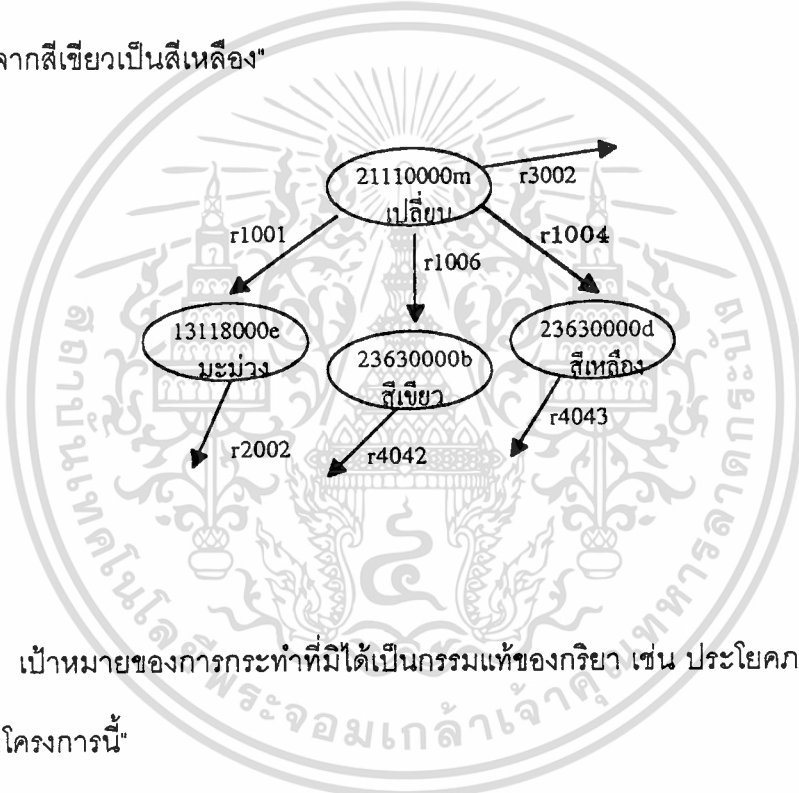
การเคลื่อนที่หรือการเปลี่ยนตำแหน่ง เช่น ประโยคภาษาไทย

"แม่วางของเล่นในกล่อง"



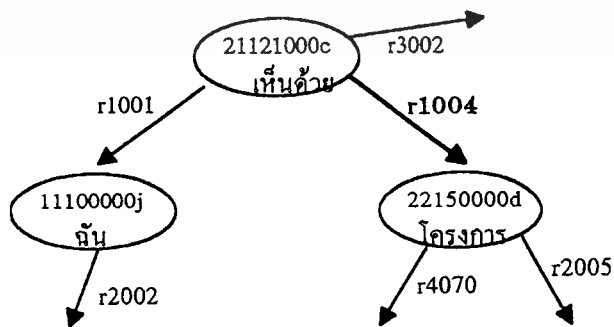
สถานะสุดท้ายของการเปลี่ยนแปลง เช่น ประโยคภาษาไทย

"มะม่วงเปลี่ยนจากสีเขียวเป็นสีเหลือง"



เป้าหมายของการกระทำที่มีได้เป็นกรรมแท้ของกริยา เช่น ประโยคภาษาไทย

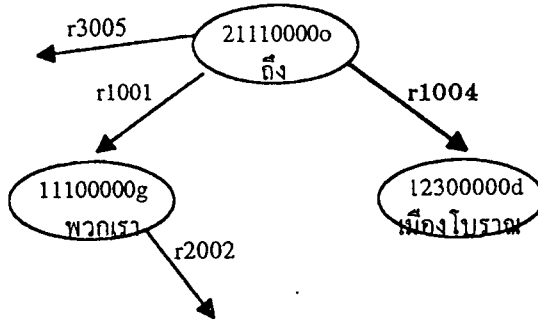
"ฉันเห็นด้วยกับโครงการนี้"



*

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา 52 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

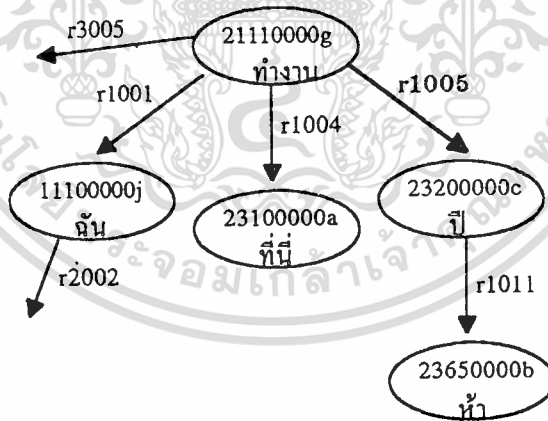
จุดหมายที่เป็นสถานที่ เช่น ประโยคภาษาไทย "พวกเราถึงเมืองโบราณแล้ว"



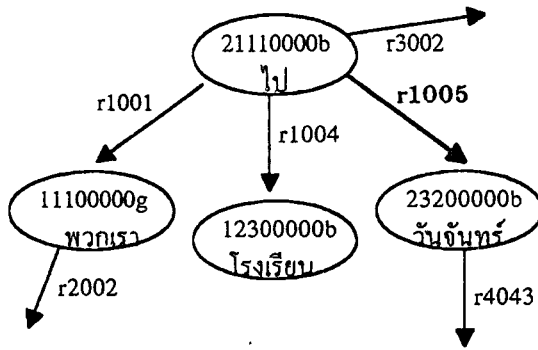
5) เวลา (Time) : จะใช้รหัส r1005

ความสัมพันธ์ระหว่างคำนามกับคำกริยาในลักษณะที่กล่าวถึงเวลาที่เกิดเหตุการณ์หรือการกระทำ
ช่วงเวลาที่ใช้ในกริยานั้น ๆ

ช่วงระยะเวลาที่ใช้ เช่น ประโยคภาษาไทย "ฉันทำงานที่นี่ 5 ปีแล้ว"



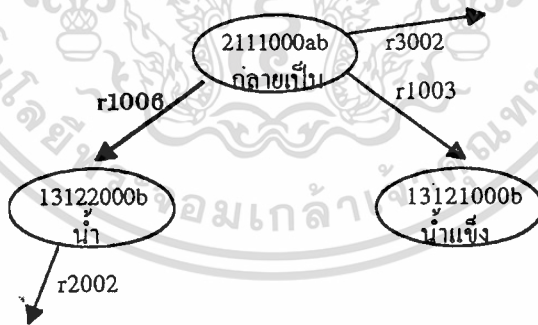
สภาพที่แสดงโดยกริยา เช่น ประโยคภาษาไทย "พวกเราไปโรงเรียนในวันจันทร์"



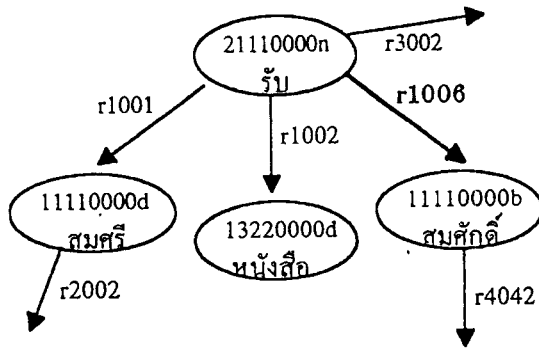
6) แหล่งเดิม (Source) : จะใช้รหัส r1006

ความสัมพันธ์ระหว่างคำนามกับคำกริยาในลักษณะบอกจุดตั้งต้นของการกระทำ ซึ่งแสดงโดยกริยาหรือที่มาของสิ่งของที่ส่งจากผู้ให้ไปยังผู้รับ หรือที่มาของเหตุการณ์ ตำแหน่งเริ่มต้นของสิ่งนั้นในเหตุการณ์ การเคลื่อนที่นี้จะเป็นการเคลื่อนที่ที่เกี่ยวกับระยะทางของสิ่งนั้น การบอกทิศทางของการเคลื่อนที่

จุดเริ่มต้นของการเปลี่ยนคุณสมบัติ เช่น ประโยคภาษาไทย "น้ำกลายเป็นน้ำแข็ง"



แหล่งที่มาของสิ่งนั้น เช่น ประโยคภาษาไทย "สมศรีรับหนังสือจากสมศักดิ์"



จุดเริ่มต้นของการเปลี่ยนตำแหน่ง เช่น ประโยคภาษาไทย

"ศักดิ์ชัยออกจากบ้านไปมหาวิทยาลัย"



7) การเชื่อมขนาน (Parallel) : จะใช้รหัส r1007

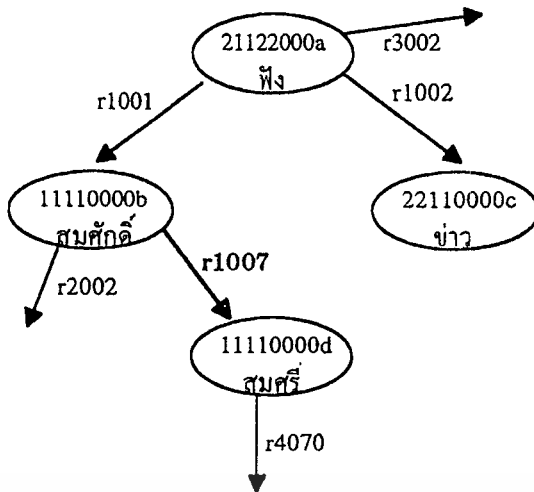
ความสัมพันธ์ระหว่างคำนามมากกว่าหนึ่ง กับคำกริยาในลักษณะที่กระทำอาการร่วมกัน หรือคำนามที่สามารถกระทำอาการมากกว่าหนึ่งอาการในเวลาเดียวกัน หรือเหตุการณ์ที่เกิดต่อเนื่องกัน หรือเหตุการณ์ในช่วงเวลาเดียวกัน หรือคำนามกับคำนามทำอาการร่วมกัน

การกระทำที่ร่วมกันของคำนามมากกว่าหนึ่ง เช่น ประโยคภาษาไทย

"สมศักดิ์และสมศรีฟังข่าว"

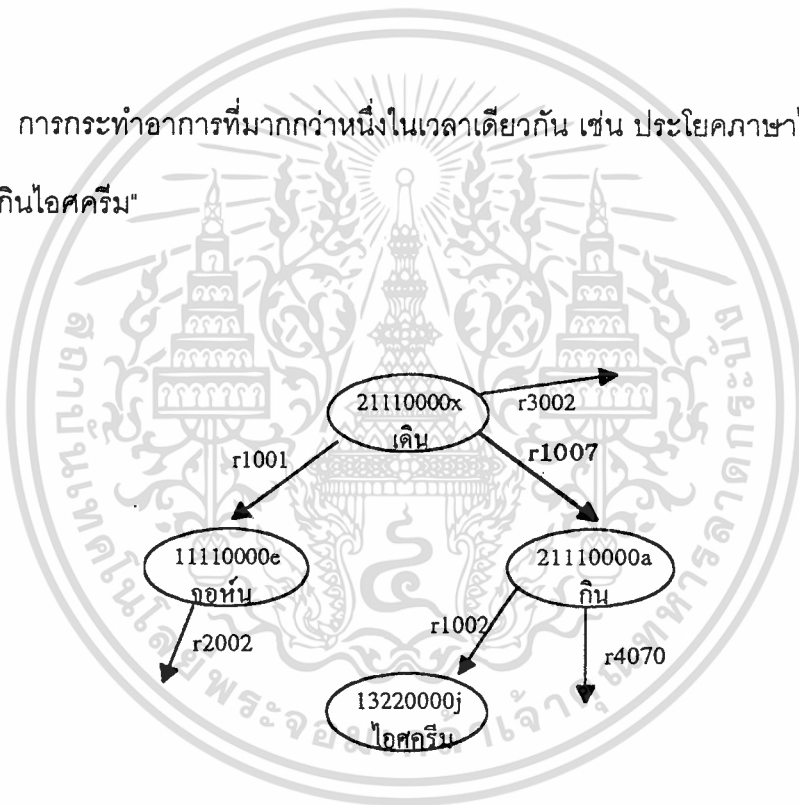
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



การกระทำอาการที่มากกว่าหนึ่งในเวลาเดียวกัน เช่น ประโยคภาษาไทย

"จอห์นเดินและกินไอศกรีม"

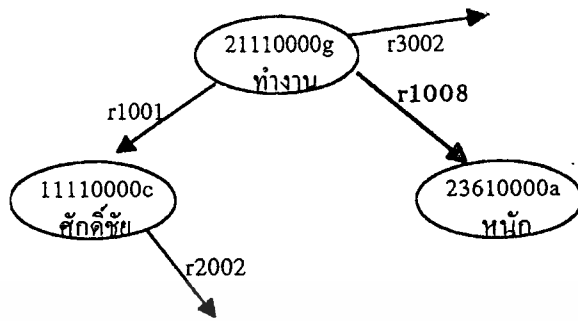


8) ส่วนขยาย (Modifier) : จะใช้รหัส r1008

ความสัมพันธ์ระหว่างคำกริยากับคำที่เป็นส่วนขยายกริยา หรือเป็นส่วนขยายคำนาม ใช้แทนส่วนขยายของกริยา คุณศัพท์ และบอกคุณสมบัติ หรือข้อจำกัดของสิ่งนั้น หรืออาการของกริยาทั้งหมด

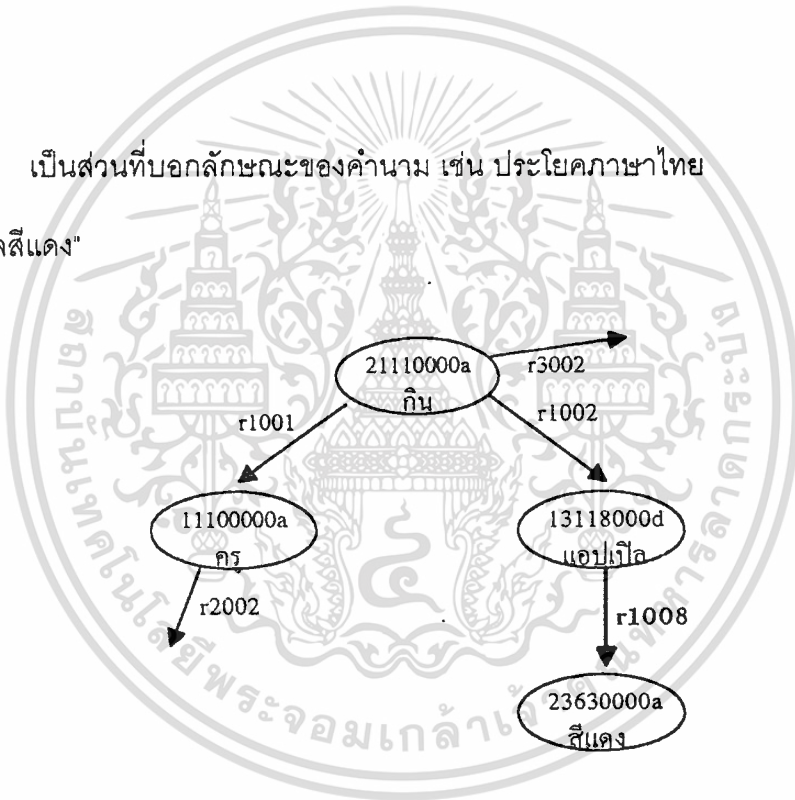
เป็นส่วนขยายกริยาโดยเป็นคุณศัพท์ที่ทำหน้าที่เหมือนกริยา เช่น ประโยคภาษาไทย

"ศักดิ์ชัยทำงานหนัก"



เป็นส่วนที่บอกลักษณะของคำนาม เช่น ประโยคภาษาไทย

"ครูกินแอปเปิลสีแดง"



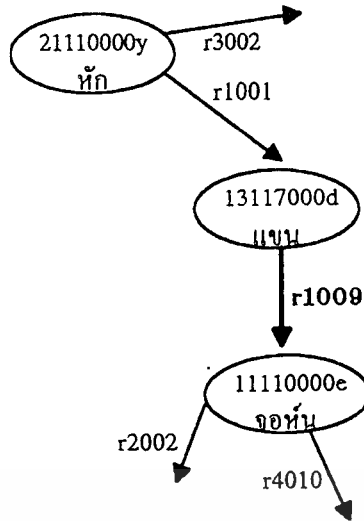
9) ส่วนประกอบ (Element) : จะใช้รหัส r1009

ความสัมพันธ์ระหว่างคำนามกับคำนามที่แสดงเป็นส่วนประกอบบางส่วน หรือหลายส่วน หรือส่วนหนึ่งของทั้งหมด หรือเป็นส่วนย่อยของสิ่งนั้น รวมถึงส่วนประกอบของอวัยวะ

เป็นส่วนหนึ่งของอวัยวะ เช่น ประโยคภาษาไทย "แขนของจอห์นหัก"

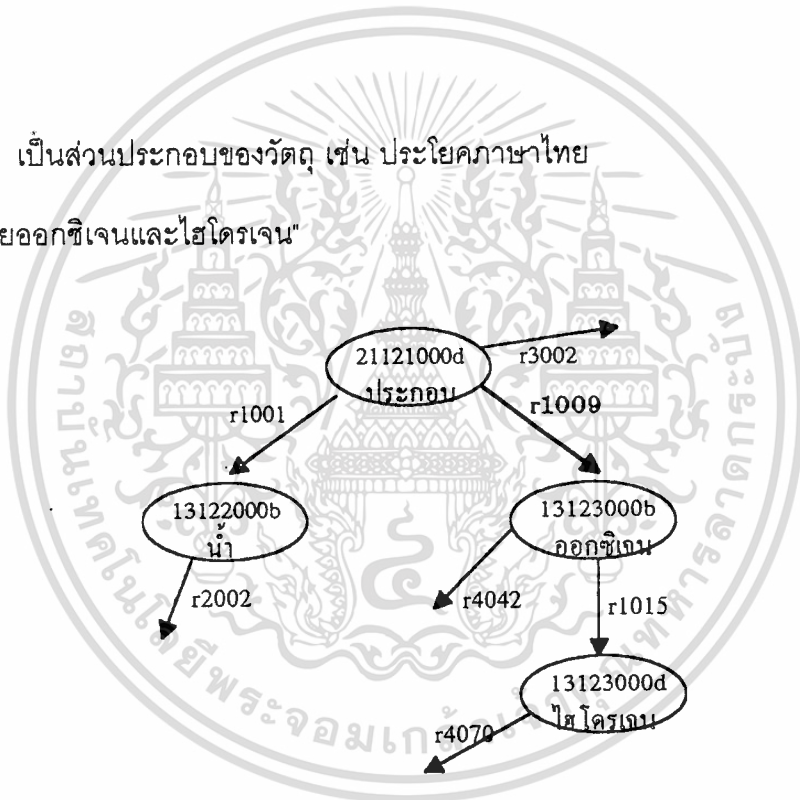
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 57 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เป็นส่วนประกอบของวัตถุ เช่น ประโยคภาษาไทย

"น้ำประกอบด้วยออกซิเจนและไฮโดรเจน"

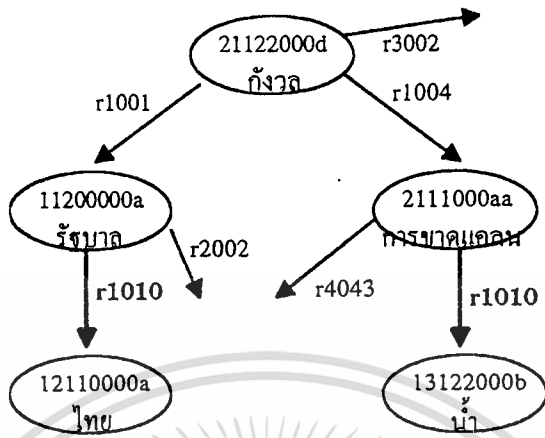


10) ส่วนขยายนาม (Noun Modifier) : จะใช้รหัส r1010

ความสัมพันธ์ระหว่างคำที่อยู่หน้าคำนามซึ่งเป็นส่วนขยายคำนามเป็นการให้ความหมายของนามทั่วไปให้ชัดเจนยิ่งขึ้น หรือ เป็นชื่อที่เรียกสิ่งนั้นเฉพาะลงไป หรือเป็นการบอกส่วนย่อยในส่วนใหญ่

เป็นส่วนขยายค่านาม เช่น ประโยคภาษาไทย

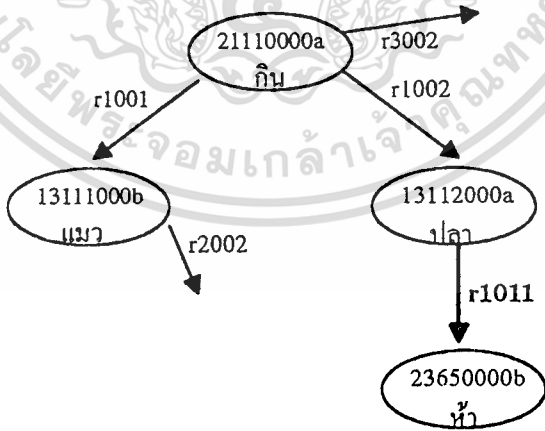
"รัฐบาลไทยกังวลเกี่ยวกับการขาดแคลนน้ำ"



11) ปริมาณ (Value) : จะใช้รหัส r1011

ความสัมพันธ์ระหว่างค่าที่บ่งบอกถึงปริมาณหรือ จำนวนของค่านามนั้น ๆ

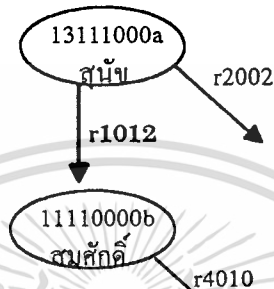
การบอกจำนวน ปริมาณของสิ่งต่างๆ เช่น ประโยคภาษาไทย "แมวกินปลา 5 ตัว"



12) เจ้าของ (Possessor) : จะใช้รหัส r1012

การแสดงความเป็นเจ้าของ อาจอยู่ในรูป "someone have something" ในลักษณะที่แสดงความเป็นเจ้าของในสิ่งนั้น

การแสดงความเป็นเจ้าของในสิ่งต่างๆ เช่น ประโยคภาษาไทย "สุนัขของสมศักดิ์"

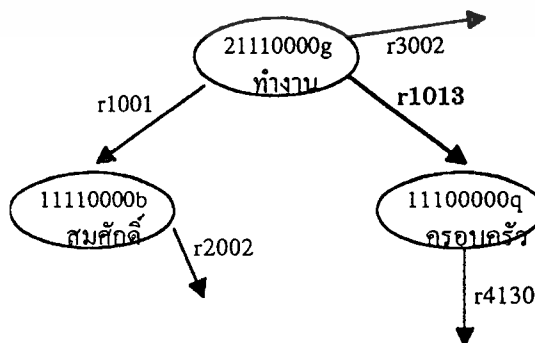


13) ผู้รับประโยชน์ (Beneficiary) : จะใช้รหัส r1013

การแสดงลักษณะของผู้ได้รับผลประโยชน์จากเหตุการณ์ หรือการกระทำนั้นซึ่งได้รับผลหรือได้เปรียบควรจะเป็นสิ่งของที่เกิดจากความตั้งใจของมนุษย์ รวมถึงสัตว์และองค์กรด้วย

บุคคลหรือองค์กรผู้ที่ได้รับผลประโยชน์ เช่น ประโยคภาษาไทย

"สมศักดิ์ทำงานเพื่อครอบครัว"

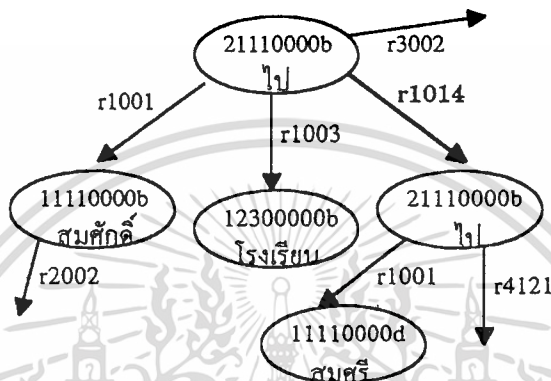


14) เหตุผล (Reason) : จะใช้รหัส r1014

เป็นเหตุผลของการกระทำรวมทั้งที่เกิดเองตามธรรมชาติ และเหตุผลที่มนุษย์สร้างหรือกำหนดขึ้น

การบอกหรือกำหนดเหตุผลของบุคคล เช่น ประโยคภาษาไทย

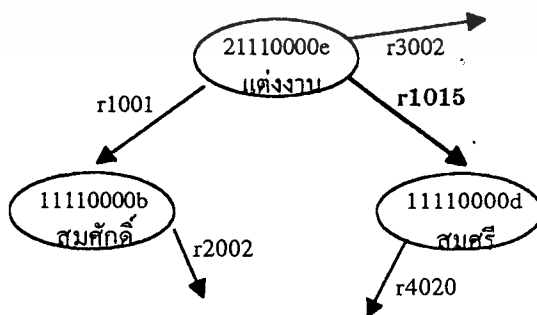
"สมศักดิ์ไปโรงเรียนเพราะสมศรีไป"



15) ผู้มีส่วนร่วม (Participant) : จะใช้รหัส r1015

เป็นการกระทำที่ต้องทำร่วมกัน ถ้าไม่ทำร่วมกันไม่สามารถเกิดเหตุการณ์ หรือการกระทำนั้น ไม่ว่าจะการกระทำนั้นจะต้องประกอบด้วยผู้ร่วมกระทำจำนวนมากหรือน้อย

การกระทำร่วมกันของสองสิ่ง เช่น ประโยคภาษาไทย "สมศักดิ์แต่งงานกับสมศรี"



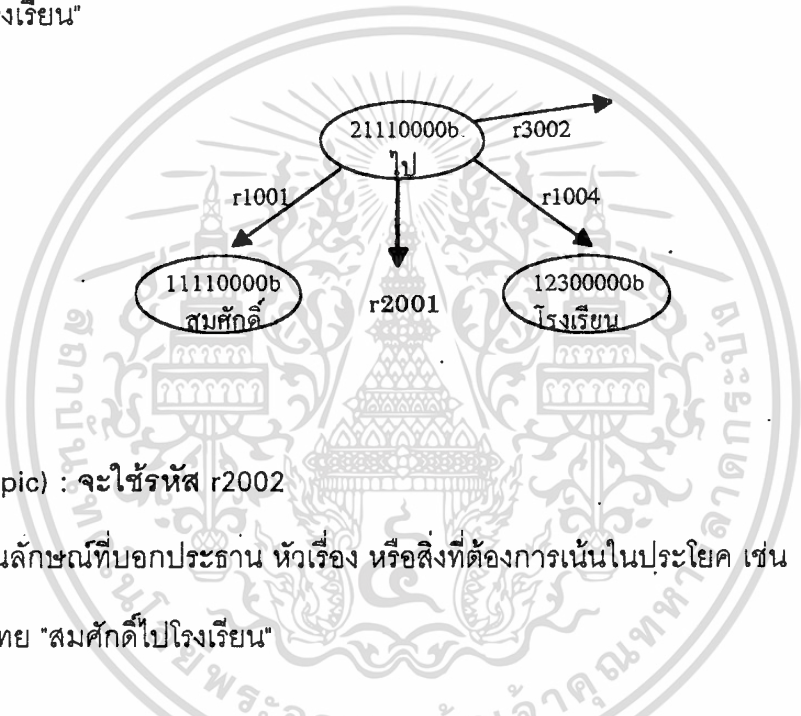
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.5 กลุ่มความสัมพันธ์คุณสมบัติ (Attributes Relation)

เป็นความสัมพันธ์ที่จะบ่งบอกคุณสมบัติของหน่วยความหมายให้ชัดเจนและเป็นการปฏิเสธของหน่วยความหมายที่เป็นคำกริยาได้วิจยออกแบบไว้ทั้งหมด 8 หน่วยความสัมพันธ์

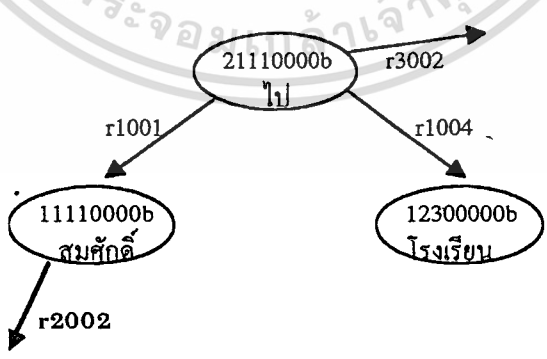
1) ปฏิเสธ (not) : จะใช้รหัส r2001

เป็นคุณลักษณะที่บอกปฏิเสธของการกระทำนั้น เช่น ประโยคภาษาไทย "สมศักดิ์ไม่ไปโรงเรียน"



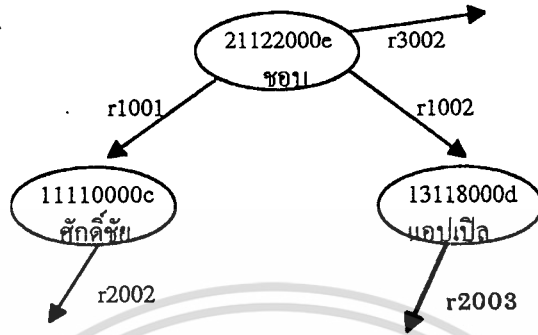
2) หัวเรื่อง (topic) : จะใช้รหัส r2002

เป็นคุณลักษณะที่บอกประธาน หัวเรื่อง หรือสิ่งที่ต้องการเน้นในประโยค เช่น ประโยคภาษาไทย "สมศักดิ์ไปโรงเรียน"



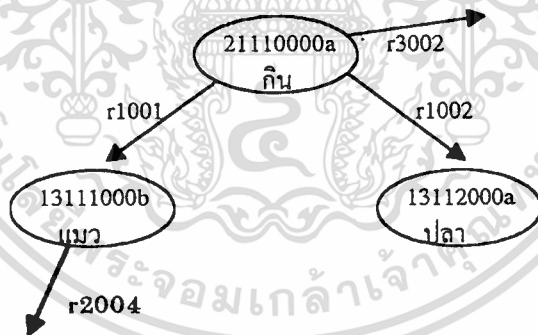
3) ความทั่วไป (generic) : จะใช้รหัส r2003

