



ระบบแปลภาษาด้วยคอมพิวเตอร์ (ส่วนวิเคราะห์ระบบ)  
Machine Translation (Analysis)



โดย  
นาย วรุต อนุชชาติคม 36013169  
นาย วิชัย วงษ์ศรี 36013170

อาจารย์ที่ปรึกษา  
อาจารย์ ประสาร ดังคีสานนท์

วัน เดือน ปี... ๑1 ก.ค. ๒5๖๐  
เลขทะเบียน... ๐๓๗๐๗๓  
เลขเรียกหนังสือ... T ๑๘๑๖๖. ๑๓๖4 ๕

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต  
ภาควิชาวิศวกรรมคอมพิวเตอร์  
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง  
ปีการศึกษา 2538

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

037073

ปริญญาโทปีการศึกษา 2538

ภาควิชา วิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง ระบบแปลภาษาด้วยคอมพิวเตอร์ (ส่วนวิเคราะห์ภาษา)  
( Machine Translation : Analysis )

ผู้จัดทำ

1. นายวรุฒ อนุชชาลาคม รหัสประจำตัว 36 013 169

2. นายวิชัย วงษ์ตรี รหัสประจำตัว 36 013 170

.....อาจารย์ที่ปรึกษา  
(.....ประสาร ตั้งติสานนท์.....)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## ระบบแปลภาษาด้วยคอมพิวเตอร์ (ส่วนวิเคราะห์ภาษา)

นายวรุศ อนุชลาตม

นายวิชัย วงษ์ตรี

อาจารย์ประสาร ตั้งศิษานนท์ ที่ปรึกษา

ปีการศึกษา 2538

### บทคัดย่อ

ระบบการแปลภาษาด้วยคอมพิวเตอร์นั้น มีส่วนประกอบที่สำคัญสองส่วนคือ ระบบการวิเคราะห์ภาษา (Analysis) ซึ่งทำหน้าที่ในการแปลงภาษาจากภาษาที่เป็นต้นฉบับไปสู่ภาษากลาง (Interlingua) และระบบสังเคราะห์ภาษา (Generation) ทำหน้าที่แปลงจากภาษากลางเป็นภาษาปลายทาง โครงการนี้จะกล่าวเฉพาะระบบวิเคราะห์ภาษาซึ่งวิเคราะห์ภาษาไทยเท่านั้น โดยแยกออกเป็นสามกระบวนการทำงานคือ

- 1.) กระบวนการตัดคำจากประโยค ( Morphology Analysis Process ) เนื่องจากการที่คำในภาษาไทยเขียนติดต่อกันทั้งประโยค จึงต้องมีกระบวนการตัดคำ
- 2.) กระบวนการวิเคราะห์โครงสร้างและความหมาย ( Syntactic & Semantic Analysis Process ) กระบวนการนี้จะนำผลจากการตัดคำ เพื่อมาวิเคราะห์ทางโครงสร้างและความหมายของประโยค และเนื่องจากผลลัพธ์จากการตัดคำนั้นอาจจะได้ผลลัพธ์หลายรูปแบบ กระบวนการนี้จะตัดสินใจเลือกรูปแบบที่ถูกต้อง โดยอาศัยฐานความรู้ทางด้านไวยากรณ์ของประโยคภาษาไทยฐานความรู้ด้านไวยากรณ์ภาษาไทยนี้ ยึดหลักการจัดเรียงตำแหน่งของคำศัพท์ซึ่งถูกแบ่งตามชนิดของคำ เช่น คำนาม สรรพนาม ฯลฯ และการผสมหน่วยคำศัพท์ในประโยคภาษาไทย ฐานความรู้ดังกล่าวถูกแทนในรูปของโครงข่าย M-ATN ( Modified-Augmented Transition Network) จากนั้นก็กำหนดความสัมพันธ์ของคำศัพท์ในประโยค แล้วแทนคำศัพท์และความสัมพันธ์เหล่านั้นด้วยโครงข่ายทางความหมาย (Semantic NetWork) ขึ้น

3.) กระบวนการวิเคราะห์และแทนรูปภาษากลาง ( Conceptual Analysis Process ) ทำหน้าที่แทนโครงข่ายทางความหมายของประโยค ในรูปแบบภาษากลางที่กำหนด

จากระบบวิเคราะห์ภาษาไทยนี้ จะทำให้สามารถพัฒนาระบบแปลภาษาไทย เพื่อใช้กับระบบแปลภาษาหลายๆภาษาได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## MACHINE TRANSLATION (ANALYSIS)

**Mr. WARUT ANUTCHALAKHOM**

**Mr. WICHAI VONGTREE**

**THESIS ADVISOR**

**PRASARN TANGTISANON**

**ACADEMIC YEAR 1995**

### ABSTRACT

Machine translation have two components are LANGUAGE ANALYSIS SYSTEM which have duty is transfrom Manuscript language to Interlingua ,and LANGUAGE GENERATION SYSTEM ,have duty is transfrom Interlingua to Target language.In this thesis describe LANGUAGE ANALYSIS SYSTEM part only. LANGUAGE ANALYSIS SYSTEM consist of three process.

1.Morphology Analysis Process ,Because word in Thai language sentence are written without spaces between word,Thus must Morphology analysis.

2.Syntactic & Semantic Analysis Process,Function of this process is take the result from Morphology Analysis Process to syntactic and semantic analysis.The result from Morphology Analysis Process may be more than one model,This process shall choose correct model by using knowledge base of grammar in Thai language.The knowledge base of grammar in Thai language to be based up on arrange in order position of vocabulary which divided follow type of word,such as noun pronoun etc.The knowledge base of grammar in Thai language is replaced in Modified-Augmented Transition Network (M-ATN).Next process is define relationship of word in sentence and replace word and relation of word in Semantic Network.

3.Conceptual Analysis Process,perform duty replace Semantic Network in Interlingua form.This THAI LANGUAGE ANALYSIS SYSTEM ,can develop Thai Machine Translator .

## สารบัญ

	หน้า
บทคัดย่อ .....	I
ABSTRACT .....	II
สารบัญ .....	III
สารบัญรูป .....	VI
สารบัญตาราง .....	VIII
บทที่ 1 บทนำ .....	1
1.1 ความรู้ทั่วไปเกี่ยวกับระบบแปลภาษาด้วยคอมพิวเตอร์ .....	2
1.1.1 ประเภทของระบบแปลภาษา .....	2
1.1.2 โครงสร้างและส่วนประกอบของระบบแปลภาษาด้วยคอมพิวเตอร์ .....	3
1.2 จุดประสงค์ของโครงการ .....	5
บทที่ 2 ความรู้เกี่ยวกับภาษาไทย .....	6
2.1 บทนำ .....	6
2.2 ชนิดของคำศัพท์ในภาษาไทย .....	6
2.2.1 คำนาม .....	6
2.2.2 คำสรรพนาม .....	8
2.2.3 คำกริยา .....	11
2.2.4 คำวิเศษณ์ .....	13
2.2.5 คำบุพบท .....	16
2.2.6 คำสันธาน .....	18
2.2.7 คำอุทาน .....	20
2.3 สัญลักษณ์ของกลุ่มคำภาษาไทย .....	21
2.4 โครงสร้างของวลีในภาษาไทย .....	23
2.4.1 นามวลีและส่วนขยาย .....	24
2.4.2 กริยาวลีและส่วนขยาย .....	30
2.4.3 วิเศษณ์วลีและส่วนขยาย .....	35

บทที่ 3 การแสดงรูปแทนภาษากลางสำหรับภาษาไทย .....	38
3.1. บทนำ.....	38
3.2. โครงสร้างของรูปแทนภาษากลาง.....	38
3.2.1 มโนทัศน์หลัก.....	39
3.2.2 ความสัมพันธ์ .....	39
3.2.3 คุณลักษณะ(attributes).....	48
3.3. การแทนรูปประโยคด้วยภาษากลาง.....	54
บทที่ 4 โครงสร้างและส่วนประกอบของระบบวิเคราะห์ภาษา.....	57
4.1 บทนำ.....	57
4.2 การแยกคำศัพท์จากประโยคภาษาไทย.....	58
4.2.1 การกำหนดดัชนีและจัดกลุ่มคำศัพท์.....	59
4.2.2 กระบวนการแยกคำศัพท์แบบ Fast Word Matching.....	67
4.2.3 ตัวอย่างการแยกคำศัพท์แบบ Fast Word Matching.....	70
4.3 การวิเคราะห์โครงสร้างและความหมายของประโยคภาษาไทย .....	78
4.3.1 ความรู้เบื้องต้นเกี่ยวกับระบบโครงข่าย.....	78
4.3.2 ชนิดและหน้าที่ของคำศัพท์ภาษาไทยที่ใช้ในโครงข่าย.....	88
4.3.3 ระบบ M-ATN ในการประมวลผลโครงสร้างประโยคภาษาไทย.....	90
4.4 การวิเคราะห์และแทนรูปภาษากลาง .....	101
4.4.1 ขั้นตอนการวิเคราะห์และสร้างโครงข่ายทางความหมาย.....	101
4.4.2 โครงข่ายความหมายและการแทนรูปภาษากลางของประโยค.....	104
บทที่ 5 ซอฟต์แวร์และการออกแบบระบบวิเคราะห์ภาษา .....	106
5.1 บทนำ.....	106
5.2 การออกแบบระบบวิเคราะห์ภาษาไทย.....	106
5.3 การออกแบบพจนานุกรม.....	108
5.3.1 ข้อมูลต่างๆของคำศัพท์ในพจนานุกรม.....	108
5.3.2 ระบบเพิ่มข้อมูลของพจนานุกรม.....	109
5.4 การออกแบบกระบวนการแยกคำศัพท์.....	111
5.4.1 อินพุตและเอาต์พุตของกระบวนการแยกคำศัพท์.....	111
5.4.2 โฟลว์ชาร์ตการทำงาน .....	112
5.4.3 ตัวอย่างการทำงานของกระบวนการแยกคำ .....	113

5.5 การออกแบบกระบวนการวิเคราะห์โครงสร้างประโยค.....	113
5.5.1 อินพุตและเอาต์พุตของกระบวนการวิเคราะห์โครงสร้างประโยค.....	113
5.5.2 โพลีซาร์ทการทำงานของระบบวิเคราะห์โครงสร้างประโยค.....	114
5.5.3 ตัวอย่างการทำงานของกระบวนการวิเคราะห์ประโยค.....	115
5.6 เพิ่มข้อมูลต่างๆในระบบ.....	115
<b>บทที่ 6 บทสรุปและการประเมินผล.....</b>	<b>116</b>
6.1 บทนำ.....	116
6.2 การทดลองแยกคำ.....	116
6.3 การทดลองส่วนวิเคราะห์โครงสร้างประโยค .....	117
6.4 การประยุกต์ใช้งานระบบวิเคราะห์ภาษาไทย .....	118
6.4.1 การแปลภาษาด้วยคอมพิวเตอร์ ( Machine Translation ) .....	118
6.4.2 การตรวจสอบตัวสะกด หรือคำผิดในงานจดจำอักษร.....	118
6.5 แนวทางการพัฒนาระบบวิเคราะห์ภาษาไทยด้วยคอมพิวเตอร์.....	118
6.6 บทสรุป.....	119
กิจกรรมประกาศ.....	120
หนังสืออ้างอิง.....	121
ภาคผนวก ก แบบโครงสร้างของประโยคโดยใช้คำกริยา (Verb Pattern).....	122
ภาคผนวก ข AKO (A Kind of).....	123

## สารบัญรูป

	หน้า
รูปที่ 1.1 แสดงระบบแปลภาษาแบบตรง.....	2
รูปที่ 1.2 แสดงระบบแปลภาษาแบบหลายภาษา.....	3
รูปที่ 1.3 แสดงส่วนประกอบของระบบแปลภาษาแบบหลายภาษา .....	4
รูปที่ 1.4 แสดงกระบวนการทำงานส่วนต่างๆของระบบแปลภาษาแบบหลายภาษา.....	5
รูปที่ 4.1 แสดงส่วนประกอบพื้นฐานของระบบวิเคราะห์ภาษา.....	57
รูปที่ 4.2 แสดงการทำงานพื้นฐานของการแยกคำศัพท์ .....	58
รูปที่ 4.3 แสดงความสัมพันธ์ระหว่าง $I_1(j)$ , $I_{j_2}(k)$ และ $I_{j_3}(n)$ .....	61
รูปที่ 4.4 แสดงดัชนีที่ 1 ซึ่งมี 44 ค่า.....	62
รูปที่ 4.5 แสดงการแบ่งกลุ่มย่อยตามอักขระที่ 2 ของกลุ่มอักขระที่ 1 เป็น 'ก' .....	64
รูปที่ 4.6 แสดงการกำหนดดัชนีที่ 3.....	66
รูปที่ 4.7 โพลซาร์ทการแยกคำ.....	69
รูปที่ 4.8 แสดงตัวอย่างการแยกประโยค “เธอ ชน เขา” .....	76
รูปที่ 4.9 แสดงตัวอย่างการแยกประโยค “โคลงเรื่องจนเรือโคลง” .....	77
รูปที่ 4.10 แสดงส่วนประกอบพื้นฐานของโครงข่ายในระบบต่างๆ.....	79
รูปที่ 4.11 ตัวอย่างแสดงความสัมพันธ์ทางโครงสร้างบางส่วน ของภาษาอังกฤษในระบบ TN.....	80
รูปที่ 4.12 (ก),(ข),(ค) แสดงตัวอย่างความสัมพันธ์ทางโครงสร้างบางส่วน ของประโยคภาษาอังกฤษในระบบ RTN.....	82
รูปที่ 4.13 ตัวอย่างความสัมพันธ์ทางโครงสร้างบางส่วน ของประโยคภาษาอังกฤษในระบบ ATN.....	86
รูปที่ 4.14 แสดงส่วนประกอบต่างๆ ของโครงข่ายในระบบ M-ATN.....	90
รูปที่ 4.15 แสดงโครงข่ายของประโยค ( Sentence Network : [SN] ) .....	92
รูปที่ 4.16 แสดงโครงข่ายนามวลี ( Noun Phrase : [NP] ).....	93
รูปที่ 4.17 แสดงโครงข่ายกลุ่มของ (Genitive Marker : [GM] ) .....	93
รูปที่ 4.18 แสดงโครงข่ายของกริยาวลี ( Verb Phrase : [VP] ) .....	94
รูปที่ 4.19 แสดงตัวอย่างการประมวลผลประโยค “นักเรียนไปโรงเรียน” .....	97
รูปที่ 4.20 แสดงตัวอย่างการประมวลผลประโยค “ฉันกับเธอไปกินข้าว” .....	98
รูปที่ 4.21 แสดงการวิเคราะห์ประโยค “โคลง/เรือ/จน/เรือ/โคลง” .....	99

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 4.22 แสดงการวิเคราะห์ประโยค “โค/ลง/เรือ/จน/เรือ/โคลง”.....	99
รูปที่ 4.23 แสดงการวิเคราะห์ประโยค “โคลง/เรือ/จน/เรือ/โค/ลง” .....	100
รูปที่ 4.24 แสดงการวิเคราะห์ประโยค “โค/ลง/เรือ/จน/เรือ/โค/ลง”.....	100
รูปที่ 4.25 แสดงรูปแบบโครงข่ายทางความหมายของประโยค.....	104
รูปที่ 5.1 แสดงส่วนประกอบต่างๆและข้อมูลเข้าออกของแต่ละส่วน .....	107
รูปที่ 5.2 แสดงโครงสร้างเพิ่มข้อมูลของพจนานุกรม.....	110
รูปที่ 5.3 แสดงตัวอย่างเพิ่มข้อมูลคำศัพท์ “ฉันทน์”.....	111
รูปที่ 5.4 โพลซาร์ทการแยกคำ.....	112
รูปที่ 5.5 แสดงตัวอย่างการทำงานของกระบวนการแยกคำ.....	113
รูปที่ 5.6 แสดงโพลซาร์ทของกระบวนการวิเคราะห์โครงสร้างประโยค .....	114
รูปที่ 5.7 ตัวอย่างการทำงานของกระบวนการวิเคราะห์ประโยค.....	115
รูปที่ 6.1 แสดงกราฟผลลัพธ์การแยกคำศัพท์.....	116



## สารบัญตาราง

	หน้า
ตารางที่ 2.1 เป็นการเปรียบเทียบระหว่าง ประพันธ์สรรพนาม กับ ประพันธ์วิเศษณ์.....	5
ตารางที่ 2.2 ตารางการใช้คำสันธาน.....	20
ตารางที่ 2.3 แสดงสัญลักษณ์ย่อชนิดของคำ.....	22
ตารางที่ 4.1 แสดงอักษรที่สามารถเป็นอักษรแรกของคำศัพท์.....	62
ตารางที่ 4.2 แสดงตัวอย่างการแบ่งกลุ่มคำศัพท์ตามดัชนีที่ 2 ( $I_{1,2}(k)$ ).....	65
ตารางที่ 4.3 แสดงกระบวนการทำงานของระบบ TN.....	80
ตารางที่ 4.4 แสดงกระบวนการทำงานของระบบ RTN.....	83
ตารางที่ 4.5 แสดงกระบวนการทำงานของระบบ ATN.....	87
ตารางที่ 4.6 แสดงกระบวนการการทำงานของระบบ M-ATN.....	95
ตารางที่ 5.1 แสดงข้อมูลของคำศัพท์ในพจนานุกรม.....	109



# บทที่ 1

## บทนำ

ภาษาเพื่อการสื่อความหมายระหว่างกันของมนุษย์โลก มีอยู่มากมายหลายรูปแบบ ตั้งแต่การแสดงท่าทางด้วยอาการของส่วนใดส่วนหนึ่งของร่างกาย หรือด้วยเครื่องมือบางชนิด เช่น การโบกมือ พยักหน้า หรือการใช้ควันไฟของคนในสมัยโบราณ เป็นต้น ไปจนถึงการใช้เสียงต่างๆ และยังได้มีการกำหนดสัญลักษณ์รูปแบบต่างๆเพื่อใช้ในการสื่อสารด้วยการมองเห็นในเวลาต่อมา เช่น รูปภาพ หรือตัวอักษร จนเกิดเป็นภาษาพูด และภาษาเขียน ที่มีรูปแบบเฉพาะที่แตกต่างกันออกไปตามชนชาติที่แตกต่างกัน ซึ่งขึ้นอยู่กับขนบธรรมเนียมประเพณี และสภาพแวดล้อมต่างๆ มนุษย์สามารถทำความเข้าใจและสามารถใช้ภาษานั้นๆ ได้โดยอาศัยประสบการณ์ที่สะสมมาตั้งแต่เกิด

ในการใช้ภาษาต่างๆผู้ใช้จะต้องมีความรู้ และความเข้าใจในภาษานั้น ไม่ว่าจะเป็นการพูด หรือการเขียน โดยการเปลี่ยนกระบวนการทางความคิด เพื่อเปลี่ยนประโยคที่คิดไว้และต้องการสื่อสารนั้นให้เป็นเสียง หรือเป็นสัญลักษณ์ตามที่ภาษานั้นๆได้กำหนดไว้ ซึ่งจะต้องมีรูปแบบที่ทั้งผู้ส่งสารและผู้รับสารสามารถเข้าใจได้ตรงกัน ดังนั้นทั้งการพูดและการเขียนภาษาใดๆจะต้องมีรูปแบบการพูด(Sound Pattern) และรูปแบบการเขียน(Written Pattern) หรืออาจจะกล่าวได้อีกอย่างหนึ่งว่าความสัมพันธ์ของคำศัพท์ในประโยค (Syntax)

ต่อมาเมื่อเทคโนโลยีเจริญขึ้น ก็ได้มีความพยายามที่จะทำให้คอมพิวเตอร์เข้าใจภาษามนุษย์ ซึ่งเป็นศาสตร์แขนงหนึ่งทางด้านปัญญาประดิษฐ์(Artificial Intelligence)การที่จะให้คอมพิวเตอร์เข้าใจภาษามนุษย์นั้น มนุษย์เองที่จะต้องสอนหรือให้ความรู้แก่คอมพิวเตอร์ เช่น ให้ความรู้ทางโครงสร้างภาษา และกระบวนการเพื่อการวิเคราะห์ทางความหมาย(Semantic Analysis) และได้มีการคิดค้นวิธีการที่จะประมวลผลทางโครงสร้าง ตามกฎเกณฑ์ของประโยคหรือที่เราเรียกว่า ไวยากรณ์(Grammar) โดยใช้หน้าที่และชนิดของคำศัพท์ในประโยค ทฤษฎีต่างๆเช่น Pattern Matching, Transition Network, Context-Free Grammar, Augmented Transition Network, Transformation เป็นต้น

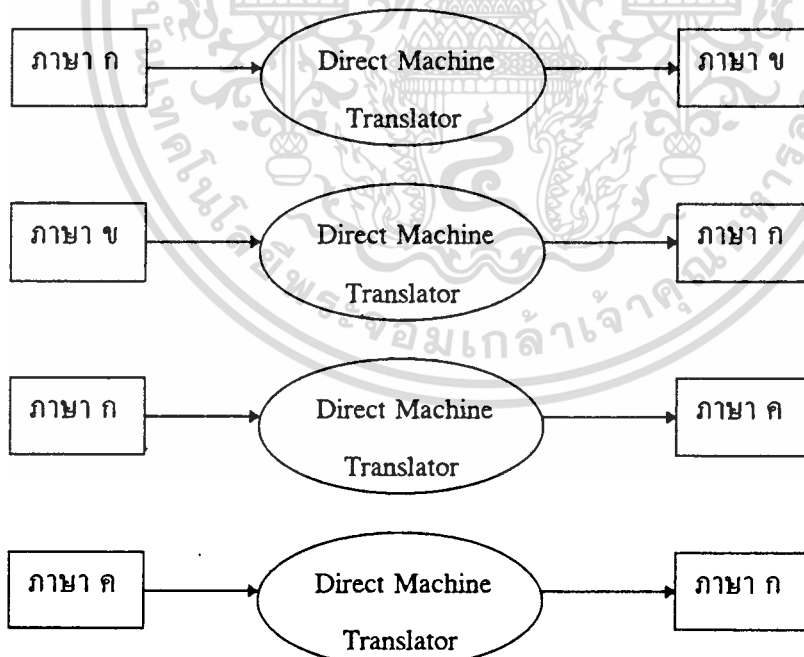
ต่อมาจนถึงยุคแห่งการสื่อสารข้อมูล การติดต่อกันระหว่างชนชาติต่างๆเป็นไปอย่างกว้างขวางมากขึ้น จึงได้มีความคิดที่จะนำคอมพิวเตอร์มาใช้ในการแปลภาษา เนื่องจากแต่ละภาษา มีความแตกต่างกัน ดังนั้นมนุษย์จึงต้องสอนคอมพิวเตอร์ ให้อรรถกถาภาษาต่างๆ ภาษาไทยก็เช่นเดียวกัน เราต้องสอนกฎเกณฑ์ต่างๆของภาษาไทยให้คอมพิวเตอร์ทราบ แต่การประมวลผลภาษาไทยมี ปัญหาอยู่มาก เช่น ต้องมีการตัดคำ เพราะประโยคภาษาไทยนั้นเขียนแบบต่อเนื่องกัน ไม่เหมือนภาษาอังกฤษ จึงจำเป็นจะต้องมีระบบการตัดหรือแยกคำศัพท์ด้วย ซึ่งจะได้กล่าวในภายหลังต่อไป

## 1.1 ความรู้ทั่วไปเกี่ยวกับระบบแปลภาษาด้วยคอมพิวเตอร์

### 1.1.1 ประเภทของระบบแปลภาษา

ประเภทของระบบแปลภาษาด้วยคอมพิวเตอร์นั้นมีอยู่ 2 ประเภทแบ่งตามวิธีการแปลคือ

ก. ระบบแปลภาษาแบบตรง (Direct Machine Translation)

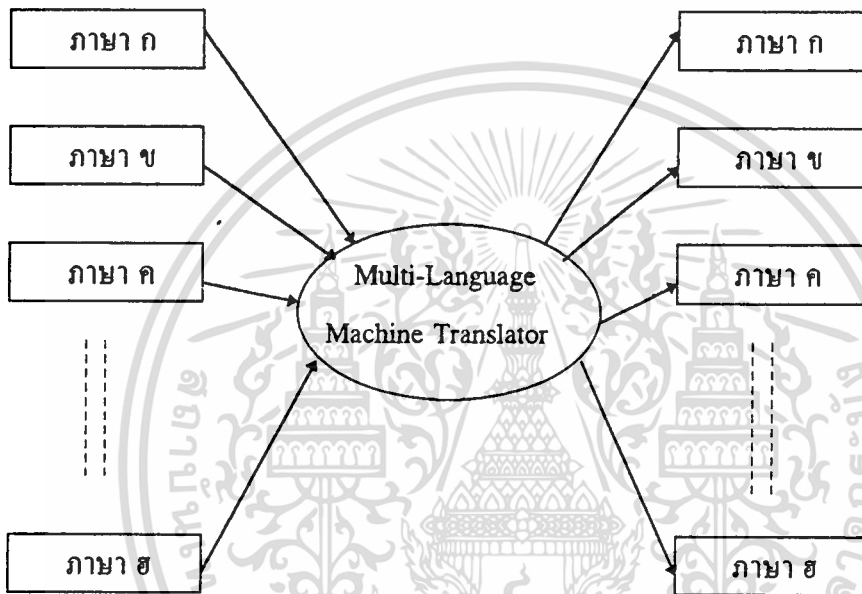


รูปที่ 1.1 แสดงระบบแปลภาษาแบบตรง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 1.1 เป็นการแปลภาษาจากภาษา ก ไปเป็นภาษา ข, จากภาษา ข ไปเป็นภาษา ก, จากภาษา ก ไปเป็นภาษา ค และจากภาษา ค ไปเป็นภาษา ก จะเห็นว่าเมื่อมีการเพิ่มภาษาที่จะแปล จะต้องมีการสร้างตัวแปลภาษาแบบตรงเพิ่มขึ้นด้วย

#### ข. ระบบแปลภาษาแบบหลายภาษา (Multi-Language Machine Translation)



รูปที่ 1.2 แสดงระบบแปลภาษาแบบหลายภาษา

จากรูปที่ 1.2 จะเห็นว่าการแปลภาษาจากภาษา ก ไปเป็นภาษา ก, ข, ค, ....ฮ นั้นทำได้โดยการสร้างระบบแปลภาษาแบบหลายภาษาเพียงระบบเดียว ซึ่งการแปลภาษาแบบนี้จำเป็นต้องอาศัยสื่อกลาง หรือภาษากลาง (InterLingua) ในการแปลจากภาษาต้นทางสู่ภาษาเป้าหมาย

#### 1.1.2 โครงสร้างและส่วนประกอบของระบบแปลภาษาด้วยคอมพิวเตอร์

ระบบการแปลภาษาด้วยคอมพิวเตอร์ประกอบไปด้วย 2 ส่วนโดยใช้ภาษากลางเป็นสื่อคือ .