เป็นคุณลักษณะแสดงความทั่วไป ความปกติ ไม่ว่าจะเป็นการกระทำหรือสิ่งของ เช่น ประโยคภาษาไทย "ศักดิ์ชัยชอบแอปเปิล"



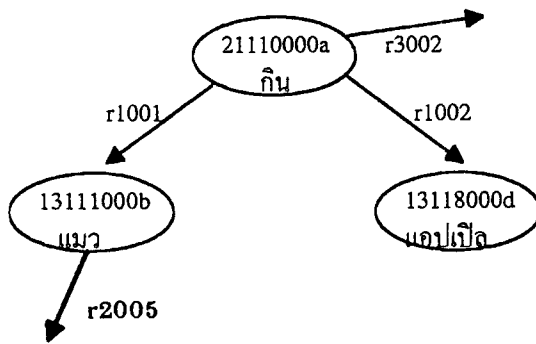
4) ทั้งหมด (all) : จะใช้รหัส r2004

เป็นคุณลักษณะบอกทุกหน่วย ทั้งหมดของกลุ่มหรือสิ่งของนั้น เช่น ประโยคภาษาไทย "แมวทุกตัวกินปลา"



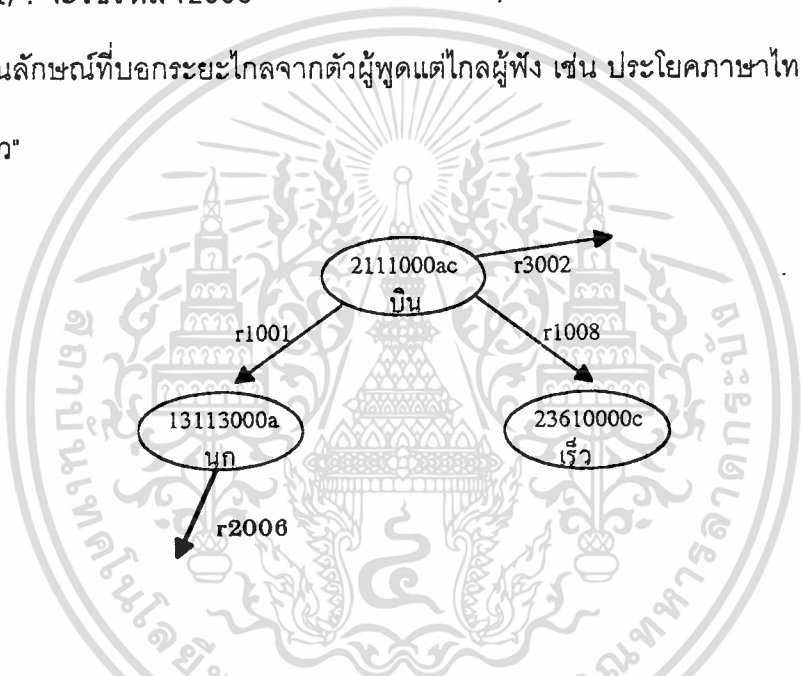
5) สิ่งนี้ (this) : จะใช้รหัส r2005

เป็นคุณลักษณะที่บอกระยะใกล้ตัวผู้พูดหรือเรื่องใกล้ตัว เช่น ประโยคภาษาไทย "แมวด่วนนี้กินแอปเปิล"



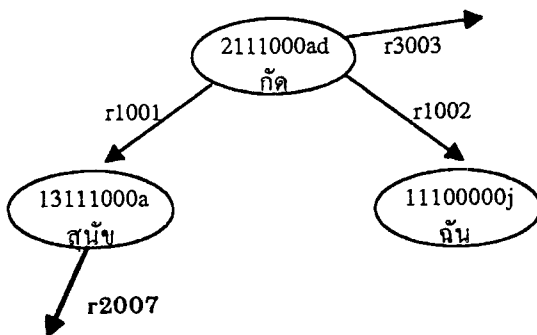
6) สิ่งนั้น (that) : จะใช้รหัส r2006

เป็นคุณลักษณะที่บอกระยะไกลจากตัวผู้พูดแต่ไกลผู้ฟัง เช่น ประโยคภาษาไทย "นกตัวนั้นบินเร็ว"



7) สิ่งโน้น (those) : จะใช้รหัส r2007

เป็นคุณลักษณะที่บอกระยะไกลจากตัวผู้พูดและผู้ฟัง เช่น ประโยคภาษาไทย "สุนัขตัวโน้นจะกัดฉัน"

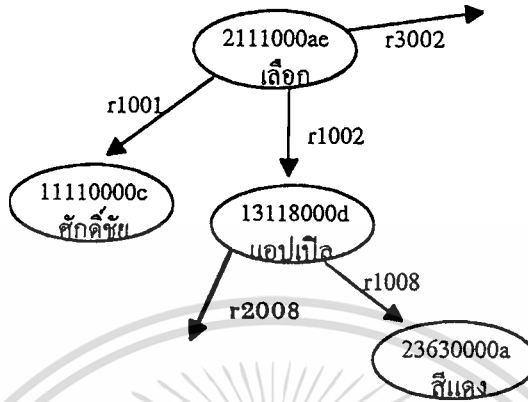


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา 64 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

8) ซีเฉพาะ (the) : จะใช้รหัส r2008

เป็นคุณลักษณะที่บอกกลุ่มคำที่ระบุหรือซีเฉพาะ เช่น ประโยคภาษาไทย

"ศักดิ์ชัยเลือกแอปเปิลสีแดง"



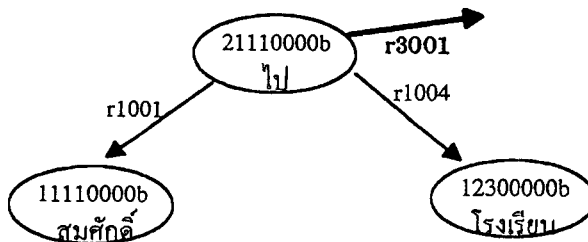
4.6 กลุ่มความสัมพันธ์สถานการณ์ (Event Relation)

เป็นกลุ่มความสัมพันธ์ที่บอกเวลาของสถานการณ์ของประโยค โดยส่วนใหญ่ความสัมพันธ์นี้จะใช้กับคำกริยาได้ใจยออกแบบไว้ทั้งหมด 8 ความสัมพันธ์

1) อดีต (past) : จะใช้รหัส r3001

เป็นสถานการณ์ที่ผ่านมาแล้ว หรือเกิดขึ้นในอดีต เช่น ประโยคภาษาไทย

"สมศักดิ์ไปโรงเรียนแล้ว"



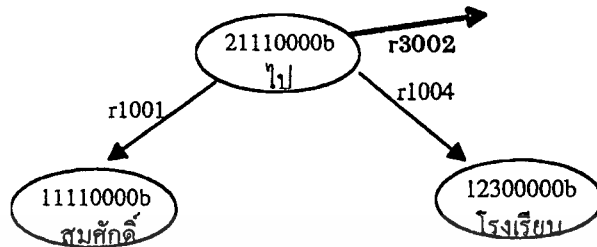
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2) ปัจจุบัน (present) : จะใช้รหัส r3002

เป็นสถานการณ์ที่ดำเนินอยู่ หรือเกิดขึ้นในปัจจุบัน เช่น ประโยคภาษาไทย

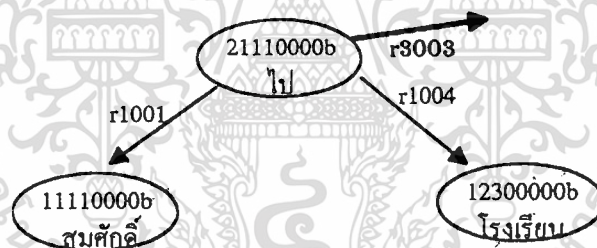
"สมศักดิ์ไปโรงเรียน"



3) อนาคต (future) : จะใช้รหัส r3003

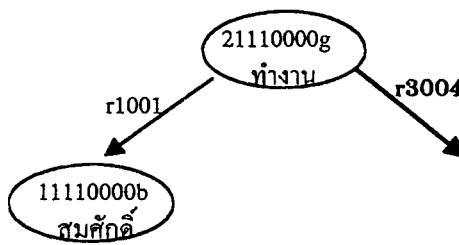
สถานการณ์ที่ยังไม่เกิดขึ้นหรือกำลังจะเกิดขึ้นในอนาคต เช่น ประโยคภาษาไทย

"สมศักดิ์จะไปโรงเรียน"



4) ยังไม่ (yet) : จะใช้รหัส r3004

เป็นจุดที่บอกว่าสิ่งนั้นยังไม่เริ่มต้นกระทำ เช่น ประโยคภาษาไทย "สมศักดิ์ยังไม่เริ่มทำงาน"

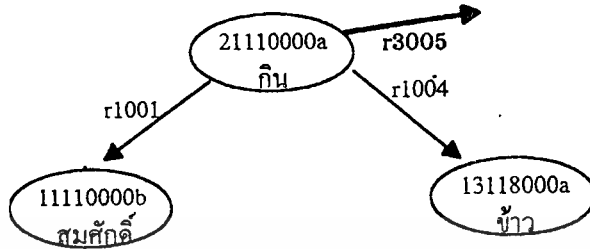


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 66 จะต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5) เสร็จแล้ว (already) : จะใช้รหัส r3005

เป็นจุดที่บอกว่าสิ่งนั้นกระทำเรียบร้อยแล้ว เสร็จสิ้นแล้ว เช่น ประโยคภาษาไทย

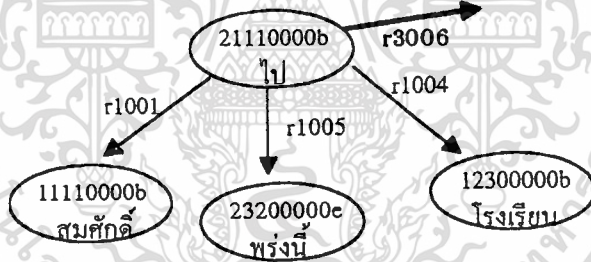
"สมศักดิ์กินข้าวเสร็จแล้ว"



6) จะเริ่ม (soon) : จะใช้รหัส r3006

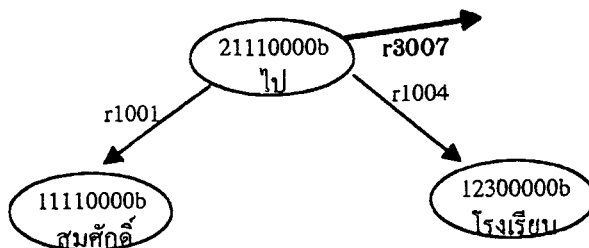
เป็นจุดที่บอกว่าสิ่งนั้นกำลังจะเริ่มต้นในช่วงระยะเวลาสั้นๆ เช่น ประโยคภาษาไทย

"สมศักดิ์จะไปโรงเรียนพรุ่งนี้"



7) เพิ่งเริ่ม (just) : จะใช้รหัส r3007

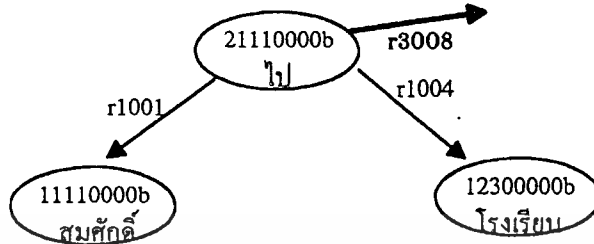
เป็นจุดที่บอกว่าสิ่งนั้นเพิ่งเริ่มต้น เช่น ประโยคภาษาไทย "สมศักดิ์เพิ่งไปโรงเรียน"



8) ดำเนิน (progress) : จะใช้รหัส r3008

เป็นจุดที่บอกว่าสิ่งนั้นกำลังดำเนินการกระทำอยู่ เช่น ประโยคภาษาไทย

"สมศักดิ์กำลังไปโรงเรียน"

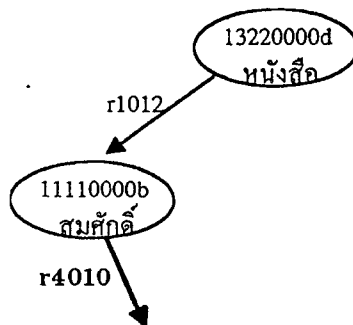


4.7 กลุ่มความสัมพันธ์คำเชื่อม (Conjunction Relation)^[6]

เป็นความสัมพันธ์ที่เสริมความสัมพันธ์ของหน่วยความหมายกับความสัมพันธ์หลัก ซึ่งหน่วยความหมายจะต้องอาศัยการแบ่งลำดับย่อย (Hierarchical Structure) ด้วยเพื่อที่จะสามารถเข้าใจความหมาย และสามารถที่จะใช้คำเชื่อมได้อย่างถูกต้อง โดยจำแนกคำเชื่อมตามความหมายได้เป็น 1-7 ประเภทดังต่อไปนี้

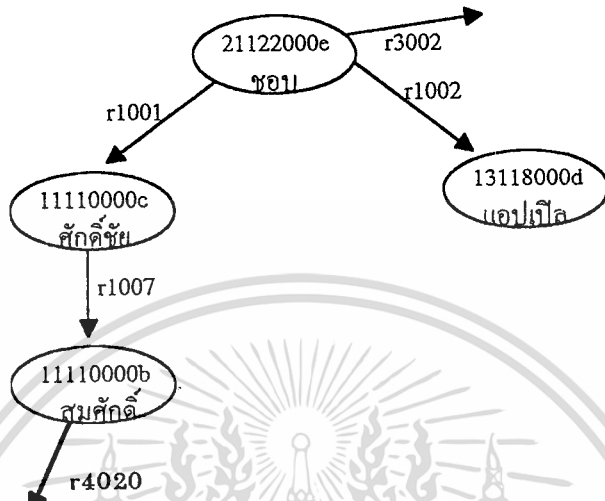
1) คำเชื่อมบอกเจ้าของ (Possessor Conjunction) : จะใช้รหัส r4010

เป็นคำเชื่อมความสัมพันธ์ของคำนามกับคำนามที่แสดงความเป็นเจ้าของ (r1012) เช่น ประโยคภาษาไทย "หนังสือของสมศักดิ์"



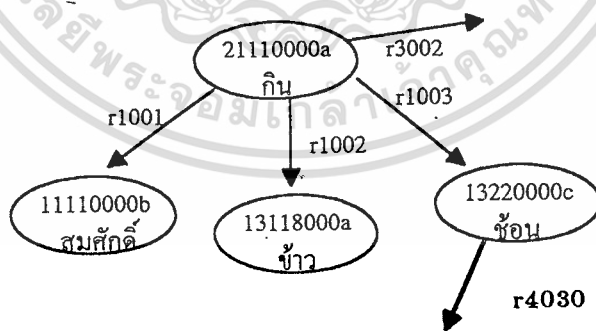
2) คำเชื่อมบอกร่วมกัน (Parallel Conjunction) : จะใช้รหัส r4020

เป็นคำเชื่อมของการกระทำร่วมกันของคำนาม หรือการกระทำอย่างต่อเนื่องกับการกระทำอื่น เช่น ประโยคภาษาไทย "ศักดิ์ชัยและสมศักดิ์ชอบแอปเปิ้ล"



3) คำเชื่อมบอกเครื่องมือ (Instrument Conjunction) : จะใช้รหัส r4030

เป็นคำเชื่อมความสัมพันธ์เกี่ยวกับเครื่องมือ หรือวิธีการของการกระทำนั้น เช่น ประโยคภาษาไทย "สมศักดิ์กินข้าวด้วยช้อน"



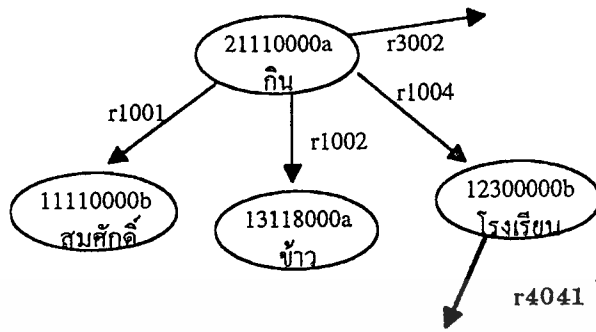
4) คำเชื่อมบอกแหล่ง (Place Conjunction) : จะใช้รหัส r4040

คำเชื่อมบอกแหล่งหรือสถานที่ของการกระทำ โดยลักษณะของคำเชื่อมนี้จะแบ่งออกเป็นหน่วยย่อยได้ 3 หน่วยย่อยได้แก่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา 69 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

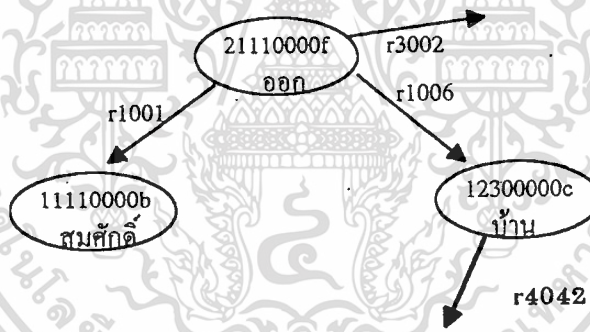
4.1) คำเชื่อมนอกแหล่งที่มีสิ่งต่างๆ จะใช้รหัส r4041

เช่น ประโยคภาษาไทย "สมศักดิ์กินข้าวภายในโรงเรียน"



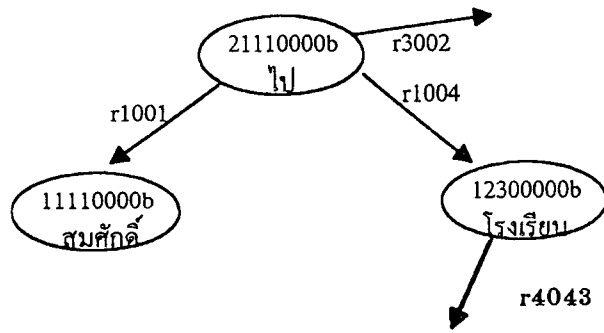
4.2) คำเชื่อมนอกแหล่งเดิมของสิ่งต่างๆ จะใช้รหัส r4042

เช่น ประโยคภาษาไทย "สมศักดิ์ออกจากบ้าน"



4.3) คำเชื่อมบอกจุดมุ่งหมายของสิ่งต่างๆ จะใช้รหัส r4043

เช่น ประโยคภาษาไทย "สมศักดิ์ไปที่โรงเรียน"

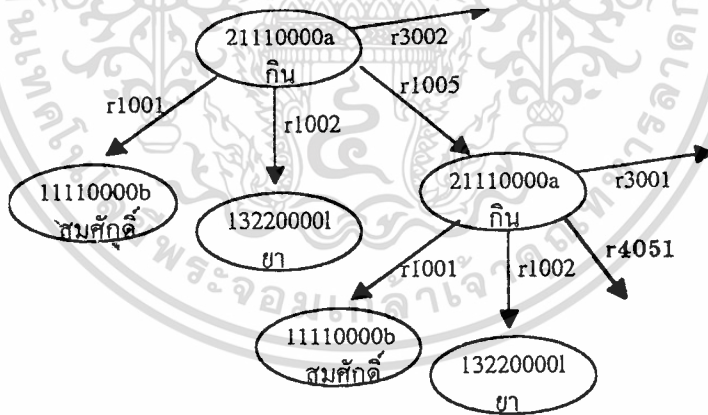


5) คำเชื่อมบอกเวลา (Time Conjunction) : จะใช้รหัส r4050

เป็นคำเชื่อมที่เกี่ยวเนื่องกับกาลเวลาที่เกิดขึ้นไม่ว่าจะเป็นอดีต ปัจจุบัน หรืออนาคต โดยจะแบ่งคำเชื่อมบอกเวลานี้ออกเป็น 5 ส่วนย่อยคือ

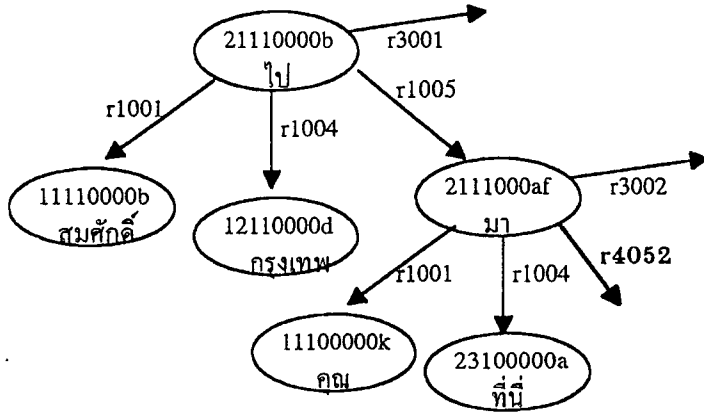
5.1) คำเชื่อมบอกเวลาก่อน จะใช้รหัส r4051

เช่น ประโยคภาษาไทย "สมศักดิ์กินยานหลังจากที่สมศักดิ์กินอาหารแล้ว"



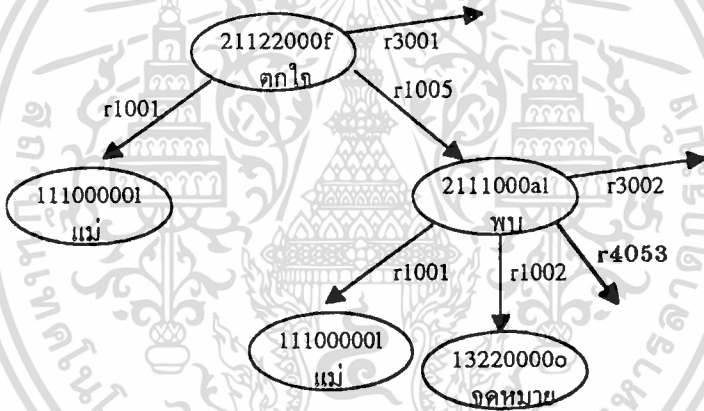
5.2) คำเชื่อมบอกเวลาภายหลัง จะใช้รหัส r4052

เช่น ประโยคภาษาไทย "สมศักดิ์ไปกรุงเทพก่อนคุณมาที่นี่"



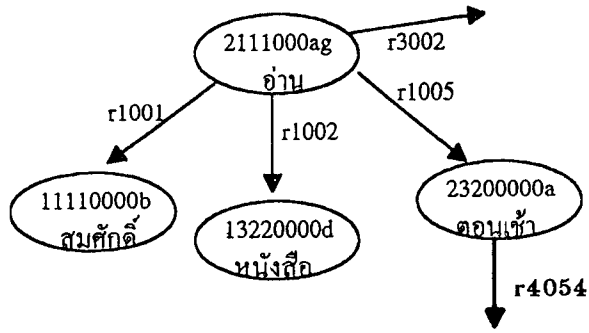
5.3) คำเชื่อมบอกเวลาเดียวกัน จะใช้รหัส r4053

เช่น ประโยคภาษาไทย "แม่ตกใจเมื่อแม่พบจดหมาย"



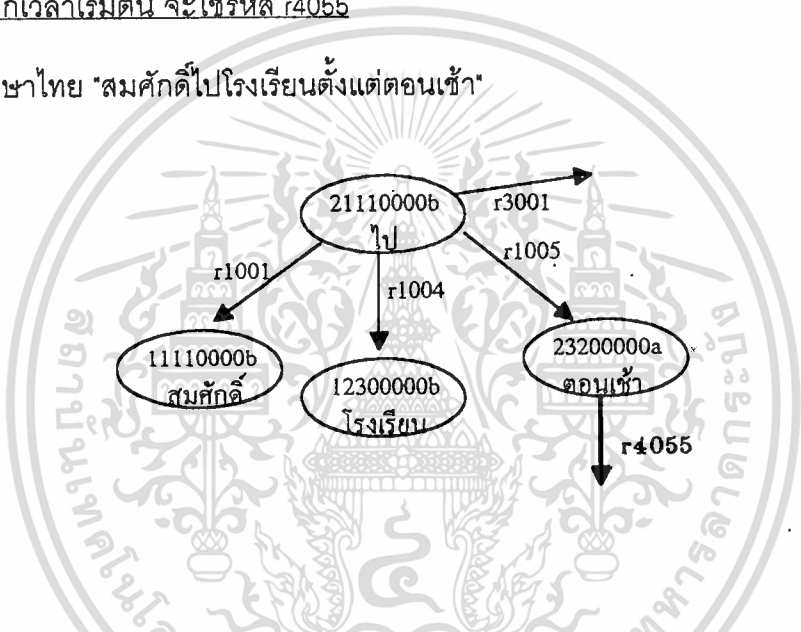
5.4) คำเชื่อมบอกเวลาสิ้นสุด จะใช้รหัส r4054

เช่น ประโยคภาษาไทย "สมศักดิ์อ่านหนังสือจนถึงตอนเช้า"



5.5) คำเชื่อมบอกเวลาเริ่มต้น จะใช้รหัส r4055

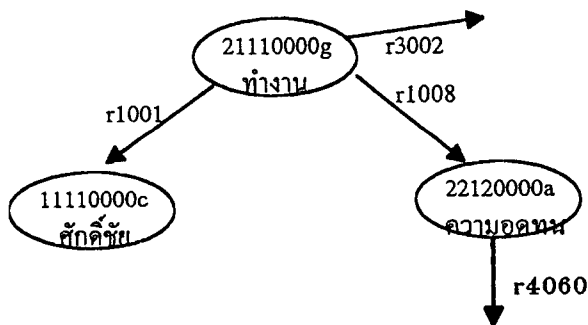
เช่น ประโยคภาษาไทย "สมศักดิ์ไปโรงเรียนตั้งแต่ตอนเช้า"



6) คำเชื่อมบอกลักษณะ (Attribute Conjunction) : จะใช้รหัส r4060

เป็นคำเชื่อมความสัมพันธ์ของส่วนขยายให้ชัดเจน โดยบอกลักษณะของส่วนที่มาเสริม

เช่น ประโยคภาษาไทย "ศักดิ์ชัยทำงานด้วยความอดทน"

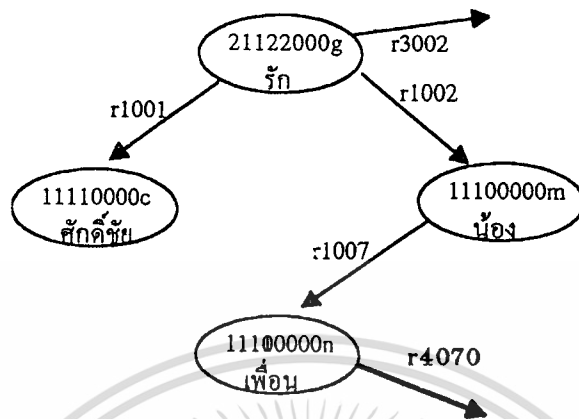


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

7) คำเชื่อมบอกความคล้ายตาม (Parallel Conjunction) : จะใช้รหัส r4070

เป็นคำเชื่อมที่บอกความคล้ายตามของคำนามหรือคำกริยาที่บอกความหมายในทำนองเดียวกัน เช่น ประโยคภาษาไทย "ศักดิ์ชัยรักน้องและเพื่อน"

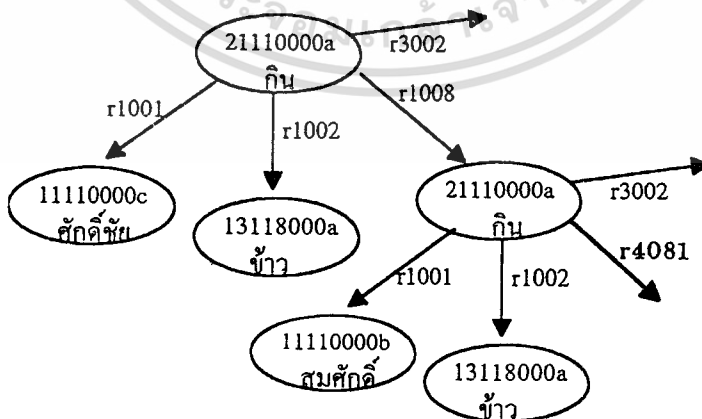


8) คำเชื่อมบอกการเปรียบเทียบ (Compare Conjunction) : จะใช้รหัส r4080

เป็นคำเชื่อมที่บอกการเปรียบเทียบของคำนามในระดับต่างๆสามารถแบ่งออกเป็น 3 หน่วยย่อยดังต่อไปนี้

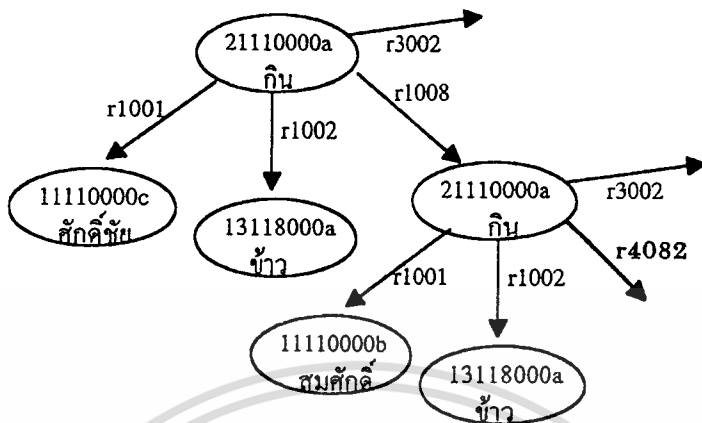
8.1) คำเชื่อมบอกการเปรียบเทียบสูงกว่า มากกว่า จะใช้รหัส r4081

เช่น ประโยคภาษาไทย "ศักดิ์ชัยกินข้าวมากกว่าสมศักดิ์กินข้าว"



8.2) คำเชื่อมบอกการเปรียบเทียบต่ำกว่า น้อยกว่า จะใช้รหัส r4082

เช่น ประโยคภาษาไทย "ศักดิ์ชัยกินข้าวน้อยกว่าสมศักดิ์กินข้าว"