ก. ระบบวิเคราะห์ภาษา (Analysis) ซึ่งเป็นเนื้อหาของโครงการนี้

ข. ระบบสังเคราะห์ภาษา (Generation)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะวิธีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 1.3 แสดงส่วนประกอบของระบบแปลภาษาแบบหลายภาษา

### ระบบวิเคราะห์ภาษา

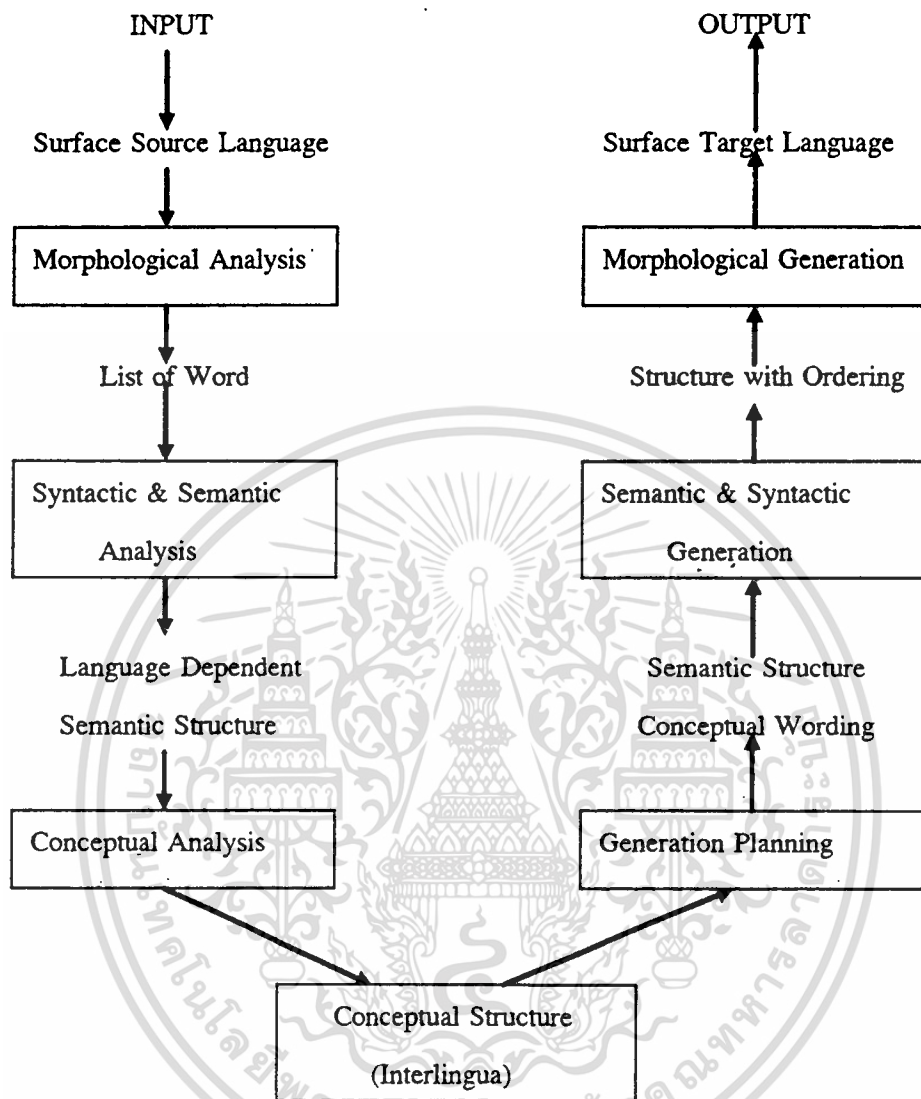
ระบบวิเคราะห์ภาษาแบ่งการทำงานออกเป็น 3 ส่วนคือ

- ส่วนวิเคราะห์หน่วยคำ (Morphological Analysis) ทำหน้าที่แยกคำศัพท์ในประโยค โดยในโครงงานนี้จะใช้พจนานุกรมในการแยกหน่วยคำ
- ส่วนวิเคราะห์ไวยากรณ์และความหมาย (Syntactic & Semantic Analysis) วิเคราะห์ความสัมพันธ์ของคำศัพท์ และความหมายของประโยค
- ส่วนวิเคราะห์เพื่อสร้างภาษากลาง (Conceptual Analysis) ทำหน้าที่สร้างภาษากลางขึ้นในลักษณะโครงข่าย ตามความหมายของประโยค

### ระบบสังเคราะห์ภาษา

ระบบสังเคราะห์ภาษาแบ่งการทำงานออกเป็น 3 ส่วนคือ

- ส่วนเตรียมการข้อมูล (Generation Planning) ทำหน้าที่เตรียมข้อมูลที่มากับภาษากลาง และข้อมูลเกี่ยวกับคำศัพท์พจนานุกรม
- ส่วนสังเคราะห์ภาษากลาง (Semantic & Syntactic Generation) บนพื้นฐานของไวยากรณ์ และความหมายของภาษาเป้าหมาย
- ส่วนจัดการหน่วยคำของประโยค (Morphological Generation) มีหน้าที่จัดเรียงคำศัพท์ให้เป็นประโยคของภาษาเป้าหมาย ตามหลักไวยากรณ์



รูปที่ 1.4 แสดงกระบวนการทำงานส่วนต่างๆของระบบแปลภาษาแบบหลายภาษา

## 1.2 จุดประสงค์ของโครงการ

จุดประสงค์ของการทำโครงการนี้คือ การพัฒนาระบบแยกหน่วยคำของประโยค ระบบการประมวลผลโครงสร้างประโยค และระบบการสร้างรูปแบบภาษากลางของประโยคภาษาไทย รวมแล้วเป็นระบบวิเคราะห์ภาษาเพื่อแปลภาษาต้นทาง(ภาษาไทย) ไปเป็นภาษากลาง ซึ่งเป็นส่วนประกอบส่วนหนึ่งของระบบแปลภาษาด้วยคอมพิวเตอร์ แนวทางในโครงการนี้อาจจะสามารถนำไปประยุกต์ใช้กับงานอื่นๆเช่น ระบบตรวจสอบตัวสะกดคำศัพท์ เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 2

### ความรู้เกี่ยวกับภาษาไทย -

#### 2.1 บทนำ

ในการทำงานด้านการแปลภาษาด้วยคอมพิวเตอร์นั้น สิ่งที่สำคัญอย่างยิ่งคือ ความรู้เกี่ยวกับหลักภาษาไทย ทั้งการจัดประเภทของคำศัพท์ และความสัมพันธ์ของคำศัพท์ในประโยค รวมไปถึงการจัดเรียงคำศัพท์ในประโยค ตามหลักไวยากรณ์ของภาษาไทย ในบทนี้จะได้อธิบายอย่างละเอียดเกี่ยวกับชนิดของคำศัพท์ ตลอดจนหลักไวยากรณ์ของภาษาไทยเพื่อใช้เป็นฐานความรู้ในการวิเคราะห์ประโยคภาษาไทย ซึ่งจะได้อธิบายโดยละเอียดต่อไป

#### 2.2 ชนิดของคำศัพท์ในภาษาไทย

คำ หมายถึง อักษรที่ประสมกันแล้วมีความหมาย ซึ่งแบ่งออกเป็น 7 ชนิดได้แก่

- 1) คำนาม
- 2) คำสรรพนาม
- 3) คำกริยา
- 4) คำวิเศษณ์
- 5) คำบุพบท
- 6) คำสันธาน
- 7) คำอุทาน

##### 2.2.1 คำนาม

คำนาม คือ คำที่เป็นชื่อของคน สัตว์ สถานที่ สิ่งของ และกริยาอาการต่าง ๆ แบ่งย่อยออกเป็น 5 ชนิด คือ

- ก) สามานยนาม คือ คำที่เป็นชื่อของคน สัตว์ สถานที่ สิ่งของ และกริยาอาการทั่วไป เช่น
  - นักเรียน ไป โรงเรียน
  - วิว ชอบกิน กล้วย
  - หนังสือเรียน อยู่ใน กระเป๋า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข) วิสามานยนาม คือคำนามที่เป็นชื่อเฉพาะของ คน สัตว์ สถานที่ และสิ่งของที่กำหนดขึ้นสำหรับใช้เรียก และชื่อเฉพาะเจาะจงลงไปว่าเป็นใคร หรืออะไร เช่น

- จารุณี เป็นคนสวย
- นาย ปราโมทย์ และ อดิสร เป็นพี่น้องกัน
- เขาทำงานที่ กระทรวงพาณิชย์

ค) ลักษณะนาม คือ คำนามที่ทำหน้าที่ประกอบนามขึ้น เพื่อแสดงรูปลักษณะ ขนาด หรือประมาณของนามนั้นให้ชัดเจนยิ่งขึ้น เช่น

- นอน 1 อัน
- สมุด 2 เล่ม
- วัวป่า ฝูง หนึ่ง กำลัง กิน หญ้า

ง) สมุหนาม คือ

1) คำนามที่ทำหน้าที่แสดงหมวดหมู่ของสามานยนามและวิสามานยนาม เช่น

- ฝูง ลิง วิ่ง เล่น บน ต้นไม้
- คณะ กรรมการ กำลัง ประชุม ใน ห้อง ประชุม
- กอง ทัพ ทหาร ตั้ง อยู่ที่ เชิง เขา

2) คำนามที่เป็นชื่อสถานที่ หรือองค์การต่าง ๆ แต่สมมติ ให้เป็นบุคคลขึ้นตามความนิยมของภาษา คือ แทนที่จะให้ความหมายถึงสถานที่ หรือองค์การต่างๆ ตามรูปศัพท์แต่เปลี่ยนความหมายให้เป็นกลุ่มคนผู้มีหน้าที่รับผิดชอบในสถานที่หรือองค์การนั้น ๆ เช่น

- ประชาชน ชาวไทย ยินดี ต้อนรับ ชาว ต่างชาติ
- บริษัท ต้องการ พนักงาน เพิ่ม
- สถาบัน ประกาศ เลื่อน สอบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จ) อาการนาม คือ คำนามซึ่งเกิดจากคำกริยาหรือคำวิเศษณ์ที่มีคำว่า “การ” หรือ “ความ” นำหน้า คำนามชนิดนี้มีลักษณะติดกับคำนามชนิดอื่น คือใช้คำประสมทั้งสิ้น เช่น การยืม การเดิน การนอน ความรู้ ความชั่ว ความมั่งง่าย ฯลฯ

### 2.2.2 คำสรรพนาม

คำสรรพนาม คือ คำที่ใช้แทนนาม หรือข้อความที่กล่าวนามหรือข้อความนั้นซ้ำอีกเพราะภาษาต้องการความไพเราะและความหมัดจดเกลี้ยงเกลา ถ้าต้องการกล่าวคำหรือข้อความซ้ำกันอยู่บ่อย ๆ ก็ขาดความสละสลวย สรรพนามแบ่งย่อยออกเป็น 6 ชนิดคือ

ก) บุรุษสรรพนาม คือสรรพนามที่ใช้แทนชื่อผู้พูด ผู้ที่พูดด้วย และผู้ที่พูดถึง แบ่งออกเป็น 3 ชนิดคือ

- 1) ถ้าใช้แทนชื่อผู้พูด เช่นคำว่า ฉัน ผม ข้าพเจ้า ดิฉัน อาตมา เป็นต้น เรียกว่า บุรุษที่ 1
- 2) ถ้าใช้แทนชื่อผู้พูดด้วย เช่นคำว่า ท่าน เธอ ได้เท่า ฝ่าพระบาท เป็นต้น เรียกว่า บุรุษที่ 2
- 3) ถ้าใช้แทนชื่อผู้ที่พูดถึง เช่นคำว่า เขา มัน ใคร ผู้ใด เป็นต้น เรียกว่า บุรุษที่ 3

ข) ประพันธสรรพนาม คือสรรพนามที่ใช้แทนนาม หรือแทนสรรพนาม ที่อยู่ติดต่อกันข้างหน้า ได้แก่คำว่า ผู้ ที่ ซึ่ง อัน ดัง ผู้ที่ ผู้ซึ่ง ดังตัวอย่าง ต่อไปนี้

- คน ที่ เป็นทหารต้องมีความอดทน
- ความรัก ซึ่ง ทำให้คนตาบอด

ประพันธสรรพนาม ตามตัวอย่างนั้น ทำหน้าที่ในคราวเดียวกัน 2 ประการ คือ

- 1) ใช้แทนคำนาม หรือ สรรพนามที่อยู่ข้างหน้า
- 2) เป็นตัวขยายนาม หรือ สรรพนามที่อยู่ข้างหน้า

คำประพันธ์สรรพนามบางคำ คือ ที่ ซึ่ง อัน ใ้ซ้ำกับคำประพันธ์วิเศษณ์ แต่มีวิธีใช้ต่างกันคือ ถ้าเป็นสรรพนามต้องเรียงติดต่อกับนามหรือ สรรพนาม ถ้าเป็นประพันธ์วิเศษณ์ต้องเรียงติดต่อกับ คำกริยาหรือคำวิเศษณ์ ดังตัวอย่างต่อไปนี้

- คน ที่ ยืนอยู่นั้นเป็นทหาร (ที่ เป็นสรรพนาม)
- เขาเป็นคนดี ที่ ฉลาด พึ่งปรารถนา (ที่ เป็นวิเศษณ์)
- ของ ซึ่ง วางอยู่ในห้องหายไปไหน (ซึ่ง เป็นสรรพนาม)
- มันเป็นของดี ซึ่ง ควรแก่การรักษา (ซึ่ง เป็นวิเศษณ์)

ค) วิภาคสรรพนาม คือสรรพนามที่ไ้แทนนามหรือสรรพนามที่แยกออกเป็น แต่ละคน แต่ละสิ่ง หรือแต่ละพวก ได้แก่คำว่า ต่าง บาง กัน เช่น

- นักเรียน ต่าง ตั้งใจเรียน
- คนงาน ต่าง ก็ขนของขึ้นจากเรือ
- ทั้งสองคนสูงเท่า กัน
- เราเข้าใจซึ่ง กัน และกัน

คำ ต่าง บาง กัน ที่ เป็นวิภาคสรรพนามนั้น จะต้องใช้แทนนามหรือ สรรพนาม ดังตัวอย่างที่แสดงไว้นั้น ถ้าทำหน้าที่ประกอบนาม สรรพนาม กริยา หรือ วิเศษณ์ ต้องนับว่าเป็นคำวิเศษณ์ เช่น

- 1) ประกอบนาม ต่าง คน ต่าง ใจ, ต่าง จิต ต่าง ใจ, คน บ้าง สัตว์ บ้าง
- 2) ประกอบสรรพนาม ต่าง เขา ต่าง เรา, เขา บ้าง เรา บ้าง
- 3) ประกอบกริยา เขาพูด ต่าง กับฉัน, ทำ บ้าง หยุค บ้าง, เขา พยายาม กัน มาก
- 4) ประกอบวิเศษณ์ เขาเป็นคนดี ต่าง กัน, ทำมาก บ้าง ทำน้อย บ้าง

คำ ต่าง และ กัน ยังใช้เป็นคำชนิดอื่นได้อีก เช่น

- ไทย ต่าง กับลาว (ต่าง เป็นกริยา)
- ต่าง หนีขึ้นบนหลังคา (ต่าง เป็นนาม)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- กัน มีลูกสามคน (กัน เป็นบุรุษสรรพนาม)
- เขา กัน เสบียงไว้ส่วนหนึ่ง (กัน เป็นกริยา)

ง) นิยมสรรพนาม คือสรรพนามที่ชี้แทนนาม หรือ ข้อความที่กล่าวมาแล้วหรือที่ ปรากฏ อยู่ เฉพาะหน้า เป็นสรรพนามชี้เฉพาะเพื่อบ่งบอกความให้ชัดเจน ได้แก่ คำต่อไปนี้คือ นี้ โน่น นี่ นั่น โน้น ทั้งนี้ ทั้งนั้น เช่นนี้ เช่นนั้น อย่างนี้ อย่างนั้น อย่างโน้น เช่น

- นี้ คือ อาหารที่ฉันชอบ
- นั่น เป็นความคิดของคนฉลาด
- เขาอยู่ นี้ สบายกว่าอยู่ โน่น
- นี้ เป็นความเห็นของนายกฯ
- โน้น คือภูเขา

จ) อนิยมสรรพนาม คือสรรพนามที่ชี้แทนนามทั่ว ๆ ไป ไม่ใช่เฉพาะเจาะจงเหมือนนิยมสรรพนาม ได้แก่คำต่อไปนี้คือ ใคร ไหน ผู้ใด อื่น ผู้อื่น ผู้ใดผู้หนึ่ง ผู้หนึ่งผู้ใด เช่น

- ใคร จะมากับฉันก็ได้
- ฉันไม่ทราบว่า อะไร เป็น อะไร แล้ว
- ผู้ใด เป็นคนดีเราควรคบผู้นั้น

ฉ) ปฤจฉาสรรพนาม คือสรรพนามที่ชี้แทนนามแต่มีเนื้อความเป็นคำถาม ได้แก่ ใคร อะไร ไหน ผู้ใด ฯลฯ คำเหล่านี้มีใช้อยู่แล้วในอนิยมสรรพนาม แต่ต่างกับอนิยมสรรพนาม เพราะใช้เป็นคำถาม ส่วนอนิยมสรรพนามใช้เป็นคำแทนชื่อที่ไม่ชี้เฉพาะหาได้ใช้เป็นคำถามไม่ เช่น

- เมื่อเช้า ใคร มาหาฉัน (เป็นปฤจฉาสรรพนาม)
- ใคร จะมาหาฉันก็ได้ (เป็นอนิยมสรรพนาม)
- อะไร อยู่ในกระเป๋าค่ะ (เป็นปฤจฉาสรรพนาม)
- ฉันจะรับประทาน อะไร ก็ได้ (เป็นอนิยมสรรพนาม)
- ไหน คือบ้านของท่าน (เป็นปฤจฉาสรรพนาม)
- ฉันจะอยู่ ไหน ก็ได้ (เป็นอนิยมสรรพนาม)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 2.2.3 คำกริยา

คำกริยา คือคำที่แสดงอาการของนามและสรรพนาม แบ่งย่อยออกเป็น 5 ชนิด

ก) สกรรมกริยา คือคำกริยาที่ต้องมีกรรมมารับจึงจะให้ใจความสมบูรณ์ครบถ้วนตาม กระแสความ เช่น

- ฉัน กิน ข้าว
- แม่ หุง ข้าว
- ชาวสวน รดน้ำ ต้นไม้
- พ่อค้า ขาย ของ

จะเห็นว่า คำกริยาทั้ง 5 ประโยคนั้น ถ้าไม่มีคำนามมาเป็นกรรมรับข้างหลังจะไม่ได้ ความครบสมบูรณ์

ข) อกรรมกริยา คือกริยาที่มีความหมายครบถ้วนในตัวเอง โดยไม่ต้องมีกรรมมารับก็ได้ ความหมายชัดเจน เช่น

- คน เดิน ที่ถนน
- นก บิน ในอากาศ
- เธอ นั่ง บนเตียง
- อาจารย์ ยืน ในห้องเรียน
- ต้นไม้ โค่น ข้างถนน

ค) วิตรรกกริยา คือกริยาที่ไม่สำเร็จความหมายในตัวเอง และจะให้เป็นกริยาของ ประธานตามลำพังตัวเองก็ไม่ได้ จะต้องมีคำนาม คำสรรพนาม หรือคำวิเศษณ์ มาขยาย จึง จะได้ความ กริยาพวกนี้ได้แก่คำว่า เป็น เหมือน เท่า คล้าย คือ เสมือน คู่จ คู่ ประจุ ประหนึ่ง ราวกับ เปรียบเสมือน เช่น

## 1) ใช้คำนามขยาย

- นายแดง เป็น ครูใหญ่
- กิ่งก่า คือ สัตว์เลี้ยงคลานชนิดหนึ่ง

## 2) ใช้คำสรรพนามขยาย

- ถ้าเธอ เป็น ฉัน เธอจะรู้สึก
- เขา กลัว ฉันมาก

## 3) ใช้คำวิเศษณ์ขยาย

- เขาพูดอะไร เป็น ตลกไปหมด

ง) กริยานุเคราะห์ คือคำกริยาที่ทำหน้าที่ช่วยกริยาชนิดอื่นให้แสดงความหมายออกมา เป็น กาล มาลา หรือวางก ต่าง ๆ เพราะคำกริยาในภาษาไทยมีรูปคงที่ไม่เปลี่ยนแปลงไปตาม กาล มาลาหรือวางก เหมือนภาษาที่มีวิภัติปัจจัย เช่น ภาษาบาลี สันสกฤต ภาษาอังกฤษ เป็นต้น การที่จะรู้ว่าเป็นกาลมาลา หรือวางกนั้น นอกจากสังเกตความหมายของรูป ประโยคแล้ว จำเป็นจะต้องอาศัยกริยานุเคราะห์ที่เป็นเครื่องช่วยด้วย คำกริยานุเคราะห์ที่มี อยู่มากชนิดด้วยกัน จะจำแนกออกเป็นชนิดตามที่ทำหน้าที่บอกกาลมาลา และวางก ดังต่อไปนี้

## 1) ชนิดบอกกาล

- |                    |  |
|--------------------|--|
| 1.1 บอกอนาคตกาล    | ได้แก่คำ ขอม                           |
| 1.2 บอกปัจจุบันกาล | ได้แก่คำ อยู่ ชัง-ชัง กำลัง กำลัง-อยู่ |
| 1.3 บอกอดีตกาล     | ได้แก่คำ ได้ เคย                       |
| 1.4 บอกอนาคตกาล    | ได้แก่คำ จะ จัก                        |
| 1.5 บอกกาลสมบุรณ์  | ได้แก่ แล้ว เสร็จ                      |

## 2) ชนิดบอกมาลา

- |                     |                                      |
|---------------------|--------------------------------------|
| 2.1) บอกปริกัลปมาลา | ได้แก่คำ ชะรอย รอย เห็นจะ ที่จะ ทำจะ |
| 2.2) บอกศักติมาลา   | ได้แก่คำ คง คงจะ ต้อง                |

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- 2.3) บอกอานัติมาลา ได้แก่คำ จง อย่า จงอย่า ของง ขออย่า อย่า เห่ง  
อย่าเพื่อ ชี นะ เอะ เถิด เทอญ

### 3) ชนิดบอกวาจก

- 3.1) บอกรรควาจก ได้แก่คำ ให้  
3.2) บอกรรนวาจก ได้แก่คำ ถูก ถูก-ให้ ถูกให้  
3.3) บอกรรตวาจก ได้แก่คำ ถูก ถูก-ให้ ถูกให้

จ) กริยาสภาวมาลา คือคำกริยาที่ทำหน้าที่คล้ายกับนาม อาจจะเป็นประธาน เป็นกรรมหรือเป็นบทขยายส่วนใดส่วนหนึ่งของประโยคก็ได้ เช่น

- นอน เป็นการพักผ่อนที่ดี
- เขาชอบ เล่น กีฬา

### 2.2.4 คำวิเศษณ์

คำวิเศษณ์ คือคำที่ทำหน้าที่ประกอบคำนาม สรรพนาม คำกริยา และคำวิเศษณ์ด้วยกันให้ได้ใจความชัดเจนยิ่งขึ้น ตัวอย่าง เช่น

#### 1) ประกอบคำนาม

- คน อ้วน เป็นเพื่อนกับคน ผอม
- โต๊ะ กลม ทาสี เหลือง
- คน ดี มีวาจก ไพเราะ

#### 2) ประกอบคำสรรพนาม

- ท่าน ทั้งหลาย อยากดูใคร บ้าง
- เขา ทั้งหมด จะมาหาเรา ทั้งสอง

#### 3) ประกอบคำกริยา

- เขาพูด เพราะ
- ม้าวิ่ง เร็ว
- ผมมีปากกาคำเดียว ขอรับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

#### 4) ประกอบคำวิเศษณ์

- เขากินอาหาร จุ มาก (ประกอบ จุ)
- คนอ้วน ดูดี วิ่งช้า : (ประกอบ อ้วน)
- เขารักน้องมาก เหลือเกิน (ประกอบ มาก)

คำวิเศษณ์แบ่งย่อยออกเป็น 10 ชนิด คือ

ก) **ลักษณะวิเศษณ์** คือคำวิเศษณ์ที่ประกอบบอกลักษณะต่าง ๆ เช่น บอกชนิดสี สัตว์ฐาน ขนาด เสียง อากาณ กลิ่น รส ความรู้สึก เป็นต้น เช่น ดี ชั่ว ขาว ดำ กลม แบน ใหญ่ เล็ก โครม เปรี๊ยะง เร็ว ช้า หอม เหม็น เปรี๊ยะง หวาน เย็น ร้อน ฯลฯ คำนามและคำกริยาที่นำมาใช้เป็นคำวิเศษณ์ ดดยมากเป็นลักษณะวิเศษณ์ และคำวิเศษณ์ใด ๆ ที่จัดเข้าในคำวิเศษณ์ชนิดอื่น ๆ ไม่ได้ต้งนับว่าเป็นลักษณะวิเศษณ์ทั้งสิ้น

ข) **กาลวิเศษณ์** คือคำวิเศษณ์ที่ประกอบบอกเวลา เช่น เช้า สาย บ่าย เทียง เย็น ค่ำ อืดอนาคต ปัจจุบัน ฯลฯ

ค) **สถานวิเศษณ์** คือคำวิเศษณ์ที่ประกอบบอกสถานที่ เช่น บน ล่าง เหนือ ใต้ ไกล ใกล้ บกน้ำ บ้าน ป่า ฯลฯ และคำวิเศษณ์ชนิดนี้ ถ้ามีคำนามหรือสรรพนามมารับข้างหลังจะกลายเป็นคำบุพบท

ง) **ประมาณวิเศษณ์** คือคำวิเศษณ์ที่ประกอบบอกจำนวน แบ่งออกเป็น 2 ชนิดคือ

- 1) บอกจำนวนนับ เช่น หนึ่ง สอง สาม ที่หนึ่ง ที่สอง ฯลฯ จะใช้เป็นตัวเลขหรือตัวหนังสือก็ได้
- 2) บอกจำนวนประมาณ คือมิได้บ่งบอกชัดเจนไปว่าเท่านั้นเท่านี้ เป็นต้น กำหนดว่ามากหรือน้อยซึ่งพอจะรู้ความหมายได้โดยประมาณ เช่น มาก น้อย หลาย ทั้งหลาย จุ ทั้งปวง ทั้งหมด บรรดา คนละ สิ่งละ ต่าง บาง กิน ฯลฯ

จ) **นิยมวิเศษณ์** คือคำวิเศษณ์ที่ประกอบบอกความชี้เฉพาะหรือจำกัดลงไปว่าเป็นเช่นนั้น เช่นนี้ หรือสิ่งนั้นสิ่งนี้ เช่น นี้ นั้น โนน นี้ นั้น โนน ทั่วนี้ ทั้งนั้น อย่างนี้ อย่างนั้น ดังนี้ ดังนั้น แท้จริง เฉพาะ เอง ดอก แนนอน ที่เดียว เจียว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ฉ) อนิยมวิเศษณ์ คือคำวิเศษณ์ที่ประกอบบอกบอกความไม่ชี้เฉพาะ หรือไม่จำกัดลงไป ว่าเป็นเช่นนั้นเป็นเช่นนี้ หรือเป็นสิ่งนั้นเป็นสิ่งนี้ เช่น ไค ไร ไหน ก็อะไร ทำไม ฉนั้นไค เช่นไร ฯลฯ แต่คำนวนนี้จะต้องไม่ใช่ในข้อความที่เป็นคำถาม หรือสงสัยจึงจะเป็นอนิยมวิเศษณ์ ถ้าใช้ในข้อความที่เป็นคำถามหรือสงสัย เรียกว่า ปฤจฉาวิเศษณ์

ช) ปฤจฉาวิเศษณ์ คือคำวิเศษณ์ที่ประกอบบอกเนื้อความเป็นคำถาม หรือความสงสัยเช่น ไค ไหน ก็ อะไร ทำไม ฉนั้นไค เช่นไร ไหม อันไค อย่างไร เท่าไร ไช หรือ ฯลฯ

ซ) ประดิษฐานวิเศษณ์ คือคำวิเศษณ์ที่ประกอบบอกเสียงเรียกร้องและเสียงขานรับ เพื่อแสดงความสละสลวยของภาษา และแสดงความเป็นกันเองในระหว่างผู้พูดเช่น จำ ขอรับ เว้ย โวย โวยและบรรดาคำรับต่าง ๆ

ฌ) ประดิษฐานวิเศษณ์ คือคำวิเศษณ์ที่ประกอบบอกความปฏิเสธ หรือไม่ยอมรับ เช่น ไม มิ ไม่ใช่ มิใช่ บ บ่ บมิ ฤา มิได้ หาไม่ หามิได้ ฯลฯ

ฎ) ประพันธ์วิเศษณ์ คือคำวิเศษณ์ที่ประกอบบอกคำกริยาหรือคำวิเศษณ์ เพื่อเชื่อมประโยคให้มีข้อความเกี่ยวข้องกัน เช่น ที่ ซึ่ง อัน อย่างที่ ชนิดที่ ให้ ว่า ที่ว่า คือ เพื่อ ให้ เพื่อว่า ซึ่ง มีคำบางคำเหมือนกันกับประพันธ์สรรพนามเช่น ที่ ซึ่ง อัน ซึ่งข้อแตกต่างระหว่างประพันธ์สรรพนามกับประพันธ์วิเศษณ์ดังตารางที่ 2.1

ประพันธ์สรรพนาม	ประพันธ์วิเศษณ์
<ol style="list-style-type: none"> <li>ใช้แทนนามหรือสรรพนามที่อยู่ข้างหน้า</li> <li>ต้องเรียงอยู่ติดกับนาม หรือสรรพนามที่มันแทน</li> </ol>	<ol style="list-style-type: none"> <li>ประกอบคำกริยา หรือคำวิเศษณ์</li> <li>ต้องเรียงอยู่ข้างหลังคำอื่น นอกจากคำนามและสรรพนาม</li> </ol>

ตารางที่ 2.1 เป็นการเปรียบเทียบระหว่าง ประพันธ์สรรพนาม กับ ประพันธ์วิเศษณ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 2.2.5 คำบุพบท

คำบุพบท คือคำที่แสดงความสัมพันธ์ระหว่างคำหรือประโยค เพื่อให้รู้ว่าคำหรือประโยคที่อยู่คำบุพบทนั้นมีหน้าที่เกี่ยวกับคำ หรือประโยคที่อยู่ข้างหน้าอย่างไร ซึ่งหน้าที่ของคำบุพบทมีดังนี้คือ

#### ก) นำหน้านาม เช่น

- หนังสือ ของ บิดาหาย
- เขาไป กับ เพื่อน
- นักเรียนอยู่ ใน ห้อง

#### ข) นำหน้าสรรพนาม เช่น

- ปากกา ของ ฉันอยู่ ที่ เขา
- ฉันคิด ถึง ท่านมาก
- ฉันจะอยู่ ใกล้ เธอ

#### ค) นำหน้ากริยา เช่น

- เขากิน เพื่อ อยู่
- เขาก็กินอาหารไว้ สำหรับ รับประทานอาหาร
- เขาทำงาน กระทั่ง ตาย

#### ง) นำหน้าคำวิเศษณ์ เช่น

- เขาทำ ดี ดี
- ถ้าพูดกัน ตาม จริงแล้ว เขาต้องมาหาฉัน โดย เร็ว

#### จ) นำหน้าประโยค เช่น

- เขายืน ใกล้ กับ นายคำทำงาน
- เขามา ตั้งแต่ ฉันตื่นนอน
- เขาให้รางวัล เฉพาะ คนแต่งโคลงได้ทีหนึ่ง

จากหน้าที่ของคำบุพบท เราจะแบ่งคำบุพบทได้เป็น 5 ชนิดดังต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ก) บุพบทที่แสดงความสัมพันธ์ระหว่าง นามกับนาม นามกับสรรพนาม หรือ นามกับคำกริยา เช่น

- ฉันชอบนาฬิกาของนายแดง
- โต๊ะในห้องเรียนมีสามสิบห้าตัว
- มือของเขาสวยมาก

ข) บุพบทที่แสดงความสัมพันธ์ระหว่างสรรพนามกับนาม สรรพนาม กับสรรพนาม หรือ สรรพนามกับกริยา เช่น

- ท่านจะต้องการอะไรที่ นายแดง
- เธอพบใคร ใน ห้องนั้น

ค) บุพบทที่แสดงความสัมพันธ์ระหว่างกริยากับนาม กริยากับสรรพนาม กริยากับกริยา หรือกริยาวิเศษณ์ บุพบทชนิดนี้อาจนำ

- เขาเห็นแก่หน้ากัน
- เขาอยู่ในบ้าน
- ฉันไปกับเธอ
- เขาไปโดยเร็ว

ง) บุพบทที่ไม่แสดงความสัมพันธ์กับบทอื่น ได้แก่ คำ ดูกร คุณนะ ดูรา ข้าแต่ ยัง

จ) บุพบทที่แสดงความสัมพันธ์ระหว่างประโยคกับระโยค เช่น

- เขาอนตั้งแต่ฝนตกครั้งแรก
- เขาเก็บอาหารไว้สำหรับคนขายหญ้าที่สนาม

### 2.2.6 คำสันธาน

คำสันธาน คือคำที่ทำหน้าที่เชื่อมคำกับคำ ประโยคกับประโยค ข้อความกับข้อความ หรือเชื่อมความให้สละสลวย ชนิดของคำสันธานแบ่งออกเป็น 11 ชนิด

ก) เชื่อมความที่คล้ายตามกัน เช่น และ พอ ถ้า ก็ กับ จึง เช่น ว่า ให้ คือ ทั้ง ก็ดี ก็ได้ เท่า กับ ฯลฯ เช่น

- เขาพบครูและนักเรียน
- บิดามารดาและบุตรไปเที่ยว
- นายดำ กับ นายแดงเป็นเพื่อนกัน

ข) เชื่อมความที่ขัดแย้งกัน เช่น แต่ แต่ว่า แต่ทว่า กว่า...ก็ ถึง...กัน ฯลฯ เช่น

- เขาชอบแกงเผ็ด แต่ ฉันชอบแกงจืด

ค) เชื่อมความที่เป็นเหตุเป็นผลกัน เช่น เพราะ ด้วย จึง ฉะนั้น ฉะนั้น คำที่ ด้วย ว่า เหตุเพราะ เหตุว่า เพราะว่า ฉะนั้น...จึง เพราะฉะนั้น เหตุฉะนั้น ฯลฯ เช่น

- เขาเรียนหนังสือเก่ง เพราะ เขาอ่านหนังสือมาก

ง) เชื่อมความที่เลือกเอา เช่น หรือ มิฉะนั้น ไม่ก็ ไม่เช่นนั้น มิฉะนั้น หรือมิฉะนั้น

- เธอจะไป หรือ ไม่ไป

จ) เชื่อมความแสดงลักษณะอาการ เช่น ให้ ว่า ทั้ง

- เขาพูด ว่า นายแดงเป็นทหาร



ด) เชื่อมความแสดงประมาณ เช่น ตลอด จน จน

- ฉับร่อ จน คนเหล่านั้นกลับหมด

ช) เชื่อมความแสดงเวลา เช่น จน เมื่อ

- เก็บอาหารไว้ จน ปีนึงผ่านไป

- เขานอน เมื่อ นาฬิกาตี 11 ที

ซ) เชื่อมความแสดงเหตุ เช่น เพราะ เพราะว่า

- น้ำท่วม เพราะ ฝนตก

- ต้นไม้โค่น เพราะว่า โคนคนตัด

ณ) เชื่อมความแสดงผล เช่น จน

- เขา ดู จนถูกรองไห้

- ลมพัด จน หน้าต่างเปิด

ญ) เชื่อมความแสดงเปรียบเทียบ เช่น เหมือน ราวกับ กว่า อย่าง

- เขาเดินเร็ว เหมือน ม้าวิ่ง

- เขาทำได้ดี กว่า ผมทำ

ฎ) เชื่อมความให้สะดวก เช่น อัน อันว่า อย่างไรก็ตาม อย่างไรก็ดี ถึงกระนั้นก็ดี คำสันธาน จะเรียงไว้หน้าคำ ข้างหน้าประโยค ข้างหลังคำ ข้างหลังประโยค หรือในระหว่างประโยคก็ได้ดังตารางที่ 2.2

ตำแหน่ง	ตัวอย่าง
หน้าคำ	ฉันเห็นนายคำ และ นายแดง ทานชอบข้าวสวย หรือ ข้าวต้ม
หน้าประโยค	อย่างไรก็ตาม ผมจะมาก่อน 8.00 น. เสมอ อัน ความดีนั้น ใครๆก็ต้องการ
หลังคำ	คน ก็ตาม สัตว์ ก็ตาม ต้องการอาหารทั้งนั้น เสือ ก็ดี กวางก็ดี ก็ดี เราควรเก็บให้เป็นที่
หลังประโยค	เขาจะเป็นใคร ก็ตาม ฉันยินดีรับรองเขาทั้งนั้น
ระหว่างประโยค	เขาเดินมา แต่ ฉันเดินไป เขามาลำบาก เพราะว่า ผันคกหนัก

## ตารางที่ 2.2 ตารางการใช้คำสันธาน

### 2.2.7 คำอุทาน

คำอุทาน คือคำที่แสดงถึงเสียงที่เปล่งออกมาในเวลาใดใจ เสียใจ ตกใจ ประหลาดใจ หรือ  
กริ่งใจ เป็นต้น หรือเป็นคำที่ใช้ต่อถ้อยเสริมบทให้บริบูรณ์ยิ่งขึ้น คำอุทานแบ่งออกเป็น 2 ชนิดคือ

#### ก) อุทานบอกอาการ สำหรับ

1) ใช้แสดงความรู้สึกต่าง ๆ ในการพูดเช่น คำ ชิ ชิๆ ชิชะ ไร๋ บ๊ะ วะ วา อือหือ  
เหม่ แหม่ อนิจจัง อ๊ะ อือ อนิจา อุบ๊ะ เอ เอ๊ะ เอ็ส เอ็อ เฮอ โอ โอย โอัย ฮะ ฮ้า เฮ้  
เฮ้ย เฮ้ว เฮ้อ ไฮ เป็นต้น

2) ใช้เป็นคำขึ้นต้นประโยคในคำประพันธ์ เพื่อแสดงความรำพึง วิงวอน หรือ  
ปลอบโยน เป็นต้น ได้แก่ คำ โอ้ โอ้ว

ข) อุทานเสริมบท คืออุทานที่ใช่เป็นคำสร้อยหรือคำเสริมบทต่าง ๆ เพื่อให้มีคำครบถ้วนตามที่ต้องการ หรือให้ความกระชับ หรือให้สละสลวยขึ้น มี 3 ชนิดคือ

1) อุทานเสริมบทที่ใช่เป็นคำสร้อย คือคำอุทานใช่เป็นสร้อยของโคลงและร่าย หรือใช่เป็นคำลงท้ายในคำประพันธ์ แสดงว่าจบข้อความบริบูรณ์ เช่น ฮา เฮย แอ แล นารา เอย ฯลฯ

2) อุทานเสริมบทที่ใช่เป็นคำแทรก คือคำอุทานที่ใช่แทรกในระหว่างคำ หรือข้อความมี 2 ชนิดคือ

2.1) ใช่เป็นบทบุรุษ คือ เป็นคำที่ทำให้บทประพันธ์มีพยางค์ครบถ้วนตามฉันทลักษณ์ได้แก่คำ นู จี ลี นิ เป็นต้น

2.2) ใช่ประกอบข้างท้ายให้ความกระชับหรือสละสลวยขึ้นได้แก่คำ นา เอย เอย เอย โวย ฯลฯ เช่น

3) อุทานเสริมบทที่ใช่เป็นคำเสริม คือคำอุทานที่ใช่ตัดย่อยเสริมคำให้ยาวยืมนอกไป แต่ไม่ต้องการความหมายที่เสริมนั้น เช่น นาง วา สารา รี กระเดี้ยว ฯลฯ

### 2.3 สัตยลักษณะของกลุ่มคำภาษาไทย

ประโยคภาษาไทยมีการกำหนดตำแหน่งของคำที่ปรากฏในประโยคพื้นฐาน (Simple Sentence) ทัว ๆ ไป ซึ่งมักประกอบด้วยส่วนของนามวลี (Noun Phrase : NP) กับกริยาวลี (Verb Phrase : VP) ซึ่งบางส่วนอาจถูกละไว้เช่น วิเศษณ์วลี (Adverbial Phrase : ADVP) เป็นต้น ที่มักจะปรากฏในประโยคที่มีโครงสร้างซับซ้อน ซึ่งเกิดจากการนำประโยคธรรมดามาเรียงต่อกัน ทำให้ประโยคในภาษาไทย มีได้หลายรูปแบบของประโยค โดยอักษรย่อของคำแต่ละคำในบทนี้จะมีความหมายตามตารางที่ 2.3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คำย่อ	ชนิด	ความหมาย
ADJ	Adjective	คำคุณศัพท์
ADVP	Adverb Phrase	คำวิเศษณ์วลี
AUX	Auxiliary Verb	คำกริยาช่วย
CL	Classifier	คำลักษณนาม
CONJ	Conjunction	คำสันธาน
CPD	Compound Sentence	ประโยคประสม
CPX	Complex Sentence	ประโยคซับซ้อน
DET	Determinative	คำกำหนด
G.M.	Genitive Marker	“ของ”
LINK	Linker	คำเชื่อม
MARK	Marker	คำหมาย
N	Noun	คำนาม
NUM	Numeral	คำนับ
NEG	Negator	คำปฏิเสธ
OPT	Optional Transformation	คำให้เลือก
PCL	Particle	คำลงท้าย
PRET	Preposition	คำบุพบท
PROH.NEG	Prohibitive Negator	คำปฏิเสธห้าม
REL.LINK	Relative Linker	คำเชื่อมประโยค
S	Sentence	ประโยค
V	Verb	คำกริยา
VP	Verb Phrase	กริยาวลี

ตารางที่ 2.3 แสดงสัญลักษณ์ย่อชนิดของคำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 2.4 โครงสร้างของวลีในภาษาไทย

ในประโยคธรรมดาทั่ว ๆ ไปอาจจะประกอบด้วยวลีเพียงอย่างเดียว หรือกริยาวลีเพียงอย่างเดียวก็ได้ ซึ่งเราก็จะเรียกเป็นประโยคเหมือนกัน แต่โดยทั่วไปแล้วประโยคภาษาไทยจะมีโครงสร้างพื้นฐานที่ประกอบเรียงตามลำดับได้ดังนี้

Sentence -----> (ADVP) NP (ADVP) + (ADVP)

แสดงโครงสร้างพื้นฐานของประโยคของภาษาไทย ซึ่งสามารถที่จะยกตัวอย่างตามโครงสร้างพื้นฐานนี้ได้เช่น

(ADVP)	NP	(ADVP)	VP	(ADVP)
วานนี้	ฝน	ที่เชียงใหม่	ตก	หนัก
วานนี้	ฝน		ตก	หนักที่เชียงใหม่
ที่	เชียงใหม่		ฝนตก	หนักวานนี้

แต่ประโยคที่ง่ายที่สุดของภาษาไทยนั้น ในส่วนของนามวลีจะประกอบด้วยคำนาม (Noun : N) เพียงคำเดียว และในส่วนของกริยาวลีจะประกอบไปด้วยคำกริยา (Verb :: V) เพียงคำเดียวเท่านั้น ไม่มีส่วนขยายอื่นมาเพิ่มเติมก็ได้ดังนี้

NP -----> N

VP -----> V

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับคำวิเศษณ์วลีนั้น ส่วนมากมักจะได้แก่ เวลาและสถานที่ ซึ่งจะอยู่ในตำแหน่งต่างๆของประโยคได้ โดยสามารถที่จะอยู่หน้าหรือหลังนามวลี ก็มีความหมายเหมือนกับวางอยู่หลังกริยวลี ดังตัวอย่างต่อไปนี้

N	V
รถ	หยุด
ฝน	ตก
เขา	นอน
เขา	เดิน

เมื่อเราทราบถึงโครงสร้างของประโยคอย่างคร่าวๆ แล้ว ต่อไปจะกล่าวถึงส่วนประกอบโดยละเอียดของ นามวลี กริยวลี และวิเศษณ์วลี

#### 2.4.1 นามวลีและส่วนขยาย

ในส่วนของนามวลีที่มีโครงสร้างง่ายที่สุดจะประกอบด้วยคำนามเพียงตัวเดียว จะไม่มีส่วนขยายใดๆ จนกระทั่งโครงสร้างที่ยาก ๆ ในประโยคผลม (Compound or Complex Sentence) นั้นสามารถที่จะสรุปออกมาเป็นรูปแบบต่างๆ ได้ดังนี้

ก) เป็นนามวลีที่ประกอบไปด้วยคำนาม และตามด้วยคำลักษณนาม (Classifier) หรือคำคุณศัพท์ (Adjective) หรือคำลักษณนาม หรือคำกำหนด (Determinative : DET) (หมายเหตุ คำว่า “และ” หมายความว่า จะต้องมีคำชนิดนั้นๆ ไปประโยคไม่มีไม่ได้ “หรือ” หมายความว่า จะมีหรือไม่มีคำชนิดนั้นๆ ในประโยคก็ได้)

NP -----> N + (CL) (ADJ) (CL) (DET)

(หมายเหตุ ชนิดของคำในเครื่องหมายวงเล็บจะมีหรือไม่มีก็ได้)

จากรูปแบบของประโยคนี้สามารถที่จะยกตัวอย่างประโยคได้ดังนี้

N	(CL)	(ADJ)	(CL)	(DET)
หนังสือ				
หนังสือ เล่ม				นี้
ไก่		ดำ		
ไก่	ตัว	ดำ		
ไก่	ตัว	ดำ		นี้
ไก่	ตัว	ใหญ่	ตัว	นี้

ข) นามวลีที่ประกอบด้วย นามวลี และหรือ “ของ” และนามวลี และหรือ “ของ” และนามวลี และหรือ คำลักษณะนาม หรือ คำกำหนด

NP -----> NP + (G.M) NP + (G.M) NP+ (CL) (DET)

ตัวอย่างเช่น

NP	(G.M)	NP	(G.M)	NP + (CL)	(DET)
บ้าน		เพื่อน			
บ้าน		ของ		เพื่อน	
บ้าน	ของ	เพื่อน	ของ	ผม	
แมว	ของ	ป่า	ของ	ผม	ตัว นี้

ค) นามวลีที่ประกอบด้วย คำนาม และหรือ คำลักษณะนาม และคำคุณศัพท์ และหรือ “ของ” และนามวลี

NP -----> N (CL) + ADJ (G.M) + NP

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## ตัวอย่างเช่น

N	(CL)	ADJ	(G.M)	NP
เสื้อ	ตัว	ใหม่		
เสื้อ		สีดำ	ของ	แม่
เสื้อ	ตัว	ใหม่	ของ	แม่
รองเท้า	คู่	เก่า	ของ	พ่อ
นาฬิกา	เรือน	ใหม่	ของ	ผม

ง) นามวลีที่ประกอบด้วย คำนาม และหรือ คำคุณศัพท์ และนับ และ คำลักษณะนาม และหรือ คำกำหนด

NP -----> N (ADJ) + NUM CL (DET)

ตัวอย่างเช่น

N	(ADJ)	NUM	CL	(DET)
คน		สอง	คน	
คน	แก่	สอง	คน	
บ้าน	ใหญ่	สอง	หลัง	นั้น
คน	หลาย	คน		
คน	บาง	คน		

จ) นามวลีที่ประกอบด้วย คำนาม และหรือ คำคุณศัพท์ และ ตัวลักษณะนาม และ คำนับ และหรือ คำกำหนด

NP -----> N (ADJ) + CL + NUM (DET)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวอย่างเช่น

N	(ADJ)	CL	NUM	(DET)
บ้าน		หลัง	เดียว	
คน		คน	เดียว	
บ้าน	ใหญ่	หลัง	เดียว	
บ้าน	ใหญ่	หลัง	หนึ่ง	
บ้าน	ใหญ่	หลัง	ที่หนึ่ง	
บ้าน	ใหญ่	หลัง	ที่สอง	นั้น

ฉ) นามวลีที่ประกอบด้วย คำนาม และ คำเชื่อม และคำนามคำนับ และหรือ คำลักษณนาม คำหมาย หรือ คำกำหนด

NP → N + LINK + NUM (CL) (DET)

MARK

ตัวอย่างเช่น

N	LINK	N	NUM	(CL)	(DET)
คน	ทั้ง	บ้าน			
นก	ทั้ง			ฝูง	
แมว	ทั้ง	สอง		ตัว	นั้น
คน	ทั้ง	สอง			
ห้อง	ทั้ง	สาม		ห้อง	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข) นามวลีที่ประกอบด้วย คำหมาย และ คำนาม หรือ คำสรรพนาม และหรือ คำหมาย

NP -----> MARK + N / PRO (MARK)

ตัวอย่างเช่น

	MARK	N /	PRO	(MARK)
	พวก	ครู		
	พวก		เรา	
	พวก		เรา	ทั้งหลาย
	พวก	ตำรวจ		ทั้งหลาย
ข) นามวลีที่ประกอบด้วย คำนาม และ คำสรรพนาม		คำนาม	และ คำสรรพนาม	
	NP ----->	N +	PRO	
ตัวอย่างเช่น		N	PRO	
		พ่อ	ท่าน	
		แมว	มัน	

ฉ) นามวลีที่ประกอบด้วย คำนาม และ คำเชื่อม และหรือ คำสัณฐาน และ คำนาม

NP -----> N + LINK (CONJ) + N

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## ตัวอย่างเช่น

N	LINK	(CONJ)	N
พ่อ	กะ		แม่
พ่อ	กับ		แม่
โต๊ะ		หรือ	เก้าอี้
สมุด		และ	หนังสือ

ญ) นามวลีที่ประกอบด้วย คำนาม และหรือ คำลักษณนาม และ คำคุณศัพท์ และ คำลักษณนาม และหรือ คำกำหนด คำนับ

				DET
	NP	----->	N (CL) + ADJ + CL	
				NUM
ตัวอย่างเช่น				
N	(CL)	ADJ	CL	DET / NUM
บ้าน		ใหญ่	หลัง	นั้น
บ้าน	หลัง	ใหญ่	หลัง	นั้น
บ้าน	หลัง	ใหญ่	หลัง	ที่สอง

ฎ) นามวลีที่ประกอบด้วย คำเชื่อม และ คำนาม และ คำเชื่อม และ คำนาม

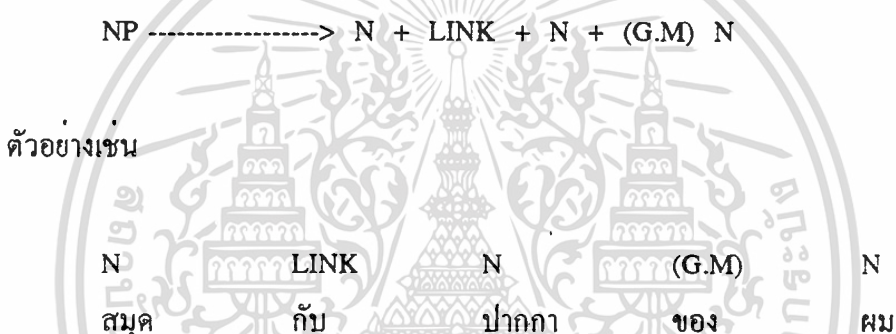
NP -----> LINK + N + LINK + N

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวอย่างเช่น

LINK	N	LINK	N
ทั้ง	พ่อ	ทั้ง	แม่
ทั้ง	นก	และ	หนู
ทั้ง	พี่	และ	น้อง

ฎ) นามวลีที่ประกอบด้วย คำนาม และ คำเชื่อม และ คำนาม และหรือ “ของ” และ คำนาม



#### 2.4.2 กริยาวลีและส่วนขยาย

ส่วนประกอบของกริยาวลีที่น้อยที่สุด จะประกอบด้วยคำกริยาเพียงตัวเดียว แต่เมื่อรวมส่วนขยายต่างๆ ทำให้กริยาวลีมีโครงสร้างที่แตกต่างกันออกไป ซึ่งแสดงได้ดังนี้

ก) กริยาวลีที่ประกอบด้วย คำกริยา และหรือ คำนาม/คำสรรพนาม และหรือ คำขยาย และหรือ คำลงท้าย

VP -----> V + N/PRON (MARK) (PCL)



ตัวอย่างเช่น

V	(N / PRO)	MARK	NP
ทำงาน		เพื่อ	เงิน
กิน	ข้าว	ด้วย	มือ
เห็น		กับ	ตา

ง) กริยาวลีที่ประกอบด้วย คำกริยา และ นามวลี และหรือ คำกริยา และ คำนับ และ คำลักษณนาม

ตัวอย่างเช่น

VP -----> V + NP (V) + NUM + CL

V	NP	(V)	NUM	CL
ส่ง	จดหมาย	ไป	สอง	ฉบับ
มี	เงิน	อยู่	ร้อย	บาท
ให้	หนังสือ	ไป	สอง	เล่ม

จ) กริยาวลีที่ประกอบด้วย คำกริยาช่วย และ คำกริยา และหรือ คำนาม/คำสรรพนาม และ หรือ คำทวน และหรือ คำลงท้าย

VP -----> AUX = V (N/PRO) (MARK) (PCL)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## ตัวอย่างเช่น

AUX	V	(N / PRO)	(MARK)	(PCL)
กำลัง	อ่าน	จดหมาย		
จะ	ไป		แล้ว	นะ

ฉ) กริยาวลีที่ประกอบด้วย คำกริยาช่วย และคำกริยา และหรือ คำบุรพบท และ นามวลี และหรือ คำหมาย และ คำกริยา และหรือ คำนาม/คำสรรพนาม

VP -----> (AUX) V + (PREP) NP + (MARK) V (N/PRO)

## ตัวอย่างเช่น

(AUX)	V	(PREP)	NP	(MARK)	V	(N/PRO)
จะ	ไป	ที่	ตลาด	เพื่อ	ซื้อ	ปลา
จะ	ไป	ที่	บ้าน	เพื่อ	ทาน	ข้าว

ช) กริยาวลีที่ประกอบด้วย คำกริยาช่วย และ คำกริยา และหรือ คำนาม/คำสรรพนาม และ คำกริยา และหรือ คำบุรพบท และ คำนาม

VP -----> (AUX) V (N/PRO) + V + (PREP) N/PRO

## ตัวอย่างเช่น

(AUX)	V	(N / PRO)	V	(PREP)	N / PRO
จะ	เอา	เงิน	ให้		ผม
	ซื้อ	ปากกา	ให้		น้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ซ) กริยาวลีที่ประกอบด้วย คำกริยาช่วย และ คำกริยา และหรือ นามวลี และ คำเชื่อม และ นามวลี

VP -----> (AUX) V (NP) + LINK + NP

ตัวอย่างเช่น

(AUX)	V	(NP)	LINK	NP
จะ	ไป	วัด	กับ	แม่
	ซื้อ	ผัก	กับ	หมู
จะ	ฟัง	วิทยุ	กับ	น้อง

ฅ) กริยาวลีที่ประกอบด้วย คำกริยาช่วย และ คำกริยา และหรือ นามวลี และ คำเชื่อม และ ประโยค