8.3) คำเชื่อมบอกการเปรียบเทียบ เท่ากัน เหมือนกัน จะใช้รหัส r4083

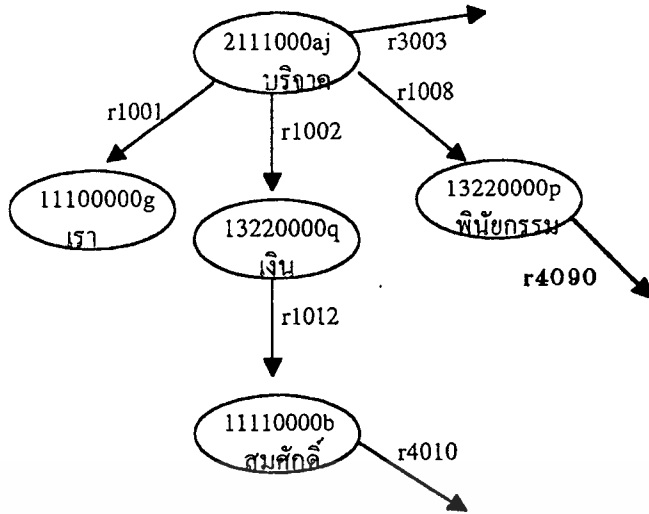
เช่น ประโยคภาษาไทย "เขาคือสมศักดิ์"



9) คำเชื่อมบอกสิ่งกำหนด (Define Conjunction) : จะใช้รหัส r4090

เป็นคำเชื่อมที่กำหนดให้กริยาเป็นเช่นนั้น

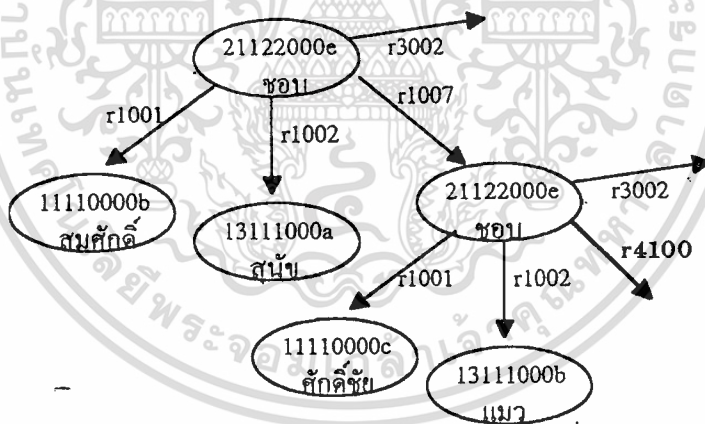
เช่น ประโยคภาษาไทย "เราจะบริจาคเงินของสมศักดิ์ตามพินัยกรรม"



10) คำเชื่อมบอกความขัดแย้ง (Contradict Conjunction) : จะใช้รหัส r4100

เป็นเชื่อมบอกความขัดแย้งของประโยคที่มีความหมายขัดแย้งกัน หรือตรงกันข้ามกัน

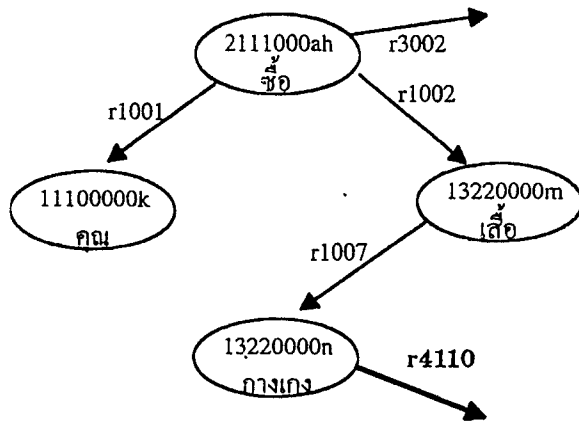
เช่น ประโยคภาษาไทย "สมศักดิ์ชอบสุนัขแต่ศักดิ์ชัยชอบแมว"



11) คำเชื่อมบอกความให้เลือกเอา (Choose Conjunction) : จะใช้รหัส r4110

เป็นคำเชื่อมที่ให้เลือกเอาอย่างใดอย่างหนึ่งไม่ว่าจะเป็นคำนามหรือคำกริยา

เช่น ประโยคภาษาไทย "คุณจะซื้อเสื้อหรือกางเกง"

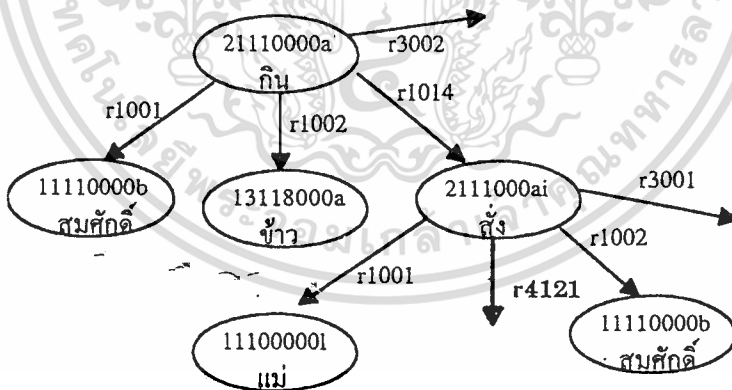


12) คำเชื่อมบอกเหตุ (Reason Conjunction) : จะใช้รหัส r4120

เป็นคำเชื่อมที่แสดงเหตุของประโยค ซึ่งคำเชื่อมในลักษณะนี้จะแบ่งออกเป็น 3 หน่วยย่อยดังต่อไปนี้

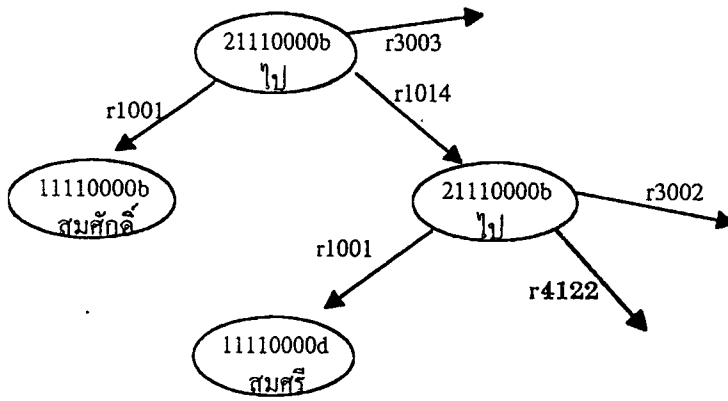
12.1) คำเชื่อมแสดงเหตุผล จะใช้รหัส r4121

เช่น ประโยคภาษาไทย "สมศักดิ์กินข้าวเพราะแม่สั่งสมศักดิ์"



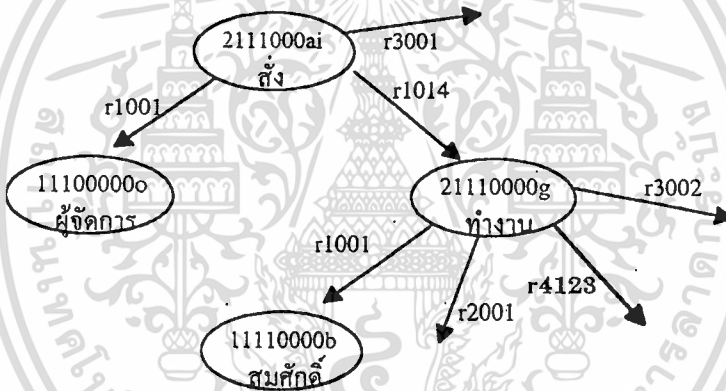
12.2) คำเชื่อมแสดงเหตุซึ่งบังเกิดผลคล้ายตาม จะใช้รหัส r4122

เช่น ประโยคภาษาไทย "สมศักดิ์จะไปก็ต่อเมื่อสมศรีไป"



12.3) คำเชื่อมแสดงเหตุซึ่งบังเกิดผลขัดแย้ง จะใช้รหัส r4123

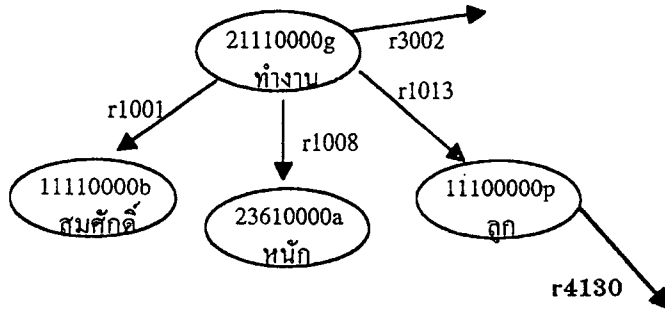
เช่น ประโยคภาษาไทย "ทั้งที่ผู้จัดการสั่งสมศักดิ์ก็ไม่ทำ"



13) คำเชื่อมบอกจุดหมาย (Goal Conjunction) : จะใช้รหัส r4130

คำเชื่อมลักษณะนี้จะเป็นการแสดงจุดหมายของการกระทำ

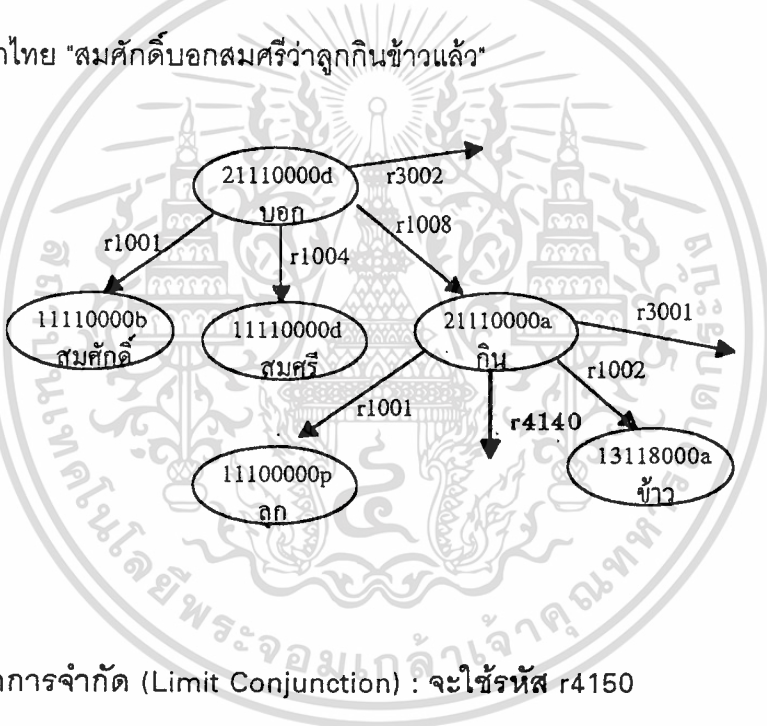
เช่น ประโยคภาษาไทย "สมศักดิ์ทำงานหนักเพื่อลูก"



14) คำเชื่อมบอกเนื้อความ (Substance Conjunction) : จะใช้รหัส r4140

คำเชื่อมลักษณะนี้จะเป็นการกำหนดเนื้อความของประโยคนั้น

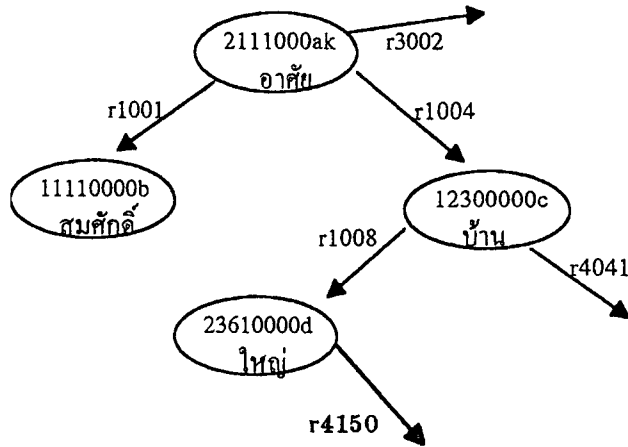
เช่น ประโยคภาษาไทย "สมศักดิ์บอกสมศรีว่าลูกกินข้าวแล้ว"



15) คำเชื่อมบอกการจำกัด (Limit Conjunction) : จะใช้รหัส r4150

คำเชื่อมนี้จะใช้จำกัดความหมายของคำนามหรือคำกริยาให้แคบลง

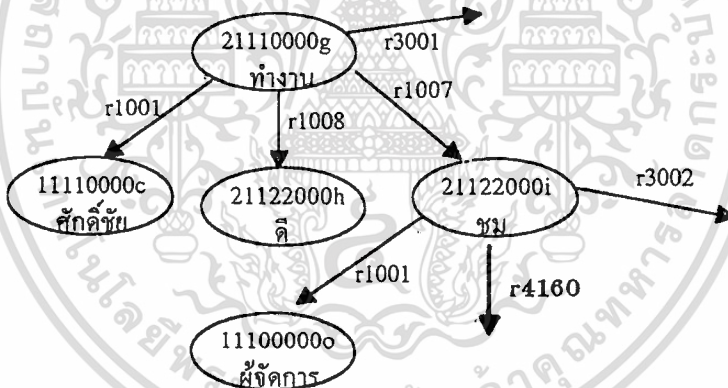
เช่น ประโยคภาษาไทย "สมศักดิ์อาศัยในบ้านที่ใหญ่"



16) คำเชื่อมบอกรวมผล (Aftereffect Conjunction) : จะใช้รหัส r4160

เป็นคำเชื่อมบอกรวมผลกระทบที่ต่อเนื่องมาจากการกระทำ

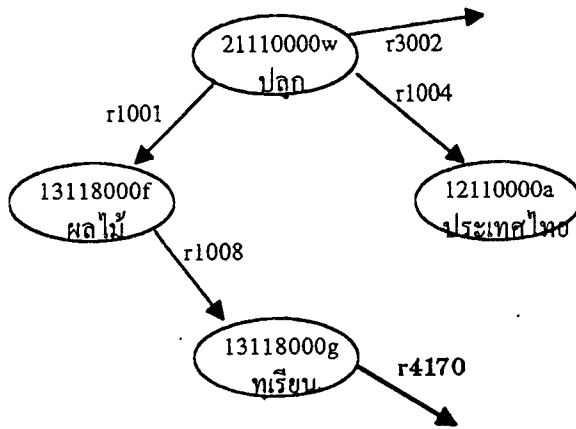
เช่น ประโยคภาษาไทย "ศักดิ์ชัยทำงานดีจนผู้จัดการชม"



17) คำเชื่อมแสดงตัวอย่าง (Example Conjunction) : จะใช้รหัส r4170

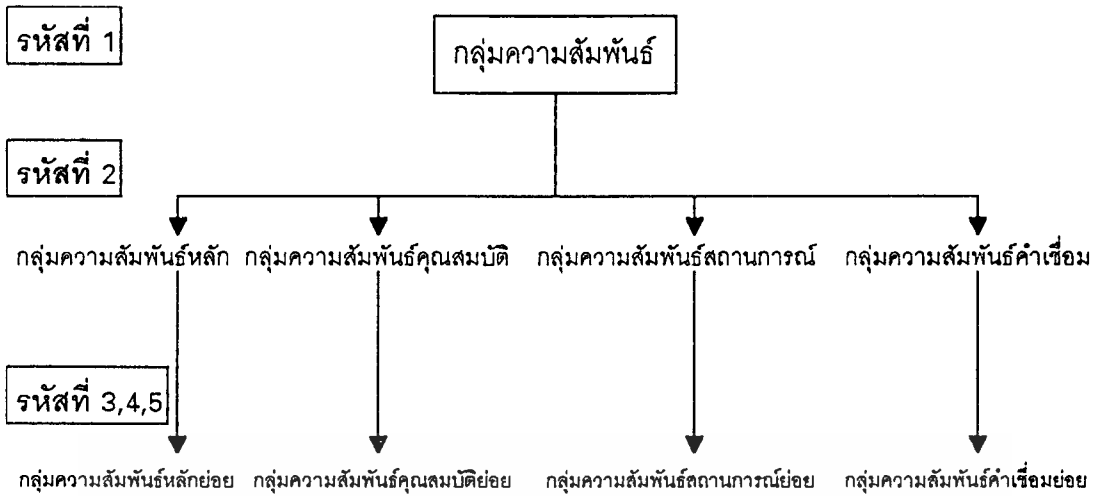
เป็นคำเชื่อมที่ใช้แสดงตัวอย่างของประโยคหรือของคำที่ได้กล่าวถึง

เช่น ประโยคภาษาไทย "ผลไม้ปลูกในประเทศไทยเช่น ทุเรียน"



4.8 การกำหนดรหัสความสัมพันธ์

การกำหนดรหัสความสัมพันธ์นี้มีลักษณะใกล้เคียงกับการกำหนดรหัสแยกแยะความหมาย โดยการกำหนดรหัสแยกแยะความหมายจะแบ่งเป็นกลุ่มประเภทของความหมายของคำ แต่การกำหนดรหัสความสัมพันธ์จะแบ่งตามลักษณะความสัมพันธ์ของกลุ่มหน่วยความหมาย และส่วนเสริมให้ประโยคสมบูรณ์ โดยจะแบ่งรหัสความสัมพันธ์เป็นความสัมพันธ์แบบลดหลั่น 5 รหัสตัวอักษร รหัสตัวแรกจะเป็นการกำหนดประเภทว่าเป็นประเภทของความสัมพันธ์จะกำหนดเป็นตัวอักษร "r" เพื่อให้ทราบว่า เป็นประเภทความสัมพันธ์ซึ่งมาจากคำว่า "Relation" ส่วนกลุ่มรหัสตัวที่สองเป็นการกำหนดกลุ่ม ความสัมพันธ์โดยจะแบ่งออกเป็น 4 กลุ่มความสัมพันธ์คือ กลุ่มความสัมพันธ์หลักเป็นเลข "1" กลุ่มความสัมพันธ์คุณสมบัติเป็นเลข "2" กลุ่มความสัมพันธ์สถานการณ์เป็นเลข "3" และกลุ่มความสัมพันธ์คำเชื่อมเป็นเลข "4" โดยที่กลุ่มความสัมพันธ์หลักเป็นองค์ประกอบหลักของภาษากลางในวิทยานิพนธ์นี้ และจะมีกลุ่มของความสัมพันธ์คุณสมบัติเป็นความสัมพันธ์ที่กำหนดข้อมูลเฉพาะของหน่วยความหมายที่มีความสัมพันธ์กับกลุ่มความสัมพันธ์หลัก และกลุ่มความสัมพันธ์สถานการณ์จะเป็นกลุ่มความสัมพันธ์ที่กำหนดกาลของหน่วยความหมายที่เป็นการกระทำ และกลุ่มความสัมพันธ์คำเชื่อมจะเป็นกลุ่มคำที่เสริมให้ประโยคเป้าหมายสมบูรณ์ยิ่งขึ้น และกลุ่มความสัมพันธ์คำเชื่อมนี้จะแปรผันตามความสัมพันธ์หลักและกลุ่มความสัมพันธ์นี้จะแบ่งย่อยเป็นรหัสที่ 3 รหัสที่ 4 และรหัสที่ 5 ลงไปอีก เพื่อให้การกำหนดข้อมูลของประโยคสมบูรณ์ซึ่งทั้งหมดนี้เป็นส่วนเสริมความหมายของหน่วยความหมายในประโยคและทำให้ภาษากลางมีข้อมูลเพียงพอที่จะสามารถนำไปสังเคราะห์เป็นภาษาเป้าหมายได้ การกำหนดรหัสที่กล่าวมานี้สามารถแสดงดังรูปที่ 4.2



รูปที่ 4.2 การกำหนดรหัสความสัมพันธ์

จากที่ได้กล่าวมาทั้งหมดสามารถสรุปกลุ่มความสัมพันธ์หลักในตารางที่ 4.1 กลุ่มความสัมพันธ์คุณสมบัติในตารางที่ 4.2 กลุ่มความสัมพันธ์สถานการณ์ในตารางที่ 4.3 และกลุ่มความสัมพันธ์ค่าเชื่อมในตารางที่ 4.4 ดังต่อไปนี้

ความสัมพันธ์หลัก	รหัสของความสัมพันธ์
ผู้กระทำ (Agentive)	r1001
ผู้ถูกกระทำ (Object)	r1002
เครื่องมือ (Instrumental)	r1003
จุดหมาย (Target)	r1004
เวลา (Time)	r1005
แหล่งเดิม (Source)	r1006
การเชื่อมขนาน (Parallel)	r1007
ส่วนขยาย (Modifier)	r1008
ส่วนประกอบ (Element)	r1009
ส่วนขยายนาม (Noun Modifier)	r1010
ปริมาณ (Value)	r1011
เจ้าของ (Possessor)	r1012
ผู้รับประโยชน์ (Beneficiary)	r1013
เหตุผล (Reason)	r1014
ผู้มีส่วนร่วม (Participant)	r1015

ตารางที่ 4.1 กลุ่มความสัมพันธ์หลัก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ความสัมพันธ์คุณสมบัตื	รหัสของความสัมพันธ์
ปฏิเสธ (not)	r2001
หัวเรื่อง (topic)	r2002
ความทั่วไป (generic)	r2003
ทั้งหมด (all)	r2004
สิ่งนี้ (this)	r2005
สิ่งนั้น (that)	r2006
สิ่งโน้น (those)	r2007
ชี้เฉพาะ (the)	r2008

ตารางที่4.2 กลุ่มความสัมพันธ์คุณสมบัตื

ความสัมพันธ์สถานการณ์	รหัสของความสัมพันธ์
อดีต (past)	r3001
ปัจจุบัน (present)	r3002
อนาคต (future)	r3003
ยังไม่ (yet)	r3004
เสร็จแล้ว (already)	r3005
จะเริ่ม (soon)	r3006
เพิ่งเริ่ม (just)	r3007
ดำเนิน (progress)	r3008

ตารางที่4.3 กลุ่มความสัมพันธ์สถานการณ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ความสัมพันธ์คำเชื่อม		รหัสของความสัมพันธ์
คำเชื่อมบอกเจ้าของ (Possessor Conjunction)		r4010
คำเชื่อมบอกส่วนร่วม (Parallel Conjunction)		r4020
คำเชื่อมบอกเครื่องมือ (Instrument Conjunction)		r4030
คำเชื่อมบอกแหล่ง (Place Conjunction)		r4040
	คำเชื่อมบอกแหล่งที่มีสิ่งต่างๆ	r4041
	คำเชื่อมบอกแหล่งเดิมของสิ่งต่างๆ	r4042
	คำเชื่อมบอกจุดมุ่งหมายของสิ่งต่างๆ	r4043
คำเชื่อมบอกเวลา (Time Conjunction)		r4050
	คำเชื่อมบอกเวลาก่อน	r4051
	คำเชื่อมบอกเวลาภายหลัง	r4052
	คำเชื่อมบอกเวลาเดียวกัน	r4053
	คำเชื่อมบอกเวลาสิ้นสุด	r4054
	คำเชื่อมบอกเวลาเริ่มต้น	r4055
คำเชื่อมบอกลักษณะ (Attribute Conjunction)		r4060
คำเชื่อมบอกความคล้ายตาม (Parallel Conjunction)		r4070
คำเชื่อมบอกการเปรียบเทียบ (Compare Conjunction)		r4080
	คำเชื่อมบอกการเปรียบเทียบสูงกว่า มากกว่า	r4081
	คำเชื่อมบอกการเปรียบเทียบต่ำกว่า น้อยกว่า	r4082
	คำเชื่อมบอกการเปรียบเทียบเท่ากัน เหมือนกัน	r4083
คำเชื่อมบอกสิ่งกำหนด (Define Conjunction)		r4090
คำเชื่อมบอกความขัดแย้ง (Contradict Conjunction)		r4100
คำเชื่อมบอกความให้เลือกเอา (Choose Conjunction)		r4110
คำเชื่อมบอกเหตุ (Reason Conjunction)		r4120
	คำเชื่อมบอกเหตุผล	r4121
	คำเชื่อมบอกเหตุซึ่งบังเกิดผลคล้ายตาม	r4122

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 85 และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คำเชื่อมบอกเหตุซึ่งบังเกิดผลขัดแย้ง	r4123
คำเชื่อมบอกจุดหมาย (Goal Conjunction)	r4130
คำเชื่อมบอกเนื้อความ (Substance Conjunction)	r4140
คำเชื่อมบอกการจำกัด (Limit Conjunction)	r4150
คำเชื่อมบอกผล (Aftereffect Conjunction)	r4160
คำเชื่อมบอกตัวอย่าง (Example Conjunction)	r4170

ตารางที่ 4.4 กลุ่มความสัมพันธ์คำเชื่อม

4.9 สรุป

ในบทนี้เป็นการกำหนดความสัมพันธ์ของหน่วยความหมายที่กล่าวมาแล้วในบทที่ 3 โดยกล่าวถึงความเป็นมาของการวิจัยเกี่ยวกับความสัมพันธ์ของหน่วยคำจากนักภาษาศาสตร์ เพื่อที่ใช้เป็นแนวทางในการกำหนดความสัมพันธ์ของหน่วยคำเพื่อการแปลภาษาด้วยคอมพิวเตอร์ให้กับวิทยานิพนธ์นี้ โดยสามารถสรุปกลุ่มความสัมพันธ์ออกมาได้ 4 กลุ่มคือ

- กลุ่มความสัมพันธ์หลัก ซึ่งกลุ่มความสัมพันธ์นี้เป็นกลุ่มความสัมพันธ์ที่ต้องเกิดกับหน่วยคำเสมอสามารถแบ่งได้ทั้งหมด 15 ความสัมพันธ์
- กลุ่มความสัมพันธ์คุณสมบัติ เป็นกลุ่มความสัมพันธ์เสริมหน่วยคำให้ชัดเจนโดยแบ่งความสัมพันธ์ของกลุ่มนี้ได้ทั้งหมด 8 ความสัมพันธ์
- กลุ่มความสัมพันธ์สถานการณ์ เป็นกลุ่มความสัมพันธ์ที่กาลเวลาของหน่วยคำกริยาโดยแบ่งความสัมพันธ์ของกลุ่มนี้ได้ 8 ความสัมพันธ์
- กลุ่มความสัมพันธ์คำเชื่อม เป็นกลุ่มความสัมพันธ์ที่เสริมหน่วยคำกับหน่วยความสัมพันธ์หลัก ให้ได้ความหมายที่สมบูรณ์ โดยแบ่งกลุ่มความสัมพันธ์นี้ออกเป็น 17 ความสัมพันธ์

บทที่ 5

พจนานุกรมเพื่อการแปลหลายภาษาด้วยคอมพิวเตอร์

5.1 บทนำ

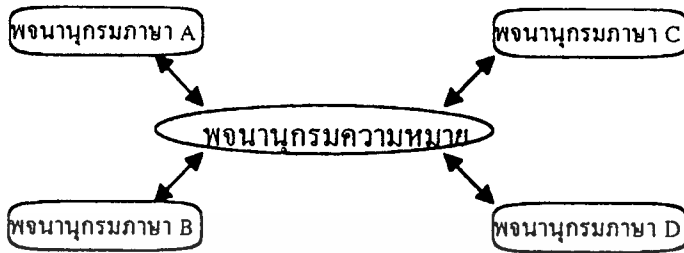
การแปลภาษาธรรมชาติด้วยคอมพิวเตอร์นั้น พจนานุกรมเป็นสิ่งจำเป็นพื้นฐานที่สำคัญของการแปลไม่ว่าจะเป็นการแปลโดยตรง การแปลโดยวิธีการถ่ายถอด หรือการแปลโดยใช้ภาษากลาง จะต้องมีฐานความรู้และข้อมูลซึ่งเรียกว่า พจนานุกรม เพื่อใช้ในการประมวลผลการแปลทั้งนั้น ซึ่งข้อมูลและฐานความรู้จะต้องครอบคลุมเนื้อหาที่สามารถจะทำการแปลได้ ประกอบไปด้วยข้อมูลต่างๆ ของหน่วยคำ (Morphological Information) ข้อมูลทางด้านหน้าที่ทางไวยากรณ์ของคำ (Syntactic Information) และรายละเอียดทางด้านความหมาย (Semantic Information) ข้อมูลต่างๆ เหล่านี้จะอยู่ในรูปโค้ด (code) และสัญลักษณ์ (symbol) ต่างๆ แต่การที่จะกำหนดขอบเขตของความหมายของคำทั้งหมดในพจนานุกรมเป็นไปได้ยาก เพราะฉะนั้นการเก็บข้อมูลในพจนานุกรมที่ใช้สำหรับการแปลภาษาด้วยคอมพิวเตอร์นั้นก็จะต้องเก็บตามความจำเป็น และข้อมูลที่จัดเก็บจะมีขอบเขตของความหมายที่แน่นอนและชัดเจนเพื่อลดความกำกวม

วิทยานิพนธ์นี้ได้ทำการวิจัยออกแบบโครงสร้างของพจนานุกรมคำศัพท์ให้สอดคล้องและประยุกต์ใช้ร่วมกับระบบภาษากลาง หรือระบบการแปลหลายภาษาด้วยคอมพิวเตอร์โดยผ่านสื่อภาษากลางของงานวิจัยนี้ โดยเริ่มแรกจะจัดแบ่งความหมายของคำที่มีความหมายระดับแก่นเดียวกันเข้าไปในกลุ่มความหมายเดียวกัน เพื่อลดความซ้ำซ้อนและความกำกวมที่อาจจะเกิดขึ้นดังที่ได้กล่าวมา และเพื่อให้สอดคล้องกับโครงสร้างภาษากลางดังที่ได้กล่าวมาแล้วเช่นกัน โดยกลุ่มความหมายของคำแต่ละกลุ่มจะมีหน่วยความหมาย (CP) เฉพาะของตนเองซึ่งจะเป็นตัวแทนของคำให้กับกลุ่มความหมายนั้น