VP -----> (AUX) V (NP) + LINK + S

ตัวอย่างเช่น

(AUX)	V	(NP)	LINK	S
จะ	ทาน	ข้าว	แล้ว	ไปดูหนัง

ฉ) กริยาวลีที่ประกอบด้วย กริยาช่วย และ คำกริยา และหรือ นามวลี และ คำกริยา และ หรือ บุพบทวลี และหรือ นามวลี และหรือ คำพยาน

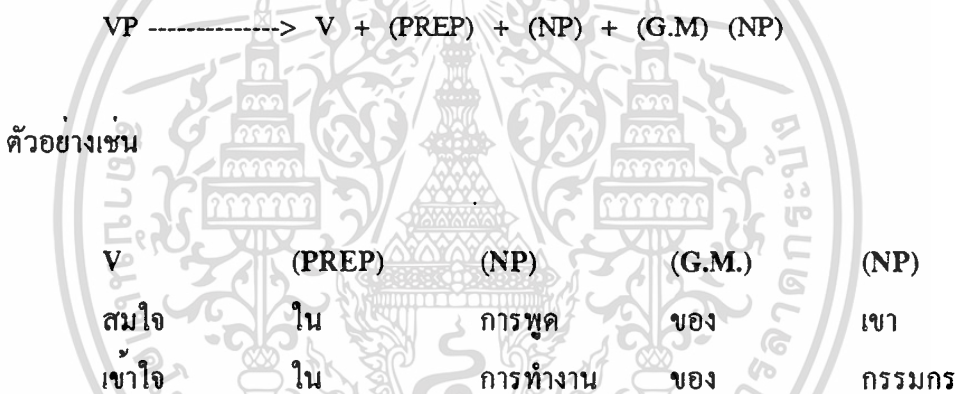
VP -----> (AUX) V (NP) + V (PREP P) (NP) (MARK)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวอย่างเช่น

(AUX)	V	(NP)	V	(PREP)	(NP)	(MARK)
เคย	เขียน	จดหมาย	ไป	ที่	เขา	แล้ว
จะ	พา	เขา	ไป	ที่	บ้านพรุ่งนี้	

ฎ) กริยาวลีที่ประกอบด้วย คำกริยา และหรือ คำบุรพบท และ คำนาม และหรือ “ของ” และ นามวลี



#### 2.4.8 วิเศษณ์วลีและส่วนขยาย

ในส่วนของวิเศษณ์วลีนี้ จะเป็นกลุ่มคำที่แสดงความสัมพันธ์ระหว่างคำ หรือประโยค โดยจะบ่งบอกเกี่ยวกับ สถานที่ ลักษณะ กาลเวลา ซึ่งคำเหล่านี้สามารถที่จะอยู่ต้น หรืออยู่ท้าย ประโยคก็ได้ และในบางครั้งก็อาจจะปรากฏอยู่กลางประโยคก็ได้ โดยไม่ทำให้ความหมายโดยรวมเปลี่ยนไป ในที่นี้กำหนดให้คำเหล่านี้เป็นคำวิเศษณ์ในภาษาไทยทั้งหมด ซึ่งมีรูปแบบที่ปรากฏในภาษาได้ดังนี้

ก) วิเศษณ์ที่ประกอบด้วย คำวิเศษณ์ และ คำกำหนด

ADVP -----> ADV + DET

ตัวอย่างเช่น

ADV	DET
เร็ว	นี้

ข) วิเศษณ์วลีที่ประกอบด้วย คำบุพบท และ นามวลี และหรือ คำกำหนด

ADVP -----> PREP + NP (DET)

ตัวอย่างเช่น

PREP	NP	(DET)
ใน	บ้าน	
ที่	โรงเรียน	
ใน	วันนี้	
โดย	เครื่องบิน	
ใน	เดือน	นี้

ค) วิเศษณ์วลีที่ประกอบด้วย คำปฏิเสธ และ คำคุณศัพท์ หรือ คำวิเศษณ์ หรือ คำกริยา

ADVP -----> NEG + ADJ / ADV / V

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## ตัวอย่างเช่น

NEG	ADJ	/	ADV	/	V
ไม่	ดี				
ไม่	สวย				
ไม่	สะดวก				
ไม่	เลว				

ง) วิเศษณ์วลีที่ประกอบด้วย คำวิเศษณ์ และ คำวิเศษณ์ และหรือ คำวิเศษณ์



## บทที่ 8

### การแสดงรูปแทนภาษากลางสำหรับภาษาไทย

#### 8.1. บทนำ

รูปแทนภาษากลางเป็นโครงข่ายทางความหมาย (semantic network) ของประโยครูปแบบ และข้อกำหนดภาษากลางออกแบบโดย อุชิดะ (Uchida : 1991) และพัฒนาโดยนักวิจัยจาก 5 ประเทศที่เข้าร่วมโครงการเรื่องการเครื่องแปลภาษาสำหรับภาษาในเอเชีย ซึ่งได้แก่ประเทศญี่ปุ่น จีน มาเลเซีย และไทย ทุกประเทศที่เข้าร่วมโครงการใช้ข้อกำหนดนี้เพื่อแสดงรูปแทนทางความหมาย (semantic representation) สำหรับเครื่องแปลภาษาของประเทศนั้น ๆ ในโครงงานนี้ขอเสนอ การแทนรูปประโยคภาษาไทยด้วยโครงข่ายความหมายตามโครงสร้างของภาษากลาง

#### 8.2. โครงสร้างของรูปแทนภาษากลาง

รูปแทนภาษากลางแสดงความสัมพันธ์ของมโนทัศน์ 2 มโนทัศน์ด้วยการกัมพันธ์ และคุณลักษณะในรูปของ tuple form ดังนี้

<มโนทัศน์>	::=<มโนทัศน์หลัก> <มโนทัศน์ประกอบ>
<มโนทัศน์>	::= {<ความสัมพันธ์แบบ ไบนารี> [<ความสัมพันธ์แบบ ไบนารี>...] [<ความสัมพันธ์แบบยูนารี>] [<ความสัมพันธ์แบบยูนารี>[<ความสัมพันธ์แบบ ไบนารี>...] [<ความสัมพันธ์แบบยูนารี>]}
	(<คุณลักษณะ>, รูปว่าง, <มโนทัศน์>, <มโนทัศน์>)
<ความสัมพันธ์แบบ ไบนารี>	::=(<ความสัมพันธ์>, <มโนทัศน์>, <มโนทัศน์>)
<ความสัมพันธ์แบบยูนารี>	::=(<คุณลักษณะ>, <มโนทัศน์>, รูปว่าง)
<มโนทัศน์หลัก>	::=รูปที่แทนมโนทัศน์
<ความสัมพันธ์>	::=รูปสำหรับแทนความสัมพันธ์ระหว่างมโนทัศน์
<คุณลักษณะ>	::=รูปสำหรับแทนคุณลักษณะของมโนทัศน์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 3.2.1 มโนทัศน์หลัก

มโนทัศน์หลักจะอธิบายมโนทัศน์ประโยค มโนทัศน์หลักที่ใช้ในภาษากลางนี้เป็นหลักที่พัฒนาขึ้นโดย EDR(Japan Electronic Dictionary Research Institute, Ltd.) ประเทศญี่ปุ่น ที่จริงการกำหนดมโนทัศน์ของแต่ละภาษาเป็นผลมาจากการที่เจ้าของภาษารับรูปร่างโครงสร้างของการแยกแยะ (conceptualization) ต่าง ๆ ในโลก ซึ่งแต่ละชาติและแต่ละภาษาจะต่างกันออกไป เพราะวัฒนธรรมและการรับรู้สิ่งต่าง ๆ ในโลก แต่ในที่นี้มโนทัศน์หลักจะกำหนดไว้โดยละเอียดและครอบคลุมเพื่อที่จะให้ทุกภาษาสามารถติดต่อกันได้ด้วยมโนทัศน์หลักนี้

### 3.2.2 ความสัมพันธ์

ความสัมพันธ์แสดงว่ามโนทัศน์สองมโนทัศน์มีความสัมพันธ์กันอย่างไร ความสัมพันธ์ดังกล่าวแบ่งออกเป็นแบบต่าง เช่น การกลับพันความสัมพันธ์ของเหตุการณ์ ความสัมพันธ์ทางความหมาย ความสัมพันธ์เทียม เป็นต้น

ต่อไปนี้เป็นรายละเอียดของความสัมพันธ์ที่ใช้แสดงความหมายในภาษาไทย

#### ก.) การกลับพัน

เป็นความสัมพันธ์ที่แสดงการกลับพันระหว่างมโนทัศน์ 2 มโนทัศน์

agent : ใช้แสดงความสัมพันธ์ว่าเป็นผู้กระทำเหตุการณ์อย่างตั้งใจ ผู้กระทำที่มีความตั้งใจ ได้แก่ มนุษย์ สัตว์ องค์กร คอมพิวเตอร์ ปราบกฏการณ์ ธรรมชาติ และสิ่งที่เคลื่อนที่ด้วยเครื่องยนต์

ตัวอย่างที่ 1. ผู้กระทำที่เป็นองค์กรเราถือว่องค์กรเป็นนิติบุคคลที่ทำหน้าที่เหมือนบุคคลตามกฎหมาย

ป.ย 1 บริษัทให้พนักงานมาทำงาน 8 โมง

company<--agent--c#order--object-->c#come-agent->c#employee

└--purpose-->c#work

c#3<--nuber--c#hour<--time

ตัวอย่างที่ 2. ผู้กระทำที่เป็นปรากฏการณ์ธรรมชาติ ในที่นี้เราถือว่าปรากฏการณ์ธรรมชาติเกิดขึ้นเองโดยไม่มีผู้ใดจงใจเหตุการณ์ธรรมชาติทำให้เกิดเหตุการณ์โดยความตั้งใจ

ป.ย.2 ฝนตกมาก

c#rain<--agent--c#fall--manner--c#hard

ตัวอย่างที่ 8. ผู้กระทำที่เป็นสิ่งเคลื่อนที่ได้ด้วยเครื่องยนต์ เช่นรถยนต์ที่ขับเคลื่อนด้วยเครื่องยนต์ หรือโปรแกรม

ป.ย.3 รถไฟวิ่งเร็วมาก

c#rain <--agent-- c#run--manner--c#fast<--object--c#very

Object : แสดงความสัมพันธ์ที่เป็นผู้กระทำ หรือผู้ทรงสภาพ

ตัวอย่างที่ 1. ผู้ถูกกระทำหมายถึง ผู้รับผลจากการเปลี่ยนแปลงของเหตุการณ์

ป.ย.4 เราซื้อหนังสือที่ร้านหนังสือ

c#we<--agent--c#buy--object-->c#  
...  
--play-->c#bookshop

ตัวอย่างที่ 2. ผู้ทรงสภาพ หมายถึงการเป็นเงื่อนไขของเหตุการณ์ หรือบอกการคงอยู่ของสิ่งหนึ่ง

ป.ย.5 มีกระดาษบนโต๊ะ

c#exist--object-->c#paper<--object--c#nil--place-->c#table

ตัวอย่างที่ 3. ผู้รองรับความสัมพันธ์ บางครั้งเราต้องการแสดงความมากน้อยของคุณลักษณะ เราจะใช้ 'object' เพื่อแสดงความสัมพันธ์แบบนี้

ป.ย.6 ดอกไม้สวยมาก

c#flower<--a-object--c#beautiful<--object--c#very

a-object : คุณลักษณะของวัตถุ ความสัมพันธ์แบบนี้แสดงถึงคุณสมบัติของวัตถุที่หมายรวมถึง รูป

ร่าง ลักษณะ สี รสชาติ ธรรมชาติ และการประเมิณผล รวมทั้งการบรรยาย

ตัวอย่างที่ 1. รูปทรง

ป.ย.8 ลูกบอลกลม

c#ball<--a-object--sphere

ข. สี

ป.ย. 9 ลูกบอลแดง

c#ball<--a-object--c#red

ค. การประเมีนผล

ป.ย.10 พ่อเป็นเซลล์แมน

c#father<--a-object--c#salesman

**manner** : วิธีการกระทำหรือปรากฏการณ์เปลี่ยนแปลง ใช้แสดงวิธีการเปลี่ยนแปลง ผลของการกระทำ หรือช่วงเวลาความถี่ของเหตุการณ์

ตัวอย่างที่ 1. วิธีการของการกระทำ

ป.ย.11 นักเรียนเรียนอย่างตั้งใจ

c#student<--agent--c#learn--manner-->intentionally

ข. ผลของการกระทำ

ป.ย.12 นักเรียนเรียนหนักมาก

c#student<--agent--c#learn--manner>c#hard<--object--c#very

ค. ช่วงเวลาของความถี่ของเหตุการณ์

ป.ย.13 นักเรียนอยู่ที่ห้อง 3 ชั่วโมง

c#student<--agent--c#stay--manner-->hour---number--c#3

place

c#room

**implement** : เครื่องมือหรือวิธีการ ในการกระทำที่ทำให้เหตุการณ์เปลี่ยนแปลง

ตัวอย่างที่ 1. เครื่องมือหมายถึงอุปกรณ์ที่ผู้กระทำใช้กระทำเหตุการณ์

ป.ย.14 เขาขว้างก้อนหินใส่ผู้คน

c#he<--agent--c#throw--implement-->c#stone

goal

c#c#people

ตัวอย่างที่ 2. วิธีการหมายถึงการที่ใช้ในการกระทำรวมถึง วิธีการ ขบวนการหรือสิ่งของที่ใช้ในการกระทำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ป.ย.15 ฉันทันเดินทางด้วยรถไฟ

c#I<--agent--c#travel--implement-->c#train

material : วัสดุที่ใช้เป็นส่วนประกอบหรือองค์ประกอบของสิ่งของหรือสภาพหนึ่งความสัมพัทธ์นี้ใช้เพื่อแสดงองค์ประกอบของสิ่งของ

ตัวอย่างที่ 1. วัสดุที่ใช้ประกอบการสร้างสิ่งใหม่ ในภาษาไทยส่วนใหญ่จะใช้กริยาวลีว่า "ทำด้วย"

ป.ย.16 เก้าอี้ทำด้วยพลาสติก

c#chair<--object--c#made of-- material-->c#plastic

ข. องค์ประกอบของวัตถุหมายถึงวัตถุนั้นประกอบไปด้วยองค์ประกอบอะไร

ป.ย.17 น้ำประกอบด้วยไฮโดรเจนและออกซิเจน

c#compose---material-->(--main-->c#hydrogen--and-->c#oxygen)--object-->  
water

time : เวลาที่เหตุการณ์นั้นเกิด ใช้ในความหมายทั้งเวลาที่เกิดขึ้นจริง และเวลาสัมพัทธ์  
ตัวอย่างที่ 1. เวลาจริง หมายถึงเวลาที่เหตุการณ์นั้นเกิดขึ้นตามเวลาจริง เช่น เวลาที่บ่งได้เป็น วันที่  
เดือน ปี ฤดู หรือวันเทศกาล

ป.ย.18 เขาเกิดวันที่ 7 สิงหาคม 1960

main----->

<---main

c#born--time-->(c#date--modifier-->7

and

#month--modifier-->c#august

and

c#year--modifier-->c#1960

ตัวอย่างที่ 2.ข. เวลาสัมพัทธ์ หมายถึงเวลาที่สัมพันธ์กับเหตุการณ์อื่นหรือสถานที่อื่น

ป.ย.19 นักเรียนเล่นกันตอนครูสอนหนังสือ

c#play---time-->c#teach--agent-->c#teacher

..... agent-->c#student

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



condition : เงื่อนไขการเกิดเหตุการณ์ ให้แสดงความสัมพันธ์แบบเงื่อนไขหรือการก่อให้เกิดเหตุการณ์

ตัวอย่าง ก. เงื่อนไขของเหตุการณ์ (conditional event)

ป.ย.26 คุณต้องทำงานให้เสร็จภายในเดือนนี้

c#do--purpose-->c#complete--condition-->c#month---->this--agent-->c#you

ข. เหตุการณ์เป็นเหตุเป็นผลกัน (causal event)

ป.ย.27 ฝนตกหนักมากน้ำจึงท่วม

c#flood--object-->c#water

--condition-->c#fall--manner-->c#hard

--agent-->#rain

cooccurrence : การเกิดพร้อมกันแสดงความสัมพันธ์ที่เหตุการณ์ 2 เหตุการณ์เกิดขึ้นพร้อมกัน เช่น

ป.ย.28 ฉันนั่งกินไอศกรีม

c#I<--agent--c#sit--cooccurrence-->c#eat--object-->c#icecream

sequence : เวลาที่ต่อเนื่อง หมายถึงการแสดงความสัมพันธ์ที่เกิดขึ้นต่อเนื่องกัน เช่น

ป.ย.29 ฉันนั่งลงแล้วก็วาดการ์ตูน

c#I<--agent--c#sit down<--sequence--c#object-->c#cartoon

quantity : หมายถึงปริมาณของจำนวนสิ่งของที่เปลี่ยนไปในเหตุการณ์หรือจำนวนปริมาณของปรากฏการณ์ บางอย่าง

ตัวอย่าง ก. ปริมาณสิ่งของ

ป.ย.30 ส้ม 1 กิโล

c#orange--quantity-->c#kilogram--number-->c#1

ข. ปริมาณของการเปลี่ยนแปลง

ป.ย.31 ฉันทิ้งน้ำหนักลดลง 3 กิโล

c#I<--agent--c#lose--quantity-->c#kilogram--number-->c#3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

number : จำนวนนับของสิ่งของ ใช้เพื่อแสดงปริมาณของสิ่งของด้วยจำนวนนับ

ป.ย.32 นกบินมา 3 ตัว

c#fly--agent-->c#bird--number-->c#3

basis : ฐานของความสัมพันธ์ ใช้เพื่อแสดงการเปรียบเทียบ

ป.ย.33 บ้านฉันสวยกว่าบ้านเธอ

c#beautiful<--object--c#more--basis->c#house---possessor->c#you

and : การเชื่อมมโนทัศน์หรือเหตุการณ์ในทางคล้ายคลึงกัน

ตัวอย่างที่ 1. การเชื่อมเหตุการณ์ 2 เหตุการณ์

ป.ย.34 วันนี้อากาศเย็นและลมแรง

(--main-->c#cloudy--and-->c#windy)--a-object-->c#weather

ตัวอย่างที่ 2: การเชื่อมมโนทัศน์ 2 มโนทัศน์

ป.ย.35 ฉันซื้อหนังสือและสมุดมากมาย

(main---->c#book--and-->c#notebook)<--a-object--c#many

or : การเชื่อมแบบเลือกอย่างใดอย่างหนึ่ง ใช้แสดงเพื่อการเลือกอันเนื่องมาจากความต่างกันของมโนทัศน์ 2 มโนทัศน์ หรือ เหตุการณ์ 2 เหตุการณ์

ตัวอย่างที่ 1. การเลือกระหว่างมโนทัศน์ 2 มโนทัศน์

ป.ย.36 คุณ ชอบ นก หรือ แมว

c#like--object-->(main---->c#bird--or-->c#cat)

--agent-->c#you

ตัวอย่างที่ 2. คุณ จะ ไป หรือ ไม่ ไป

main---->(c#go--or-->c#go---->not)--agent-->c#you

purpose : จุดประสงค์ของการกระทำ หมายถึงการแสดงถึงการกระทำที่เป็นจุดประสงค์  
ของผู้ทำว่ามุ่งที่จะทำ หรือ ไม่ทำเหตุการณ์นั้น

ป.ย.38 พ่อต้องทำงานหนักเพื่อเลี้ยงลูก

c#work--purpose-->c#take care--object-->c#children

agent-->c#father

modifier : การขยายมโนทัศน์ ใช้เพื่อแสดงการที่มโนทัศน์มีส่วนที่ขยายเป็นอีกมโนทัศน์  
หนึ่ง

ตัวอย่างที่ 1. การขยายมโนทัศน์

ป.ย.39 ความก้าวหน้าของชาติ

c#progress---modifier-->c#nation

ตัวอย่างที่ 2. การแสดงประเภทเดียวกันของมโนทัศน์

ป.ย.40 มะม่วงเป็นผลไม้

c#fruit--modifier-->c#mango

ตัวอย่างที่ 3. การขยายด้วยจำนวนนับ

ป.ย.41 วันที่ 5

c#day--modifier-->c#5

ข.) Pseudo relation หมายถึงความสัมพันธ์ที่ไม่ได้เกิดขึ้นจริง แต่เป็นการตีความ มีอยู่ 3  
แบบดังนี้

possessor : การเป็นเจ้าของ ใช้เพื่อแสดงความสัมพันธ์ระหว่างเจ้าของและสิ่งของ

ป.ย.42 หนังสือของฉัน

c#book---possessor-->c#I

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ป.ย.43 โทรทัศน์ช่อง 5 เป็นของทหาร

c#T.V.---possessor-->c#military

---modifier-->c#chanel---modifier-->c#5

from to : ขอบเขตของวัตถุและเหตุการณ์ แสดงถึงขอบเขตของเวลา และระยะทางของ  
มโนทัศน์ 2 มโนทัศน์

ตัวอย่างที่ 1. ขอบเขตของเวลา

ป.ย.44 เราทำงานตั้งแต่ 5 โมงถึง 7 โมง

c#work--manner-->(main---->c#o'clock--from-to-->c#o'clock

--agent-->c#we-----modifier-->c#7-----modifier--c#5)

ตัวอย่างที่ 2. ช่วงของระยะทาง

ป.ย.45 เราเดินทางจากกรุงเทพถึงโตเกียว

c#travel--modifier-->(main-->c#tokyo---from-to-->c#bangkok)

--agent-->c#we

unit : หน่วยของสิ่งของและเหตุการณ์ ใช้เพื่อแสดงเกณฑ์มาตรฐานของการวัด เช่น เมื่อ  
เราต้องการบอกมาตรฐานของสิ่งของตามหน่วยมาตรวัด

ป.ย.46 เราซื้อมะม่วงกิโลละ 20 บาท

(main---->c#bath--number-->c#20)<--unit--c#kilogram--number-->c#1

c#buy--manner-->:

beneficiary : บุคคลที่ได้รับประโยชน์จากการกระทำใช้เพื่อแสดงความสัมพันธ์ของ  
บุคคลที่ได้รับประโยชน์หรือเสียผลประโยชน์จากการกระทำอย่างใดอย่างหนึ่ง

ป.ย.47 พ่อทำงานเพื่อลูก

c#work---beneficiary-->c#children

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ค.) Semantic relation เป็นความสัมพันธ์ทางความหมายของมโนทัศน์ ได้แก่

part-of : เป็นการแสดงเรื่องความสัมพันธ์แบบส่วนใหญ่-ส่วนย่อย โดยทั่วไปความสัมพันธ์แบบนี้จะปรากฏอยู่ในพจนานุกรมประเภท concept classification แต่ถ้าพบความสัมพันธ์นี้ปรากฏอยู่ในโครงสร้างประโยคให้ใช้ "modifier" แทน

ตัวอย่างที่ 1. ความสัมพันธ์ทางความหมาย

ป.ย.48 หลังคาบ้านผุแล้ว

c#roof--part of-->c#house

ตัวอย่างที่ 2. ในโครงข่ายภาษากลางที่ปรากฏ

หลังคาบ้านผุแล้ว

c#house--modifier-->roof

kind of : เป็นความสัมพันธ์แบบชนิดหนึ่ง บอกความสัมพันธ์ที่เป็น upper class และ lower class ใช้อธิบายความสัมพันธ์ทาง paradigmatic ในเรื่องลำดับของมโนทัศน์ ความสัมพันธ์ชนิดนี้จะเก็บไว้ในพจนานุกรม เช่น c#human--kind of-->c#animal

### 3.2.3 คุณลักษณะ(attributes)

คุณลักษณะแสดงถึงคุณสมบัติของมโนทัศน์ เกี่ยวกับการบอกระยะไกล ไกล เวลา การณ์ ลักษณะ การปฏิเสธ ความเห็นและเจตนาของผู้พูด การบอกจำนวน เป็นต้น รูปแบบของการแสดงคุณลักษณะ มี 2 รูปแบบคือเป็นความสัมพันธ์ที่ชี้ออก หรือชี้ข้าม โนทัศน์ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- 1.) มโนทัศน์---คุณลักษณะ---->
- 2.) หมายถึง<----คุณลักษณะ----

ก.) คุณลักษณะที่แสดงบอกระยะทางไกลใกล้

that : บอกระยะไกล

ป.ย.49 นกตัวโน้นสวย

c#beautiful--a-object-->c#bird---->that

this : บอกระยะใกล้

ป.ย.50 แมวตัวนี้กินปลาไป

c#eat--agent-->c#cat---->this

ข.) คุณลักษณะที่บอกจำนวนที่กำหนดไม่ได้ ความสัมพันธ์ที่มี 3 ความสัมพันธ์ คือ

all : แทนทุกหน่วยที่เป็นของมโนทัศน์

ป.ย.51 แมวทุกตัวชอบกินปลา

c#like-agent-->c#cat---->all

generic : แสดงความทั่วไปของมโนทัศน์

ป.ย.52 เขาชอบแอปเปิ้ล

c#he<--agent--c#like--object-->c#apple---->generic

ค.) คุณลักษณะที่แสดงอาการปฏิเสธ มีไวยากรณ์ 1 แบบ คือ

not : ใบบอกการปฏิเสธมโนทัศน์ หรือ ปฏิเสธเหตุการณ์ หรือปฏิเสธคุณลักษณะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นิยมนำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ป.ย.53 เขาไม่ชอบหมอฟัน

c#he<--agent--c#like--object-->c#dentist-->not

ง.) คุณลักษณะใช้แสดงกาล คุณลักษณะที่กำหนดขึ้นแทนกาล แบ่งตามมุมมองของผู้พูดเป็น 3 คุณลักษณะด้วยกันคือ

past : แสดงเหตุการณ์ที่เกิดขึ้นในเวลาที่ผ่านไป

present : แสดงเหตุการณ์ที่เกิดขึ้นในเวลาปัจจุบัน

future : แสดงเหตุการณ์ที่เกิดขึ้นในเวลาที่ยังมาไม่ถึง

จ.) คุณลักษณะที่แสดงการณลักษณะ (aspect)

รูปแบบการณลักษณะของภาคแสดง

ตัวอย่างภาษาไทย

-การณลักษณะบอกเหตุการณ์ที่ยังไม่เริ่มต้น

ยังไม่ได้ กริยา

(c#pred--begin-->)--yet-->

--การณลักษณะบอกเหตุการณ์ก่อนการเริ่มต้น

กำลังจะ กริยา

(c#pred--begin-->)--soon-->

เกือบจะ กริยา

กำลังเริ่มจะ กริยา

กำลังจะเริ่ม กริยา

-การณลักษณะบอกเหตุการณ์การเริ่มต้น

เริ่ม กริยา

c#pred--begin-->

-การณลักษณะบอกเหตุการณ์หลังเริ่มต้น

เพิ่ง กริยา

(c#pred--begin-->)--just-->

เพิ่งเริ่ม กริยา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เพิ่งจะ กริยา

-การณัฏกษณะบอกเหตุการณ์ที่กำลังดำเนินอยู่

กำลัง กริยา

c#pred--progress-->

กำลัง กริยา อยู่

-การณัฏกษณะบอกเหตุการณ์ที่ต่อเนื่อง

ยัง กริยา อยู่

c#pred--continue-->

ทุกวัน

-การณัฏกษณะบอกเหตุการณ์ก่อนจะเสร็จสิ้น

กริยา เกือบ จะ เสร็จ

(c#pred--end--->)--soon-->

กริยา เกือบ เสร็จ แล้ว

-การณัฏกษณะบอกเหตุการณ์เสร็จสิ้น

กริยา เสร็จ แล้ว

(c#pred--end-->)--already-->

-การณัฏกษณะบอกเหตุการณ์สมบูรณ์

กริยา จบ

c#pred--complete-->

กริยา ครบ

กริยา สำเร็จ

-การณัฏกษณะบอกเหตุการณ์เพิ่งจะเสร็จสิ้น

เพิ่ง กริยา เสร็จ

(c#pred--end-->)--just-->

-การณัฏกษณะบอกเหตุการณ์ที่บอกสภาพ

กริยา อยู่

c#pred--state-->

กริยา ไว้

-การณัฏกษณะบอกเหตุการณ์ที่ยังไม่เสร็จ

ยัง กริยา ไม่ เสร็จ

(c#pred--end-->)--yet-->

กริยา ยัง ไม่ เสร็จ

-การณัฏกษณะบอกเหตุการณ์ยังไม่หยุดสิ้น

กริยา ไม่ จบ

(c#pred--complete-->)--yet-->

กริยา ไม่ ครบ

กริยา ไม่ สำเร็จ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ด.) คุณลักษณะที่แสดงความตั้งใจและเจตนาของผู้พูด คุณลักษณะดังกล่าวเป็นคุณลักษณะที่ขยายประโยคทั้งประโยค ดังนั้นรูปแทนของประโยคที่ประกอบไปด้วยคุณลักษณะแสดงเจตนาจะแสดงดังนี้

ตัวอย่าง

ป.ย.54 ผมต้องจอดรถไว้ข้างล่างต้อง แสดงคุณลักษณะ "duty"

state---->---place--->c#downstair

c#I<--agent--c#park--object-->c#car

st

c#statement---->duty

คุณลักษณะ	รูป	ประเภทของคำ	ตัวอย่าง
1.OBLIGATION			
-IMPERATIVE	จง	AUX	จงกาทบาทหน้าข้อที่ถูก
	ห้าม	AUX	ห้ามสูบบุหรี่
-REQUEST	กรุณา	AUX	กรุณาทำให้ฉันด้วย
	ช่วย	AUX	ช่วยวักให้ที
-INVITE	กันเถอะ	ADV	เราไปกินข้าวกันเถอะ
	มา (กริยา) กัน	AUX	มาร้องเพลงกัน
2.THREAT			
-NATURAL	นิยม	AUX	คนไทยนิยมกินหมาก
THING	มัก	AUX	คนแก่มักไปวัดในวันพระ
	ชอบ	AUX	คนญี่ปุ่นชอบกินปลาดิบ
3.GRANT			
-GRANT	ยอมให้	V	เขายอมให้เธอไป
	อนุญาตให้	V	ผู้จัดการอนุ
ญาติให้ลูกน้องลาได้			

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

-CONSENT	ยินดี	V	คุณยินดีที่จะไปไหม
-SUFFICIENCY	ก็พอ	ADV	ไปเยี่ยมก็พอ
	เพียง	ADV	เพียงโทรศัพท์มาก็พอ
-GRANT NOT	ไม่ยอมให้	V	คุณไม่ยอมให้เธอไป
	ไม่อนุญาตให้	V	คุณไม่อนุญาตให้เธอไป

#### 4.AGREEMENT

-NATURAL	จึง	AUX	ฝนตกน้ำจึงท่วม
RESULT	เพราะ	CONJ	น้ำท่วมเพราะฝนตก
	ก็เลย	AUX	ฝนตกน้ำก็เลยท่วม

#### 5.REQUIRE-AGREEMENT

-REQUIRE	จริงไหม	EITT	ที่นี่หนาวจริงไหม
AGREEMENT	ใช่ไหม	EITT	ที่นี่หนาวใช่ไหม

#### 6.JUDGEMENT

-IF	หาก	CONJ	หากเขาไม่มา ฉันต้องตาย
-THOUGHT	ถ้า	CONJ	ถ้าฉันเป็นนก ฉันจะบินไปให้ไกล

#### 7.PROBABILITY

-MAYBE	บางที	ADV	บางทีเขาคงมา
	อาจ	AUX	เขาอาจไม่รู้
-SEEM	คาดว่า	V	คาดว่าพรุ่งนี้ฝนจะตก
-SURE	แน่	ADV	เขามาแน่
-RUMOR	ลือกันว่า	V	ลือกันว่าเธอสวยมาก
	เล่ากันว่า	V	เล่ากันว่าพี่จีสวยมาก
-APPEARANCE	ทำท่า	V	ฝนทำท่าจะตก
-CONCLUDE	จึง	V	นักเรียนไม่ศิครจึงตี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 9.FEELING

-BLAME	ไม่น่า	AUX	เขาไม่น่าทำเล่น
	ไม่ควร	AUX	เขาไม่ควรทำแบบนี้
-EXCLAMATION	อะไรอย่างนี้	AUX	สวยอะไรอย่างนี้
-PITY	น่าสงสารที่	AUX	น่าสงสารที่เขาไม่ได้ไป
-REGENT	ถึงแม้ว่า	CONJ	ถึงแม้เขาจะรวย เขาไม่มีความสุข
-UNDER	แค่	ADV	แค่เดินไปเท่านั้น

## ESTIMATE

-UNEXPECTED	ไม่นึก ไม่ฝัน	V	ไม่นึกไม่ฝันว่าจะเดินได้
-------------	---------------	---	--------------------------

## 10.FEELINGS FOR BENEFIT

-GIVE BENEFIT	ให้	AUX	เขาทำการบ้านให้
	ช่วย	AUX	เขาช่วยเธอทำงาน
-GET BENEFIT	ถูก	AUX	หลอนถูกใจให้ไปซื้อของ

## 11.INTERROGATION

-INTERROGATION	ทำไม	PRON	ทำไมเขาไม่มา
----------------	------	------	--------------

## 12.RESPECT

-RESPECT	ครับ ค่ะ	PARTICLE	จะกลับแล้วหรือครับ
-POLITE	โปรด	AUX	โปรดออดรองเท้า

## 3.8.การแทนรูปประโยคด้วยภาษากลาง

โครงสร้างภาษากลางแสดงโครงสร้างทางความหมายของประโยคได้โดยใช้นิพจน์ c#statement เชื่อมกับกริยาเพื่อบอกความเป็นประโยค ทั้งยังเป็นที่ยึดข้อมูลที่ขยายความทั้งประโยคได้ นอกจากนี้ยังมีความสัมพันธ์ที่แสดงการเชื่อมต่อระหว่างประโยค เช่น การเป็น sub statement หรือ co-statement ของ ประโยค 2 ประโยค เช่น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### การแสดง statement

ป.ย.55 บ้านนี้สวย

```
<--this--c#house<--a-object--c#beautiful
      ↑
      st
      |
      c#statement
```

### .การแสดงการเชื่อมโยงด้วย sub-statement

ป.ย. 56 เขาไปทำงานที่บริษัท

```
main---->place-->company
he<--agent--c#move--purpose-->work
go<--
      ↑
      st
      |
      c#statement
```

```
---sub--st---->
```