นอกจากนั้นพจนานุกรมเพื่อการแปลหลายภาษาด้วยคอมพิวเตอร์ที่ทำการวิจัยนี้เป็นการเก็บฐานข้อมูลแบบรีเลชันแนล (Relational Database) มีลักษณะเป็นพจนานุกรมรวมภาษา (Multilingual Dictionary) ดังแสดงในรูปที่ 5.1 โดยจะแบ่งออกเป็น 2 ส่วนคือ

พจนานุกรมความหมาย (Conceptual Dictionary)

พจนานุกรมภาษา (Language Dictionary)



รูปที่ 5.1 พจนานุกรมรวมภาษา

5.2 พจนานุกรมความหมาย (Conceptual Dictionary)

พจนานุกรมความหมายเป็นพจนานุกรมที่บ่งบอกถึงความหมายของ CPID หรือคำนั้น เพื่อให้ทราบความหมายที่แท้จริงโดยความหมายนั้นจะเป็นความหมายเดียวต่อหนึ่ง CPID หรือคำ เพื่อให้การแปลภาษาด้วยคอมพิวเตอร์มีประสิทธิภาพมากขึ้น ลดความกำกวมของคำในประโยค เช่น

The coach lost a set.^[4]

จะเห็นว่าเมื่อพิจารณาจากฐานข้อมูลพจนานุกรมภาษาอังกฤษ^[12]

coach เป็นได้ทั้งคำนามและคำกริยาที่มีความหมายที่ได้จากพจนานุกรมดังตารางที่ 5.1

lost เป็นคุณศัพท์ที่มีความหมายที่ได้จากพจนานุกรมดังตารางที่ 5.2

set เป็นได้ทั้งนาม กริยา และคุณศัพท์ที่มีความหมายที่ได้จากพจนานุกรมดังตารางที่ 5.3

SURFACE	CATEGORY	CONCEPTUAL MEANING
coach ^[1]	noun	1) a large enclosed 4-wheeled horse-drawn carriage, used esp. in former times or in official ceremonies
		2) a bus used for long-distance travel or touring: We went by coach
		3) a railway passenger carriage, esp. for day travel:
		4) a person a) who trains sportsmen for games, competitions: a football coach b) employed privately to train a pupil or student for an examination: an English coach
		5) a farm cow used for attracting wild cattle into a trap
	verb	6) to train or teach (a person or a group of people); give instruction or advice to (a person or a group of people): I coach people for English examinations
		7) to act or be employed as a COACH ^[4a]

ตารางที่ 5.1 พจนานุกรมความหมายของคำว่า "coach"^[12]

[*] Longman Dictionary of Contemporary English, Paul Procter

SURFACE	CATEGORY	CONCEPTUAL MEANING
lost ^[1]	adj	1) no longer possessed: It's no use thinking about one's lost youth
		2) that cannot be found: a lost dog/lost keys
		3) unable to find the way: lost in the mists
		4) not used, obtained, or won: a lost chance
		5) destroyed, ruined, killed, drowned: the boat and all its men were lost in the storm
		6) not behaving in a way acceptable to society or a religion
		7) a) no longer belonging to: My son was lost to me when he married b) no longer possible for (somebody) c) not noticing: He was reading his book, completely lost to the world
		8) having no influence on: Good advice is lost on him

ตารางที่ 5.2 พจนานุกรมความหมายของคำว่า "lost"^[12]

[*] Longman Dictionary of Contemporary English, Paul Procter

SURFACE	CATEGORY	CONCEPTUAL MEANING
set ¹	verb	1) to put (to stay) in place: Set a lamp on the tables
		2) to make to be in a stated condition: Opened the cage and set the bird free
		3) to fix or determine (a rule, time, standard, number, etc): Set a wedding day
		4) to give (a piece of work) for (someone) to do: Who set (the question for) the examination ?
		5) to put into a position, esp. into order for use: Set the clock (by the time given on the radio)
		6) a) to put (a broken bone or limb) into a fixed position for proper joining b) (of a broken bone or limb) to become joined in a fixed position
		7) to fix firmly (a part of the body, esp. regarded as showing one's intention, feelings, etc.): He set his jaw and refused to agree to anything I said
		8) to put into action: Your remarks have set me thinking
		9) a) to cause (a liquid, paste, soft material, etc.) to become solid: Set the jelly by putting it in a cold place b) (of such materials) to harden or become solid
		10) a) to fix (a colour) against being changed as by water b) (of a colour or colouring matter) to become fixed: This chemical is used to set a dye
		11) to put (a bird) on eggs to hatch them
		12) (of a heavenly body) to pass downwards out of sight: In the water the sun sets early
		13) to arrange ((in) metal letters (type) of printing: today most books are set (up) by machine

		14) a) to arrange (hair) when wet to give the desired style when dry b) (of hair) to dry after being arranged in this way: straight hair that doesn't set easily
		15) to write or provide (music) for a poem or other words to be sung: Has the poem ever been set (to music) ?
		16) to fix (a precious stone) into (a piece of jewelry): Set a diamond in a ring
		17) (of a plant) to form and develop seed or fruit: Our apple trees set well last year even though there was a water shortage
		18) to give a particular setting to (a story, play, etc.): The (action of) the book is set in 17th century Spain
		19) (of a dog) to point out the position of (a bird) with the nose while keeping still
	adj	1) placed; fixed: a city set on a hill/eyes set deep in his head
		2) determined; of a fixed intention: He's very set on going, and I can't make him see that it's a bad idea
		3) fixed: wages set by law/I must study at set hours each day
		4) given or fixed for study: The examination will have questions on the set subjects/books
		5) (of part of the body, manner, state of mind, etc.) fixed in position ; unmoving: She greeted her 50 guests with a set smile
		6) ready; prepared: Are you all set?
		7) (of a restaurant meal) complete and at a fixed price: a set dinner
	noun	1) a group of naturally connected things; group forming a whole: a set of gardening tools
		2) tech (in MATHEMATICS) a formal object which is a collection of members: the set of all numbers greater than 3
		3) a group of people of a social type: He goes around with a rather wild set

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา 92 จะต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

	<p>4) a natural position of part of the body: From the set of her shoulders it was clear that she was tired</p>
	<p>5) the way in which (part of) a garment fits the body: I don't like the set of this collar</p>
	<p>6) a direction, as of a) movement, pointing, bending, etc.: The wind had a western set b) thinking, opinion, etc.: The set of public opinions is against bulding up the army</p>
	<p>7) the hardening of a liquid, paste, soft solid, etc: You won't get a good set if you don't keep the jelly cold</p>
	<p>8) an electric apparatus, esp. a radio or television: a colour television set / a crystal set Is your set (television) working?</p>
	<p>9) a) something built and provided with furniture, scenery, etc., to represent the scene of (part of) the action of a play b) a place (of this kind or in natural surroundings) where a film is acted: Everyone must be on the set ready to begin filming at 10 o'clock</p>
	<p>10) a part of a tennis match which may be won by a SCORE of 6 games to 0,1, 2, 3, etc.</p>
	<p>11) a young plant to be SET OUT:potato sets</p>
	<p>12) the act of seting by a hunting dog</p>
	<p>13) a type of square stone used for floors, roads, etc.</p>
	<p>14) an act or result of setting hair: I'd like a shampoo and set, please</p>
	<p>15) poet the going down in the sky of a heavenly body: That will be before set of sum</p>

		<p>16) a) to combine to attack (someone): The newspapers have made a dead set at this politician for using his official car for private purposes</p> <p>b) to try to gain the favour of (someone of the opposite sex): Jean's making a dead set at Charles</p>
--	--	--

ตารางที่ 5.3 พจนานุกรมความหมายของคำว่า "set"¹²¹

[*] Longman Dictionary of Contemporary English, Paul Procter

จากตารางแสดงความหมายของคำในประโยคนี้ จะเห็นว่าความเป็นไปได้ของความหมายของ คำว่า "coach" ทั้งหมด 7 ความหมาย ความหมายของคำว่า "lost" มีทั้งหมด 8 ความหมาย และความหมายของคำว่า "set" มีทั้งหมด 19+7+16 = 42 ความหมาย ความเป็นไปได้ของประโยคที่จะนำความหมายมาใช้เป็นไปได้ถึง 2352 แบบ และปัญหานี้เองที่ต้องกำหนดให้เป็น CPID และหนึ่ง CPID มีหนึ่งความหมายก็สามารถที่จะลดขั้นตอนและความกำกวมได้

ดังนั้นงานวิทยานิพนธ์นี้ได้วิจัยออกแบบโครงสร้างของพจนานุกรมความหมายให้สอดคล้องกับการกำหนดหน่วยความหมายของภาษากลาง โดยให้มีการบรรจุข้อมูลทางภาษาระดับแก่นหรือความหมายระดับแก่นไว้ให้สมบูรณ์ที่สุด และลดความซ้ำซ้อนหรือความกำกวม โดยกำหนดให้ประกอบไปด้วยข้อมูลทางภาษา 5 ส่วนได้แก่

1. CPID (Conceptual Primitive Identification)
2. CPDC (Conceptual Primitive Description)
3. CAT (Category)
4. SUBCAT (Subcategory)
5. TYPE

โครงสร้างการเก็บบันทึกข้อมูลคำศัพท์ต่อ 1 หน่วยคำในพจนานุกรมความหมาย แสดงในรูปที่

5.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 94 และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

CPID	CPDC	CAT	SUBCAT	TYPE
------	------	-----	--------	------

รูปที่ 5.2 Data Structure ของพจนานุกรมความหมายใน 1 record ของ 1 รหัส

5.2.1 CPID (Conceptual Primitive Identification)

เป็นกลุ่มของสัญลักษณ์ที่กำหนดกลุ่มความหมายของคำนั้น ดังที่ได้กล่าวในบทที่ 3.2

5.2.2 CPDC (Conceptual Primitive Description)

กลุ่มของคำอธิบายความหมายของ CPID นั้น ดังที่ได้กล่าวในบทที่ 3.3

5.2.3 CAT (Category)

เป็นกลุ่มที่บอกชนิดของคำที่วิเคราะห์ตามหน้าที่ของคำในประโยค ในการแบ่งวิเคราะห์ชนิดและหน้าที่ของ CPID นั้น โดยจะเน้นให้เครื่องคอมพิวเตอร์เข้าใจภาษาเพื่อการประมวลผลอย่างมีประสิทธิภาพ จึงไม่ได้แบ่งชนิดและหน้าที่ของคำตามหนังสือหลักไวยากรณ์ทั่วไป แต่นำมาจัดชนิดและหน้าที่ของคำใหม่เพื่อที่จะให้เครื่องคอมพิวเตอร์ประมวลผลได้อย่างมีประสิทธิภาพมากขึ้น

1 คำนาม (Noun ใช้สัญลักษณ์ย่อ N) คือ คำที่ใช้เรียกคน สัตว์ สิ่งของ

2 สรรพนาม (Pronoun ใช้สัญลักษณ์ย่อ PRON) คือ คำที่ใช้แทนคำนาม และอาจแทนด้วยคำบ่งชี้ที่มีความหมายรวมถึงนามที่ละไว้ด้วย

3 กริยา (Verb ใช้สัญลักษณ์ย่อ V) คือ คำที่แสดงกริยาอาการ หรือสภาพของประธานในประโยค และถือเป็นคำหลักของประโยค

4 คำกริยาวิเศษณ์ (Adverb ใช้สัญลักษณ์ย่อ ADV) คือ คำขยายกริยาเพื่อบอกว่าเหตุการณ์หรือสิ่งที่เกิดขึ้นนั้น เกิดเมื่อไหร่ ที่ไหน หรืออย่างไร

ส่วนลักษณะของชนิดคำอื่นๆ เช่น คำสันธาน (Conjunction), คำบุรพบท (Preposition) ฯลฯ จะรวมอยู่ในกลุ่มของ Relation เพื่อความสะดวกในการคำนวณของคอมพิวเตอร์ และเพิ่มประสิทธิภาพในการแปลด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา⁹⁵ และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.2.4 SUBCAT (Subcategory)

เป็นกลุ่มของชนิดย่อยของคำเป็นการแบ่งประเภทของคำให้ย่อยลงไปอีก เพื่อแสดงลักษณะเฉพาะ หรือความหมายที่แคบลงของคำเพื่อให้ระบบวิเคราะห์ทำการแยกแยะความแตกต่างระหว่างคำต่างๆ ได้ในอีกระดับหนึ่ง ดังต่อไปนี้

1.Noun แบ่งเป็นชนิดย่อยได้ 7 ประเภท คือ

- 1.1 Proper Noun ใช้สัญลักษณ์ย่อ NPRP หมายถึง วิสามัญนามซึ่งเป็นคำนามที่เฉพาะเป็นชื่อเฉพาะของคน หรือสิ่งใดสิ่งหนึ่ง ซึ่งแสดงความสัมพันธ์ในแง่การที่เป็นชื่อเฉพาะ และไม่ได้มีหน้าที่เป็นคำขยาย เช่น ประเทศ
- 1.2 Common Noun ใช้สัญลักษณ์ย่อ NCMN หมายถึง สามัญนามซึ่งเป็นคำนามโดยทั่วไปที่ใช้เรียก คน สัตว์ สิ่งของ เช่น บ้าน, แมว, สุนัข ฯ รวมทั้งคำนามที่เป็นประเภทย่อยของคำนามนั้น ซึ่งไม่ได้เป็นชื่อ แต่เป็นการแสดงการขยาย attribute เท่านั้น เช่น รถเก๋ง ดอกไม้ป่า และคำนามที่มาจากกรเติมคำอุปสรรค (Prefix) เช่น ความรัก
- 1.3 Cardinal Number ใช้สัญลักษณ์ย่อ NCNM หมายถึง คำนามที่แสดงจำนวนนับ เช่น หนึ่ง, สอง, สาม
- 1.4 Ordinal Number ใช้สัญลักษณ์ย่อ NONM หมายถึง คำนามที่บอกลำดับที่ ก่อน, หลัง เช่น ที่หนึ่ง, ที่สอง, ที่สาม
- 1.5 Adverb Like Noun ใช้สัญลักษณ์ย่อ NALN หมายถึง คำนามที่ทำหน้าที่ได้เช่นเดียวกับคำกริยาวิเศษณ์ (adverb) มักจะเป็นคำนามที่ใช้แสดงเวลา เช่น วันอาทิตย์, วันจันทร์
- 1.6 a Noun which is Conceptually an Adjective ใช้สัญลักษณ์ย่อ NAJN หมายถึง คำนามที่แสดงการขยายคำอื่นๆ ได้เช่นเดียวกับคำวิเศษณ์ (Adjective) เช่น ส่วนใหญ่
- 1.7 Label ใช้สัญลักษณ์ย่อ NLBL หมายถึง คำนามที่เป็นเครื่องหมาย หรือสัญลักษณ์ และทำหน้าที่ได้เหมือนคำนามทุกๆ ไป เช่น 1,2,+,...

2 Pronoun แบ่งเป็นชนิดย่อยได้ 6 ประเภท

- | | |
|---------------------------|--|
| 2.1 Personal Pronoun | ใช้สัญลักษณ์ย่อ PPRS หมายถึง บุรุษสรรพนามซึ่งเป็นคำสรรพนามที่แทนตัวผู้พูด ผู้ที่พูดด้วย หรือผู้ที่กล่าวถึง เช่น ฉัน, ผม, คุณ |
| 2.2 Indefinite Pronoun | ใช้สัญลักษณ์ย่อ PIND หมายถึง สรรพนามไม่ชี้เฉพาะที่ใช้แทนคนหรือสิ่งต่างๆ ไปไม่มีขอบเขตที่แน่นอน เช่น บางส่วน, มากมาย |
| 2.3 Interrogative Pronoun | ใช้สัญลักษณ์ย่อ PNTR หมายถึง สรรพนามที่แสดงการถามใช้แทนสิ่งที่ผู้พูดต้องการทราบคำตอบ เช่น ที่ไหน, อะไร, ใคร |
| 2.4 Demonstrative Pronoun | ใช้สัญลักษณ์ย่อ PDMN หมายถึง สรรพนามบ่งชี้ที่ใช้แทนสิ่งที่คุณพูดชี้เฉพาะเจาะจงถึง เช่น นี่คือน้ำที่คุณต้องการ |
| 2.5 Definite Pronoun | ใช้สัญลักษณ์ย่อ PDEF หมายถึง สรรพนามชี้เฉพาะที่ใช้แทนเพื่อบอกขอบเขตของสิ่งที่พูดถึง เช่น เหล่านี้ ทั้งหมด |
| 2.6 Relative Pronoun | ใช้สัญลักษณ์ย่อ PREL หมายถึง ประพันธสรรพนามที่ใช้แทนนามและทำหน้าที่เชื่อมประโยคเข้าด้วยกัน เช่น ที่ ซึ่ง อัน |

3 Verb แบ่งออกเป็นชนิดย่อยได้ 3 ประเภท

- | | |
|--------------------|--|
| 3.1 Active Verb | ใช้สัญลักษณ์ย่อ VACT หมายถึง คำกริยาที่แสดงอาการ สามารถเห็นการกระทำได้โดยจะรวมทั้งกริยาที่มีกรรมมารับ และกริยาที่ไม่มีกรรมมารับ เช่น รเรียน, วิ่ง, กิน |
| 3.2 Stative Verb | ใช้สัญลักษณ์ย่อ VSTA หมายถึง คำกริยาแสดงสภาพ ความรู้สึกของประธาน หรือแสดงผลจากการกระทำอย่างใดอย่างหนึ่ง เช่น ราคา ลำบาก |
| 3.3 Attribute Verb | ใช้สัญลักษณ์ย่อ VATT หมายถึง คำกริยาที่บ่งบอกคุณสมบัติของคำนาม มักมาจากคำคุณศัพท์ เช่น สวย, ใหญ่ |

4. Adverb แบ่งออกเป็นชนิดย่อยได้ 4 ประเภท

- | | |
|--------------------------------|--|
| 4.1 Adverb with normal form | ใช้สัญลักษณ์ย่อ ADVN หมายถึง คำกริยาวิเศษณ์ต่างๆ ไปที่ขยายกริยา เช่น เร็ว, สม่่าเสมอ |
| 4.2 Sentential adverb | ใช้สัญลักษณ์ย่อ ADVS หมายถึง คำกริยาวิเศษณ์ที่ขยายประโยคโดยจะอยู่ต้นประโยค หรือท้ายประโยค เช่น โดยปกติ, ธรรมดา |
| 4.3 Adverb with prefixed form | ใช้สัญลักษณ์ย่อ ADVP หมายถึง คำกริยาวิเศษณ์ที่ขยายกริยา เช่น อย่างรวดเร็ว, โดยเร็ว |
| 4.4 Adverb with Iterative Form | ใช้สัญลักษณ์ย่อ ADVI หมายถึง คำกริยาวิเศษณ์ที่เป็นคำซ้ำๆ เช่น เร็วๆ มัวๆ |

จากการกำหนด Category ข้างต้นสามารถสรุปได้ในตาราง 5.4

5.2.5 TYPE

เป็นกลุ่มคำที่บอกลักษณะจำนวนของสิ่งนั้นไม่ว่าจะเป็นเอกพจน์ (singular) หรือพหูพจน์ (plural) โดยแบ่งกลุ่มดังนี้

- | | |
|----------------------|---|
| 1 Singular | ใช้สัญลักษณ์ย่อ S หมายถึง กลุ่มคำที่มีสิ่งเดียวหรือเป็นเอกพจน์ |
| 2 Singular of Male | ใช้สัญลักษณ์ย่อ M หมายถึง การกำหนดกลุ่มคำที่เป็นเพศผู้และเป็นเอกพจน์ |
| 3 Singular of Female | ใช้สัญลักษณ์ย่อ W หมายถึง การกำหนดกลุ่มคำที่เป็นเพศเมียและเป็นเอกพจน์ |
| 4 Plural | ใช้สัญลักษณ์ย่อ P หมายถึง การกำหนดกลุ่มคำที่มีหลายสิ่งหรือจำนวนมากกว่าหนึ่ง หรือเรียกว่าพหูพจน์ |

ตัวอย่างของการแสดงพจนานุกรมความหมายดังตารางที่ 5.5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 98 และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

CAT		SUBCAT		EXAMPLE
ชื่อ	สัญลักษณ์	ชื่อ	สัญลักษณ์	
NOUN	N	Proper Noun	NPRP	ประเทศไทย, สมศักดิ์
		Common Noun	NCMN	คอมพิวเตอร์, รถยนต์
		Cardinal Number	NCNM	หนึ่ง, สอง, สาม
		Ordinal Number	NONM	ที่สอง, ที่สาม
		Adverb Like Noun	NALN	วันอาทิตย์, วันจันทร์
		a Noun which is Conceptually an Adjective	NAJN	ส่วนใหญ่, ทั้งหมด
		Label Noun	NLBL	1, 2, 3,...
PRONOUN	PRON	Personal Pronoun	PPRS	ฉัน, ผม, คุณ
		Idefinite Pronoun	PIND	บางส่วน, มากมาย
		Interrogative Pronoun	PNTR	ที่ไหน, อะไร, ใคร
		Demonstrative Pronoun	PDMN	นี่คือ
		Definite Pronoun	PDEF	เหล่านี้
		Relative Pronoun	PREL	ที่ ซึ่ง อัน
VERB	V	Active Verb	VACT	เรียน วิ่ง
		Stative Verb	VSTA	ราคา ลำบาก
		Attribute Verb	VATT	สวย ใหญ่
ADVERB	ADV	Adverb with normal form	ADVN	เร็ว สม่่าเสมอ
		Sentential adverb	ADVS	โดยปกติ ธรรมดา
		Adverb with prefixed form	ADVP	โดยเร็ว อย่างรวดเร็ว
		Adverb with interative form	ADVI	เร็ว ๆ

ตารางที่ 5.4 สรุปข้อมูลคำศัพท์

CPID	CPDC	CAT	SUBCAT	TYPE
1110000a	a person who has the qualification to teach someone something in school.	N	NCMN	S
1111000a	a human being name vichean.	N	NPRP	M
1211000a	the land belonging Thai nation.	N	NPRP	S
1222000a	star forming center of system of worlds or planets;central body of solar system.	N	NPRP	S
1230000a	corporate institution with power of conferring degrees and providing instruction in higher branches of larning.	N	NCMN	S
1311200a	cold-blooded animal living wholly in water and breathing through gills,with fins for swimming.	N	NCMN	S
1311500a	a plant whose seeds of rice or grain are edible.	N	NPRP	S
1311800a	pearl-white grain used as a staple food.	N	NCMN	S
1312100a	precious yellow metal used for making coins, ornaments, jewellery.	N	NCMN	S
1312200a	liquid which does not mix with water,obtained from animals, plants or found in rock underground.	N	NCMN	S
1312300a	the mixture of gases that surrounds the earth and which we breathe.	N	NCMN	S
1322000a	a vehicle with 3 or usually 4 wheels and driven by a motor.	N	NCMN	S
2111000a	take (food etc.) into the mouth and swallow it.	V	VACT	
2320000a	the first part of the day;from the time when the sun rises	N	NALN	
2365000a	in a manner of being repeated many times.	ADV	ADVN	

ตารางที่ 5.5 ตัวอย่างพจนานุกรมความหมาย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 100% ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.3 พจนานุกรมภาษา (Language Dictionary)

พจนานุกรมภาษาเป็นพจนานุกรมสองภาษา (Bilingual Dictionary) ระหว่างภาษากลางกับภาษาอื่น ๆ อีกหนึ่งภาษา พจนานุกรมนี้ จะมีความสัมพันธ์กับพจนานุกรมความหมาย โดยมีตัวแปรร่วมกันคือ CPID เพื่อเป็นดัชนี (index) ที่ไปยัง surface ของภาษาอื่น ๆ ซึ่งข้อมูลในพจนานุกรมความหมายจะใช้ร่วมกับพจนานุกรมภาษา

งานวิจัยได้ออกแบบให้มีโครงสร้างประกอบด้วย

1. CPID (Conceptual Primitive Identification)
2. WORD
3. CLASS (Classifier)
4. TYPE
5. VMP (Verb Map Pattern)
6. NID

CPID	WORD	CLASS		TYPE	VMP	NID
		Class1	Class2			

รูปที่ 5.2 Data Structure ของพจนานุกรมภาษาใน 1 record ของ 1 รหัส

รายละเอียดและนิยามของข้อมูลทางภาษาในส่วนต่างๆ มีดังต่อไปนี้

5.3.1 CPID (Conceptual Primitive Identification)

เป็นกลุ่มของสัญลักษณ์ที่กำหนดกลุ่มความหมายของคำนั้น ดังที่กล่าวในบทที่ 3.2

5.3.2 WORD

เป็น surface ของภาษาต่างๆเพื่อแทนความหมายใน CPDC

5.3.3 CLASS (Classifier)

เป็นกลุ่มของลักษณะนามของ CPID ที่เป็นคำนาม สามารถที่จะมีได้หลายลักษณะนามก็ได้ ขึ้นอยู่กับรูปแบบของประโยค และคำนามนั้น เช่น ตัว, ใบ ฯ ซึ่งในบางภาษาอาจไม่มีลักษณะนามก็ได้

5.3.4 TYPE

เป็นส่วนที่บอกความหมายเฉพาะกลุ่มของคำกริยา ซึ่งบางภาษาอาจจะไม่จำเป็นต้องใช้ เพราะในกริยาบางภาษาต้องมีการเปลี่ยนแปลงคำเพื่อที่จะสามารถนำไปใช้กับกลุ่มคำนั้น หรือประโยคนั้น เช่น

ภาษาญี่ปุ่น

กริยาแบ่งออกเป็น 3 ประเภทคือ

- 1.กริยาประเภทที่ลงท้ายด้วย i (v1)
- 2.กริยาประเภทที่ลงท้ายด้วย e (v2)
- 3.กริยาประเภทที่ลงท้ายด้วย shimasu (v3)

คุณศัพท์แบ่งออกเป็น 2 ประเภทคือ

- 1.คุณศัพท์ i (a1)
- 2.คุณศัพท์ na (a2)

5.3.5 VMP (Verb Map Pattern)

เป็นการกำหนดแบบโครงสร้างของประโยคโดยใช้คำกริยาเป็นหลัก แบบโครงสร้างนี้จะแสดงความสัมพันธ์กับรหัสแยกแยะความหมาย (CPID) แบบโครงสร้างของประโยคนั้นจะใช้โครงสร้างของประโยคความเดียว (ประโยคที่มีหน่วยกริยา 1 หน่วย มีหน่วยคำนาม 1 หน่วยหรือมากกว่านั้นและอาจจะมีหน่วยเสริมอยู่ด้วยหรือไม่ก็ได้) โดยที่หน่วยคำนามที่อยู่ในประโยคความเดียวจะทำหน้าที่เป็น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 102 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประธาน กรรมตรง หรือกรรมรอง ประโยคความเดียวที่ได้ทำการวิจัยนี้สามารถเปรียบเทียบกับหน่วยความสัมพันธ์ที่ได้กำหนดไว้ในวิทยานิพนธ์นี้ โดยมีรูปแบบโครงสร้างของประโยคที่ประกอบด้วยลักษณะของหน่วยคำต่างๆ ที่สัมพันธ์กับคำกริยา แบ่งออกเป็นกลุ่มๆ ดังต่อไปนี้

กลุ่มที่ 1 เป็นกลุ่มที่มี 1 หน่วยคำที่สัมพันธ์กับกริยาโดยเป็นประธานของคำกริยา

กลุ่มที่ 2 เป็นกลุ่มที่มี 2 หน่วยคำที่สัมพันธ์กับกริยาโดยเป็นประธาน และกรรมของคำกริยา

กลุ่มที่ 3 เป็นกลุ่มที่มี 3 หน่วยคำที่สัมพันธ์กับกริยาโดยเป็นประธาน กรรมตรง และกรรมรอง หรือส่วนเสริมของคำกริยา

กลุ่มที่ 4 เป็นกลุ่มที่มี 4 หน่วยคำที่สัมพันธ์กับกริยาโดยเป็นประธาน กรรมตรง กรรมรอง และส่วนเสริมของคำกริยา

กลุ่มทั้ง 4 กลุ่มสามารถแบ่งเป็นกลุ่มย่อยๆ ได้ดังต่อไปนี้

	VMP.no	Verb Map Pattern
กลุ่มที่ 1	1a	SUB=r1001
กลุ่มที่ 2	2a	SUB=r1001, DOB=r1002
	2b	SUB=r1001, DOB=r1003
	2c	SUB=r1001, DOB=r1004
	2d	SUB=r1001, DOB=r1005
	2e	SUB=r1001, DOB=r1006
	2f	SUB=r1001, MOD=r1008
	2g	SUB=r1001, DOB=r1009
	2h	SUB=r1001, DOB=r1013
	2i	SUB=r1001, DOB=r1015
	2j	SUB=r1006, DOB=r1003
กลุ่มที่ 3	3a	SUB=r1001, DOB=r1002, IDO=r1003
	3b	SUB=r1001, DOB=r1002, IDO=r1004

- 3c SUB=r1001,DOB=r1002,IDO=r1005
- 3d SUB=r1001,DOB=r1002,IDO=r1006
- 3e SUB=r1001,DOB=r1002,MOD=r1008
- 3f SUB=r1001,DOB=r1003,IDO=r1005
- 3g SUB=r1001,DOB=r1004,IDO=r1003
- 3h SUB=r1001,DOB=r1004,IDO=r1005
- 3i SUB=r1001,DOB=r1004,MOD=r1008
- 3j SUB=r1001,DOB=r1004,IDO=r1013
- 3k SUB=r1001,DOB=r1006,IDO=r1004
- 3l SUB=r1001,DOB=r1006,IDO=r1005
- 3m SUB=r1001,DOB=r1006,IDO=r1007
- 3n SUB=r1001,DOB=r1005,MOD=r1008
- 3o SUB=r1001,DOB=r1006,MOD=r1008
- 3p SUB=r1001,DOB=r1013,MOD=r1008
- 3q SUB=r1001,DOB=r1015,MOD=r1008
- 3r SUB=r1001,DOB=r1013,IDO=r1005
- 3s SUB=r1001,DOB=r1015,IDO=r1005
- 3t SUB=r1001,DOB=r1013,MOD=r1008