### .การแสดงการเชื่อมโยงด้วย co-statement

ป.ย. 57 น้องสาวสะสมรูปภาพดาราดำแต่ที่สาวสะสมแสตมป์

```
<--main --modifier
c#sister<--agent--c#collect--object-->c#photograph
      |
      --st-->c#statement<
      |
      or
      |
      --st-->c#statement
      |
      c#elder sister<--agent--c#collect--object-->c#stamp
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



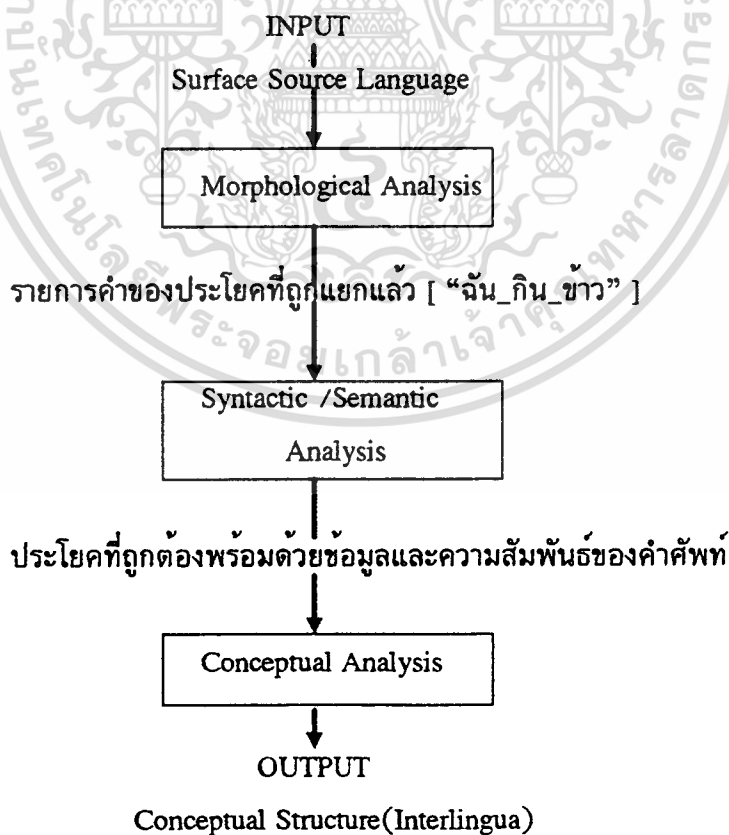
## บทที่ 4

### โครงสร้างและส่วนประกอบของระบบวิเคราะห์ภาษา

#### 4.1 บทนำ

ดังที่ได้กล่าวมาแล้วว่า ในโครงการนี้จะกล่าวถึงเฉพาะส่วนระบบวิเคราะห์ภาษาเท่านั้น คือเป็นส่วนที่เริ่มทำงานจาก การแยกคำศัพท์ของประโยคอินพุท โดยวิธีการแยกจะใช้การเปรียบเทียบเพื่อค้นหาคำศัพท์ในพจนานุกรม จากนั้นก็นำคำศัพท์เหล่านั้นไปตรวจสอบหาความสัมพันธ์ตามหลักไวยากรณ์ของภาษาไทยที่ถูกต้อง แล้วจึงจะสามารถตรวจสอบหาความหมายได้โดยอาศัยข้อมูลของคำศัพท์เหล่านั้น เช่น ชนิด, รูปแบบของกริยาที่มีอยู่ในประโยค , Akind of เป็นต้น ขั้นตอนสุดท้ายคือ นำคำศัพท์ และข้อมูลของคำศัพท์เหล่านั้น ผ่านกระบวนการสร้างเป็นภาษากลางต่อไป

ส่วนประกอบของระบบวิเคราะห์ภาษาแสดงดังในรูป 4.1



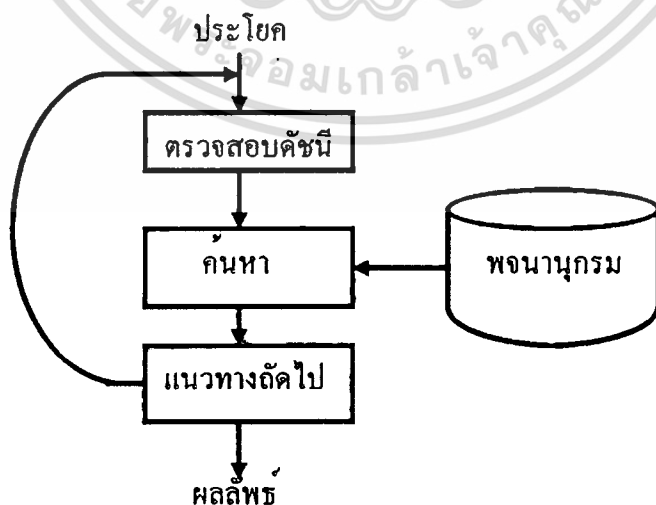
รูปที่ 4.1 แสดงส่วนประกอบพื้นฐานของระบบวิเคราะห์ภาษา

เอกสารนี้เป็นเอกสารที่ลิขสิทธิ์สงวนสิทธิ์ในชื่อผู้จัดทำเท่านั้น ไม่อนุญาตให้เผยแพร่ไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 4.2 การแยกคำศัพท์จากประโยคภาษาไทย

การแยกคำศัพท์ภาษาไทยนั้น ได้มีการคิดค้นกันมานานแล้วซึ่งทำให้เกิดแนวความคิดที่แตกต่างหลากหลายวิธีการด้วยกัน เช่น การแยกคำศัพท์โดยการประมวลผลจากการผสมพยัญชนะกับสระ คือการนำตัวอักษรและสระมาผสมกัน เพื่อให้เกิดเป็นคำศัพท์ที่มีความหมายในภาษาไทย อีกวิธีการหนึ่งคือ การค้นหาและเลือกคำศัพท์ที่ยาวที่สุด(จำนวนอักขระมากที่สุด) หรือเรียกว่า Longest Mapping วิธีนี้อาศัยข้อมูลคำศัพท์ในพจนานุกรม ในการเปรียบเทียบเพื่อหาคำศัพท์ที่ยาวที่สุด แต่ทั้งสองวิธีเป็นวิธีการที่ไม่มีประสิทธิภาพ เนื่องจากความล่าช้า และความถูกต้องน้อย เช่นในวิธีการ Longest Mapping หาคำศัพท์ที่ยาวที่สุด ซึ่งในประโยคอาจจะมีคำศัพท์ที่มีความหมายถูกต้องกว่า แต่มีความยาวน้อยกว่า

ดังนั้นในโครงการนี้จึงขอเสนอวิธีการแยกคำศัพท์ในประโยค โดยการใช้ระบบดัชนี (Index) ในการเข้าถึงกลุ่มคำศัพท์ที่ต้องการได้อย่างรวดเร็ว และยังค้นหาเพื่อให้ได้ผลการแยกคำศัพท์ทุกแบบที่เป็นไปได้ในประโยค เรียกว่า Fast Word Matching ในระบบนี้จะต้องมีฐานข้อมูลคือ พจนานุกรมคำศัพท์ภาษาไทย ซึ่งคำศัพท์ทุกคำเหล่านั้นจะถูกจัดแบ่งออกเป็นหมวดหมู่ โดยแบ่งกลุ่มตามระบบดัชนีซึ่งจะกล่าวถึงต่อไป การทำงานของการแยกคำศัพท์ภาษาไทยแบบ Fast Word Matching นั่นคือ การค้นหาคำศัพท์ทุกคำที่เป็นไปได้ในพจนานุกรม แล้วรวบรวมทางหรือรูปแบบการแยกที่เป็นไปได้ทุกทางเป็นผลลัพธ์ ดังรูปที่ 4.2



รูปที่ 4.2 แสดงการทำงานพื้นฐานของการแยกคำศัพท์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

#### 4.2.1 การกำหนดดัชนีและจัดกลุ่มคำศัพท์

คำศัพท์ภาษาไทยที่มีอยู่ในพจนานุกรมทั้งหมด จะถูกจัดแบ่งออกเป็นหมวดหมู่ ตามค่าของดัชนี ในระบบนี้เราจะใช้ดัชนี 3 ตัวด้วยกันคือ

- ก.) ดัชนีที่ 1 (First Index) ใช้ค่าแอสกี (ASCII) ของอักขระตัวที่ 1
- ข.) ดัชนีที่ 2 (Second Index) เป็นกลุ่มดัชนีย่อยของดัชนีที่ 1 ใช้ค่าแอสกีของอักขระที่ 2
- ค.) ดัชนีที่ 3 (Third Index) เป็นกลุ่มดัชนีย่อยของดัชนีที่ 2 ใช้ค่าความยาว (จำนวนอักขระ)ของคำศัพท์

ดังนั้นกลุ่มคำศัพท์ตามดัชนีที่ 1 จะถูกแบ่งเป็นกลุ่มย่อยๆลงไปอีกตามจำนวนของดัชนีที่ 2 ของแต่ละกลุ่ม และคำศัพท์แต่ละกลุ่มที่ถูกแบ่งตามดัชนีที่ 2 ก็จะถูกแบ่งย่อยลงไปอีกตามดัชนีที่ 3 เช่น

แบ่งกลุ่มตามดัชนีที่ 1 ได้เช่น

กลุ่มคำศัพท์ที่มีอักขระตัวแรกเป็น 'ก'

กลุ่มคำศัพท์ที่มีอักขระตัวแรกเป็น 'ข'

กลุ่มคำศัพท์ที่มีอักขระตัวแรกเป็น 'ค'

กลุ่มคำศัพท์ที่มีอักขระตัวแรกเป็น 'ง'

ฯลฯ

แบ่งกลุ่มตามดัชนีที่ 2 ได้เช่น

กลุ่มคำศัพท์ที่มีอักขระตัวแรกเป็น 'ก' คำศัพท์กลุ่มนี้จะถูกแบ่งออกเป็นกลุ่มย่อยตามอักขระตัวที่ 2 ดังนี้

กลุ่มคำศัพท์ที่มีอักขระตัวแรกเป็น 'ก' , อักขระที่สองเป็น 'า'

กลุ่มคำศัพท์ที่มีอักขระตัวแรกเป็น 'ก' , อักขระที่สองเป็น 'ค'

กลุ่มคำศัพท์ที่มีอักขระตัวแรกเป็น 'ก' , อักขระที่สองเป็น 'น'

กลุ่มคำศัพท์ที่มีอักขระตัวแรกเป็น 'ก' , อักขระที่สองเป็น 'ร'

ฯลฯ

แบ่งกลุ่มตามดัชนีที่ 3 ได้เช่น

กลุ่มคำศัพท์ที่มีอักษรตัวแรกเป็น 'ก' คำศัพท์กลุ่มนี้จะถูกแบ่งออกเป็นกลุ่มย่อยตามอักษรตัวที่ 2 ได้ตามจำนวนอักษรที่ 2 ที่มีได้ของแต่ละกลุ่มใหญ่ และแต่ละกลุ่มย่อยที่ถูกแบ่งตามดัชนีที่ 2 ยังถูกแบ่งเป็นกลุ่มย่อยๆ อีกตามค่าความยาวของคำศัพท์ ดังนี้

กลุ่มคำศัพท์ที่มีอักษรตัวแรกเป็น 'ก' , อักษรที่สองเป็น 'า' , ความยาวเท่า 2

กลุ่มคำศัพท์ที่มีอักษรตัวแรกเป็น 'ก' , อักษรที่สองเป็น 'า' , ความยาวเท่า 3

กลุ่มคำศัพท์ที่มีอักษรตัวแรกเป็น 'ก' , อักษรที่สองเป็น 'า' , ความยาวเท่า 4

กลุ่มคำศัพท์ที่มีอักษรตัวแรกเป็น 'ก' , อักษรที่สองเป็น 'า' , ความยาวเท่า 5

กลุ่มคำศัพท์ที่มีอักษรตัวแรกเป็น 'ก' , อักษรที่สองเป็น 'า' , ความยาวเท่า 6

ฯลฯ

ให้  $W$  แทน จำนวนคำศัพท์ในพจนานุกรมทั้งหมด  
 $I_1, I_2, I_3$  แทน ดัชนีที่ 1, ดัชนีที่ 2, ดัชนีที่ 3  
 $j, k, n$  แทน ลำดับของกลุ่มคำศัพท์ตามค่าดัชนีต่างๆ

$I_{j,1}$  คือดัชนีที่ 1 เมื่อ  $j$  คือลำดับกลุ่มของดัชนีที่ 1 ( $j = 1, 2, 3, \dots$ ) เช่น  
 $I_{1,1}$  คือดัชนีที่ 1 ของกลุ่มคำศัพท์ที่ 1  
 $I_{1,2}$  คือดัชนีที่ 1 ของกลุ่มคำศัพท์ที่ 2  
 $I_{1,3}$  คือดัชนีที่ 1 ของกลุ่มคำศัพท์ที่ 3  
 $I_{1,4}$  คือดัชนีที่ 1 ของกลุ่มคำศัพท์ที่ 4 เป็นต้น

$I_{j,2}(k)$  คือดัชนีที่ 2 เมื่อ  $k$  คือลำดับกลุ่มของดัชนีที่ 2 ( $k = 1, 2, 3, \dots$ ) เช่น  
 $I_{j,2}(1)$  คือดัชนีที่ 2 ของกลุ่มคำศัพท์ที่ 1 ซึ่งเป็นกลุ่มย่อยของคำศัพท์ตามดัชนีที่ 1 กลุ่มที่  $j$   
 $I_{j,2}(2)$  คือดัชนีที่ 2 ของกลุ่มคำศัพท์ที่ 2 ซึ่งเป็นกลุ่มย่อยของคำศัพท์ตามดัชนีที่ 1 กลุ่มที่  $j$   
 $I_{j,2}(3)$  คือดัชนีที่ 2 ของกลุ่มคำศัพท์ที่ 3 ซึ่งเป็นกลุ่มย่อยของคำศัพท์ตามดัชนีที่ 1 กลุ่มที่  $j$  เป็นต้น

$I_{j,k,3}(n)$  คือดัชนีที่ 3 เมื่อ  $n$  คือลำดับกลุ่มของดัชนีที่ 3 ( $n=1,2,3,\dots$ ) เช่น

$I_{j,k,3}(1)$  คือดัชนีที่ 3 ของกลุ่มคำศัพท์ที่ 1 ซึ่งเป็นกลุ่มย่อยของคำศัพท์ตามดัชนีที่ 2 กลุ่มที่  $k$  (ซึ่งเป็นกลุ่มย่อยของคำศัพท์ตามดัชนีที่ 1 กลุ่มที่  $j$ )

$I_{j,k,3}(1)$  คือดัชนีที่ 3 ของกลุ่มคำศัพท์ที่ 2 ซึ่งเป็นกลุ่มย่อยของคำศัพท์ตามดัชนีที่ 2 กลุ่มที่  $k$  (ซึ่งเป็นกลุ่มย่อยของคำศัพท์ตามดัชนีที่ 1 กลุ่มที่  $j$ )

$I_{j,k,3}(1)$  คือดัชนีที่ 3 ของกลุ่มคำศัพท์ที่ 3 ซึ่งเป็นกลุ่มย่อยของคำศัพท์ตามดัชนีที่ 2 กลุ่มที่  $k$  (ซึ่งเป็นกลุ่มย่อยของคำศัพท์ตามดัชนีที่ 1 กลุ่มที่  $j$ ) เป็นต้น

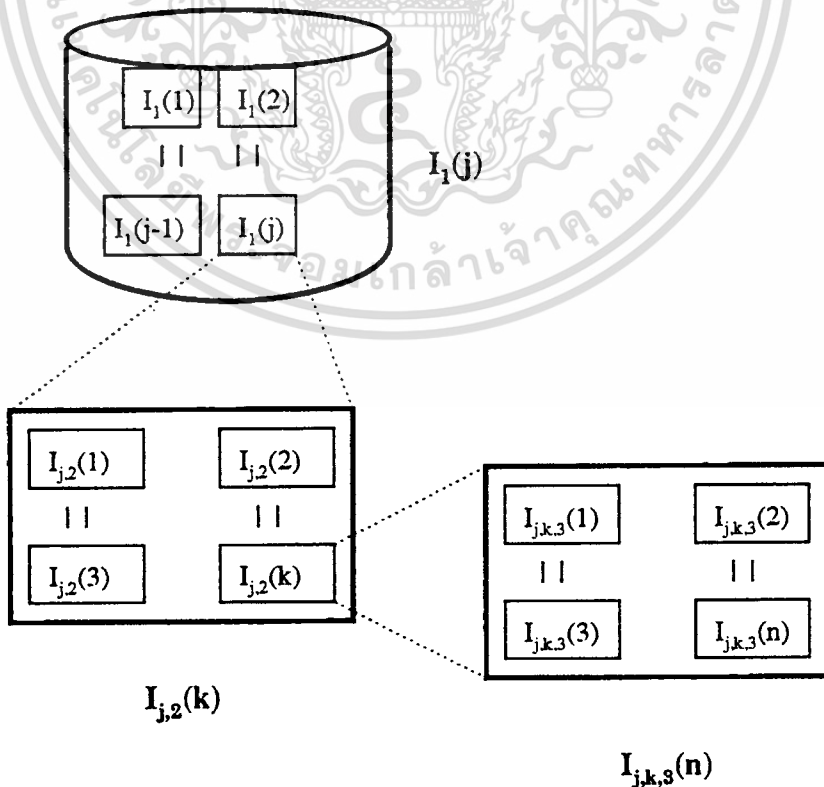
และ  $I_1(j)$  = ค่าแอสกีของอักขระที่ 1

$I_{j,2}(k)$  = ค่าแอสกีของอักขระที่ 2

$I_{j,k,3}(n)$  = ค่าความยาวของคำศัพท์

ดังนั้นจะได้ว่า

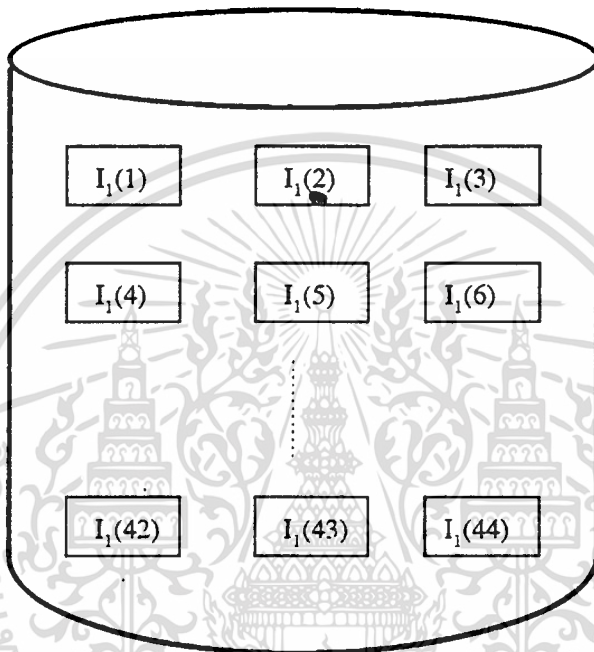
$$W = \{ I_1(j) \{ I_{j,2}(k) \{ I_{j,k,3}(n) \} \} \}$$



รูปที่ 4.3 แสดงความสัมพันธ์ระหว่าง  $I_1(j)$ ,  $I_{j,2}(k)$  และ  $I_{j,k,3}(n)$

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากการศึกษาคำศัพท์ภาษาไทย พบว่าอักขระที่จะเป็นอักขระตัวแรกของคำศัพท์นั้นมีอยู่ 44 ตัว เป็นพยัญชนะ 39 ตัว และสระอีก 5 ตัวดังตารางที่ 4.1 ดังนั้นดัชนีที่ 1 จะมี 44 คำ หรือ  $I_1(j)$  มีค่า 44 คำ ( $j = 1, 2, 3, 4, \dots, 43, 44$ ) จึงทำให้ดัชนีของกลุ่มคำศัพท์กลุ่มใหญ่นี้เป็น  $I_1(1), I_1(2), \dots, I_1(44)$  ดังรูปที่ 4.4



รูปที่ 4.4 แสดงดัชนีที่ 1 ซึ่งมี 44 คำ

j	อักขระ	ค่าแอสกี
1	ก	161
2	ข	162
3	ค	164
4	ฅ	166
5	ง	167
6	จ	168
7	ฉ	169
8	ช	170
9	ซ	171

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้ในการเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้เผยแพร่โดยไม่ได้รับอนุญาตจากฝ่ายวิชาการ

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

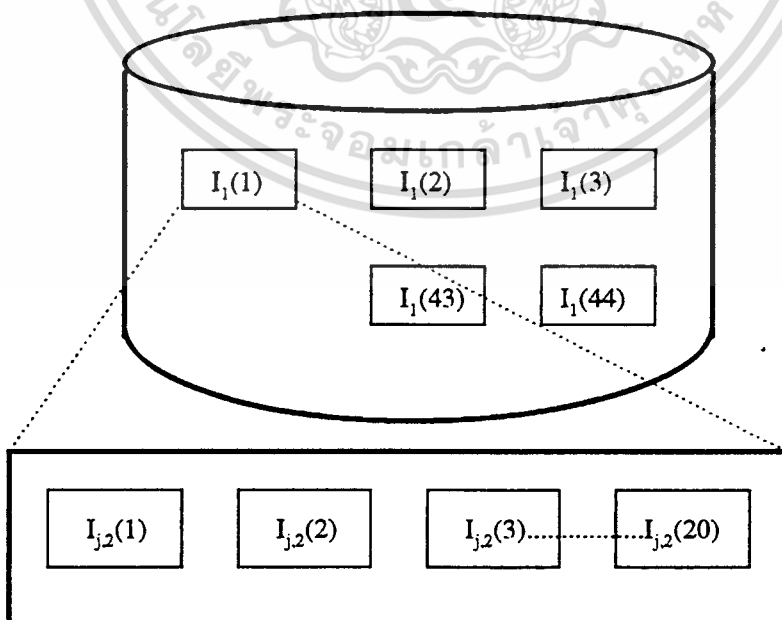
j	อักษร	ค่าแอสกี
10	ณ	172
11	ญ	173
12	ฎ	174
13	ฏ	175
14	ฐ	176
15	ณ	179
16	ด	180
17	ต	181
18	ถ	182
19	ท	183
20	ธ	184
21	น	185
22	บ	186
23	ป	187
24	ผ	188
25	ฝ	189
26	พ	190
27	ฟ	191
28	ภ	192
29	ม	193
30	ย	194
31	ร	195
32	ฤ	196
33	ล	197
34	ว	199
35	ศ	200
36	ส	202
37	ห	203

เอกสารนี้เป็นเอกสารที่สงวนเวลาหรับการเชิงงานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

j	อักขระ	ค่าแอสกี
38	อ	205
39	ฮ	206
40	เ	224
41	แ	225
42	โ	226
43	ใ	227
44	ไ	228

ตารางที่ 4.1 แสดงอักขระที่สามารถเป็นอักขระแรกของคำศัพท์

ในแต่ละกลุ่มคำศัพท์ตามดัชนีที่ 1 สามารถแบ่งแยกออกเป็นหลายกลุ่มย่อยๆ ตามค่าแอสกีของอักขระตัวที่ 2 ของคำศัพท์ สำหรับอักขระที่สามารถเป็นอักขระที่ 2 ของคำศัพท์ในแต่ละกลุ่มของดัชนีที่ 1 นั้นจะมีไม่เท่ากัน เช่น กลุ่มคำศัพท์ 'ก' จะมีอักขระที่สามารถตามหลัง 'ก' ได้ 20 ตัว แสดงในตารางที่ 4.2 ดังนั้น ถ้า  $I_1(1) = \text{ASCII}('ก')$   $I_{j,2}(k)$  จะมีค่า 20 ค่าคือ  $I_{j,2}(1), I_{j,2}(2), \dots, I_{j,2}(20)$  ดังแสดงในรูปที่ 4.5



รูปที่ 4.5 แสดงการแบ่งกลุ่มย่อยตามอักขระที่ 2 ของกลุ่มอักขระที่ 1 เป็น 'ก'

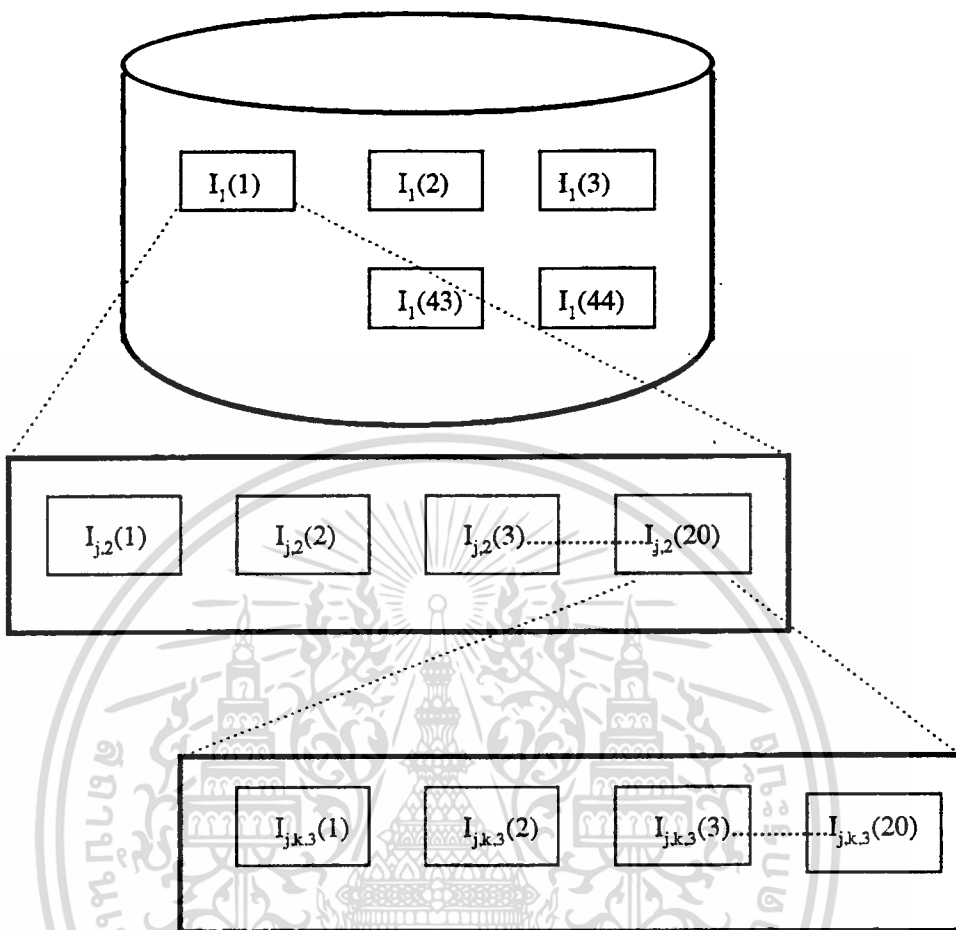
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับกํารใช้งานเพื่อการศึกษาเท่านั้น ไม่นอญูญได้หน้าไปไซ่ประโยชน์ดํานการคํ่า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

k	อักขระที่ 1 (ก) และที่ 2	แอสกีของ อักขระตัวที่ 2	ตัวอย่าง
1	ก	226	กก
2	ก	225	กน
3	ก	224	กง
4	กี	223	กี
5	กั	221	กั , กัด
6	กิ	218	กีฬา
7	กิ	217	กิน
8	กู	216	กู
9	กู	215	กูด
10	ก้า	207	ก้าดั่ง
11	กา	206	กาง
12	กะ	204	กะทะ
13	กอ	202	กอง
14	กค	195	กคม
15	กร	193	กรรไกร
16	กบ	185	กบ
17	กค	178	กค
18	กฏ	173	กฏ
19	กง	167	กง
20	กก	161	กก

ตารางที่ 4.2 แสดงตัวอย่างการแบ่งกลุ่มคำศัพท์ตามดัชนีที่ 2 ( $I_{1,2}(k)$ )

และเช่นเดียวกันกลุ่มคำศัพท์ตามดัชนีที่ 2 ก็สามารถแยกออกเป็นหลายๆกลุ่มย่อย ตามดัชนีที่ 3 หรือความยาวของคำศัพท์ ดังที่ได้แสดงในรูปที่ 4.6

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.6 แสดงการกำหนดดัชนีที่ 3

คำศัพท์ในภาษาไทยนั้นจะมีความยาวตั้งแต่ 2 ขึ้นไป ดังนั้นค่าของดัชนีที่ 3 จะมีได้ตั้งแต่ 2 ขึ้นไปเช่นกัน และสำหรับในโครงการนี้ค่าของดัชนีที่ 3 จะมีได้ตั้งแต่ 2 ถึง 50 นั่นคือคำศัพท์ที่ยาวที่สุดในที่นี้คือมีความยาว 50 อักขระ ดังนั้น  $I_{j,k,3}(n)$  มีค่า 49 ค่า ( $n = 1, 2, 3, \dots, 49$ ) และ

กำหนดให้  $I_{j,k,3}(1) = 50$

$$\cdot I_{j,k,3}(2) = 49$$

$$I_{j,k,3}(3) = 48$$

“

$$I_{j,k,3}(48) = 3$$

$$I_{j,k,3}(49) = 2$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

#### 4.2.2 กระบวนการแยกคำศัพท์แบบ Fast Word Matching

สิ่งจำเป็นสำหรับกระบวนการแยกคำศัพท์แบบ Fast Word Matching คือ ข้อมูลคำศัพท์ภาษาไทยในพจนานุกรม ที่ถูกจัดแบ่งออกเป็นหมวดหมู่ตามระบบดัชนีดังที่ได้กล่าวมาแล้ว หลักการทำงานของกระบวนการแยกคำศัพท์นั้น มีขั้นตอนพื้นฐาน 3 ขั้นตอนดังนี้

- 1.) กำหนดหาดัชนีที่ 1 ดัชนีที่ 2 และดัชนีที่ 3
- 2.) หากจุดเริ่มต้นสำหรับการค้นหาคำศัพท์ โดยใช้ดัชนีที่ 1 ดัชนีที่ 2 และดัชนีที่ 3
- 3.) เปรียบเทียบเพื่อค้นหาคำศัพท์ในพจนานุกรม

ให้	S	คือประโยคอินพุทภาษาไทย
	W	คือคำศัพท์ในพจนานุกรม
	C <sub>n</sub>	คืออักขระในประโยคตัวที่ n (n = 1,2,3,4, .....)
	D <sub>n</sub>	คืออักขระในคำศัพท์ในพจนานุกรมตัวที่ n (n = 1,2,3,4, .....)
	L	คือความยาวของประโยค
	I <sub>1</sub>	คือตัวแปรที่เก็บค่าดัชนีที่ 1
	I <sub>2</sub>	คือตัวแปรที่เก็บค่าดัชนีที่ 2
	I <sub>3</sub>	คือตัวแปรที่เก็บค่าดัชนีที่ 3
	Mark <sub>n</sub>	คือรูปแบบของการแยกคำศัพท์ที่ n (n = 1,2,3,4, .....)

สำหรับการทำงานของกระบวนการแยกคำศัพท์ในรูปแบบของโฟลว์ชาร์ต (Flow Chart) นั้นได้แสดงไว้ในรูปที่ 4.7

#### ขั้นตอนการแยกคำศัพท์แบบ Fast Word Matching

1. กำหนดหาดัชนีที่ 1 และดัชนีที่ 2 ซึ่งก็คือรหัสแอสกีของอักขระตัวที่ 1 และ 2 ตามลำดับ และหาความยาว (จำนวนตัวอักษร) ของประโยคเพื่อกำหนดให้เป็นดัชนีที่ 3

$$I_1 = \text{ASCII}(C_1), I_2 = \text{ASCII}(C_2), I_3 = L = \text{Long}(S)$$

2. นำประโยคอินพุทตามข้อ 1. ไปเปรียบเทียบกับคำศัพท์ในพจนานุกรม

##### 2.1. ค้นหาคำศัพท์ในพจนานุกรมตามดัชนีที่ 1 และดัชนีที่ 2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2. ใช้ดัชนีที่ 3 (ความยาวของประโยค) ในการเริ่มต้นเปรียบเทียบ

2.2.1. ถ้าความยาวของประโยคมากกว่าหรือเท่ากับความยาวที่มากที่สุดของคำศัพท์ในกลุ่มคำศัพท์ตามดัชนีที่ 1 และ 2 กำหนดไว้ ให้เริ่มทำการเปรียบเทียบที่กลุ่มคำศัพท์นี้

2.2.2. ถ้าความยาวของประโยคน้อยกว่า ความยาวของคำศัพท์ให้เลื่อนเพื่อเริ่มการเปรียบเทียบลงไปอีกตามความยาวของคำศัพท์(ความยาวของคำศัพท์ลดลง)จนพบว่า ความยาวของประโยคมากกว่าหรือเท่ากับ ความยาวของคำศัพท์

3. เริ่มทำการเปรียบเทียบโดยการเปรียบเทียบจะกระทำแบบอักขระต่ออักขระ คือเปรียบเทียบอักขระที่อยู่ตำแหน่งเดียวกัน เช่น



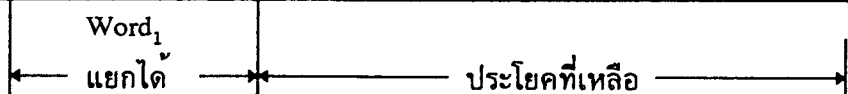
3.1. ผลการเปรียบเทียบ ถ้าตำแหน่งใดที่เหมือนกัน ให้ทำเครื่องหมายไว้ เช่น '0'

3.2. ถ้าในทุกตำแหน่งของคำศัพท์ในพจนานุกรมสามารถทำเครื่องหมาย '0' ได้ครบ ถือว่าแยกคำนั้นได้ ให้เก็บบันทึกคำศัพท์คำนั้นไว้ ( $Mark_1, Word_1$ ) และค้นหาต่อไปอีกในคำศัพท์กลุ่มเดียวกันจนหมด ถ้าพบให้เก็บบันทึกคำนั้นไว้ ( $Mark_n, Word_n$ )

3.3. ถ้าไม่สามารถทำเครื่องหมาย '0' ได้ครบทุกตำแหน่ง ให้ทำการเปรียบเทียบกับคำศัพท์คำอื่นๆ ในกลุ่มคำศัพท์เดียวกัน จนกระทั่งถึงคำศัพท์ที่มีความยาวน้อยที่(2)

3.4. ถ้าไม่พบคำศัพท์เลย คือไม่สามารถทำเครื่องหมาย '0' ได้ครบทุกตำแหน่งของทุกคำศัพท์ในกลุ่มเดียวกัน ถือว่าไม่สามารถแยกหน่วยคำของประโยคนั้นได้(ต้องเพิ่มคำศัพท์ในพจนานุกรม)

S	เ	ช	า	ไ	ป	ท	ะ	เ	ล
W	เ	ช	า						
ผล	0	0	0	1	1	1	1	1	1
Mark <sub>1</sub>			1						
	เ	ช	า						



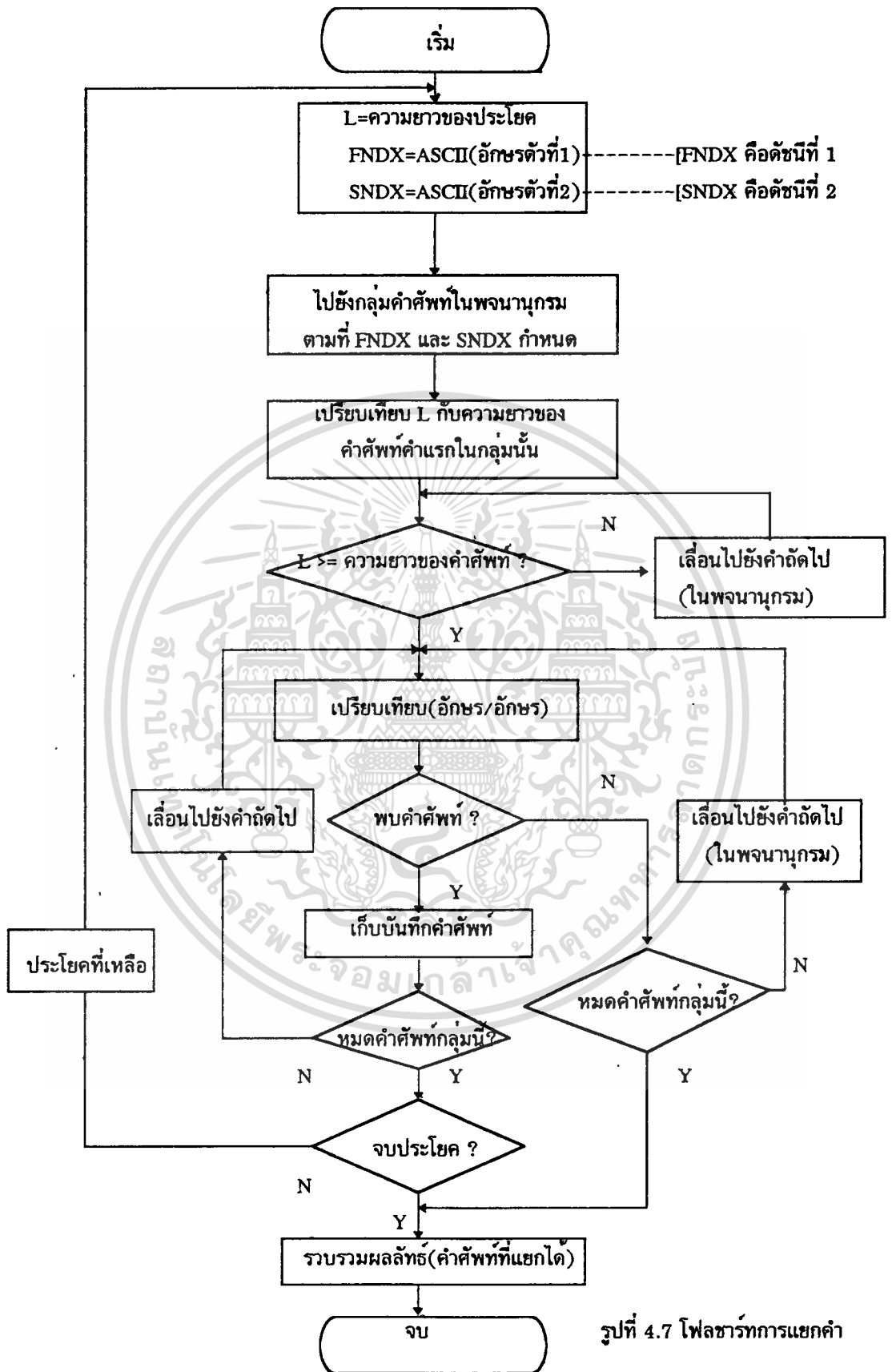
4. ให้นำประโยคที่เหลือ(ถ้ามี) ไปผ่านกระบวนการตามข้อ 1, 2 และ 3 อีก จนกระทั่งจบประโยค และเก็บบันทึกคำศัพท์ที่สามารถแยกได้ทั้งหมด ผลลัพธ์เช่น

Mark<sub>1</sub> { Word<sub>1</sub> / Word<sub>2</sub> / ..... Word<sub>n</sub> }

Mark<sub>n</sub> { Word<sub>1</sub> / Word<sub>2</sub> / ..... Word<sub>n</sub> }

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.7 โฟลชาร์ทการแยกคำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 4.2.3 ตัวอย่างการแยกคำศัพท์แบบ Fast Word Matching

ต่อไปนี้จะขอยกตัวอย่างการทำงานของกระบวนการแยกคำศัพท์ในประโยคภาษาไทยด้วยวิธีการ Fast Word Matching

ประโยคตัวอย่างที่ 1 ประโยค---->“เธอ ชน เขา”

ลำดับที่ 1. คำนวณหาดัชนีที่ 1 และดัชนีที่ 2 ซึ่งก็คือรหัสแอสกีของอักขระตัวที่ 1 และ 2 ตามลำดับ และหาความยาว(จำนวนตัวอักษร)ของประโยคเพื่อกำหนดให้เป็นดัชนีที่ 3

$$I_1 = \text{ASCII}(C_1), I_2 = \text{ASCII}(C_2), I_3 = L = \text{Long}(S)$$

ตำแหน่ง	1	2	3	4	5	6	7	8
S	เ	อ	ช	น	เ	ช	า	

$$I_1 = \text{ASCII}(เ) = 224$$

$$I_2 = \text{ASCII}(อ) = 184$$

$$I_3 = L = \text{ความยาวประโยค} = 8$$

ลำดับที่ 2. นำประโยคอินพุตตามข้อ 1. ไปเปรียบเทียบกับคำศัพท์ในพจนานุกรม

2.1. ค้นหาคำศัพท์ในพจนานุกรมตาม  $I_1, I_2$

2.2. ใช้  $I_3$  (ความยาวของประโยค) ในการเริ่มต้นเปรียบเทียบ

2.2.1. ถ้าความยาวของประโยคมากกว่าหรือเท่ากับความยาวที่มากที่สุดของคำศัพท์ในกลุ่มคำศัพท์ตามดัชนีที่ 1 และ 2 กำหนดไว้ ให้เริ่มทำการเปรียบเทียบที่กลุ่มคำศัพท์นี้

2.2.2. ถ้าความยาวของประโยคน้อยกว่า ความยาวของคำศัพท์ให้เลื่อนเพื่อเริ่มการเปรียบเทียบลงไปอีกตามความยาวของคำศัพท์(ความยาวของคำศัพท์ลดลง)จนพบว่า ความยาวของประโยคมากกว่าหรือเท่ากับ ความยาวของคำศัพท์  $L \geq \text{ความยาวของคำศัพท์}$

ลำดับที่ 3. เริ่มทำการเปรียบเทียบโดยการเปรียบเทียบจะกระทำแบบอักขระต่อ

อักขระ คือเปรียบเทียบอักษรที่อยู่ตำแหน่งเดียวกัน เช่น

คำศัพท์ในพจนานุกรม  $D_1 \quad D_2 \quad D_3 \dots D_n$

คำศัพท์ในประโยค  $C_1 \quad C_2 \quad C_3 \dots C_n$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

S	เ	ธ	อ	ช	น	เ	ช	า
W	เ	ธ	อ					
ผลการเปรียบเทียบ	0	0	0	1	1	1	1	1
Mark <sub>1</sub>			1					
	เ	ธ	อ					

พบคำศัพท์คำว่า “เธอ” ดังนั้นผลลัพธ์คือ

Word<sub>1</sub>

Mark<sub>1</sub> { เธอ/ }

จากนั้นเปรียบเทียบกับคำศัพท์ถัดไปในกลุ่มเดียวกัน (ความลดย่อยลงเรื่อยๆ) จนถึงคำศัพท์ที่มีความยาวน้อยที่สุดคือ 2 ปรากฏว่าไม่พบคำศัพท์อีก

ลำดับที่ 4. ให้นำประโยคที่เหลือ(ถ้ามี) ไปผ่านกระบวนการตามข้อ 1,2 และ 3 อีกจนกระทั่งจบประโยค และเก็บบันทึกคำศัพท์ที่สามารถแยกได้ทั้งหมด ผลลัพธ์ขณะนี้คือ

Word<sub>1</sub>

S

Mark<sub>1</sub> { เธอ/ }

ชนเขา }

ลำดับที่ 1. คำนวณหาดัชนีที่ 1 และดัชนีที่ 2 ซึ่งก็คือรหัสแอสกีของอักขระตัวที่ 1 และ 2 ตามลำดับ และหาความยาว(จำนวนตัวอักษร)ของประโยคเพื่อกำหนดให้เป็นดัชนีที่ 3

$$I_1 = \text{ASCII}(C_1), I_2 = \text{ASCII}(C_2), I_3 = L = \text{Long}(S)$$

ตำแหน่ง	1	2	3	4	5
S	ช	น	เ	ช	า

$$I_1 = \text{ASCII}(ช) = 170$$

$$I_2 = \text{ASCII}(น) = 185$$

$$I_3 = L = \text{ความยาวประโยค} = 5$$

ลำดับที่ 2. นำประโยคอินพุตตามข้อ 1. ไปเปรียบเทียบกับคำศัพท์ในพจนานุกรม

2.1. ค้นหาคำศัพท์ในพจนานุกรมตาม  $I_1 ; I_2$

2.2. ใช้  $I_3$  (ความยาวของประโยค) ในการเริ่มต้นเปรียบเทียบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- 2.2.1. ถ้าความยาวของประโยคมากกว่าหรือเท่ากับความยาวที่มากที่สุดของคำศัพท์ในกลุ่มคำศัพท์ตามลำดับที่ 1 และ 2 กำหนดไว้ ให้เริ่มทำการเปรียบเทียบที่กลุ่มคำศัพท์นี้
- 2.2.2. ถ้าความยาวของประโยคน้อยกว่า ความยาวของคำศัพท์ให้เลื่อนเพื่อเริ่มการเปรียบเทียบลงไปอีกตามความยาวของคำศัพท์(ความยาวของคำศัพท์ลดลง)จนพบว่า ความยาวของประโยคมากกว่าหรือเท่ากับ ความยาวของคำศัพท์  $L \geq \text{ความยาวของคำศัพท์}$

ลำดับที่ 3. เริ่มทำการเปรียบเทียบโดยการเปรียบเทียบจะกระทำแบบอักขระต่ออักขระ คือเปรียบเทียบอักขระที่อยู่ตำแหน่งเดียวกัน เช่น



S	ช	น	เ	ช	า
W	ช	น			
ผลการเปรียบเทียบ	0	0	1	1	1
Mark <sub>1</sub>		1			
	ช	น			

พบคำศัพท์คำว่า “ชน” ดังนั้นผลลัพธ์คือ

Word<sub>1</sub>

Mark<sub>1</sub> { ชน/ }

จากนั้นเปรียบเทียบกับคำศัพท์ถัดไปในกลุ่มเดียวกัน (ความลดน้อยลงเรื่อย ๆ) จนถึงคำศัพท์ที่มีความยาวน้อยที่สุดคือ 2 ปรากฏว่าไม่พบคำศัพท์อีก

ลำดับที่ 4. ให้นำประโยคที่เหลือ(ถ้ามี) ไปผ่านกระบวนการตามข้อ 1, 2 และ 3 อีกจนกระทั่งจบประโยค และเก็บบันทึกคำศัพท์ที่สามารถแยกได้ทั้งหมด ผลลัพธ์ขณะนี้คือ

Word<sub>1</sub> Word<sub>2</sub> S  
Mark<sub>1</sub> { เธอ/ ชน/ เขา }

ลำดับที่ 1. คำนวณหาดัชนีที่ 1 และดัชนีที่ 2 ซึ่งก็คือรหัสแอสกีของอักขระตัวที่ 1 และ 2 ตามลำดับ และหาความยาว(จำนวนตัวอักษร)ของประโยคเพื่อกำหนดให้เป็นดัชนีที่ 3

$$I_1 = \text{ASCII}(C_1), I_2 = \text{ASCII}(C_2), I_3 = L = \text{Long}(S)$$

ตำแหน่ง	1	2	3
S	เ	ช	า

$$I_1 = \text{ASCII}(เ) = 224$$

$$I_2 = \text{ASCII}(ช) = 162$$

$$I_3 = L = \text{ความยาวประโยค} = 3$$

ลำดับที่ 2. นำประโยคอินพุตตามข้อ 1. ไปเปรียบเทียบกับคำศัพท์ในพจนานุกรม

2.1. ค้นหาคำศัพท์ในพจนานุกรมตาม  $I_1, I_2$

2.2. ใช้  $I_3$  (ความยาวของประโยค) ในการเริ่มต้นเปรียบเทียบ

2.2.1. ถ้าความยาวของประโยคมากกว่าหรือเท่ากับความยาวที่มากที่สุดของคำศัพท์ในกลุ่มคำศัพท์ตามดัชนีที่ 1 และ 2 กำหนดไว้ ให้เริ่มทำการเปรียบเทียบที่กลุ่มคำศัพท์นี้

2.2.2. ถ้าความยาวของประโยคน้อยกว่า ความยาวของคำศัพท์ให้เลื่อนเพื่อเริ่มการเปรียบเทียบลงไปอีกตามความยาวของคำศัพท์(ความยาวของคำศัพท์ลดลง)จนพบว่า ความยาวของประโยคมากกว่าหรือเท่ากับ ความยาวของคำศัพท์  $L \geq \text{ความยาวของคำศัพท์}$

ลำดับที่ 3. เริ่มทำการเปรียบเทียบโดยการเปรียบเทียบจะกระทำแบบอักขระต่ออักขระ คือเปรียบเทียบอักษรที่อยู่ตำแหน่งเดียวกัน เช่น

คำศัพท์ในพจนานุกรม	$D_1$	$D_2$	$D_3 \dots$	$\dots D_n$
	↑	↑	↑	↑
คำศัพท์ในประโยค	$C_1$	$C_2$	$C_3 \dots$	$\dots C_n$

S	เ	ช	า
W	เ	ช	า
ผลการเปรียบเทียบ	0	0	0
Mark <sub>1</sub>			1
	เ	ช	า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษเท่านั้น เมื่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

พบคำศัพท์คำว่า “เขา” ดังนั้นผลลัพธ์คือ

Word<sub>1</sub>

Mark<sub>1</sub> { เขา/ }

จากนั้นเปรียบเทียบกับคำศัพท์ถัดไปในกลุ่มเดียวกัน (ความลดน้อยลงเรื่อย ๆ) จนถึงคำศัพท์ที่มีความยาวนานที่สุดคือ 2 ปรากฏว่าไม่พบคำศัพท์อีก

- ลำดับที่ 4.      ให้นำประโยคที่เหลือ(ถ้ามี) ไปผ่านกระบวนการตามข้อ 1,2 และ 3 อีกจนกระทั่งจบประโยค และเก็บบันทึกคำศัพท์ที่สามารถแยกได้ทั้งหมด ผลลัพธ์ขณะนี้คือ

Word<sub>1</sub>    Word<sub>2</sub>    Word<sub>2</sub>    S

Mark<sub>1</sub> { เธอ/    ชน/    เขา    }

และพบว่าจบประโยคแล้ว ได้ผลลัพธ์คือ

Mark<sub>1</sub> { เธอ/ชน/เขา } ซึ่งเป็นรูปแบบเดียวที่สามารถแยกคำได้

ประโยคตัวอย่างที่ 2    ประโยค----->“โคลงเรือจนเรือโคลง”

- ลำดับที่ 1.      คำนวณหาดัชนีที่1 และดัชนีที่2 ซึ่งก็คือรหัสแอสกีของอักขระตัวที่ 1 และ 2 ตามลำดับ และหาความยาว(จำนวนตัวอักษร)ของประโยคเพื่อกำหนดให้เป็นดัชนีที่ 3

$$I_1 = \text{ASCII}(C_1), I_2 = \text{ASCII}(C_2), I_3 = L = \text{Long}(S)$$

ตำแหน่ง	1	2	3	4	5	.....18
---------	---	---	---	---	---	---------

S	โ	ค	ล	ง	เ	.....ง
---	---	---	---	---	---	--------

$$I_1 = \text{ASCII}(\text{โ}) = 226$$

$$I_2 = \text{ASCII}(\text{ค}) = 164$$

$$I_3 = L = \text{ความยาวประโยค} = 18$$

- ลำดับที่ 2.      นำประโยคอินพุตตามข้อ 1. ไปเปรียบเทียบกับคำศัพท์ในพจนานุกรม

2.1. ค้นหาคำศัพท์ในพจนานุกรมตาม  $I_1, I_2$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2. ใช้  $I_3$  (ความยาวของประโยค) ในการเริ่มต้นเปรียบเทียบ

2.2.1. ถ้าความยาวของประโยคมากกว่าหรือเท่ากับความยาวที่มากที่สุดของคำศัพท์ในกลุ่มคำศัพท์ตามลำดับที่ 1 และ 2 กำหนดไว้ ให้เริ่มทำการเปรียบเทียบที่กลุ่มคำศัพท์นี้

2.2.2. ถ้าความยาวของประโยคน้อยกว่า ความยาวของคำศัพท์ให้เลื่อนเพื่อเริ่มการเปรียบเทียบลงไปอีกตามความยาวของคำศัพท์(ความยาวของคำศัพท์ลดลง)จนพบว่า ความยาวของประโยคมากกว่าหรือเท่ากับ ความยาวของคำศัพท์  $L \geq$  ความยาวของคำศัพท์

ลำดับที่ 3. เริ่มทำการเปรียบเทียบโดยการเปรียบเทียบจะกระทำแบบอักษรต่ออักษร คือเปรียบเทียบอักษรที่อยู่ตำแหน่งเดียวกัน เช่น



S	โ	ค	ล	ง	เ	...	ล	ง
$W_1$	โ	ค	ล	ง				
$W_2$	โ	ค						
ผลการเปรียบเทียบ <sub>1</sub>	0	0	0	0	1	...	1	1
ผลการเปรียบเทียบ <sub>2</sub>	0	0	1	1	1	...	1	1
Mark <sub>1</sub>				1		...		
Mark <sub>2</sub>		1				...		
	โ	ค	ล	ง		...		

Word<sub>1</sub>

Word<sub>2</sub>

พบคำศัพท์คำว่า “โคลง” ดังนั้นผลลัพธ์คือ

Word<sub>1</sub>

Mark<sub>1</sub> { โคลง/ }

จากนั้นเปรียบเทียบกับคำศัพท์ถัดไปในกลุ่มเดียวกัน (ความลดย่อยลงเรื่อยๆ) จนถึงคำศัพท์ที่มีความยาวน้อยที่สุดคือ 2 ปรากฏว่าพบคำศัพท์คำว่า “โค” อีกดังนั้นผลลัพธ์คือ

Word<sub>1</sub>

Mark<sub>2</sub> { โค/ }

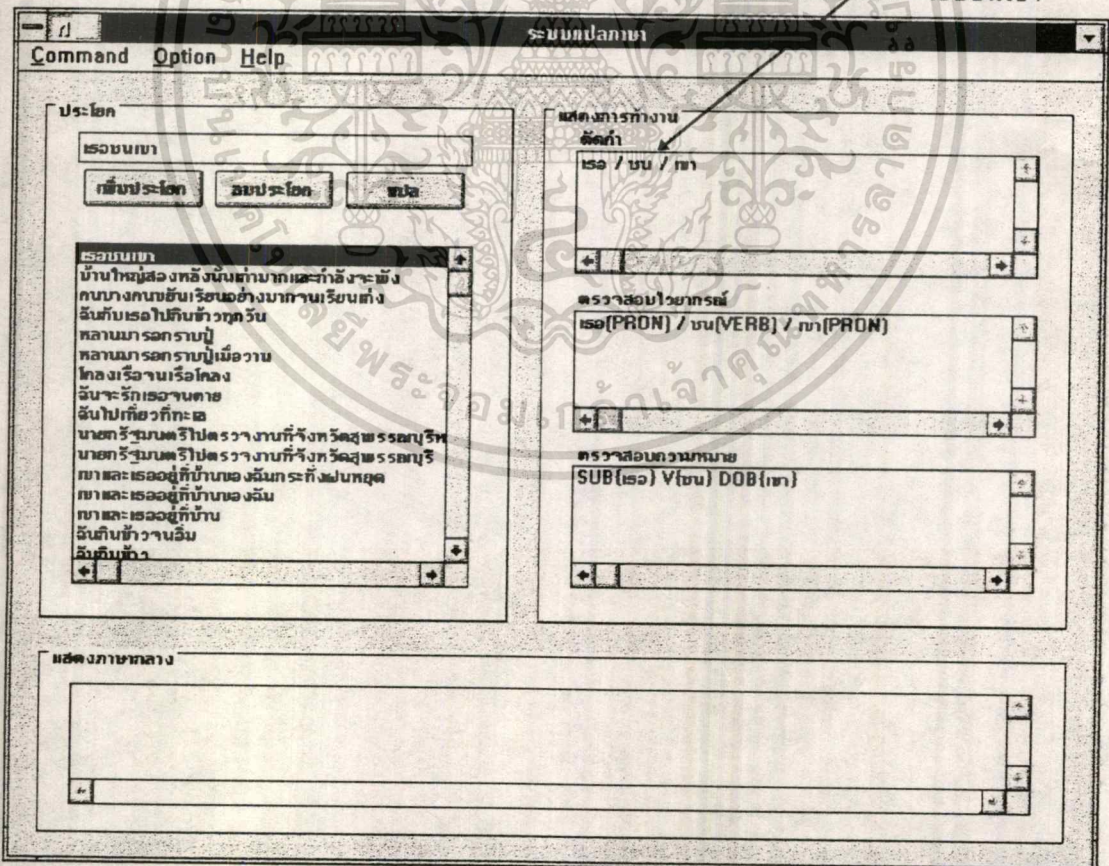
ลำดับที่ 4. ให้นำประโยคที่เหลือ(ถ้ามี) ไปผ่านกระบวนการตามข้อ 1,2 และ 3 อีกจนกระทั่งจบประโยค และเก็บบันทึกคำศัพท์ที่สามารถแยกได้ทั้งหมด ผลลัพธ์ขณะนี้คือ

Word <sub>1</sub>	S(ที่เหลือ)
Mark <sub>1</sub> { โคลง/	เรือจนเรือโคลง }
Mark <sub>2</sub> { โค/	ลงเรือจนเรือโคลง }

และเมื่อผ่านกระบวนการต่าง ๆ ตามลำดับข้างต้นแล้ว จะได้ผลลัพธ์คือ

Mark <sub>1</sub>	{ โคลง/เรือ/จน/เรือ/โคลง }
Mark <sub>2</sub>	{ โค/ลง/เรือ/จน/เรือ/โคลง }
Mark <sub>3</sub>	{ โคลง/เรือ/จน/เรือ/โค/ลง }
Mark <sub>4</sub>	{ โค/ลง/เรือ/จน/เรือ/โค/ลง }

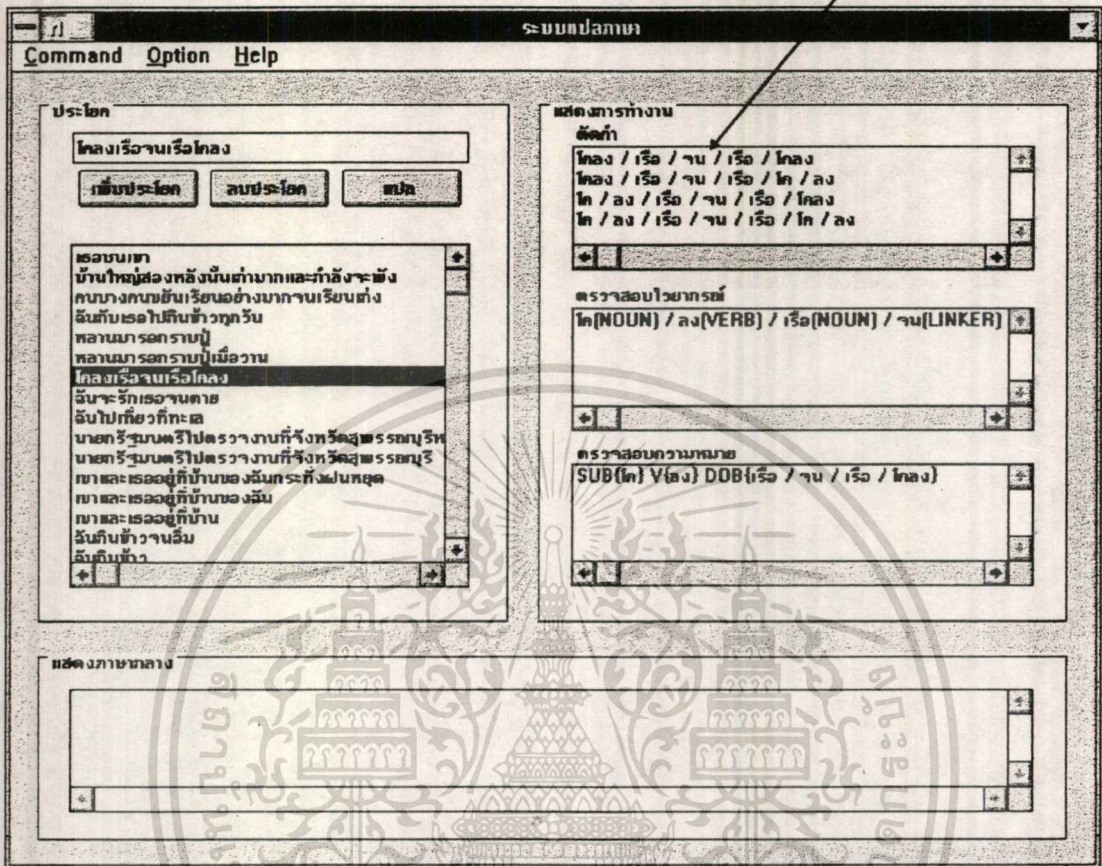
ผลการแยกประโยค “เธอชนเขา”



รูปที่ 4.8 แสดงตัวอย่างการแยกประโยค “เธอ ชน เขา”

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## ผลการแยกประโยค “โคลงเรือจนเรือโคลง”



รูปที่ 4.9 แสดงตัวอย่างการแยกประโยค “โคลงเรือจนเรือโคลง”

จะเห็นว่าประโยคภาษาไทยบางประโยคสามารถแยกคำศัพท์ได้เพียงรูปแบบเดียว แต่ก็มีบางประโยคที่สามารถแยกคำศัพท์ได้หลายรูปแบบเช่นในรูปที่ 4.9 ดังนั้นในโครงการนี้จึงมีกระบวนการเพื่อตัดสินใจเลือกเพียงประโยคเดียวที่ถูกต้องตามหลักโครงสร้างภาษา หรือหลักไวยากรณ์ภาษาไทย ซึ่งจะได้กล่าวต่อไปในหัวข้อ 4.3

### 4.3 การวิเคราะห์โครงสร้างและความหมายของประโยคภาษาไทย

ในการประมวลผลภาษาไทยนั้น ชั้นแรกจะต้องแยกคำศัพท์ในประโยคก่อนดังที่ได้อธิบายมาแล้ว จากนั้นก็ต้องตรวจสอบว่าประโยคถูกต้องตามกฎเกณฑ์ของภาษาหรือไม่ การตรวจสอบนี้ทำโดยอาศัยหลักทางโครงสร้างหรือหลักไวยากรณ์ภาษาไทยนั่นเอง การประมวลผลภาษานั้นมักจะประสบกับปัญหาเกี่ยวกับการวิเคราะห์ประโยคที่กำกวม นั่นคือถูกต้องตามหลักไวยากรณ์ทุกอย่าง แต่ความหมายผิดไป จึงจะต้องมีการวิเคราะห์ความหมายด้วย โดยอาศัยข้อมูลต่างๆ ของคำศัพท์ที่มีในประโยค และเมื่อตรวจสอบว่าประโยคนั้นถูกต้องทั้งไวยากรณ์และความหมายแล้ว จึงจะสามารถกำหนดความสัมพันธ์ของคำศัพท์ในประโยคได้ถูกต้อง ว่าคำศัพท์เหล่านั้นมีความสัมพันธ์กันอย่างไร ขั้นตอนสุดท้ายคือนำคำศัพท์และความสัมพันธ์ของคำศัพท์เหล่านั้นมาแทนในรูปของภาษากลาง

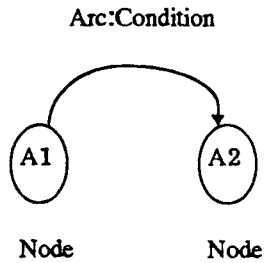
สำหรับขั้นตอนการวิเคราะห์โครงสร้างนั้น ได้มีการคิดค้นและพัฒนากันมานานเกิดเป็นทฤษฎีต่างๆ เช่น Pattern Matching, Transition Network, Context-Free Grammar, Transformational Grammar, Augmented Transition Network, Modified Augmented Transition Network (M-ATN)

สำหรับในโครงการนี้เราจะใช้ระบบซึ่งถูกพัฒนามาจากระบบ ATN นั่นคือระบบ M-ATN และถูกพัฒนาขึ้นเพื่อให้เหมาะสมกับการประมวลผลโครงสร้างภาษาไทยยิ่งขึ้น ด้วยระบบ M-ATN นี้ได้แสดงในรูปของโครงข่ายความสัมพันธ์ (Syntactical Relation Network) ของนามวลี (Noun Phrase) กริยาวลี (Verb Phrase) และวลีขยาย (Auxiliary Phrase) โดยแต่ละโครงข่ายถูกสร้างจากความรู้ทางภาษาศาสตร์เกี่ยวกับโครงสร้างภาษาไทย ที่ได้กล่าวแล้วในบทที่ 2 เราใช้โครงข่ายความสัมพันธ์นี้เป็นฐานความรู้แก่การวิเคราะห์โครงสร้างภาษาไทย ด้วยโครงข่าย M-ATN ประกอบกับกระบวนการประมวลผล M-ATN (M-ATN Process) จะสามารถวิเคราะห์โครงสร้างภาษาไทยได้อย่างถูกต้อง

#### 4.3.1 ความรู้เบื้องต้นเกี่ยวกับระบบโครงข่าย

ก.) ส่วนประกอบของโครงข่ายในระบบต่างๆ

โครงข่าย (Network) ที่ใช้ในการตรวจสอบโครงสร้างประโยคของระบบต่างๆ จะประกอบด้วย (Node) และอาร์ค (Arc) โดยโนคแต่ละโนคจะมีชื่อกำกับและจะต้องเป็นชื่อที่ไม่ซ้ำกัน โนคเหล่านี้จะถูกเชื่อมต่อกันเป็นโครงข่ายด้วยอาร์ค (หรือกลุ่มของอาร์ค) โดยอาร์คจะมีเงื่อนไข (Condition) ของการเคลื่อนย้ายที่แสดงความสัมพันธ์ทางไวยากรณ์ของโนคทั้งสองกำกับอยู่ แสดงในรูปที่ 4.10 ในรูปนี้ A1 และ A2 จะเป็นชื่อของโนค



รูปที่ 4.10 แสดงส่วนประกอบพื้นฐานของโครงข่ายในระบบต่างๆ

ในโครงงานฉบับนี้จะใช้วงกลมแทนโหนด และลูกศรแทนอาร์ค ส่วนการเคลื่อนย้าย จากโหนดหนึ่งไปยังอีกโหนดหนึ่ง จะใช้เส้นทางเดินของอาร์คเป็นตัวบ่งบอกทิศทางของการเคลื่อนที่จากโหนดที่อยู่ต้นลูกศรไปยังโหนดที่อยู่ปลายลูกศร โดยกำหนดการเคลื่อนย้ายไปตามเงื่อนไขทางไวยากรณ์ที่กำหนดไว้

นอกจากนั้นจะกำหนดให้  $S$  เป็นสัญลักษณ์แทนประโยคที่ประกอบด้วยคำที่ถูกแยกแยะได้ทั้งหมด  $n$  คำ โดยมีการเรียงลำดับของคำเป็น  $W_1, W_2, W_3, W_4, \dots, W_{n-1}, W_n$  นั่นคือ

$$S = W_1 W_2 W_3 W_4 \dots W_{n-1} W_n$$

เมื่อ  $W_i$  ( $i = 1, 2, 3, \dots, n$  : โดย  $n > 1$ ) เป็นคำที่แยกแยะได้จากประโยคด้วยระบบ Fast Work Matching ซึ่งจะมีชนิดของคำกำกับอยู่ในที่นี้  $W_1$  จะเป็นคำแรกสุดของประโยค และ  $W_n$  จะเป็นคำสุดท้ายของประโยค

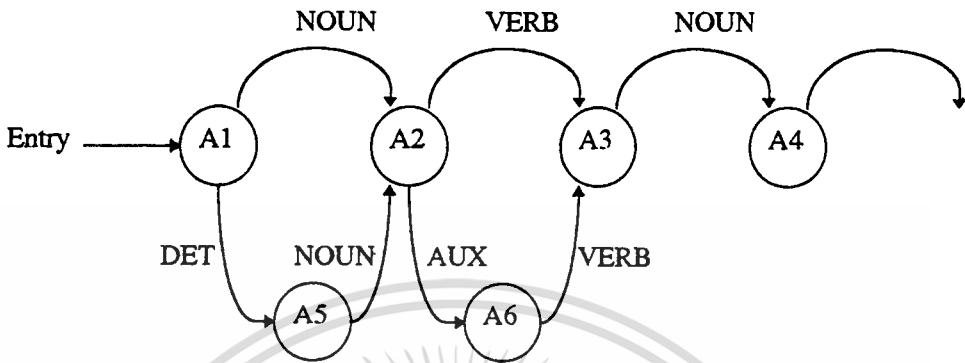
ข.) การทำงานของระบบประมวลผลโครงสร้างประโยค TN, RTN และ ATN

1.) หลักการทำงานพื้นฐานของระบบ Transition Network : TN

ระบบ TN จะประกอบด้วยกลุ่มของโหนดที่เชื่อมด้วยกลุ่มของอาร์ค ซึ่งการเชื่อมต่อเป็นไปตามเงื่อนไขความสัมพันธ์ทางไวยากรณ์ เช่น คำนาม (Noun) คำคุณศัพท์ (Adjective) คำกริยา (Verd) บุพบท (Preposition) คำช่วย (Auxiliary) เป็นต้น โดยจะมีการจัดเรียงอาร์ค

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตามหลักไวยากรณ์ภาษา ซึ่งระบบ TN จะสามารถแสดงโครงสร้างของภาษาดังตัวอย่างโครงสร้างง่าย ๆ ในภาษาอังกฤษที่แสดงในรูปที่ 4.11



รูปที่ 4.11 ตัวอย่างแสดงความสัมพันธ์ทางโครงสร้างบางส่วนของภาษาอังกฤษในระบบ TN

เมื่อนำตัวอย่างโครงข่ายรูปที่ 4.11 ของระบบ TN ไปหาชนิดของคำในประโยคคำแรกของประโยคจะถูกเริ่มประมวลผลที่โนดเริ่มต้นทุกครั้ง (ในรูปคือ Entry node ได้แก่โนด A1 นั้นเอง) จากนั้นระบบก็จะทำการตรวจสอบว่า เงื่อนไขของอาร์คใดตรงกับชนิดของคำตามหลักไวยากรณ์ที่จะประมวลผลตามหลักการทำงานดังตารางที่ 4.3

- ลำดับที่ 1 : เริ่มต้นการประมวลผลที่ Entry โนด โดยเริ่มจากคำแรกสุดของประโยคได้แก่ W1 (โดย  $I=1$ ) ซึ่งจะกำหนดให้เป็นคำที่จะประมวลคำแรกแล้วเริ่มลำดับที่ 2
- ลำดับที่ 2 : ตรวจสอบเงื่อนไขของอาร์คทั้งหมดที่ออกจากโนดนี้ว่ามีเงื่อนไขใดที่ตรงกับชนิดของคำที่จะประมวลผล (W1) หรือไม่
- 2.1) ถ้าเงื่อนไขของอาร์คใดตรงตามชนิดของคำที่จะประมวลผลให้เลือกอาร์คนั้นไว้ให้กับคำที่ W1 เลื่อนการประมวลผลไปคำที่  $W1+1$  ของประโยค พร้อมกับกำหนดให้  $W1+1$  เป็นคำที่จะประมวลผลคำใหม่ และเลื่อนการทำงานไปลำดับที่ 3