กลุ่มที่ 4

- 4a SUB=r1001,DOB=r1002,IDO=r1004,MOD=r1008
- 4b SUB=r1001,DOB=r1006,IDO=r1004,MOD=r1008
- 4d SUB=r1001,DOB=r1004,IDO=r1003,MOD=r1008
- 4e SUB=r1001,DOB=r1006,IDO=r1004,MOD=r1005
- 4f SUB=r1001,DOB=r1004,IDO=r1003,MOD=r1005

หมายเหตุ SUB = subject, DOB = direct object, IDO = indirect object, MOD = modifier

5.3.6 NID (Near IDitification) เป็นกลุ่มของสัญลักษณ์ที่มีความหมาย หรือลักษณะกลุ่มความหมายลดหลั่นใกล้เคียงหรือกลุ่มเดียวกับ CPID นั้นโดยในแต่ละภาษาอาจมีกลุ่มของ NID ที่ต่างกันได้ โดยเฉพาะในกลุ่มของภาษาท้องถิ่น (Local Language)

ตัวอย่างของการแสดงผลงานอนุกรมเฉพาะภาษาดังตารางที่ 5.2, 5.3

CPID	WORD	CLASSIFIER		TYPE	VMP	NID
		CLASS1	CLASS2			
1110000a	ครู	คน				1110000d
1111000a	วีเชียร	คน				1110000b,1110000d
1211000a	ประเทศไทย	ประเทศ				1210000a
1222000a	พระอาทิตย์	ดวง				
1230000a	มหาวิทยาลัย	แห่ง	มหาวิทยาลัย			
1311200a	ปลา	ตัว	ฝูง			
1311500a	ข้าว	ต้น	ไร่			1311800a
1311800a	ข้าว	เมล็ด	ต้น			1311500a
1312100a	ทอง	บาท	กิโลกรัม			
1312200a	น้ำมัน	ลิตร	บาเรท			
1312300a	อากาศ					1312300b,1312300d
1322000a	รถยนต์	คัน				
2111000a	กิน				2a,3a,3b,3c, 3e,4a	
2320000a	ตอนเช้า					

ตารางที่ 5.6 ตัวอย่างพจนานุกรมเฉพาะสำหรับภาษาไทย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อ 105 และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

CPID	WORD	CLASSIFIER		TYPE	VMP	NID
		CLASS1	CLASS2			
1110000a	teacher					1110000d
1111000a	vichean					1110000b,1110000d
1211000a	thailand					1210000a
1222000a	sun					
1230000a	university					
1311200a	fish					
1311500a	rice					1311800a
1311800a	rice					1311500a
1312100a	glod					
1312200a	oil					
1312300a	air					1312300b,1312300d
1322000a	car					
2111000a	eat				2a,3a,3b,3c ,3e,4a	
2320000a	morning					

ตารางที่ 5.7 ตัวอย่างพจนานุกรมเฉพาะสำหรับภาษาอังกฤษ

5.4 สรุป

ในบทนี้เป็นการออกแบบโครงสร้างการเก็บข้อมูลของพจนานุกรมเพื่อการแปลหลายภาษาด้วยคอมพิวเตอร์ โดยแบ่งพจนานุกรมออกเป็น 2 ส่วนคือพจนานุกรมความหมาย (Conceptual Dictionary) และพจนานุกรมภาษา (Language Dictionary) ซึ่งพจนานุกรมทั้ง 2 ส่วนมีความแตกต่างกัน แต่จะมีข้อมูลส่วนที่สัมพันธ์กันในลักษณะของ Relational Database เพื่อที่จะนำไปใช้ในการแปลภาษาดังที่กล่าวในบทที่ 6

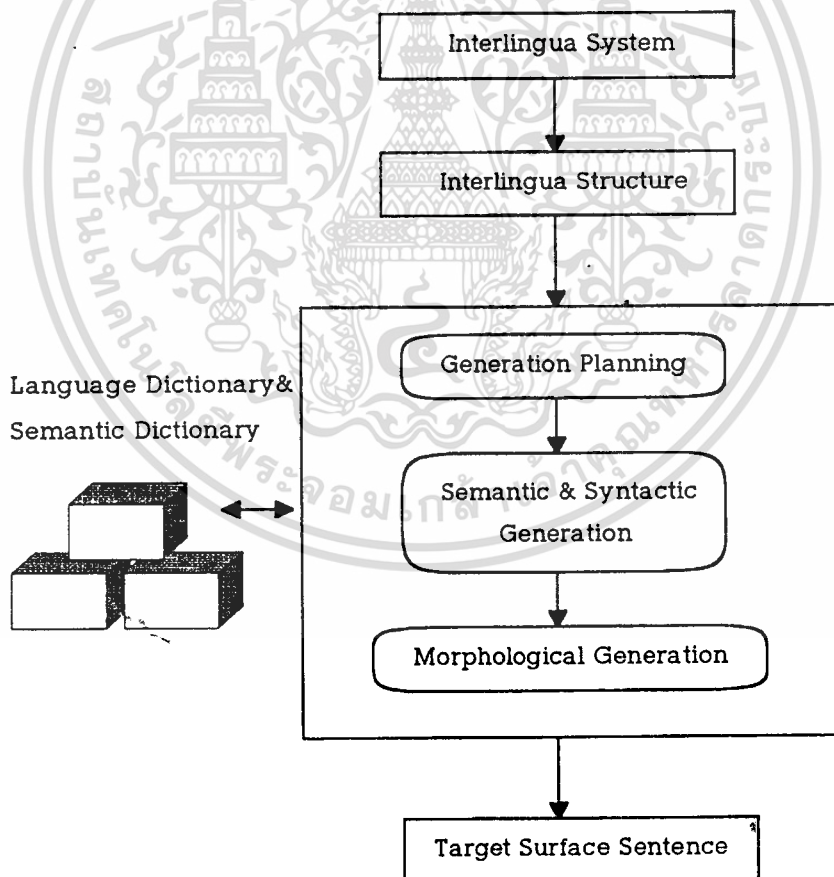


บทที่ 6

ตัวอย่างการสังเคราะห์ภาษาเป้าหมายจากภาษากลาง

6.1 บทนำ

จากบทต่างๆ ที่ได้กล่าวมาแล้วข้างต้น ได้กล่าวถึงการออกแบบและสามารถที่จะสร้างภาษากลางจากประโยคต้นแบบ ในบทนี้ได้ทำการสร้างซอฟต์แวร์ตัวอย่างการสังเคราะห์ภาษาเป้าหมายจากสื่อภาษากลางที่ออกแบบในวิทยานิพนธ์นี้ โดยอาศัยกฎทางไวยากรณ์และความหมายของภาษาและพจนานุกรมดังที่ได้กล่าวมาแล้ว โดยมีขั้นตอนการสังเคราะห์ภาษาเป้าหมายดังแสดงในรูปที่ 6.1



รูปที่ 6.1 การสังเคราะห์ภาษาเป้าหมายจากภาษากลาง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา 108 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6.2 ขั้นตอนการสังเคราะห์ภาษาเป้าหมายจากภาษากลาง

ภาษากลางที่กำหนดในวิทยานิพนธ์นี้ ได้วิเคราะห์มาจากภาษาต้นแบบด้วยมือ เมื่อวิเคราะห์ความสัมพันธ์ของคำแล้วการกำหนดรูปแบบของโครงสร้างความสัมพันธ์ให้อยู่ในรูปแบบที่คอมพิวเตอร์สามารถเข้าใจก็เป็นสิ่งสำคัญต่อการสังเคราะห์ ซึ่งมีขั้นตอนดังต่อไปนี้

6.2.1 โครงสร้างภาษากลางทางซอฟต์แวร์ (Interlingua Software Structure)

เป็นการนำโครงข่ายต้นไม้ของภาษากลางมาเก็บในรูปแบบของโครงสร้างภาษากลางทางซอฟต์แวร์ (Interlingua Software Structure) ซึ่งจะต้องมีรูปแบบการเก็บข้อมูลให้เป็นมาตรฐาน เพื่อให้คอมพิวเตอร์สามารถที่จะนำข้อมูลทั้งหมดไปใช้ได้ถูกต้อง โดยกำหนดการเก็บข้อมูลเป็นโครงสร้างอย่างง่าย ๆ เพื่อที่จะสามารถให้คอมพิวเตอร์เข้าใจความหมายต่างๆ ดังนี้

ลำดับของ node	—
id ของ CPID หรือ REL	
หมายเลข node ที่เชื่อมกับ node นี้ โดยมีทิศทางด้านหัวลูกศร	
หมายเลข node ที่เชื่อมกับ node นี้ โดยมีทิศทางด้านปลายลูกศร	

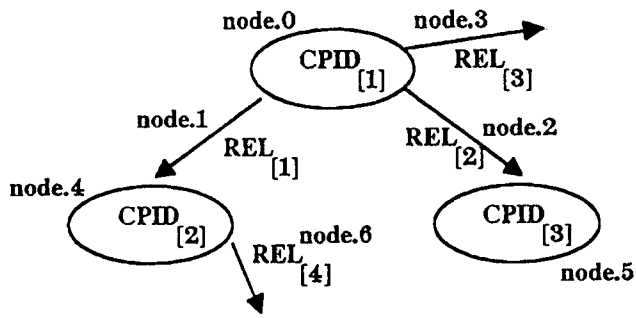
ลำดับของ node เป็นการกำหนดลำดับของ CP หรือ REL ที่มีอยู่ในประโยคภาษากลางโดยจะทำให้สามารถรู้ตำแหน่งของ node นั้นเพื่อใช้ในการเก็บข้อมูลต่างๆของ node นั้น

id ของ CPID หรือ REL เป็นตัวกำหนดความหมายของคำเพื่อให้ทราบว่า ลำดับ node นี้มีความหมายอย่างไร หรือมีความสัมพันธ์ในรูปแบบใด

หมายเลข node ที่เชื่อมกับ node นี้ โดยมีทิศทางด้านหัวลูกศรและปลายลูกศร เป็นตัวกำหนดทิศทางของความสัมพันธ์ของ CP หรือ REL เพื่อให้ทราบจุดเริ่มต้นของความสัมพันธ์เริ่มที่ใดและสิ้นสุดที่ใด

ตัวอย่างเช่นในรูปที่ 6.2 แสดงความสัมพันธ์ของ CPID และ REL

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 109 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 6.2 ตัวอย่างการกำหนดโครงสร้างภาษากลางซอฟต์แวร์

รูปแบบการเก็บข้อมูลมีลักษณะดังนี้

node.0	
id.CPID _[1]	1,2,3
	-1
node.1	
id.REL _[1]	4
	0
node.2	
id.REL _[2]	5
	0

node.3

id.REL_[3]

6

1

node.4

id.CPID_[2]

6

1

node.5

id.CPID_[3]

-1

2

node.6

id.REL_[4]

-1

4



* หมายถึง -1 หมายถึง ไม่มี node อื่นที่เชื่อมต่อ

6.2.2 การสังเคราะห์ในระดับการเตรียมข้อมูล (Generation Planning)

ขั้นตอนนี้เป็นเตรียมโครงสร้างของข้อมูล (Data Structure) เพื่อที่จะรองรับข้อมูลจากภาษากลางทางซอฟต์แวร์ (Interlingua Software Structure) พจนานุกรมรวมภาษา (Multilingual Dictionary) และจากการประมวลผลทางภาษาศาสตร์คอมพิวเตอร์ (Computational Linguistics Processing) โดยใช้กฎทางไวยากรณ์ (Syntactic) และทางความหมาย (Semantic) ของภาษา โดยมีโครงสร้างการเก็บข้อมูลดังตารางที่ 6.1

NUM	ลำดับของ node
CPID	รหัสแยกแยะความหมาย (Conceptual Primitive)
CAT	ชนิดของคำ (Category)
SUBCAT	ชนิดย่อยของคำ (Subcategory)
HLEV	node ที่อยู่ปลายลูกศรความสัมพันธ์ของ node ปัจจุบัน
LLEV	node ที่อยู่หัวลูกศรความสัมพันธ์ของ node ปัจจุบัน
VMP	Verb map pattern
WORD	คำในภาษาเป้าหมาย
CLASS	ลักษณะนามของคำในภาษาเป้าหมาย
TYPE	ประเภทของคำที่เป็นเอกพจน์หรือพหูพจน์
DEF	กลุ่มความหมายเฉพาะของแต่ละภาษา
NID	คำที่มีความหมายใกล้เคียงกับ CPID ปัจจุบัน
ORDER	ลำดับการเรียงคำในประโยคภาษาเป้าหมาย
PRG	การกำหนดความสำคัญของคำหรือ CPID ในประโยคภาษาเป้าหมาย

ตารางที่ 6.1 โครงสร้างของการเก็บข้อมูลในซอฟต์แวร์

6.2.3 การสังเคราะห์ในระดับโครงสร้างและความหมาย (Semantic and Syntactic Generation)

ในขั้นตอนนี้จะทำการดึงข้อมูลจากภาษากลางทางคอมพิวเตอร์มาเก็บในโครงสร้างที่ได้กำหนดไว้ และใช้รหัสแยกแยะความหมาย (CPID) เป็น key word หรือตัวกำหนดตำแหน่งของข้อมูลในการดึงข้อมูลจากพจนานุกรมรวมภาษาของแต่ละ CPID แล้วทำการประมวลผลทางภาษาศาสตร์คอมพิวเตอร์โดยกำหนดโครงสร้างทางไวยากรณ์ของภาษาเป้าหมาย ได้แก่นำหน้าของคำในประโยคที่มีความสัมพันธ์กับคำกริยาหลักและคำอื่นๆ โดยจะมีค่าตามโครงสร้างของประโยคคือ

คำว่า SUB สำหรับ CPID ที่เป็นประธาน (Subject) ของประโยค

คำว่า DOB สำหรับ CPID ที่เป็นกรรมตรง (Direct object) ของประโยค

คำว่า IDO สำหรับ CPID ที่เป็นกรรมรอง (Indirect Object) ของประโยค

คำว่า MOD สำหรับ CPID ที่เป็นส่วนเสริม (Modifier) ของประโยค

ต่อจากนั้นก็ทำการเรียงลำดับของคำให้อยู่ในโครงสร้างไวยากรณ์ของประโยคเป้าหมาย ก็คือการจัดลำดับของโนด CPID และโนด REL เพื่อที่จะนำมาเรียงกันในประโยคในแนวเส้นตรง โดยอาศัยข้อมูลที่กำหนดตำแหน่งของคำที่อยู่ข้างเคียงกันตามหลักโครงสร้าง (Syntax) ของภาษาเป้าหมาย โดยการเรียงลำดับคำหรือกลุ่มคำตามประเภทของคำ และหน้าที่ของคำนั้น และใช้กฎดังกล่าวเป็นฐานข้อมูลในการเปรียบเทียบอ้างอิงกับการกำหนดลำดับของโนด CPID และโนด REL ในภาษากลาง

6.2.4 การสังเคราะห์ในระดับหน่วยคำ (Morphological Generation)

ส่วนนี้เป็นการกำหนดคำที่เหมาะสมกับประโยค โดยอาศัยความสัมพันธ์ของคำหรือกลุ่มคำที่มีความสัมพันธ์กับคำ หรือกลุ่มคำอื่นๆ ในภาษานั้นๆ เช่น ในประโยคภาษาไทยคำว่า "กิน" จะใช้กับบุคคลธรรมดา แต่ถ้าใช้กับพระจะต้องใช้คำว่า "ฉัน" หรือถ้าใช้กับเชื้อพระวงศ์ก็ต้องใช้คำว่า "เสวย" โดยที่คำเหล่านี้จะมีความหมายเดียวกันคือ อากาหรที่นำอาหารเข้าปาก และกลืนลงไป เป็นต้น

6.3 ขั้นตอนการใช้โปรแกรมในการสังเคราะห์ภาษาเป้าหมายจากภาษากลาง

การใช้งานและขั้นตอนของซอฟต์แวร์ที่สร้างขึ้นมานั้นมีขั้นตอนดังนี้

1) เมื่อต้องการจะใช้ซอฟต์แวร์ให้พิมพ์คำว่า "MENU" แล้วกด ENTER เมื่อทำตามขั้นตอนนี้แล้ว หน้าจอของคอมพิวเตอร์จะแสดงเมนู สำหรับเลือกทำการแปลหรือต้องการดูข้อมูลอื่นๆ โดยการเลื่อนลูกศรขึ้นลงให้แถบไปอยู่ที่คำสั่งนั้น ดังรูปที่ 6.3 โดยระบบนี้จะมีทั้งหมด 4 คำสั่งคือ

Interlingua File

Translation

Dictionary

โปรแกรมระบบการแปลหลายภาษาด้วยคอมพิวเตอร์โดยใช้ภาษากลาง
มีผลมาโดย นายวิเชียร วัฒนศิริ
อาจารย์ที่ปรึกษา ผศ.ดร. รัดติกร วราวุฒศิริพงษ์

Interlingua File

Translation

Dictionary

Quit

คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

รูปที่ 6.3 แสดงหน้าจอการเลือกคำสั่งในการทำงาน

คำสั่ง "Interlingua File" ให้เลื่อนแถบสว่างไปที่คำสั่งนั้นแล้วกด ENTER จอภาพจะปรากฏหน้าจอเพื่อให้ป้อนชื่อ File ของ Interlingua ที่ต้องการจะแปล ดังแสดงในรูปที่ 6.4 เมื่อป้อนชื่อ File เรียบร้อยแล้ว กด ENTER ระบบจะแสดงภาพโครงสร้างของประโยคภาษากลางในรูปของโครงสร้างต้นไม้ ฟังหาเพื่อให้ทราบถึงหน่วยคำ และความสัมพันธ์ของหน่วยคำ ดังแสดงในรูปที่ 6.5 เมื่อต้องการกลับสู่เมนูหลักให้กด SPACE

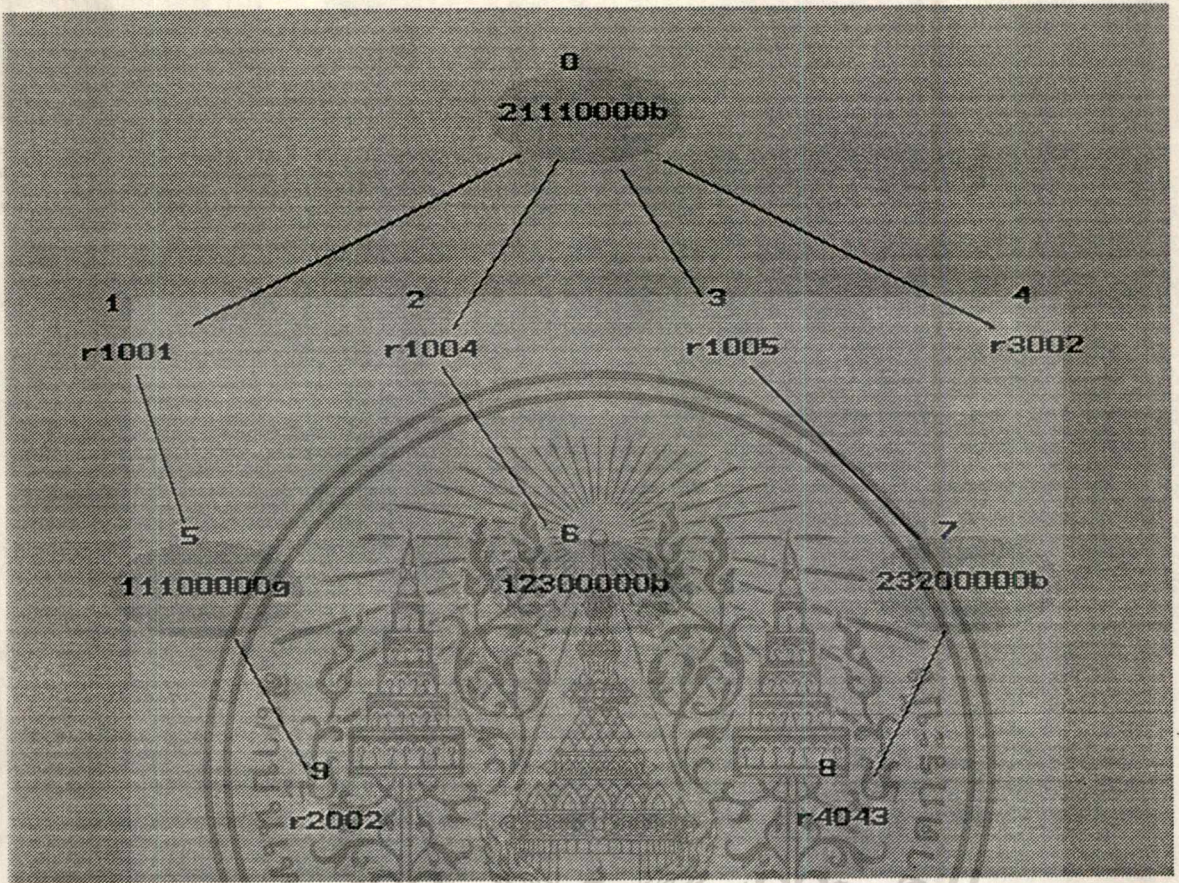
โปรแกรมระบบการแปลหลายภาษาด้วยคอมพิวเตอร์โดยใช้ภาษากลาง
พัฒนาโดย นายวิเชียร อัครธีระ
อาจารย์ที่ปรึกษา ผศ.ดร. รัตติกร วราวุฒิสิริพันธ์

Input filename:t15_



ตมขจติจกรรรมค้ำสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

รูปที่ 6.4 แสดงหน้าจอการรับข้อมูลทางแป้นพิมพ์



รูปที่ 6.5 การแสดงโครงสร้างภาษากลางในรูปโครงสร้างต้นไม้ฟังก์ชัน

คำสั่ง "Translation" เลื่อนแถบสว่างไปที่คำสั่งนั้นแล้วกด ENTER ระบบจะเข้าสู่การแปลภาษา ดังแสดงในรูปที่ 6.6 สามารถที่แปลภาษากลางไปเป็นประโยคภาษาไทย หรือภาษาอังกฤษได้ โดยการเลื่อนแถบสว่างไปที่ "Thai Sentence" เมื่อต้องการแปลไปเป็นภาษาไทย และ "English Sentence" เมื่อต้องการแปลเป็นภาษาอังกฤษ ซึ่งระบบจะให้ป้อนชื่อ File ดังแสดงในรูปที่ 6.4 เมื่อป้อนชื่อ File เสร็จแล้วกด ENTER ระบบจะแสดงประโยคภาษาเป้าหมายซึ่งสังเคราะห์จากภาษากลางดังรูปที่ 6.7

โปรแกรมระบบการแปลหลายภาษาด้วยคอมพิวเตอร์โดยใช้ภาษากลาง
มีผลมาโดย นายวิเชียร อัครวิระ
อาจารย์ที่ปรึกษา ผศ.ดร.จิตติกร วราวุศศิริพันธ์



รูปที่ 6.6 จอภาพแสดงคำสั่งในการแปลภาษากลางไปเป็นภาษาเป้าหมาย

โปรแกรมระบบการแปลหลายภาษาด้วยคอมพิวเตอร์โดยใช้ภาษากลาง
พัฒนาโดย นายวิเชียร อัสวาธีระ
อาจารย์ที่ปรึกษา ผศ.ดร.จิตติกร วราวุฒศิริพันธ์

Input filename:t15

พวกเราไปโรงเรียนในวันจันทร์_

คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

รูปที่ 6.7 จอภาพแสดงผลลัพธ์ของการสังเคราะห์ภาษาเป้าหมายจากภาษากลาง

คำสั่ง "Dictionary" เป็นคำสั่งที่สามารถแก้ไข เพิ่ม และดูข้อมูลในพจนานุกรมเพื่อประโยชน์
ในการแปล เมื่อเลื่อนแถบสว่างไปที่ คำสั่ง "Dictionary" แล้วกด ENTER จอภาพจะแสดงดังรูปที่ 6.8

โปรแกรมระบบการแปลหลายภาษาด้วยคอมพิวเตอร์โดยใช้ภาษากลาง
มีดทา โดย นายวิเชียร อิศวธีรช
อาจารย์ที่ปรึกษา ผศ.ดร. รัตติกร วราวุฒศิริพันธ์

Basic Dictionary

Thai Dictionary

English Dictionary

Quit Dictionary

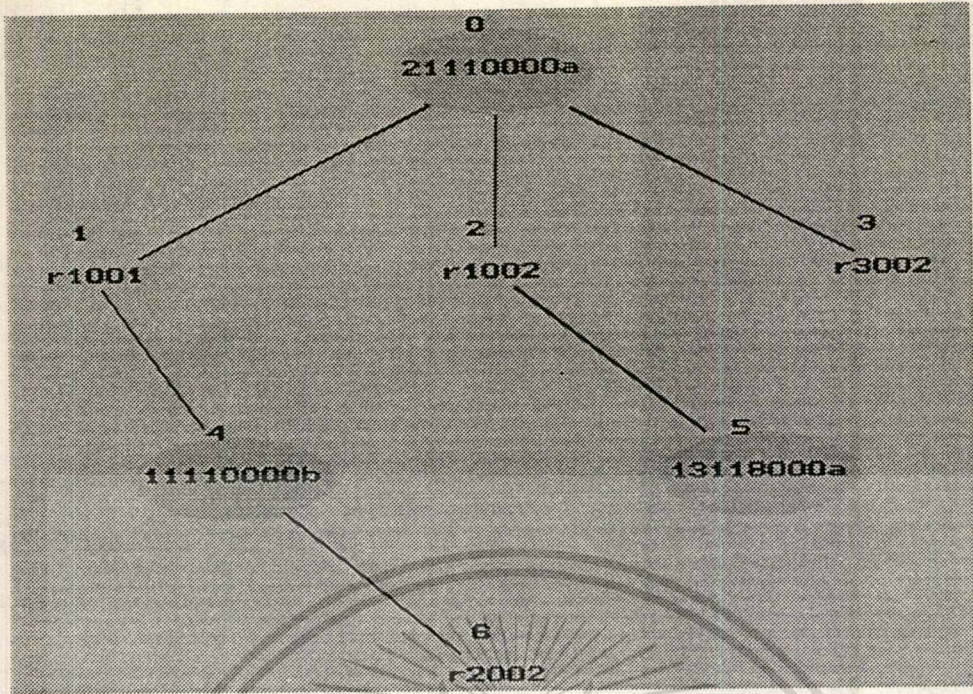
คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

รูปที่ 6.8 จอภาพแสดงพจนานุกรมรวมภาษา

คำสั่ง "QUIT" เป็นคำสั่งเมื่อต้องการเลิกใช้ซอฟต์แวร์

ตัวอย่างประโยคที่ได้ทำการวิจัยและสังเคราะห์ไปเป็นภาษาเป้าหมาย ทั้งภาษาไทยและภาษาอังกฤษ โดยการวิเคราะห์ประโยคต้นแบบทั้งภาษาไทย และภาษาอังกฤษไปเป็นภาษากลางด้วยมือ และสร้างระบบการสังเคราะห์ภาษากลางด้วยโปรแกรมคอมพิวเตอร์มาทำการสังเคราะห์ประโยคตัวอย่างนี้ ซึ่งได้ผลลัพธ์ดังแสดงในประโยคตัวอย่างทั้งหมด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 119 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ตัวอย่างที่ 1 ประโยคต้นแบบ "สมศักดิ์กินข้าว"

node.0

id.21110000a

1,2,3

-1

node.1

id.r1001

4

0

node.2

id.r1002

5

0

node.3

id.r3002

-1

0

node.4

id.11110000b

6

1

node.5

id.13118000a

-1

2

6
r2002

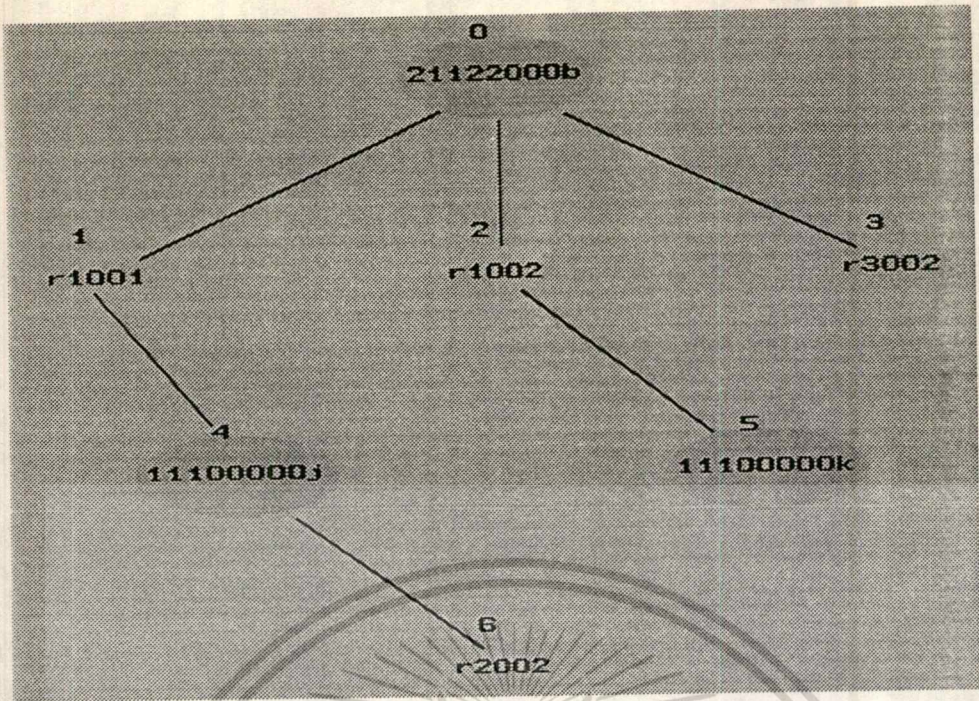
ประโยคเป้าหมาย

ประโยคภาษาไทย "สมศักดิ์กินข้าว"

ประโยคภาษาอังกฤษ "somsak eats rice"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา 120 จะต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ตัวอย่างที่ 2 ประโยคต้นแบบ "ฉันคิดถึงคุณ"

node.0
id.21122000b
1,2,3
-1
node.1
id.r1001
4
0
node.2
id.r1002
5
0

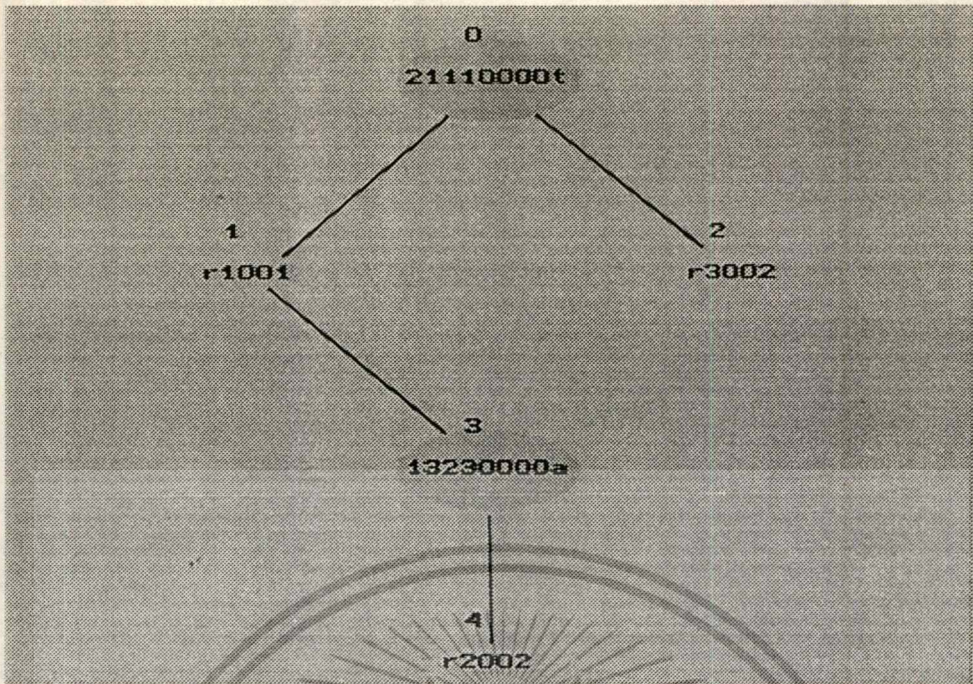
node.3
id.r3002
-1
node.4
id.11100000j
6
1
node.5
id.11100000k
-1
2

node.6
id.r2002
-1
4

ประโยคเป้าหมาย

ประโยคภาษาไทย "ฉันคิดถึงคุณ"

ประโยคภาษาอังกฤษ "I misses you"



ตัวอย่างที่ 3 ประโยคต้นแบบ "ประตูเปิด"

node.0
id.21110000t
1,2
-1
node.1
id.r1001
3
0
node.2
id.r3002
-1
0

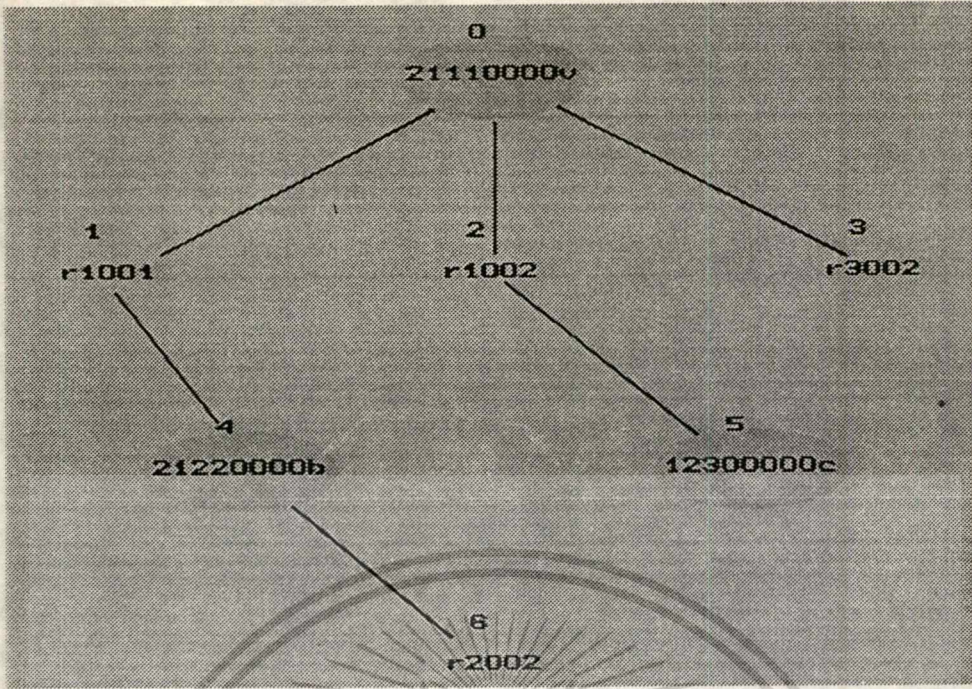
node.3
id.13230000a
4
1
node.4
id.r2002
-1
3

ประโยคเป้าหมาย

ประโยคภาษาไทย "ประตูเปิด"

ประโยคภาษาอังกฤษ "door opens"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 122 และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ตัวอย่างที่ 4 ประโยคต้นแบบ "ไฟไหม้บ้าน"

node.0
id.21110000v
1,2,3
-1
node.1
id.r1001
4
0
node.2
id.r1002
5
0

node.3
id.r3002
-1
node.4
id.21220000b
6
1
node.5
id.12300000c
-1
2

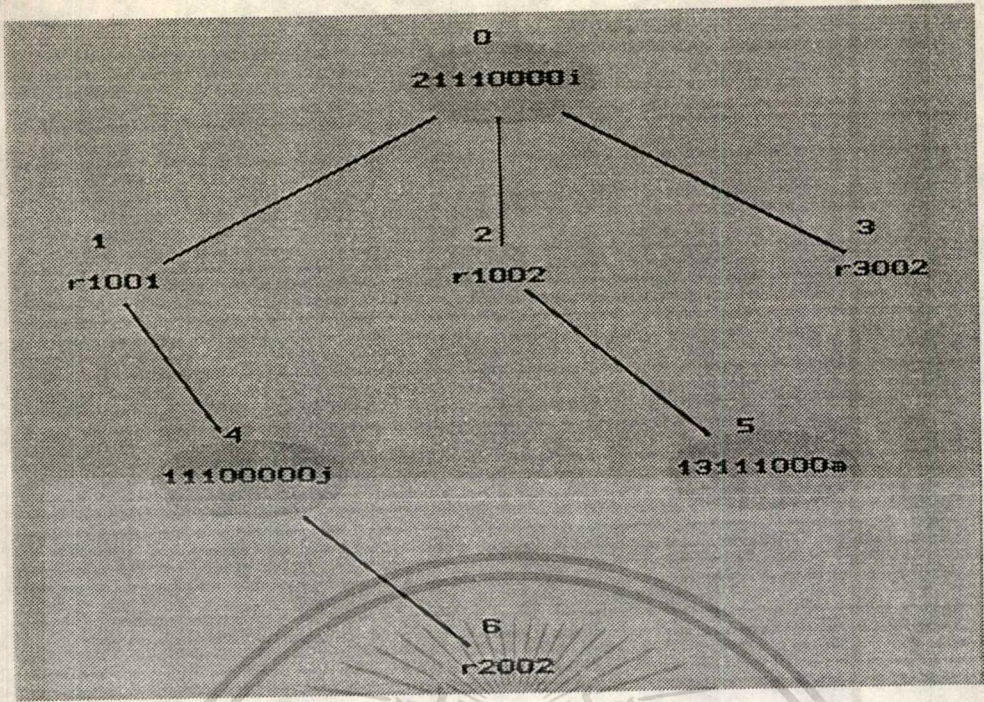
ประโยคเป้าหมาย

ประโยคภาษาไทย "ไฟไหม้บ้าน"

ประโยคภาษาอังกฤษ "fire burn home"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา ¹²³ ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ตัวอย่างที่ 5 ประโยคต้นแบบ "ฉันทิสุนัข"

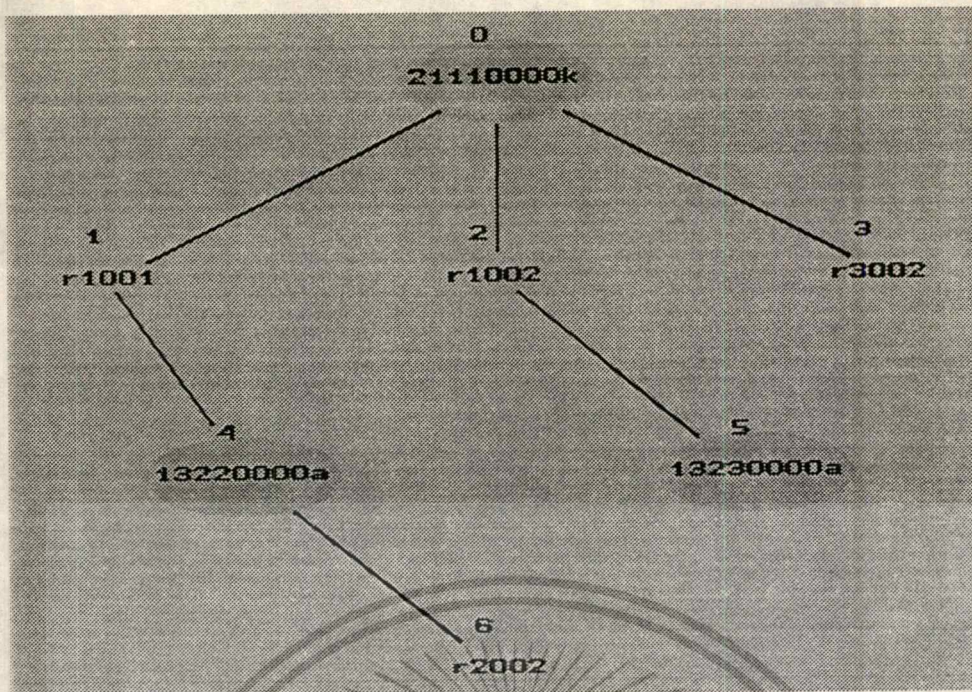
node.0	node.3	node.6
id.21110000i	id.r3002	id.r2002
1,2,3	-1	-1
-1	0	4
node.1	node.4	
id.r1001	id.11100000j	
4	6	
0	1	
node.2	node.5	
id.r1002	id.13111000a	
5	-1	
0	2	

ประโยคเป้าหมาย

ประโยคภาษาไทย "ฉันทิสุนัข"

ประโยคภาษาอังกฤษ "I hit dog"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 124 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ตัวอย่างที่ 6 ประโยคต้นแบบ "รถยนต์ชนประตู"

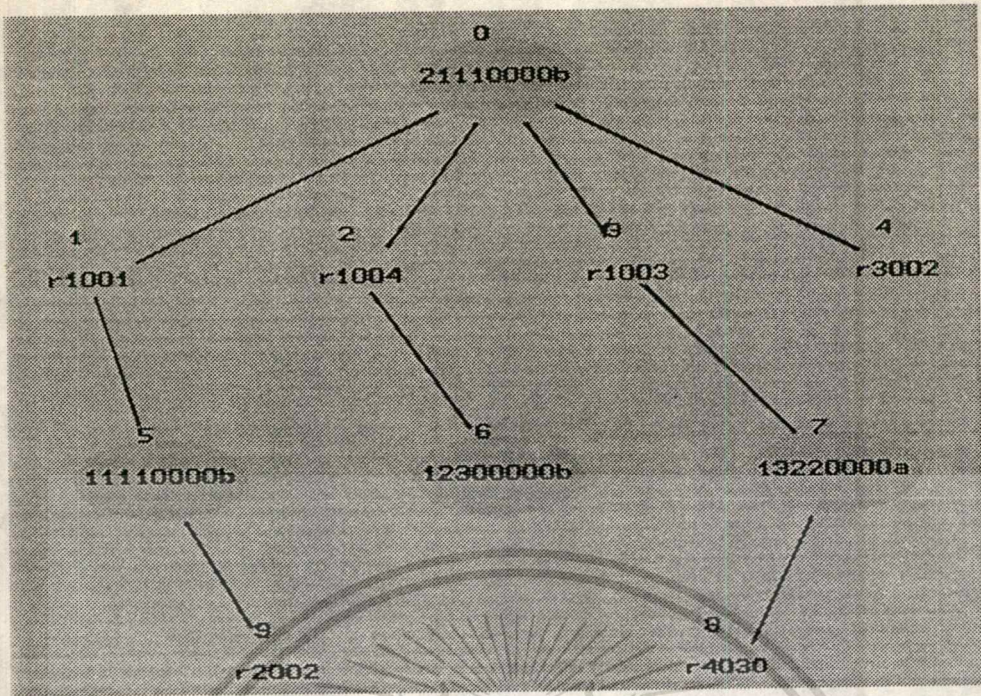
node.0
id.21110000k
1,2,3
-1
node.1
id.r1001
4
0
node.2
id.r1002
5
0

node.3
id.r3002
-1
node.4
id.13220000a
6
1
node.5
id.13230000a
-1
2

ประโยคภาษาเป้าหมาย

ประโยคภาษาไทย "รถยนต์ชนประตู"

ประโยคภาษาอังกฤษ "car bumps door"



ตัวอย่างที่ 7 ประโยคต้นแบบ "สมศักดิ์ไปโรงเรียนโดยรถยนต์"

node.0	node.3	node.6	node.9
id.21110000b	id.r1003	id.12300000b	id.r2002
1,2,3,4	7	-1	-1
-1	0	2	5
node.1	node.4	node.7	
id.r1001	id.r3002	id.13220000a	
5	-1	8	
0	0	3	
node.2	node.5	node.8	
id.r1004	id.11110000b	id.r4030	
6	9	-1	
0	1	7	

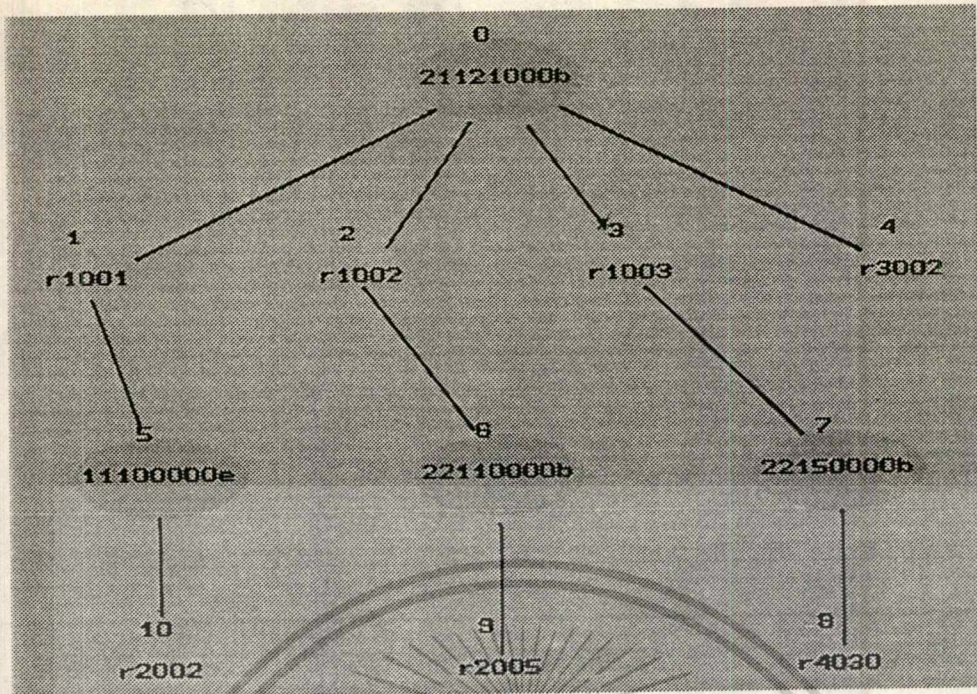
ประโยคภาษาเป้าหมาย

ประโยคภาษาไทย "สมศักดิ์ไปโรงเรียนโดยรถยนต์"

ประโยคภาษาอังกฤษ "somsak goes to school by car"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา 126 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ตัวอย่างที่ 8 ประโยคต้นแบบ "นักเรียนแก้ปัญหานี้ด้วยทฤษฎีเมตริกซ์"

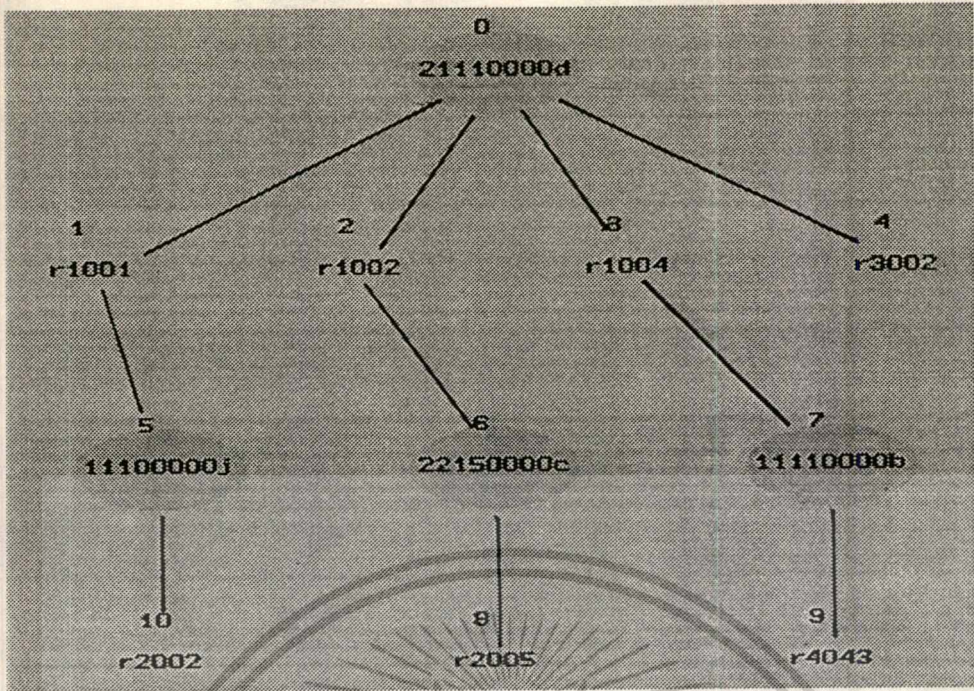
node.0	node.3	node.6	node.9
id.21121000b	id.r1003	id.22110000b	id.r2005
1,2,3,4	7	9	-1
-1	0	2	6
node.1	node.4	node.7	node.10
id.r1001	id.r3002	id.22150000b	id.r2002
5	-1	8	-1
0	0	3	5
node.2	node.5	node.8	
id.r1002	id.11100000e	id.r4030	
6	10	-1	
0	1	7	

ประโยคเป้าหมาย ประโยคภาษาไทย "นักเรียนแก้ปัญหานี้ด้วยทฤษฎีเมตริกซ์"

ประโยคภาษาอังกฤษ "student solves this problem by matric theory"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 127 จะต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ตัวอย่างที่ 9 ประโยคต้นแบบ "ฉันบอกเรื่องนี้แก่สมศักดิ์"

node.0	node.3	node.6	node.9
id.21110000d	id.r1004	id.22150000c	id.r4043
1,2,3,4	7	8	-1
-1	0	2	7
node.1	node.4	node.7	node.10
id.r1001	id.r3002	id.11110000b	id.r2002
5	-1	9	-1
0	0	3	5
node.2	node.5	node.8	
id.r1002	id.11100000j	id.r2005	
6	10	-1	
0	1	6	

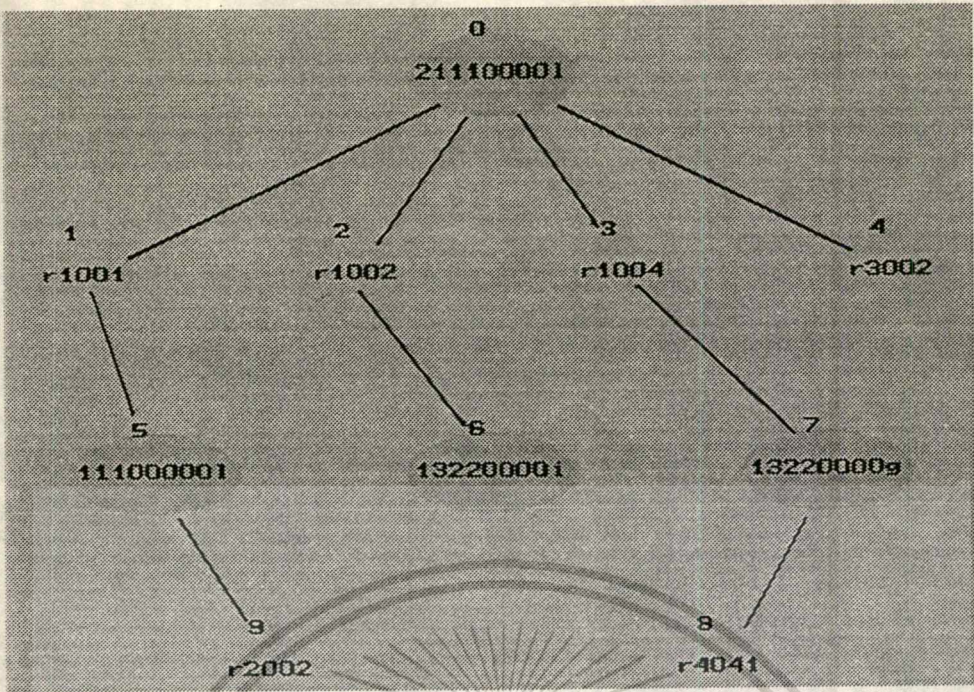
ประโยคต้นแบบ

ประโยคภาษาไทย "ฉันบอกเรื่องนี้แก่สมศักดิ์"

ประโยคภาษาอังกฤษ "i tell this story to somsak"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา 128 ละต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ตัวอย่างที่ 10 ประโยคต้นแบบ "แม่วางของเล่นในกล่อง"

node.0
id.211100001
1,2,3,4
-1
node.1
id.r1001
5
0
node.2
id.r1002
6
0

node.3
id.r1004
7
node.4
id.r3002
-1
0
node.5
id.111000001
9
1

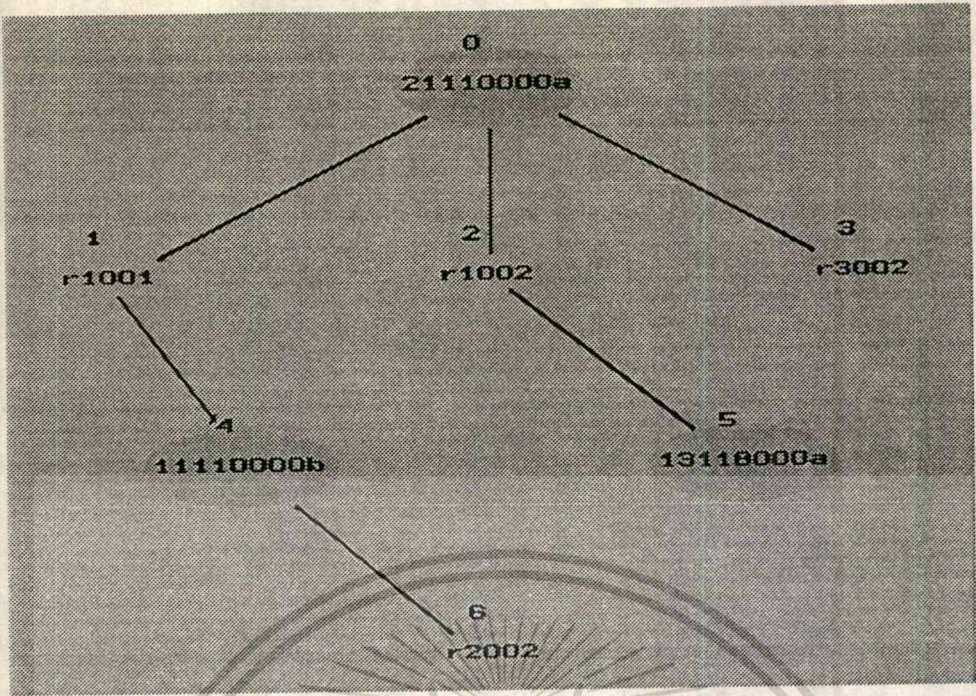
node.6
id.13220000i
-1
node.7
id.13220000g
8
3
node.8
id.r4041
-1
7

node.9
id.r2002
-1
5

ประโยคภาษาเป้าหมาย

ประโยคภาษาไทย "แม่วางของเล่นในกล่อง"

ประโยคภาษาอังกฤษ "mother puts toy in box"



ตัวอย่างที่ 11 ประโยคต้นแบบ "Somsak eats rice"

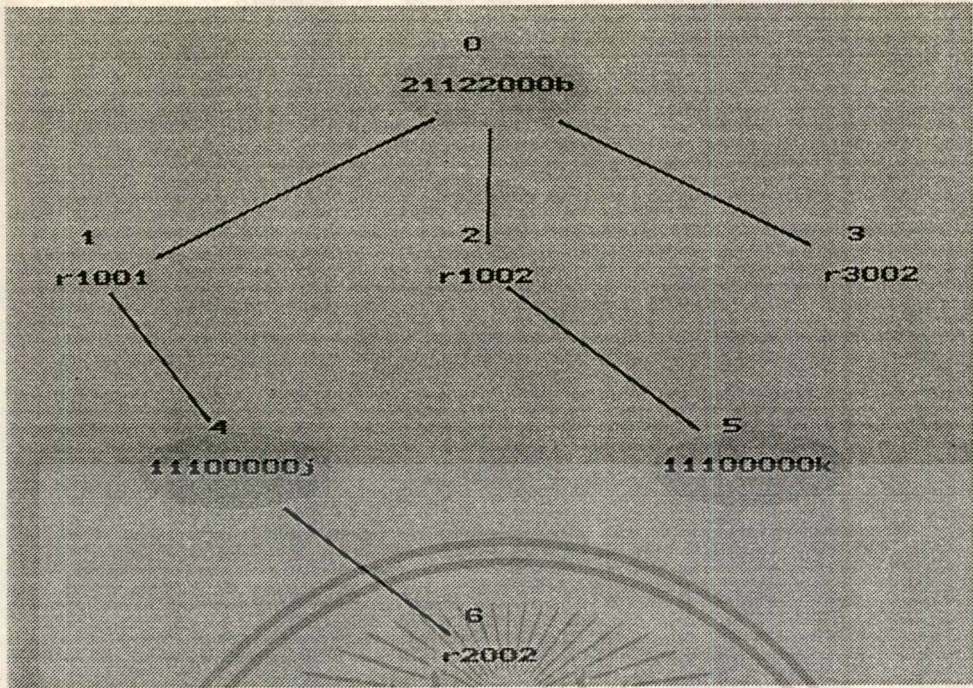
node.0	node.3	node.6
id.21110000a	id.r3002	id.r2002
1,2,3	-1	-1
-1	0	4
node.1	node.4	
id.r1001	id.11110000b	
4	6	
0	1	
node.2	node.5	
id.r1002	id.13118000a	
5	-1	
0	2	

ประโยคเป้าหมาย

ประโยคภาษาไทย "สมศักดิ์กินข้าว"

ประโยคภาษาอังกฤษ "somsak eats rice"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 130 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ตัวอย่างที่ 12 ประโยคต้นแบบ "I miss you"

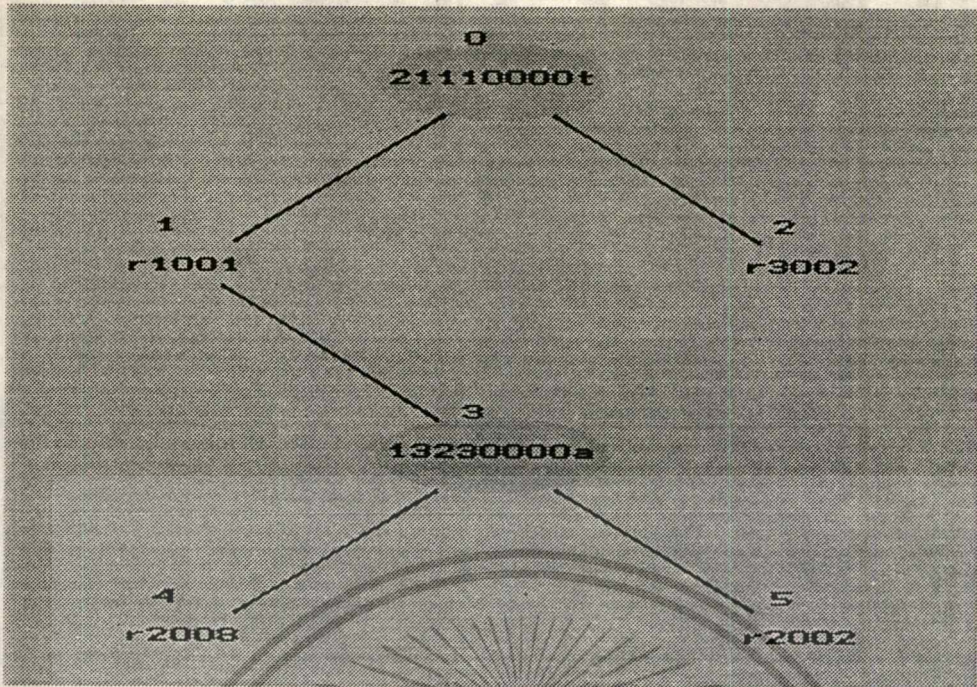
node.0
id.21122000b
1,2,3
-1
node.1
id.r1001
4
0
node.2
id.r1002
5
0

node.3
id.r3002
-1
node.4
id.11100000j
6
1
node.5
id.11100000k
-1
2

ประโยคเป้าหมาย ประโยคภาษาไทย "ฉันคิดถึงคุณ"

 ประโยคภาษาอังกฤษ "I miss you"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 131 จะต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ตัวอย่างที่ 13 ประโยคต้นแบบ "The door opens"

node.0
id.21110000t
1,2
-1
node.1
id.r1001
3
0
node.2
id.r3002
-1
0

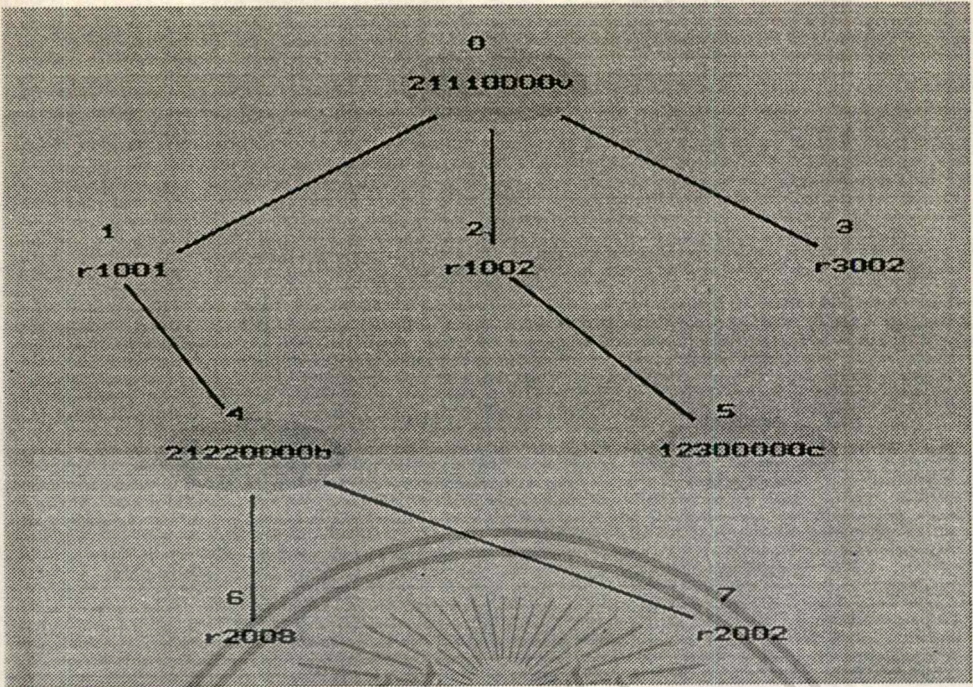
node.3
id.13230000a
4,5
1
node.4
id.r2008
-1
3
node.5
id.r2002
-1
3

ประโยคเป้าหมาย

ประโยคภาษาไทย "ประตูเปิด"

ประโยคภาษาอังกฤษ "The door opens"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 132 และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ตัวอย่างที่ 14 ประโยคต้นแบบ "The fire burn home"

node.0
id.21110000v
1,2,3
-1
node.1
id.r1001
4
0
node.2
id.r1002
5
0

node.3
id.r3002
-1
node.4
id.21220000b
6,7
1
node.5
id.12300000c
-1
2

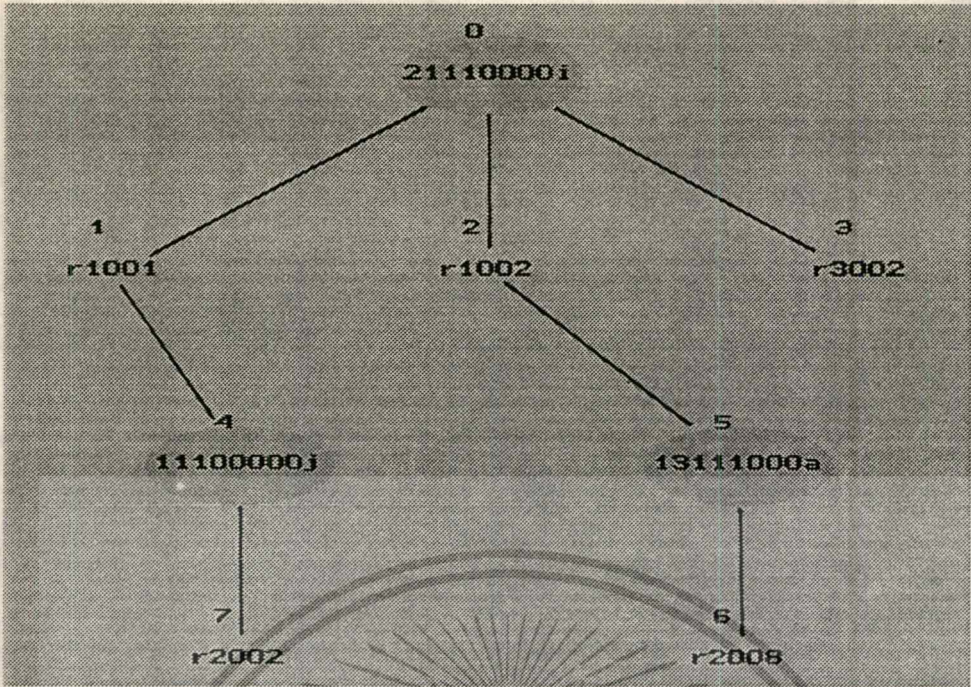
node.6
id.r2008
-1
node.7
id.r2002
-1
4

ประโยคเป้าหมาย ประโยคภาษาไทย "ไฟไหม้บ้าน"

ประโยคภาษาอังกฤษ "The fire burn home"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ตัวอย่างที่ 15 ประโยคต้นแบบ "I hit the dog"

node.0

id.21110000i

1,2,3

-1

node.1

id.r1001

4

0

node.2

id.r1002

5

0

node.3

id.r3002

-1

0

node.4

id.11100000j

7

1

node.5

id.13111000a

6

2

node.6

id.r2008

-1

5

node.7

id.r2002

-1

4

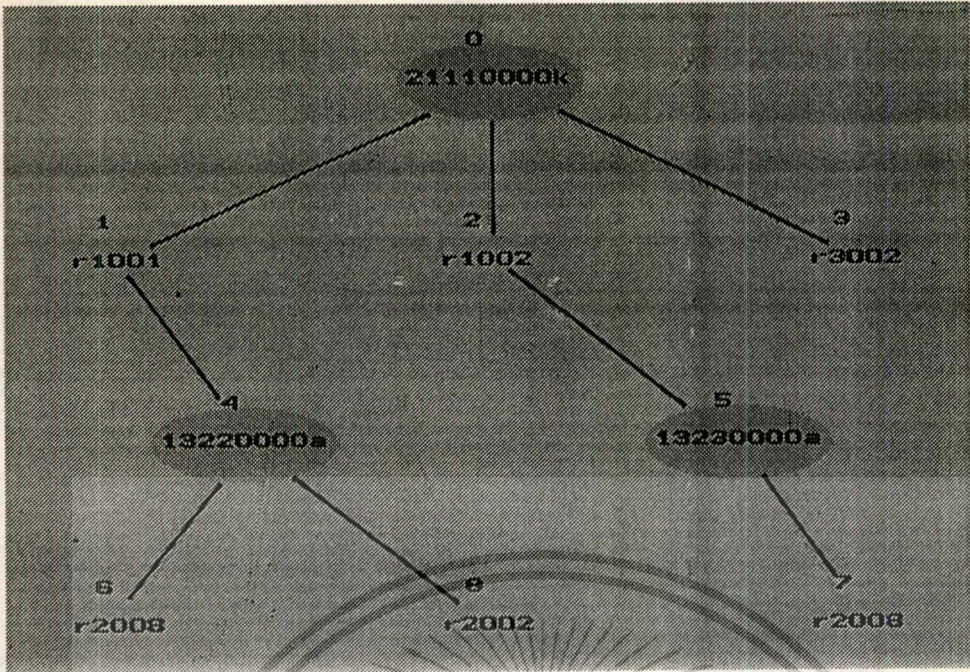
ประโยคเป้าหมาย

ประโยคภาษาไทย "ฉันตีสุนัข"

ประโยคภาษาอังกฤษ "I hit the dog"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา 134 ละต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ตัวอย่างที่ 16 ประโยคต้นแบบ "The car bumps the door"

node.0	node.3	node.6
id.21110000k	id.r3002	id.r2008
1,2,3	-1	-1
-1	0	4
node.1	node.4	node.7
id.r1001	id.13220000a	id.r2008
4	6,8	-i
0	1	5
node.2	node.5	node.8
id.r1002	id.13230000a	id.r2002
5	7	-1
0	2	4

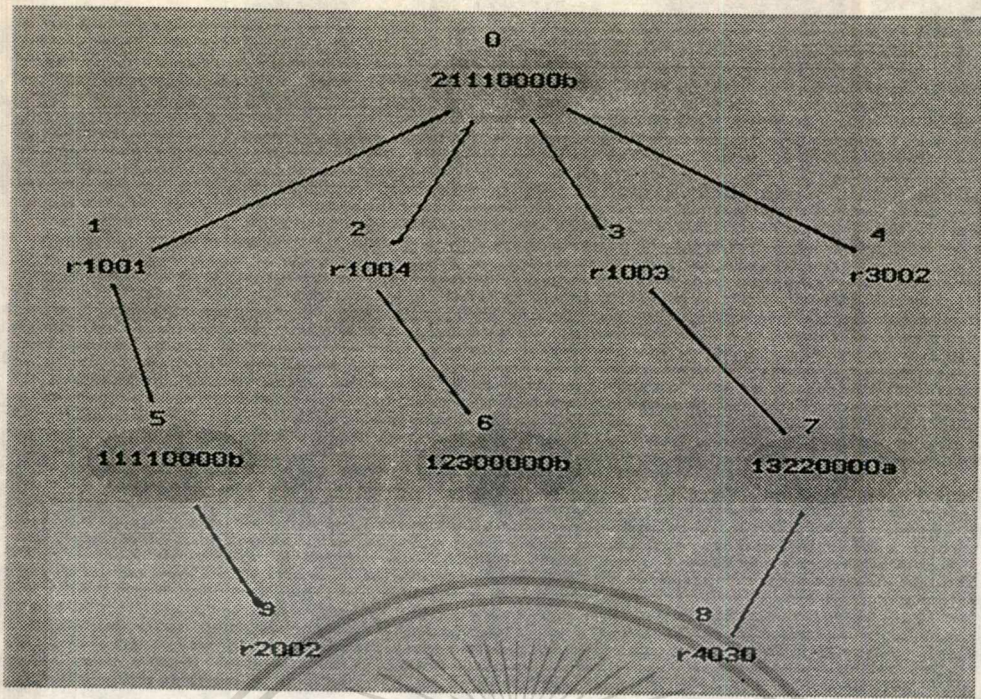
ประโยคภาษาเป้าหมาย

ประโยคภาษาไทย "รถยนต์ชนประตู"

ประโยคภาษาอังกฤษ "The car bumps the door"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ตัวอย่างที่ 17 ประโยคต้นแบบ "Somsak goes to school by car"

node.0	node.3	node.6	node.9
id.21110000b	id.r1003	id.12300000b	id.r2002
1,2,3,4	7	-1	-1
-1	0	2	5
node.1	node.4	node.7	
id.r1001	id.r3002	id.13220000a	
5	-1	8	
0	0	3	
node.2	node.5	node.8	
id.r1004	id.11110000b	id.r4030	
6	9	-1	
0	1	7	

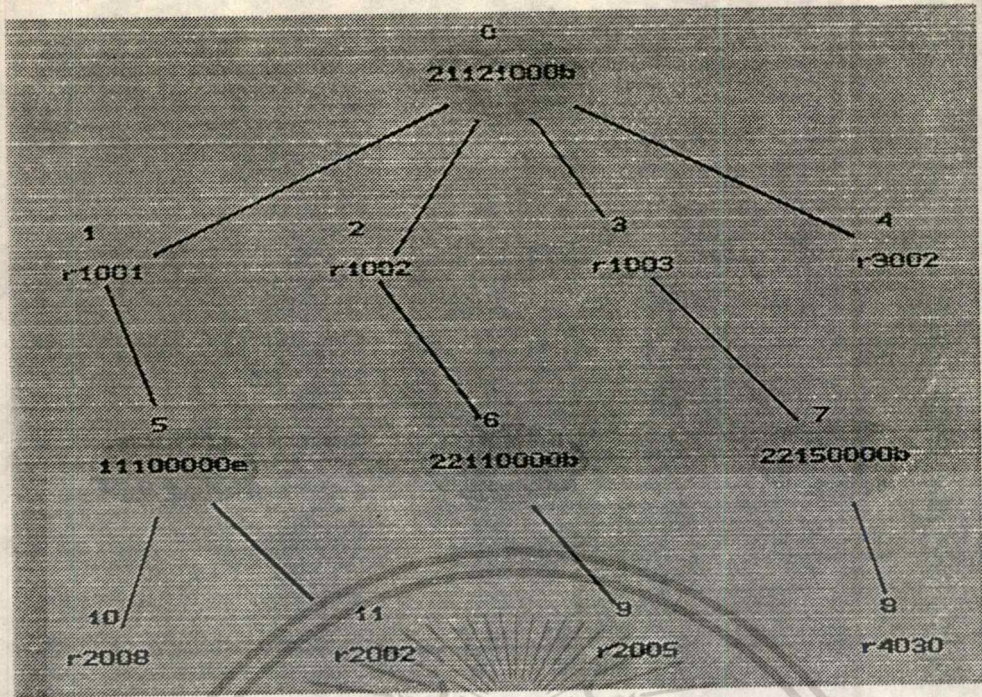
ประโยคภาษาเป้าหมาย

ประโยคภาษาไทย "สมศักดิ์ไปโรงเรียนโดยรถยนต์"

ประโยคภาษาอังกฤษ "Somsak goes to school by car"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ตัวอย่างที่ 18 ประโยคต้นแบบ "The student solves this problem by metric theory"

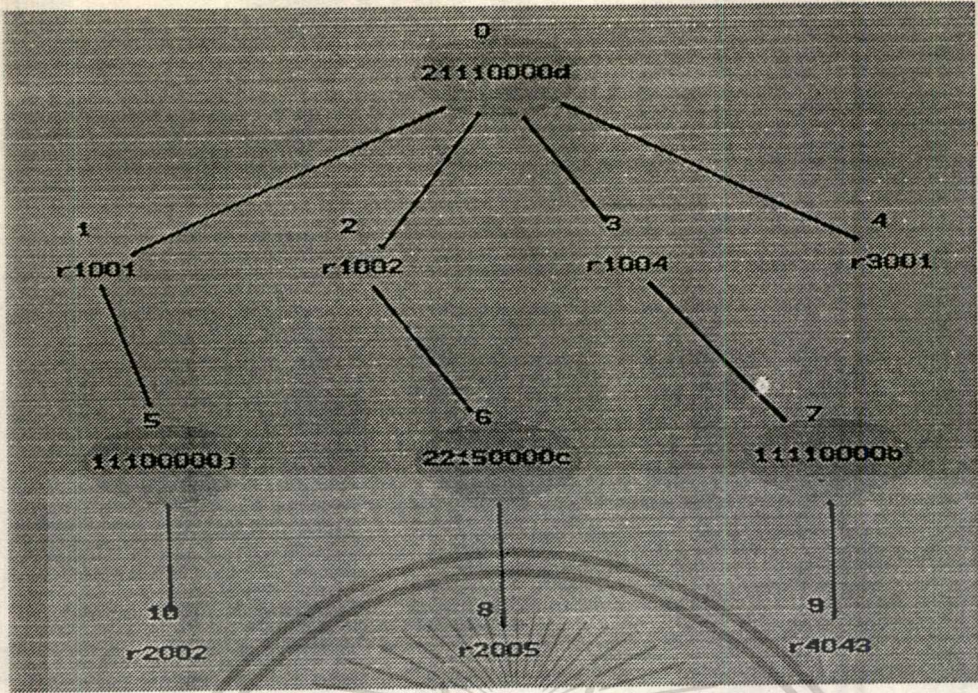
node.0	node.3	node.6	node.9
id.21121000b	id.r1003	id.22110000b	id.r2005
1,2,3,4	7	9	-1
-1	0	2	6
node.1	node.4	node.7	node.10
id.r1001	id.r3002	id.22150000b	id.r2008
5	-1	8	-1
0	0	3	5
node.2	node.5	node.8	node.11
id.r1002	id.11100000e	id.r4030	id.r2002
6	10,11	-1	-1
0	1	7	5

ประโยคเป้าหมาย

ประโยคภาษาไทย "นักเรียนแก้ปัญหานี้ด้วยทฤษฎีเมตริกซ์"

ประโยคภาษาอังกฤษ "The student solves this problem by metric theory"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 137 ละต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ตัวอย่างที่ 19 ประโยคต้นแบบ "I tell this story to Somsak"

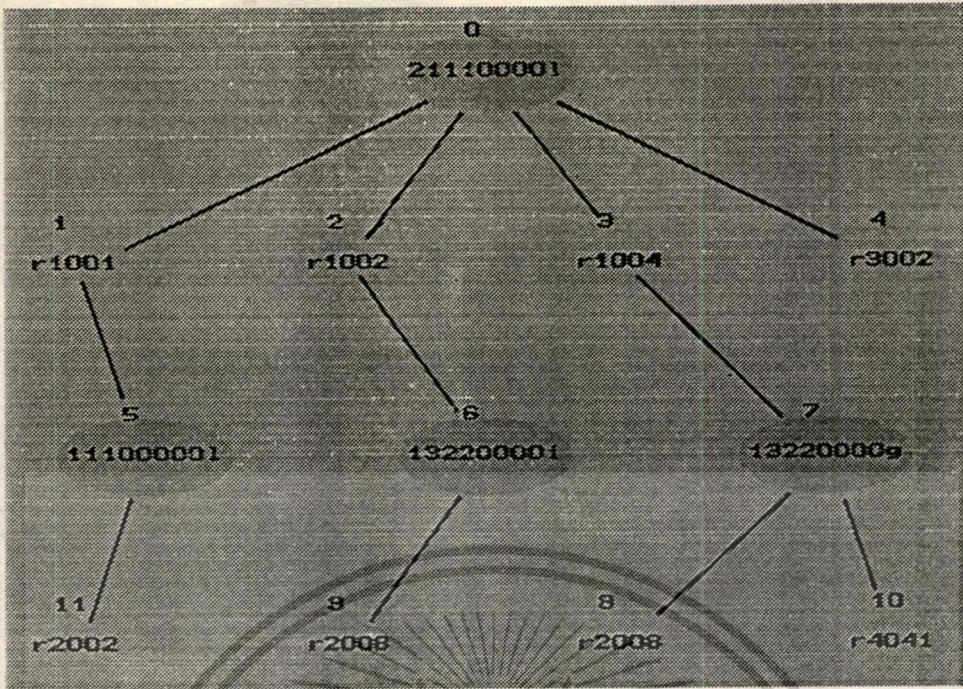
node.0	node.3	node.6	node.9
id.21110000d	id.r1004	id.22150000c	id.r4043
1,2,3,4	7	8	-1
-1	0	2	7
node.1	node.4	node.7	node.10
id.r1001	id.r3002	id.11110000b	id.r2002
5	-1	9	-1
0	0	3	5
node.2	node.5	node.8	
id.r1002	id.11100000j	id.r2005	
6	10	-1	
0	1	6	

ประโยคต้นแบบ ประโยคภาษาไทย "ฉันบอกเรื่องนี้แก่สมศักดิ์"

ประโยคภาษาอังกฤษ "I tell this story to somsak"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา 138 จะต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ตัวอย่างที่ 20 ประโยคต้นแบบ "Mother puts the toy in the box"

node.0	node.3	node.5	node.9
id.211100001	id.r1004	id.132200001	id.r2008
1,2,3,4	7	9	-1
-1	0	2	6
node.1	node.4	node.7	node.10
id.r1001	id.r3002	id.13220000g	id.r4041
5	-1	8,10	-1
0	0	3	7
node.2	node.5	node.8	node.11
id.r1002	id.111000001	id.r2008	id.r2002
6	11	-1	-1
0	1	7	5

ประโยคภาษาเป้าหมาย

ประโยคภาษาไทย "แม่วางของเล่นในกล่อง"

ประโยคภาษาอังกฤษ "Mother puts the toy in the box"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 139 จะต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6.4 สรุป

ระบบการสังเคราะห์ภาษาเป้าหมายจากภาษากลางเพื่อการแปลหลายภาษาด้วยคอมพิวเตอร์ของงานวิจัยนี้ กำหนดการทำงานไว้ 3 ส่วนคือ ส่วนแสดงโครงสร้างภาษากลางในรูปโครงสร้างต้นไม้ ส่วนการแปลไปเป็นภาษาเป้าหมายทั้งภาษาไทยและภาษาอังกฤษ ส่วนพจนานุกรมซึ่งสามารถที่จะสามารถที่จะเพิ่มหรือแก้ไข ในกรณีที่ไม่มีคำศัพท์ หรือมีการป้อนข้อมูลผิดพลาด ภาษาต้นแบบที่ทำการวิจัยนี้เป็นประโยคความเดียว ไม่มีการละคำ วิเคราะห์ให้เป็นโครงสร้างภาษากลางด้วยมือโดยอาศัยหลักไวยากรณ์ทางภาษาศาสตร์



บทที่ 7

บทสรุปและข้อเสนอแนะ

7.1 สรุป

การพัฒนากระบวนการแปลภาษาด้วยคอมพิวเตอร์โดยใช้ภาษากลางเป็นสื่อ เป็นระบบที่มีการวิจัยและพัฒนาอย่างต่อเนื่องทั้งในและนอกประเทศ จากการศึกษากระบวนการแปลโดยใช้ภาษากลางพบว่าภาษากลางจะประกอบด้วยส่วนหลัก 2 ส่วนคือ หน่วยความหมาย (Conceptual Primitive) และหน่วยความสัมพันธ์ (Relation) โดยมีลูกศรความสัมพันธ์เป็นตัวเชื่อมโยงและบอกทิศทางของความสัมพันธ์ของหน่วยความหมาย

1) หน่วยความหมาย

ในวิทยานิพนธ์นี้กำหนดหน่วยความหมายในรูปของรหัสแยกแยะความหมายที่แบ่งกลุ่มตามความสัมพันธ์ลดหลั่น (hierarchy) ซึ่งต่างจากการกำหนดหน่วยความหมายของภาษากลางที่มีการวิจัยอื่นๆ ซึ่งส่วนใหญ่กำหนดหน่วยความหมายในรูปแบบผิวของคำ (Surface Form of Word) โดยแทนรูปแบบผิวของคำให้เป็นคำในภาษาใดภาษาหนึ่ง เช่น ภาษาอังกฤษ ภาษาไทย การกำหนดในลักษณะนี้จะมีข้อดีคือ ผู้ใช้สามารถเข้าใจความหมายของรูปแบบผิวของคำได้อย่างคร่าวๆ แต่ถ้าผู้ใช้ที่ไม่รู้จักภาษานี้หรือมีความรู้เกี่ยวกับภาษานั้นน้อย ก็ไม่สามารถที่จะเข้าใจหรือเข้าใจแบบกำกวม และคำบางคำที่ไม่มีรูปแบบผิวในภาษานั้นก็ไม่สามารถที่จะแสดงรูปแบบผิวได้หรือถ้าแสดงก็อาจจะมีการเข้าใจความหมายที่ผิดพลาดได้ ส่วนหน่วยความหมายในวิทยานิพนธ์นี้กำหนดเป็นรหัสแยกแยะความหมายโดยแบ่งกลุ่มความสัมพันธ์ลดหลั่นของคำดังรูปที่ 3.2 เพื่อให้ผู้ใช้และคอมพิวเตอร์ทราบถึงประเภทของกลุ่มคำนี้ โดยไม่ต้องคำนึงถึงรูปแบบผิวของคำว่าจะเป็นภาษาใด แต่ผู้ใช้ต้องเข้าใจกลุ่มความหมายของรหัสแยกแยะ

2) หน่วยความสัมพันธ์

ในวิทยานิพนธ์นี้กำหนดหน่วยความสัมพันธ์ในรูปรหัสเพื่อแบ่งกลุ่มหน้าที่ของความสัมพันธ์ที่เกิดขึ้นในประโยคโดยแบ่งเป็น 4 กลุ่มความสัมพันธ์คือ กลุ่มความสัมพันธ์หลัก กลุ่มความสัมพันธ์เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 141 จะต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คุณสมบัติ กลุ่มความสัมพันธ์สถานการณ์ กลุ่มความสัมพันธ์คำเชื่อม ซึ่งแต่ละกลุ่มจะเป็นความสัมพันธ์เสริมของกลุ่มความสัมพันธ์หลัก และมีความสัมพันธ์ต่อกัน เพื่อขั้นตอนในการประมวลผลของคอมพิวเตอร์ โดยผู้ใช้งานต้องทราบความหมายของรหัสต่างๆ ของรหัสความสัมพันธ์

7.2 ข้อเสนอแนะ

1) การกำหนดรหัสแยะแยะหน่วยความหมายในวิทยานิพนธ์นี้ยังกำหนดให้มีความละเอียดไม่พอที่จะครอบคลุมกลุ่มความหมายของคำที่มีอยู่ทั้งหมด ควรจะมีการศึกษาจากนักภาษาศาสตร์คอมพิวเตอร์โดยตรงเพื่อกำหนดกลุ่มความหมายของคำที่สามารถนำมาประยุกต์ใช้กับระบบคอมพิวเตอร์

2) การกำหนดรหัสความสัมพันธ์ในวิทยานิพนธ์นี้ได้ศึกษาจากประโยคความเดียวของภาษาไทย ภาษาอังกฤษ และภาษาญี่ปุ่น ซึ่งยังไม่ครอบคลุมความแตกต่างทางด้านภาษาทั้งหมด และถ้าเป็นประโยคความรวม หรือประโยคความซ้อนที่มีกลุ่มของกริยาวลี (Verb Pharse) และนามวลี (Noun Pharse) หลายๆ กลุ่มก็จะทำให้ยากแก่การเข้าใจ ควรจะกำหนดความสัมพันธ์ที่บ่งบอกถึงกลุ่มของกริยาวลี และนามวลีด้วย

3) ควรสร้างระบบซอฟต์แวร์วิเคราะห์ภาษาต้นแบบให้อยู่ในรูปของภาษากลาง

4) ระบบซอฟต์แวร์สังเคราะห์ภาษาเป้าหมายควรที่จะสามารถแก้ไขภาษากลางที่เกิดจากการวิเคราะห์ภาษาต้นแบบได้ เพราะระบบซอฟต์แวร์ที่ทำการวิเคราะห์อาจจะมึข้อผิดพลาด และสามารถตรวจสอบระบบการสังเคราะห์ได้ด้วย

5) ข้อมูลในพจนานุกรมควรกำหนดให้มากพอสำหรับการวิเคราะห์และสังเคราะห์ ถ้าเป็นไปได้ ควรสร้างระบบการแปลภาษาด้วยคอมพิวเตอร์เฉพาะด้านจะทำให้การกำหนดข้อมูลคำศัพท์ครอบคลุมกลุ่มความหมายของคำที่จะแปล และจะได้ระบบที่สมบูรณ์มากขึ้น

6) โครงสร้างการเก็บข้อมูลในพจนานุกรมควรจะเป็นลักษณะโครงข่ายคล้ายกับสมองมนุษย์ที่เรียกว่านิเวรอนเน็ตเวิร์ค (Neural Network)

7.3 การนำผลงานวิจัยไปประยุกต์ใช้กับงานอื่น

ผลงานในวิทยานิพนธ์นี้สามารถที่จะนำไปประยุกต์ใช้กับงานด้านการประมวลผลภาษาธรรมชาติ (Natural Language Processing) สาขาต่างๆ ทั้งในด้านภาษาศาสตร์ และด้านการประมวลผลภาษาทางคอมพิวเตอร์ เช่น การประยุกต์ใช้กับระบบการแปลหลายภาษาด้วยคอมพิวเตอร์



เอกสารอ้างอิง

- [1] รองศาสตราจารย์ ดร.นิตยา กาญจนะวรรณ, "การแปลภาษาด้วยเครื่องทฤษฎีและวิธีการ", สำนักพิมพ์มหาวิทยาลัยรามคำแหง, 20 สิงหาคม 2534
- [2] มนทิรา ประชาฤทธิ์ภักดี, "การพัฒนารูปแบบภาษากลางเพื่อการแปลภาษาอังกฤษเป็นภาษาไทย", วิทยานิพนธ์ปริญญาโทมหาบัณฑิต คณะศิลปศาสตร์ มหาวิทยาลัยธรรมศาสตร์, 2533
- [3] นวลทิพย์ ตันติเศวตรรัตน์, วรนุช เกิดสินธุ์ชัย, กัลยา นฤดมกุล, บุญเจริญ ศิริเนาวกุล, "การแปลภาษาด้วยคอมพิวเตอร์", บทความในปริทัศน์คอมพิวเตอร์แมกกาซีน ปีที่ 1 ฉบับที่ 11 เดือนมกราคม 2533
- [4] ยืน ภู่วรรณ ชัยยงค์ วงศ์ชัยสุวัฒน์, "การประมวลผลภาษาธรรมชาติ", สถาบันเทคโนโลยีพระจอมเกล้าธนบุรี
- [5] ผศ. นวลทิพย์ ตันติเศวตรรัตน์, นางกุลวดี แหยมเกตต์, นางสาวธีรนุช โชคสุวณิช, นางสาวแก้วใจ จันท์เจริญ, นางสาวมณฑิลา บริบูรณ์, นางสาวนิศานาถ แทนนิต, "โครงการพัฒนาพจนานุกรมเพื่อการแปลภาษาด้วยคอมพิวเตอร์", บทความในการแปลภาษาด้วยคอมพิวเตอร์โครงการพัฒนา ระบบเครื่องแปลภาษาสำหรับภาษาในเอเชีย
- [6] ภานุ สังขะวร, "ความสัมพันธ์ทางอรรถศาสตร์ระหว่างคำนามกับคำกริยาในประโยคภาษาไทย", วิทยานิพนธ์ปริญญาโทมหาบัณฑิต คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2526
- [7] Hiroshi UCHIDA, Meiying ZHU, "Interlingua", Center of the Cooperation for Computriation, 1990
- [8] นววรรณ พันธุ์เมธา, "ไวยากรณ์ไทย", ศูนย์หนังสือจุฬาลงกรณ์มหาวิทยาลัย, 2527
- [9] ดร.รัตติกว วรากุลศิริพันธ์ุ, "ระบบการสร้างภาษาไทยเพื่อการแปลภาษาด้วยคอมพิวเตอร์", "การประชุมวิชาการ ครั้งที่ 2 ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติปี 2532
- [10] ดร.รัตติกว วรากุลศิริพันธ์ุ, นายวิเชียร อัครธีระ, "การวิเคราะห์ภาษาต้นแบบให้เป็นระบบภาษากลาง", สัมมนาระดับปริญญาโท คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

[11] ดร.รัตติกว วรากุลศิริพันธ์, นายวิเชียร อัสวธีระ, "การออกแบบสื่อภาษากลางเพื่อการแปลหลายภาษาด้วยคอมพิวเตอร์", บทความในประชุมใหญ่วิชาการทางวิศวกรรม ประจำปี 2534 ในวันที่ 31 ตุลาคม - 3 พฤศจิกายน 2534

[12] Longman Group, "Dictionary of Contemporary English", Richard Clay Ltd. Bungay, Suffolk, 1987

[13] สัจฉวี สายบัว, "หลักการแปล ฉบับพิมพ์ครั้งที่ 4", สำนักพิมพ์มหาวิทยาลัยธรรมศาสตร์, 2536



ภาคผนวก ก.

พจนานุกรมความหมาย

CPID	CPDC	CAT	SUBCAT	TYPE
11100000a	a person who has the qualification to teach someone something in school.	N	NCNM	S
11100000b	the male person.	N	NCMN	M
11100000c	the female person.	N	NCMN	W
11100000d	human being.	N	NCMN	S
11100000e	(undergraduate or post graduate) person who is studying at a college polytechnic or university	N	NCMN	S
11100000f	unborn or newly born human being, boy or girl, son or daughter (of any age)	N	NCMN	S
11100000g	used by a speaker or writer referring to himself and another or others.	PRON	PPRS	P
11100000h	man, woman (which are usually to be preferred) as is people for persons, is often derog except when official or impersonal	N	NCMN	S
11100000i	man who owns or manages a peasant	N	NCMN	S
11100000j	used by a speaker or writer to refer to himself	PRON	PPRS	S
11100000k	the person(s) addressed	PRON	PPRS	
11100000l	female parent; woman who has adopted a child; woman (often "housemother")	N	NCMN	S
11100000m	son of the same parents as another person	N	NCMN	M
11100000n	person, not a relation, when one knows and likes well	N	NCMN	S
11100000o	person who controls a business, a hotel, etc	N	NCMN	S
11100000p	male child of a parent	N	NCMN	M

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปเผยแพร่ภายนอก

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา 146 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1110000q	parents and children	N	NCMN	S
11110000a	a human being name vichean.	N	NPRP	M
11110000b	a human being name somsak.	N	NPRP	M
11110000c	a human being name sakchai.	N	NPRP	M
11110000d	a human being name somsri.	N	NPRP	W
11110000e	a human being name john.	N	NPRP	M
11200000a	governing power to govern, ministry, body of persons governing a state	N	NCMN	S
12100000a	large administrative division of a country	N	NCMN	S
12100000b	part of a country	N	NCMN	S
12100000c	centre of population larger than a village	N	NCMN	S
12110000a	the land belonging Thai nation.	N	NPRP	S
12110000b	the land belonging Japan nation.	N	NPRP	S
12110000c	the land belonging American nation.	N	NPRP	S
12110000d	a town which serves as the center of Thai government.	N	NPRP	S
12220000a	star forming center of system of worlds or planets; central body of solar system.	N	NPRP	S
12220000b	satellite revolving round earth in lunar month and reflecting light from sun.	N	NCMN	S
12300000a	corporate institution with power of conferring degrees and providing instruction in higher branches of learning.	N	NCMN	S
12300000b	institution for educating children or giving instruction.	N	NCMN	S
12300000c	a place where one can rest and be safe.	N	NCMN	S
12300000d	a town having existed since a very early time	N	NCMN	S
13111000a	common domestic animal, a friend of man, of which there are many breeds	N	NCMN	S

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 147 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

13111000b	small,domestic, fur-covered animal often kept as a pet,to catch mice.	N	NCMN	S
13112000a	cold-blooded animal living wholly in water and breathing through gills,with fins for swimming.	N	NCMN	S
13112000b	name piranha is tropical american freshwater fish,noted for attacking and eating live animals.	N	NPRP	S
13113000a	feathered creature with two legs and two wings, usually able to fly	N	NCMN	S
13114000a	small, cold-blooded, tail-less jumping animal living in water and on land	N	NCMN	S
13115000a	a plant whose seeds or grain are edible.	N	NCMN	S
13115000b	a plant name rose;beautiful and use sweet-smelling flower.	N	NPRP	S
13115000c	a plant	N	NCMN	S
13117000a	part of the human arm beyond the wrist	N	NCMN	S
13117000b	one of the parts of an animal's or a person's body used for walking	N	NCMN	S
13117000c	hard outer covering of bird's eggs, nuts some seeds and fruits and of some animals or parts of them	N	NCMN	S
13117000d	either of the two upper limbs of the human body, from the shoulder to the hand	N	NCMN	S
13118000a	pearl-white grain used as a staple food.	N	NCMN	S
13118000b	beautiful and use sweet-smelling flower,a part of rose tree.	N	NCMN	S
13118000c	that part of a plant and use sweet-smelling flower.	N	NCMN	S
13118000d	the fruit of an apple, skin that is red, green or yellow when it is ripe	N	NCMN	S
13118000e	the fruit of a plant named mango	N	NCMN	S

เอกสารนี้เป็นเอกสารที่ สงวนลิขสิทธิ์ไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

13118000f	that part of a plant or tree that contains the seeds and is used as food	N	NCMN	S
13118000g	the fruit of a plant named durain	N	NCMN	S
13121000a	precious yellow metal used for making coins,ornaments,jewellery.	N	NCMN	S
13121000b	frozen water;water made solid cold.	N	NCMN	S
13122000a	liquid which does not mix with water,obtained from animals,plants or found in rock underground.	N	NCMN	S
13122000b	liquid as in rivers.lakes,seas and oceans.	N	NCMN	S
13123000a	the mixture of gases that surrounds the earth and which we breathe.	N	NCMN	U
13123000b	chemical element (symbol O), gas without colour, taste or smell, present in the air, necessary to existence of all forms of life	N	NPRP	U
13123000c	gas (co ₂) produced by animal bodies and breathed out from the lungs	N	NPRP	U
13123000d	without colour, taste or smell, that combines with oxygen to form water	N	NPRP	U
13220000a	a vehicle with 3 or usually 4 wheels and driven by a motor.	N	NCMN	S
13220000b	electronic device which stores information on discs or magnetic tapes;equipment to do calculations,judgments and memorization.	N	NCMN	S
13220000c	utensil with a shallow bowl on a handle.	N	NCMN	S
13220000d	number of sheets of paper, either printed or blank, fastened together in a cover literary composition that fill a set of sheets	N	NCMN	S
13220000e	piece of furniture consisting of a flat top with (usually four supports called legs)	N	NCMN	S

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้สำหรับใช้เพื่อการศึกษาเท่านั้น ไม่อนุญาตให้เผยแพร่โดยไม่ได้รับอนุญาตจากเรา

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

13220000f	tool with a heavy metal head used for breaking things, driving in nails, etc	N	NCMN	S
13220000g	container, usually with a lid, made of wood, cardboard, plastic, metal, etc used for holding solids	N	NCMN	S
13220000h	a line of connected railroad cars drawn by an engine	N	NCMN	S
13220000i	child's plaything; small thing meant for amusement rather than for serious use; small kept as pets	N	NCMN	S
13220000j	frozen sweet of various kinds	N	NCMN	U
13220000k	piece of tasting food (usually made of suger or chocolate), fruit preserved in suger	N	NCMN	S
13220000l	substance used for medical purposed, either alone or in a mixture	N	NCMN	S
13220000m	man's loose-fitting garment for the upper part of the body (of cotton, liner, silk, etc)	N	NCMN	S
13220000n	two legged outer garment for a man or boy, reaching from the event waist to the ankles	N	NCMN	S
13220000o	written message, request account of events, etc sent by one person to another	N	NCMN	S
13220000p	used with first person (I,we) to express willingness, consent, an offer or a promise	N	NCMN	S
13220000q	coins stamped from metal (gold, copper, alloys) printed and accepted when buying and selling	N	NCMN	U
13230000a	a movable flat surface that opens and the entrance to a building, room or piece of furniture	N	NCMN	S
21110000a	take (food etc.) into the mouth and swallow it.	V	VACT	
21110000b	move,pass,from one point to another and away from the speaker.	V	VACT	
21110000c	to fall from a high place to a low place.	V	VACT	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 150 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

21110000d	to pass information from one to another by words.	V	VACT	
21110000e	to get married.	V	VACT	
21110000f	to leave a place.	V	VACT	
21110000g	activity which uses effort, especially with a special purpose, not for amusement	V	VACT	
21110000h	(of men and animals) move with quick steps, faster than when walking	V	VACT	
21110000i	give a blow or stroke to; strike (a target, an object aimed at); come against with force	V	VACT	
21110000j	move quickly by the sudden use of the muscles of the legs or (of fish) the tail, rise suddenly (from a seat, etc)	V	VACT	
21110000k	come against/into, come with a blow or knock	V	VACT	
21110000l	move something so as to be in a certain place or position	V	VACT	
21110000m	to improve something; to move something from one place to another; to give one thing in exchange for another	V	VACT	
21110000n	the act of taking possession or receiving something	V	VACT	
21110000o	to arrive at, get to	V	VACT	
21110000p	receive, have, obtain, procure, acquire	V	VACT	
21110000q	construct or produce by combining parts or putting materials together, form or shape from material, bring into existence	V	VACT	
21110000r	make use of language in an ordinary not a singing, voice	V	VACT	
21110000s	make letters or other symbols (eg ideographs) on a surface	V	VACT	
21110000t	make or cause to be open	V	VACT	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 151 จะต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

21110000u	take (liquid) into the mouth and swallow	V	VACT	
21110000v	use for the purpose of lighting or heating	V	VACT	
21110000w	develop, increase in size, height, length, etc	V	VACT	
21110000x	(of person) move by putting forward each foot in turn, not having both feet off the ground at once	V	VACT	
21110000y	to cause to become separated from the main part suddenly or violently, but not by cutting or tearing	V	VACT	
21110000z	push steadily against	V	VACT	
2111000aa	a condition having less than required	N	NCMN	
2111000ab	come or grow to be, begin to be	V	VACT	
2111000ac	move through the air as a bird does or in an aircraft	V	VACT	
2111000ad	cut into with the teeth	V	VACT	
2111000ae	pick out from a number show what or which one wants by taking	V	VACT	
2111000af	move towards or nearer to the speaker, arrive where the speaker is, was or will be	V	VACT	
2111000ag	(use in the simple tense or with can/be able) look at and (be able to) understand (written or printed)	V	VACT	
2111000ah	get in return for money, get by paying a price	V	VACT	
2111000ai	give an order	V	VACT	
2111000aj	give (eg money to a charity), contribute	V	VACT	
2111000ak	be, remain, in a place or condition	V	VACT	
2111000al	get or discover (not lost, forgotten, etc) after search, experience or effort	V	VACT	
21121000a	know the meaning, nature, explanation	V	VSTA	
21121000b	find the answer to (a problem, etc)	V	VACT	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 152 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

21121000c	one's opinion to sympathize with others, the condition of sharing the same opinion or feeling, to accept an offer from someone	V	VACT	
21121000d	made up, be composed of, be made up of	V	VACT	
21122000a	to concentrate one's attention on the person speaking or on what is being said	V	VSTA	
21122000b	realize, learn, feel regret at, the absence of	V	VSTA	
21122000c	own, have	V	VSTA	
21122000d	to feel sorry or unhappy at the absence or loss of someone	V	VSTA	
21122000e	be fond of, have a taste for, find satisfactory or agreeable	V	VSTA	
21122000f	fill with fright or terror, alarm suddenly	V	VSTA	
21122000g	have strong affection or deep tender feelings for	V	VSTA	
21122000h	having the right or desired qualities, giving satisfaction	ADV	ADV N	
21122000i	speak with approval of, say that one admires	V	VSTA	
21220000a	occasion of violent weather conditions	N	NCMN	
21220000b	condition of burning	N	NCMN	U
21240000a	in bad health, sick	V	VSTA	
22110000a	outline drawing (of or for a building) showing the relative size, positions, etc	N	NCMN	
22110000b	question to be solved or decided, difficult	N	NCMN	S
22110000c	new or fresh information, report of what has most recently happened	N	NCMN	S
22120000a	quality of tolerating opinions, beliefs, customs behaviour, etc different from one's own	N	NCMN	S
22150000a	(explanation of the) general principles of an art or science (contrasted with practice)	N	NCMN	S

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่อนำไปใช้ประโยชน์อื่นใดโดยไม่ได้รับอนุญาตให้ไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 153 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

22150000b	resoned supposition put forward to explain facts or events by matric	N	NPRP	S
22150000c	one's own past until the present	N	NCMN	S
22150000d	an organization in which people with a common interest or purpose gather and act together, a plan or suggestion offered	N	NCMN	S
23100000a	in, at, to this point of place	PRON	PDMN	
23200000a	the first part of the day;from the time when the sun rises	N	NALN	
23200000b	a day of the week named monday	N	NALN	
23200000c	time taken by the earth in making one revolution round the sun, about 365 days, period from 1 january to 31 december	N	NCMN	
23200000d	(on) the day just past, (on) the day before	N	NALN	
23200000e	(on) the day after today	N	NALN	
23610000a	with great energy, strenuous, with all one's force	ADV	ADVN.	
23610000b	at a low speed	ADV	ADVN	
23610000c	quick	ADV	ADVN	
23610000d	of large size, extent, capacity, importance	V	VATT	
23630000a	of the colour of fresh blood, rubies, human lips the tongue	V	VATT	
23630000b	green colour, green leaf vegetables, egcabbage, spinaeh	V	VATT	
23630000c	without light or almost without light, the colour of this printing-ink, opposite to white	V	VATT	
23630000d	the colour of gold or hen's egg			
23650000a	in a manner of being repeated many times.	ADV	ADVN	
23650000b	five.	N	NCNM	P

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ข.

พจนานุกรมเฉพาะภาษาไทย

CPID	WORD	CLASSIFIER		TYPE	VMP	NID
		CLASS1	CLASS2			
11100000a	ครู	คน				11100000d
11100000b	เขาผู้ชาย, เขา	คน				11100000d,11100000h, 11100000c
11100000c	เขาผู้หญิง, เขา	คน				11100000d,11100000h, 11100000b
11100000d	มนุษย์	คน				11100000h
11100000e	นักเรียน, นักศึกษา	คน				11100000d,11100000h
11100000f	เด็ก	คน				11100000d,11100000h
11100000g	เรา, พวกเรา	คน				11100000d,11100000h
11100000h	บุคคล	คน				11100000d
11100000i	ชาวนา	คน				11100000d,11100000h
11100000j	ฉัน	คน				11100000d,11100000h
11100000k	คุณ, ท่าน	คน				11100000d,11100000h
11100000l	แม่, มารดา	คน				11100000d,11100000h
11100000m	น้องชาย	คน				11100000d,11100000h, 11100000b
11100000n	เพื่อน	คน				11100000d,11100000h
11100000o	ผู้จัดการ	คน				11100000d,11100000h
11100000p	ลูกชาย	คน				11100000d,11100000h, 11100000b
11100000q	ครอบครัว	ครอบครัว				

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 155 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

11110000a	วิเชียร	คน				11100000b,11100000d, 11100000h
11110000b	สมศักดิ์	คน				11100000b,11100000d, 11100000h
11110000c	ศักดิ์ชัย	คน				11100000b,11100000d, 11100000h
11110000d	สมศรี	คน				11100000c,11100000d, 11100000h
11110000e	จอห์น	คน				11100000b,11100000d, 11100000h
11200000a	รัฐบาล	รัฐบาล				
12100000a	จังหวัด	จังหวัด				12100000c
12100000b	อำเภอ	อำเภอ				12100000c
12100000c	เมือง	เมือง				12100000a
12110000a	ประเทศไทย	ประเทศ				12100000c
12110000b	ประเทศญี่ปุ่น	ประเทศ				12100000c
12110000c	ประเทศอเมริกา	ประเทศ				12100000c
12110000d	กรุงเทพฯ	จังหวัด				12100000a
12220000a	พระอาทิตย์	ดวง				
12220000b	พระจันทร์	ดวง				
12300000a	มหาวิทยาลัย	มหาวิทยาลัย	แห่ง			
12300000b	โรงเรียน	โรงเรียน	แห่ง			
12300000c	บ้าน	หลัง				
12300000d	เมืองโบราณ	เมือง				12100000a
13111000a	สุนัข	ตัว				
13111000b	แมว	ตัว				
13112000a	ปลา	ตัว	ฝูง			

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 156 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

13112000b	ปลาปิ้งร่นย่า	ตัว	ฝูง			13112000a
13113000a	นก	ตัว	ฝูง			
13114000a	กบ	ตัว				
13115000a	ข้าว, ต้นข้าว	ต้น	ไร่			13118000a
13115000b	กุหลาบ, ต้นกุหลาบ	ต้น				13118000b
13115000c	ต้นไม้	ต้น				
13117000a	มือ	ข้าง				
13117000b	ขา	ข้าง				
13117000c	เปลือก, กระจง	อัน				
13117000d	แขน	ข้าง				
13118000a	ข้าว	เมล็ด	จาน			13115000a
13118000b	กุหลาบ	ดอก	ช่อ			13115000b
13118000c	ดอกไม้	ดอก				
13118000d	แอปเปิล	ผล	ลูก			13118000f
13118000e	มะม่วง	ผล	ลูก			13118000f
13118000f	ผลไม้	ผล				
13118000g	ทุเรียน	ผล				13118000f
13121000a	ทอง	บาท	กรัม			
13121000b	น้ำแข็ง	ก้อน				
13122000a	น้ำมัน	ลิตร				
13122000b	น้ำ	ลิตร.				
13123000a	อากาศ					13123000b,13123000d
13123000b	ออกซิเจน					
13123000c	คาร์บอนไดออกไซด์					
13123000d	ไฮโดรเจน					
13220000a	รถยนต์, รถเก๋ง	คัน				

เอก
 13220000a รถยนต์, รถเก๋ง หรับการใช้ คัน เพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

13220000b	เครื่องคอมพิวเตอร์	เครื่อง			
13220000c	ช้อน	คัน			
13220000d	หนังสือ	เล่ม			
13220000e	โต๊ะ	ตัว			
13220000f	หม้อ	ด้าม			
13220000g	กล่อง	ใบ			
13220000h	รถไฟ	ขบวน	โบกี้		
13220000i	ของเล่น	ชิ้น	อัน		
13220000j	ไอศกรีม	แท่ง			13220000k
13220000k	ของหวาน	ถ้วย			
13220000l	ยา	เม็ด			
13220000m	เสื่อ	ตัว			
13220000n	กางเกง	ตัว			
13220000o	จดหมาย	ฉบับ			
13220000p	พินัยกรรม	ฉบับ			
13220000q	เงิน	บาท			
13230000a	ประตู	บาน			
21110000a	กิน			2a,3a,3b, 3c,3e,4a	21110000u
21110000b	ไป			2c,3f,3g, 4d	21110000f, 2111000af
21110000c	ตก			1a,2c,2d, 2f	
21110000d	บอก			2a,3a,3b, 3c,4a	21110000r
21110000e	แต่งงาน			2i	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา 158 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

21110000f	ออก				2e,3k,3l, 3m,3o,4 b,4e	21110000b
21110000g	ทำงาน				1a,2b,2c, 2d,2f,2h	
21110000h	วิ่ง				1a,2b,2d ,2f,3f,3g	21110000j
21110000i	ตี				2a,3a,3b, 3c,3e,4a	21110000k
21110000j	กระโดด				1a,2b,2d ,2f,3f,3g	21110000h
21110000k	ชน				2a,3a,3b, 3c,3e,4a	21110000i
21110000l	วาง				2a,3a,3b, 3c,3e,4a	
21110000m	เปลี่ยน				2a,3a,3b, 3c,3k	21110000ab
21110000n	รับ				2a,3a,3b, 3c,3d	21110000p
21110000o	ถึง				2c,3f,3g, 4d	
21110000p	ได้รับ				2a,3a,3b, 3c,3d	21110000n
21110000q	สร้าง				2a,3a,3b, 3c,3e,4a	21110000w
21110000r	พูด				1a,2a,2c, 2d,2f,2i	21110000d
21110000s	เขียน				2a,3a,3b, 3c,3e,4a	

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์การใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 159 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

21110000t	เปิด				2a,3a,3b, 3c	
21110000u	คืบ				2a,3a,3b, 3c,3e,4a	21110000a
21110000v	ใหม่				1a,2a,2c, 2d,3b,3c	
21110000w	ปลูก				2a,3a,3b, 3c,3e,4a	21110000q
21110000x	เดิน				1a,2b,2d ,2f,3f,3g	
21110000y	หยุด				1a,2b,2d ,2f,3f,3g	
21110000z	กด				2a,3a,3b, 3c,3e,4a	
2111000aa	การขาดแคลน					
2111000ab	กลายเป็น				2j	21110000m
2111000ac	บิน				1a,2b,2c, 2d,2e,3g	
2111000ad	กัด				2a,3a,3b, 3c,3e	
2111000ae	เลือก				2a,3c,3d	
2111000af	มา				2c,3f,3g, 4d	21110000b
2111000ag	อ่าน				2a,3b,3c, 3e,4a	
2111000ah	ซื้อ				2a,3a,3b, 3c,3d	
2111000ai	ตั้ง				2a,3a,3b, 3c	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 160 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2111000aj	บริจาค				2a,3c,3d, 3e	
2111000ak	อยู่				2c,3h	
2111000al	พบ				2a,3a,3b, 3c,4a	
21121000a	เข้าใจ				2a,3e	
21121000b	แก้				2a,3a,3e	
21121000c	เห็นด้วย				2a,3c	
21121000d	ประกอบด้วย				2g	
21122000a	ฟัง				2a,3b,3c, 3d,4a	
21122000b	คิดถึง				2a,3c,3e	
21122000c	เป็นของ					
21122000d	กังวล				2c,3i	
21122000e	ชอบ				2a,3c,3e	21122000g
21122000f	ตกใจ				1a,2f	
21122000g	รัก				2a,3e	21122000e
21122000h	ดี					
21122000i	ชม, สรรเสริญ				2a,3e	
21220000a	อายุ	ลูก				
21220000b	ไฟ					
21240000a	ป่วย				1a,2f	
22110000a	แผนการ	แผนการ	แผน			
22110000b	ปัญหา	ปัญหา	เรื่อง			
22110000c	ข่าว	ข่าว				
22120000a	ความอดทน					
22150000a	ทฤษฎี	ทฤษฎี				

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

22150000b	ทฤษฎีเมตริกซ์	ทฤษฎี				22150000a
22150000c	เรื่อง, เรื่องราว	เรื่อง				
22150000d	โครงการ	โครงการ				22110000a
23100000a	ที่นี้					
23200000a	ตอนเช้า					
23200000b	วันจันทร์	วัน				
23200000c	ปี	ปี				
23200000d	เมื่อวานนี้	วัน				
23200000e	พรุ่งนี้	วัน				
23610000a	หนัก					
23610000b	ช้า					
23610000c	เร็ว					
23610000d	ใหญ่					
23630000a	สีแดง					
23630000b	สีเขียว					
23630000c	สีดำ					
23630000d	สีเหลือง					
23650000a	บ่อย					
23650000b	ห่า					

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ค.

พจนานุกรมเฉพาะภาษาอังกฤษ

CPID	WORD	CLASSIFIER		TYPE	VMP	NID
		CLASS1	CLASS2			
1110000a	teacher					1110000d
1110000b	he					1110000d,1110000h ,1110000c
1110000c	she					1110000d,1110000h ,1110000b
1110000d	human					1110000h
1110000e	student					1110000d,1110000h
1110000f	child					1110000d,1110000h
1110000g	we					1110000d,1110000h
1110000h	person					1110000d
1110000i	farmer					1110000d,1110000h
1110000j	i					1110000d,1110000h
1110000k	you					1110000d,1110000h
1110000l	mother					1110000d,1110000h
1110000m	brother					1110000d,1110000h , 1110000b
1110000n	friend					1110000d,1110000h
1110000o	manager					1110000d,1110000h
1110000p	son					1110000d,1110000h , 1110000b
1110000q	family					

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นิยมนำไปใช้ประโยชน์ทางการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา 163 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

11110000a	vichean					11100000b,11100000d ,11100000h
11110000b	somsak					11100000b,11100000d ,11100000h
11110000c	sakchai					11100000b,11100000d ,11100000h
11110000d	somsri					11100000c,11100000d ,11100000h
11110000e	john					11100000b,11100000d ,11100000h
11200000a	government					
12100000a	province					12100000c
12100000b	amphur					12100000c
12100000c	town					12100000a
12110000a	thailand					12100000c
12110000b	japan					12100000c
12110000c	american					12100000c
12110000d	bangkok					12100000a
12220000a	sun					
12220000b	moon					
12300000a	university					
12300000b	school					
12300000c	home					
12300000d	ancient					
13111000a	dog					
13111000b	cat					
13112000a	fish					

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 164 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

13112000b	piranha					13112000a
13113000a	bird					
13114000a	frog					
13115000a	rice					13118000a
13115000b	rose					13118000b
13115000c	tree					
13117000a	hand					
13117000b	leg					
13117000c	shell					
13117000d	arm					
13118000a	rice					13115000a
13118000b	rose					13115000b
13118000c	flower					
13118000d	apple					13118000f
13118000e	mango					13118000f
13118000f	fruit					
13118000g	durain					13118000f
13121000a	glod					
13121000b	ice					
13122000a	oil					
13122000b	water					
13123000a	air					13123000b,13123000d
13123000b	oxygen					
13123000c	carbon dioxide					
13123000d	hydrogen					
13220000a	car					

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 165 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

13220000b	computer					
13220000c	spoon					
13220000d	book					
13220000e	table					
13220000f	hammer					
13220000g	box					
13220000h	train					
13220000i	toy					
13220000j	icecream					13220000k
13220000k	sweetmeat					
13220000l	drug					
13220000m	shirt					
13220000n	trousers					
13220000o	letter					
13220000p	will					
13220000q	money					
13230000a	door					
21110000a	eat				2a,3a,3b,3 c,3e,4a	21110000u
21110000b	go				2c,3f,3g,4 d	21110000f,2111000af
21110000c	fall				1a,2c,2d,2 f	
21110000d	tell				2a,3a,3b,3 c,4a	21110000r
21110000e	marry				2i	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 166 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

21110000f	leave				2e,3k,3l,3 m,3o,4b,4 e	21110000b
21110000g	work				1a,2b,2c,2 d,2f,2h	
21110000h	run				1a,2b,2d,2 f,3f,3g	21110000j
21110000i	hit				2a,3a,3b,3 c,3e,4a	21110000k
21110000j	jump				1a,2b,2d,2 f,3f,3g	21110000h
21110000k	bump				2a,3a,3b,3 c,3e,4a	21110000i
21110000l	put				2a,3a,3b,3 c,3e,4a	
21110000m	change				2a,3a,3b,3 c,3k	21110000ab
21110000n	receive				2a,3a,3b,3 c,3d	21110000p
21110000o	reach				2c,3f,3g,4 d	
21110000p	get				2a,3a,3b,3 c,3d	21110000n
21110000q	make				2a,3a,3b,3 c,3e,4a	21110000w
21110000r	speak				1a,2a,2c,2 d,2f,2i	21110000d
21110000s	write				2a,3a,3b,3 c,3e,4a	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับศึกษาใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้ทำไปใช้ประโยชน์ทางการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา 167 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

21110000t	open				2a,3a,3b,3 c	
21110000u	drink				2a,3a,3b,3 c,3e,4a	21110000a
21110000v	burn				1a,2a,2c,2 d,3b,3c	
21110000w	grow				2a,3a,3b,3 c,3e,4a	21110000q
21110000x	walk				1a,2b,2d,2 f,3f,3g	
21110000y	break				1a,2b,2d,2 f,3f,3g	
21110000z	press				2a,3a,3b,3 c,3e,4a	
2111000aa	shortage					
2111000ab	become				2j	21110000m
2111000ac	fly				1a,2b,2c,2 d,2e,3g	
2111000ad	bite				2a,3a,3b,3 c,3e	
2111000ae	choose				2a,3c,3d	
2111000af	come				2c,3f,3g,4 d	21110000b
2111000ag	read				2a,3b,3c,3 e,4a	
2111000ah	buy				2a,3a,3b,3 c,3d	
2111000ai	order				2a,3a,3b,3 c	

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์การใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้拿去ใช้ประโยชน์ด้านการค้า

2111000aj	donate				2a,3c,3d,3e	
2111000ak	stay				2c,3h	
2111000al	find				2a,3a,3b,3c,4a	
21121000a	understand				2a,3e	
21121000b	solve				2a,3a,3e	
21121000c	agree				2a,3c	
21121000d	compose				2g	
21122000a	listen				2a,3b,3c,3d,4a	
21122000b	miss				2a,3c,3e	
21122000c	possess					
21122000d	worry				2c,3i	
21122000e	like				2a,3c,3e	21122000g
21122000f	frighten				1a,2f	
21122000g	love				2a,3e	21122000e
21122000h	good					
21122000i	praise				2a,3e	
21220000a	storm					
21220000b	fire					
21240000a	ill				1a,2f	
22110000a	plan					
22110000b	problem					
22110000c	news					
22120000a	tolerance					
22150000a	theory					

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์การใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา 169 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

22150000b	matric theory					22150000a
22150000c	story					
22150000d	project					22110000a
23100000a	here					
23200000a	morning					
23200000b	monday					
23200000c	year					
23200000d	yesterday					
23200000e	tomorrow					
23610000a	hard					
23610000b	slow					
23610000c	fast					
23610000d	big					
23630000a	red					
23630000b	green					
23630000c	black					
23630000d	yellow					
23650000a	often					
23650000b	five					

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้