- 2.2) ถ้าไม่มีเงื่อนไขของอาร์คใดตรงกับชนิดของคำที่ประมวลผล (W1) แสดงว่าโครงสร้างทางไวยากรณ์ของประโยคนั้นผิด ให้หยุดการประมวลผลที่ตำแหน่งนั้น

- ลำดับที่ 3 : เริ่มต้นการประมวลผลใหม่กับคำต่อมา ( $W_{i+1}$ ) ที่โนดซึ่งอยู่ต่อจากอาร์คที่เลือกไว้ โดยย้อนการทำงานไปที่ลำดับที่ 2 จนกระทั่งหมดคำที่ประมวลผล (เมื่อ  $i=n$ ) หรือ โหนดที่จะเริ่มการประมวลผลจึงจะเลื่อนการทำงานไปลำดับที่ 4
- ลำดับที่ 4 : ให้นำเงื่อนไขอาร์คที่เลือกไว้ทั้งหมด มาเรียงกันตามลำดับเพื่อแสดงโครงสร้างทาง ไวยากรณ์ที่ถูกต้องของประโยคนั้น

#### ตารางที่ 4.3 แสดงกระบวนการทำงานของระบบ TN

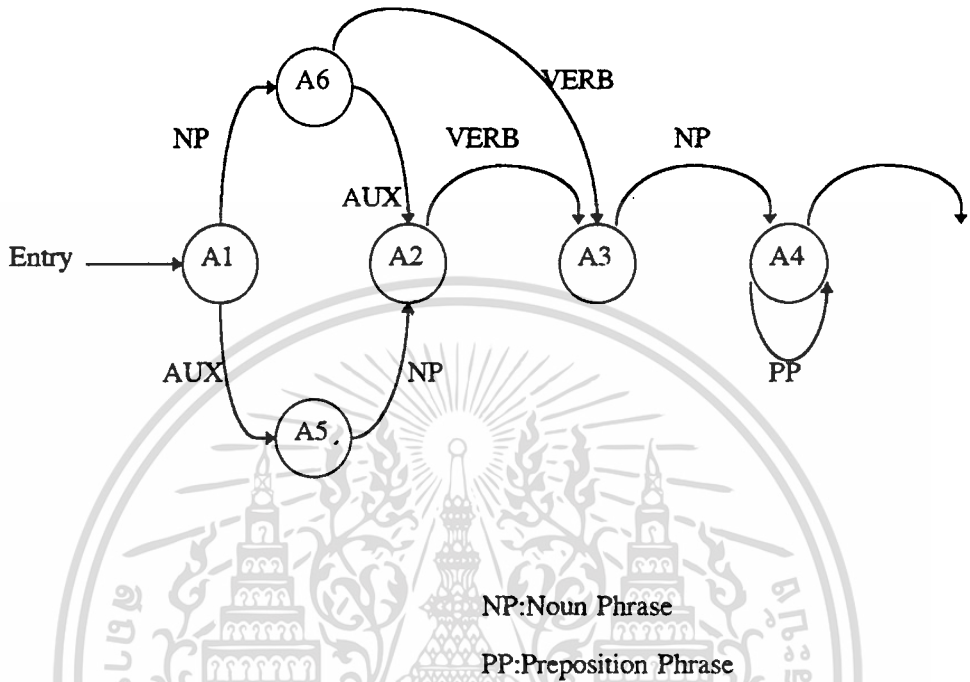
จากตัวอย่างการตรวจสอบโครงสร้างประโยคด้วยระบบ TN จะพบว่ามีความจำเป็นที่จะต้องกระจายโครงข่ายความสัมพันธ์ทางไวยากรณ์ให้ครอบคลุมทุกๆ โครงสร้างของประโยคในภาษานั้นๆ นั่นคือเงื่อนไขความสัมพันธ์ทางไวยากรณ์จะเป็นตัวเพิ่มจำนวนอาร์คให้มากขึ้นตามไปด้วย ทำให้โครงข่ายมีขนาดใหญ่และมีความซับซ้อนมากขึ้น ในการแก้ไขปรับปรุงโครงข่ายจะกระทำได้ลำบาก และถ้านำไปประยุกต์ใช้กับระบบคอมพิวเตอร์ ซอฟต์แวร์ที่สร้างขึ้นจะมีขนาดใหญ่และมีขั้นตอนการตรวจสอบที่ล่าช้าได้นี้เป็นสาเหตุทำให้มีการพัฒนาระบบ RTN ขึ้นมา

#### 2.) หลักการทำงานพื้นฐานของระบบ Recursive Transition Network : RTN

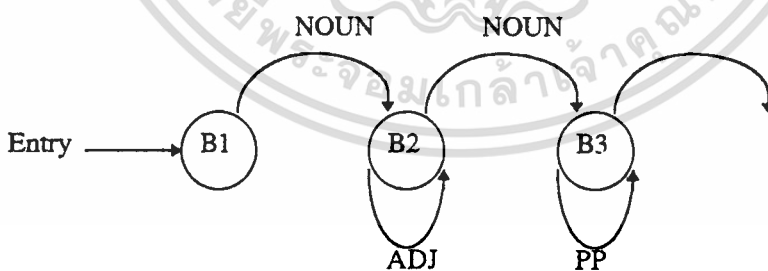
ระบบ RTN มีลักษณะโครงสร้างและหน้าที่การทำงานคล้ายๆ กับระบบ TN คือ มีจุดเริ่มต้นการประมวลผลคำเพียงโนดเดียว (Entry Node) และจะมีที่โนดสิ้นสุดก็ได้ โดยแต่ละโนดจะต้องมีชื่อกำกับและชื่อนั้นจะต้องไม่ซ้ำกัน ส่วนอาร์คจะมีเงื่อนไขที่แสดงชนิดของคำตามไวยากรณ์ในภาษา ส่วนที่แตกต่างจากระบบ TN ได้แก่การแบ่งโครงข่ายทั้งหมดที่มีในระบบออกเป็น 3 ส่วนใหญ่ ๆ ซึ่งในระบบ TN จะมีการรวมไว้เพียงโครงข่ายเดียว โครงข่ายทั้ง 3 ส่วนของระบบ RTN ได้แก่

1. Sentence Network (SN) โดยจะทำหน้าที่เป็นโครงข่ายของประโยค ซึ่งเป็นส่วนหลักของโครงข่ายทั้งหมด
2. Noun Phrase Network (NPN) เป็นโครงข่ายที่แสดงหน้าที่เกี่ยวกับคำนามและส่วนขยายนาม

3. Preposition Phrase Network (PPN) เป็นโครงข่ายที่แสดงหน้าที่ของคำบุพบทและส่วนขยายประโยค



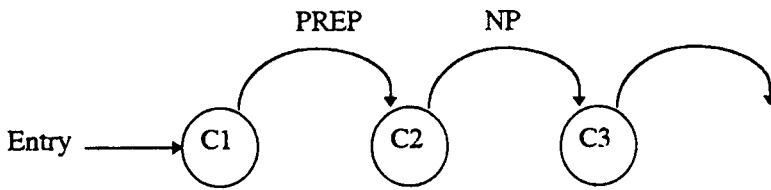
(ก) ตัวอย่างโครงข่าย SN ของภาษาอังกฤษในระบบ RTN



ADJ : Adjective

(ข) ตัวอย่างโครงข่าย NPN ของภาษาอังกฤษในระบบ RTN

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
 ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



PREP : Preposition

(ค) ตัวอย่างโครงข่าย PPN ของภาษาอังกฤษในระบบ RTN

รูปที่ 4.12 (ก),(ข),(ค) แสดงตัวอย่างความสัมพันธ์ทางโครงสร้างบางส่วนของประโยคภาษาอังกฤษ ในระบบ RTN

การทำงานของระบบ RTN ก็คล้ายกับระบบ TN คือ มีการตรวจสอบว่าคำในประโยคที่ทำการประมวลผลนั้นตรงกับเงื่อนไขของอาร์คหรือไม่ โดยจะเริ่มจากโครงข่าย SN ดังตารางที่ 4.4

- ลำดับที่ 1 : เริ่มต้นการประมวลผลที่ Entry โนดของ SN โดยเริ่มจากคำแรก ( $w_1 : I=1$ ) ของประโยคซึ่งกำหนดให้เป็นคำที่จะประมวลผลคำแรก แล้วเริ่มลำดับที่ 2
- ลำดับที่ 2 : ตรวจสอบเงื่อนไขของอาร์คทั้งหมดที่ออกจากโนดนี้ว่ามีเงื่อนไขใดที่ตรงกับชนิดของคำ ( $w_1$ ) ที่ประมวลผลหรือไม่
- 2.1) ถ้าเงื่อนไขของอาร์คใดตรงตามชนิดของคำที่ประมวลผล  $w_1$  ให้เลือกอาร์คนั้นไว้ให้กับคำนั้นแล้วเลื่อนการประมวลผลไปที่คำต่อไปของประโยค ( $w_{1+1}$ ) พร้อมกับกำหนดให้เป็นคำที่จะประมวลผลคำใหม่ และเลื่อนการทำงานไปลำดับที่ 3
  - 2.2) ถ้าไม่มีเงื่อนไขของอาร์คใดตรงกับชนิดของคำที่ประมวลผล ( $w_1$ ) ให้ย้อนกลับไปยังโนดก่อนหน้านั้นตามเงื่อนไขของอาร์คที่เข้ามา และให้นำคำศัพท์ก่อนหน้าคำศัพท์ที่ไม่สามารถประมวลผลได้ (คำที่  $w_{1-1}$ ) มาประมวลผลเงื่อนไขของอาร์คใหม่ที่ตำแหน่งของอาร์คใหม่ แต่หากประมวลผลไม่ได้แล้วย้อนกลับจนกระทั่งไปถึง Entry โนดของ SN (ที่ตำแหน่งคำศัพท์คำแรกของประโยค  $w_1$ ) แล้วไม่

สามารถประมวลผลได้อีก ให้ถือว่าประโยชน์นั้นคิดไวยากรณ์ให้หยุดการทำงาน

2.3) หากระบบทำการประมวลผลเงื่อนไขของอาร์คที่จะต้องอ้างถึงโครงข่ายอื่น (NPN, PPN) ให้ข้ามไปประมวลผลตามโครงข่ายนั้นๆ และประมวลผลเงื่อนไขตามลำดับที่ 2.1 และ 2.2 เมื่อประมวลผลเสร็จสิ้นให้ย้อนกลับไปยังโครงข่ายที่ออกมา

ลำดับที่ 3 : เริ่มต้นการประมวลผลใหม่กับค่าต่อไป ( $W_{i+1}$ ) ที่โนดซึ่งอยู่ต่อจากอาร์คที่เลือกไว้ โดยย้อนการทำงานไปที่ลำดับที่ 2 จนกว่าจะหมดค่าที่จะประมวลผล (เมื่อ  $I=1$ ) หรือโนดที่จะเริ่มการประมวลผลจึงจะเลื่อนการทำงานไปลำดับที่ 4

ลำดับที่ 4 : ให้นำเงื่อนไขของอาร์ค (ชนิดของค่า  $w$  ที่เลือกไว้ทั้งหมด มาเรียงกันตามลำดับเพื่อแสดงโครงสร้างทางไวยากรณ์ที่ถูกต้องของประโยชน์

#### ตารางที่ 4.4 แสดงกระบวนการทำงานของระบบ RTN

ข้อที่แตกต่างของการทำงานในระบบ RTN เมื่อเปรียบเทียบกับระบบ TN ก็คือ เมื่อการตรวจสอบค่ามาถึงเงื่อนไขของอาร์คที่แสดงเป็น NP หรือ PP การทำงานของระบบก็จะออกจากโครงข่าย SN ชั่วคราวเพื่อข้ามไปทำการตรวจสอบในส่วนของ NP หรือ PP ในโครงข่าย NPN และ PPN ตามลำดับ เมื่อทำการตรวจสอบต่อไป แม้ระบบ RTN จะมีความสามารถในการเพิ่มเติมหรือสร้างโครงข่ายย่อยแต่ กระบวนการทำงานยังมีปัญหา คือเมื่อไปถึงโนดใดๆ แล้วไม่มีเงื่อนไขของอาร์คที่ต่ออยู่ในโนดนั้นๆ ตรงกับค่าที่นำมาประมวลผล ระบบจะทำการย้อนกลับ (Backtack) แล้วทำการตรวจสอบเงื่อนไขของอาร์คใหม่ โดยระบบไม่มีความสามารถที่จะตรวจสอบว่าอาร์คใดที่เคยตรวจสอบแล้ว ดังนั้นจึงทำการตรวจสอบใหม่อีก ทำให้เสียเวลาโดยเปล่าประโยชน์ และระบบ RTN ยังไม่สามารถที่จะใช้กับประโยชน์ที่มีความซับซ้อนมากๆ ได้ จึงได้มีการพัฒนาระบบ Augmented Transition Network ขึ้นมา

### 3.) หลักการทำงานพื้นฐานของระบบ Augmented Transition Network : ATN

ATN มีพื้นฐานมาจาก TN (Transition Network) ซึ่งเป็นการประยุกต์ทางคณิตศาสตร์ด้านทฤษฎีกราฟ (Graph Theory) และ Finite State Machine โดยจะแสดงโครงสร้างประโยคของภาษาใดๆ ในรูปของโครงข่ายได้โดยการจัดกลุ่มของอาร์ค และโนดมาประกอบกันเป็นรูปของโครงข่าย ทั้งนี้โนดจะแสดงหน่วยคำและอาร์คจะแสดงความสัมพันธ์ระหว่างหน่วยคำที่เชื่อมติด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

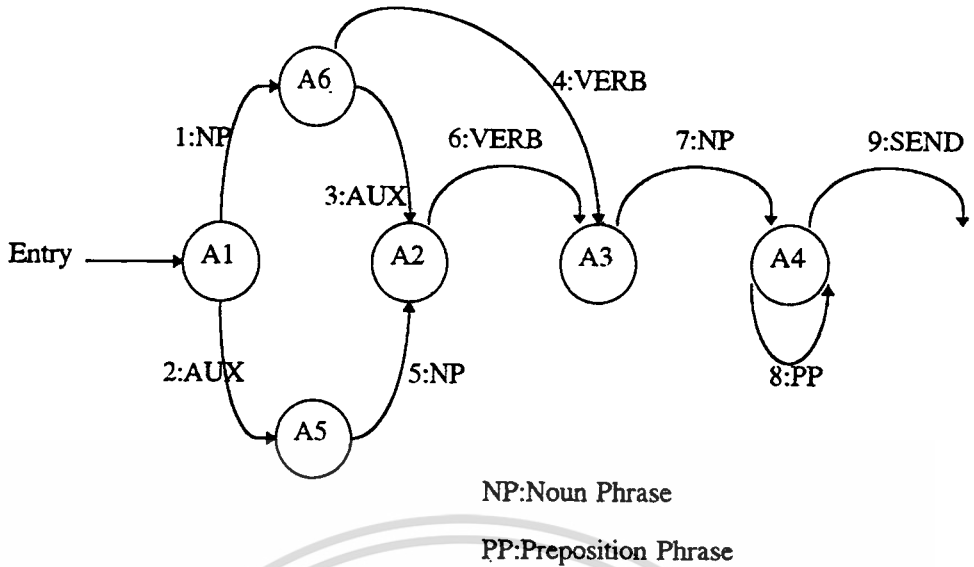
กัน ความสัมพันธ์ระหว่างหน่วยคำจะเป็นไปตามไวยากรณ์ทางด้านชนิดของคำเมื่อปรากฏในประโยค ข้อมูลทางภาษานี้จะเป็นตัวกำหนดการเชื่อมโยงโนดด้วยอาร์ค และการกำหนดความสัมพันธ์ หรือเงื่อนไขของอาร์คนั้นๆ นอกจากนั้นเมื่อนำโครงข่าย TN ไปประมวลผลโครงสร้างประโยคในภาษาใดๆ จะมีการเพิ่มเงื่อนไขสำหรับการเคลื่อนย้ายโนดและการจดจำ (Register) ว่าขั้นตอนของการประมวลผลอยู่ส่วนใดของประโยค และมีการตรวจสอบเงื่อนไขของอาร์ค ว่าเคยตรวจสอบมาแล้วหรือยัง ซึ่งในระบบ ATN มีอาร์คพื้นฐาน 4 แบบ ดังนี้

1. CATEGORY Arc(Cat Arc) หมายถึง อาร์คที่สั่งให้มีการประมวลผลตามลำดับที่กำหนด
- 2) SEEK Arc หมายถึง การให้ข้ามไปทำงานในตำแหน่งของโนดใน Network ที่กำหนดชื่อไว้เมื่อเสร็จ แล้วให้นำผลลัพธ์กลับมาด้วย เพื่อใช้ในการประมวลผลโนดต่อไป
- 3) JUMP Arc หมายถึง การทำงานของระบบเมื่อมาถึงอาร์คนี้ จะไม่ทำการตรวจสอบเงื่อนไขใดๆทั้งสิ้น ให้ข้ามไปยังโนดต่อไป หรือโนดที่กำหนด
- 4) SEND Arc หมายถึง อาร์คที่บอกถึงการสิ้นสุดของกระบวนการในโครงข่ายนั้นๆ แต่จะไม่ใช้กับโนดสุดท้ายก็ได้

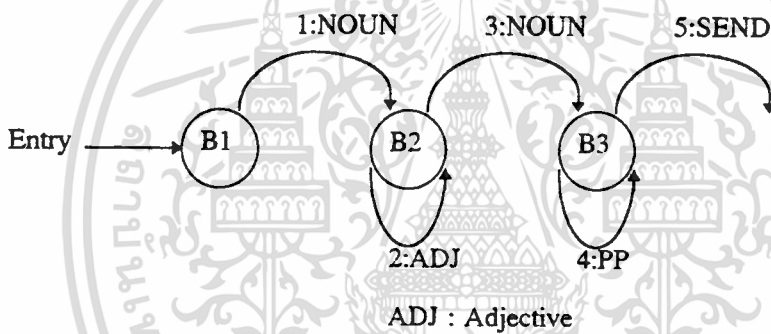
ระบบ ATN ได้แก่การแบ่งโครงข่ายทั้งหมดที่มีในระบบออกเป็น 3 ส่วนใหญ่ๆ ซึ่งโครงข่ายทั้ง 3 ส่วนของระบบ ATN ได้แก่

1. Sentence Network [SN] โดยจะทำหน้าที่เป็นโครงข่ายของประโยค ซึ่งเป็นส่วนหลักของโครงข่ายทั้งหมด
2. Noun Phrase Network [NPN] เป็นโครงข่ายที่แสดงหน้าที่เกี่ยวกับคำนาม และส่วนขยายนาม
3. Preposition Phrase Network [PPN] เป็นโครงข่ายที่แสดงหน้าที่ของคำบุพบทและส่วนขยายประโยค

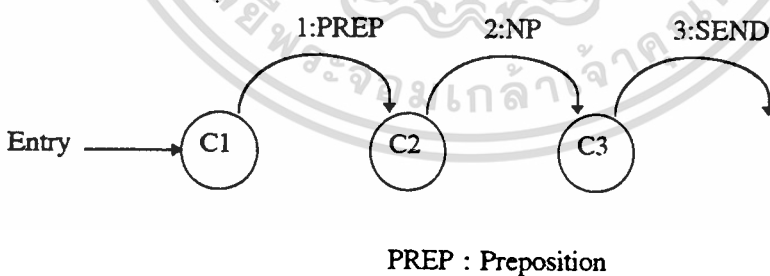
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ก) ตัวอย่างโครงข่าย SN ของภาษาอังกฤษในระบบ ATN



(ข) ตัวอย่างโครงข่าย NPN ของภาษาอังกฤษในระบบ ATN



(ค) ตัวอย่างโครงข่าย PPN ของภาษาอังกฤษในระบบ ATN

รูปที่ 4.13 ตัวอย่างความสัมพันธ์ทางโครงสร้างบางส่วนของประโยคภาษาอังกฤษในระบบ ATN

การทำงานของระบบ ATN ก็คล้ายกับระบบ RTN คือมีการตรวจสอบว่าคำในประโยคที่ทำกรการประมวลผลนั้นตรงกับเงื่อนไขของอาร์คหรือไม่ โดยจะเริ่มจากโครงข่าย SN ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ลำดับที่ 1 : เริ่มต้นการประมวลผลที่ Entry โนคของ SN โดยเริ่มจากค่าแรก ( $W1 : I=1$ ) ของประโยค ซึ่งกำหนดให้เป็นค่าที่จะประมวลผลค่าแรก แล้วเริ่มลำดับที่ 2
- ลำดับที่ 2 : ตรวจสอบเงื่อนไขของอาร์คทั้งหมดที่ออกจากโนคนี้ว่ามีเงื่อนไขใดที่ตรงกับชนิดของค่าที่ประมวลผล ( $W1$ ) หรือไม่
- 2.1) ถ้าเงื่อนไขของอาร์คใดตรงตามชนิดของค่าที่ประมวลผลให้เลือกอาร์คนั้นไว้ให้กับค่านั้น ( $W1$ ) ไว้พร้อมกับเก็บบันทึกลำดับของโนคและลำดับของอาร์คที่เลือกไว้ แล้วเลื่อนการประมวลผลไปที่ค่าต่อไป ( $W1+1$ ) ของประโยค แล้วเลื่อนการทำงานไปลำดับที่ 3
  - 2.2) ถ้าไม่มีเงื่อนไขของอาร์คใดตรงกับชนิดของค่าที่ประมวลผล ( $W1$ ) ให้ย้อนกลับไปยังโนคก่อนหน้านั้นตามเงื่อนไขของอาร์คที่เข้ามา และให้นำคำศัพท์ที่  $W1-1$  ซึ่งเป็นค่าที่อยู่ก่อนค่าที่ไม่สามารถประมวลผลได้มาประมวลผลเงื่อนไขของอาร์คใหม่ที่ตำแหน่งของอาร์คใหม่ แต่หากประมวลผลไม่ได้แล้วย้อนกลับจนกระทั่งถึง Entry โนคของ SN (ที่ตำแหน่งคำศัพท์ค่าแรก  $W1$  ของประโยค) แล้วไม่สามารถประมวลผลได้อีกให้ถือว่าประโยคที่นำมาประมวลผลนั้นผิดไวยากรณ์ ให้หยุดการทำงาน (เก็บบันทึกลำดับของโนคและลำดับของอาร์คไว้)
  - 2.3) ถ้าประมวลผลเงื่อนไขของอาร์คตรงกับชนิดของค่าที่ประมวลผล ( $W1$ ) ก่อนที่จะเลือกอาร์คนี้นี้ไว้ให้ตรวจสอบว่าเส้นทางของอาร์คที่เลือกไว้นั้นชี้ไปยังโนคที่เคยเก็บบันทึกไว้ว่า ไม่สามารถประมวลผลคำศัพท์ค่าต่อไป ( $W1+1$ ) กับเงื่อนไขของอาร์คที่ต่อออกจากโนคนั้นได้ ให้ประมวลผลเงื่อนไขของอาร์คอื่น (ถ้ามี) และเก็บบันทึกลำดับของโนคและลำดับของอาร์คไว้
  - 2.4) หากระบบทำการประมวลผลเงื่อนไขของอาร์คที่จะต้องอ้างอิงถึงโครงข่ายอื่น (NPN, PPN) ให้ข้ามไปประมวลผลตามโครงข่ายนั้น ๆ และประมวลผลเงื่อนไขตามลำดับที่ 2 เมื่อประมวลผลเสร็จสิ้นให้ย้อนกลับไปยังโครงข่ายที่ออกมา
- ลำดับที่ 3 : เริ่มต้นการประมวลผลใหม่กับค่าต่อไป ( $W1+1$ ) ที่โนคซึ่งอยู่ต่อจากอาร์คที่เลือกไว้ โดยเลื่อนการทำงานไปที่ลำดับที่ 2 จนกว่าจะหมดค่าที่จะประมวลผล (เมื่อ  $I=1$ ) หรือโนคที่จะเริ่มการประมวลผลจึงจะเลื่อนการทำงานไปลำดับที่ 4
- ลำดับที่ 4 : ให้นำเงื่อนไขของอาร์ค (ชนิดของค่า) ที่เลือกไว้ทั้งหมด มาเรียงกันตามลำดับเพื่อแสดงโครงสร้างทางไวยากรณ์ที่ถูกต้องของประโยคนั้น

#### ตารางที่ 4.5 แสดงกระบวนการทำงานของระบบ ATN

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

#### 4.8.2 ชนิดและหน้าที่ของคำศัพท์ภาษาไทยที่ใช้ในโครงการงานนี้

สำหรับชนิดของคำศัพท์ในภาษาไทยที่จะกำหนดในเงื่อนไขชนิดของคำในอาร์กนั้นได้กำหนดตามหลักไวยากรณ์ภาษาไทย โดยได้กล่าวมาแล้วในบทที่ 2 ซึ่งได้กำหนดให้เหมาะสมกับระบบ M-ATN ไว้ 16 ชนิดดังนี้

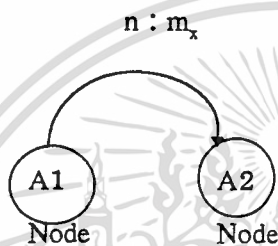
- 1) คำนาม (NOUN) ได้แก่กลุ่มคำ สามานยนาม เช่น คน สุนัข ต้นไม้ ไม้ ไร่ โรงเรียน หนังสือ โต๊ะ โทรทัศน์ นาฬิกา ถุงเท้า ฯลฯ
- 2) คำสรรพนาม (PRON) ได้แก่กลุ่มคำ บุรุษสรรพนาม เช่น ฉัน เขา เธอ หล่อน เรา ท่าน มัน คุณ ผม ฯลฯ
- 3) คำนามเฉพาะ (PROP) ได้แก่กลุ่มคำ วิสามานยนาม สมุหนาม เช่น วิชาญ นายแดง วัดพระแก้ว จังหวัดน่าน ฯลฯ
- 4) คำวิเศษณ์ (ADV) หมายถึงคำซึ่งมีหน้าที่ประกอบคำกริยา คำนาม ได้แก่กลุ่มคำ ลักษณะวิเศษณ์ กาลวิเศษณ์ ประพันธ์วิเศษณ์ นิยมวิเศษณ์ เช่น อย่างนี้ อย่างไม่ อย่างนั้น อย่างมาก อย่างรวดเร็ว อย่างช้า ดี ชั่ว เดียวนี้ ค้างนั้น ฯลฯ
- 5) คำกริยา (VERB) หมายถึงคำที่ทำหน้าที่แสดงกริยาอาการ ได้แก่กลุ่มคำวิกรรรถกริยา สกรรมกริยา อกรรมกริยา กริยาสภาวะมาลา เช่น กิน เดิน นอน นั่ง ยืน เท้า คล้าย เสมือน เป็น คือ ฯลฯ
- 6) คำกริยาช่วย (AUX) หมายถึงคำที่มาช่วยให้กริยามีความสมบูรณ์ หรือให้ความหมายดียิ่งขึ้น ได้แก่กลุ่มคำกริยานุเคราะห์ เช่น กำลัง จะ ควร กำลังจะ ต้อง ฯลฯ
- 7) คำคุณศัพท์ (ADJ) หมายถึงคำขยายคำนาม สรรพนาม และสิ่งของ ได้แก่กลุ่มคำ ลักษณะวิเศษณ์ สถานวิเศษณ์ กาลวิเศษณ์ นิยมวิเศษณ์ เช่น บน ล่าง คำ อ้วน ผอม สูง เตี้ย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- 8) คำลักษณะนาม (CLASS) หมายถึงการบ่งบอกลักษณะของ คำนาม สรรพนาม คน สัตว์ และ สิ่งของ ได้แก่กลุ่มคำลักษณะนาม สมุหนาม เช่น อัน หลัง ตัว ค้าม คน เครื่อง เล่ม ไบ
- 9) คำกำหนด (DET) หมายถึงคำที่ชี้แทน คน สัตว์ สิ่งของ เช่น นั่น โน้น
- 10) คำหมายของ (GM) หมายถึงคำเชื่อม “ของ”
- 11) คำเชื่อม (LINKER) หมายถึงคำที่ไ้เชื่อม คำนามกับคำนามได้แก่กลุ่มคำสันธาน เช่น กะ กับ และ หรือ ทั้ง ตั้ง ฯลฯ
- 12) คำนับ (NUM) หมายถึงคำที่แสดงถึงจำนวนของสิ่งของได้แก่กลุ่มคำประมาณวิเศษณ์ เช่น หนึ่ง สอง แด่ ละ บาง หลาย ทุก ก็ ฯลฯ
- 13) คำปฏิเสธ (NDG) หมายถึงคำที่มีความหมายในทางปฏิเสธได้แก่กลุ่มคำประติเสธวิเศษณ์ เช่น ไม่ได้ ไม่ได้ ฯลฯ
- 14) คำลงท้าย (END) หมายถึงคำที่อยู่ท้ายประโยคเพื่อทำให้ประโยคมีความหมายดียิ่งขึ้น เช่น ครับ ค่ะ นะ ฯลฯ
- 15) คำบุพบท (PREP) หมายถึงคำที่แสดงความสัมพันธ์ระหว่างคำ หรือ ประโยค ได้แก่คำ สำหรับ กับ ตาม โดย เพื่อ ที่ กระทั่ง ดังนั้น
- 16) คำแสดงคำถาม (QT) หมายถึงคำที่แสดงการถามได้แก่กลุ่มคำปฤจฉาวิเศษณ์ เช่น หรือ ไหม อะไร ซึ่งโครงข่ายที่กำหนดขึ้นนี้จะมีข้อกำหนดในการทำงานเป็นข้อๆ

### 4.3.3 ระบบ M-ATN ในการประมวลผลโครงสร้างประโยคภาษาไทย

ระบบ M-ATN ในที่นี้ได้พัฒนาขึ้นมาจากระบบ ATN โดยอาศัยพื้นฐานมาจากทฤษฎีกราฟที่ว่าโครงข่ายจะต้องประกอบด้วยโนด และอาร์คโดยที่โนดแต่ละโนดจะต้องมีชื่อไม่ซ้ำกัน และจะถูกเชื่อมต่อกันด้วยอาร์ค หรือกลุ่มของอาร์คกลายเป็นโครงข่าย โดยที่อาร์คแต่ละอาร์คจะมีเงื่อนไขปรากฏอยู่ดังแสดงในรูปที่ 4.14



รูปที่ 4.14 แสดงส่วนประกอบต่างๆ ของโครงข่ายในระบบ M-ATN

จากรูป สัญลักษณ์  $n : m_x$  มีความหมายดังนี้

$n$  : เป็นตัวเลขบอกลำดับของการตรวจสอบในแต่ละอาร์ค

$m_x$  คือเงื่อนไขทางไวยากรณ์ที่ใช้ในการตรวจสอบ

อาร์คจะมีทิศทางตามสูตร ดังนั้นการเคลื่อนย้ายจากโนด A ไปยังโนด B จึงประกอบด้วยเงื่อนไขตามที่กำหนด และทิศทางของการเคลื่อนย้ายจะเป็นไปตามทิศลูกศรของอาร์ค การประมวลผลโครงสร้างประโยคด้วยวิธี M-ATN นี้เป็นการประมวลผลที่ได้พัฒนาให้มีประสิทธิภาพดียิ่งขึ้นกว่าระบบก่อนๆ และให้เหมาะสมกับภาษาไทย เพราะทำให้สามารถหาชนิดของคำในประโยคได้อย่างถูกต้อง โดยเฉพาะการค้นหาคำกริยาหลัก (Main Predicate) ในประโยคซึ่งเป็นสิ่งจำเป็นอย่างมากในการหาโครงสร้างลึกของประโยค เพื่อนำไปสู่การแปลภาษาด้วยเครื่องคอมพิวเตอร์ต่อไป

ในโครงการนี้เราได้กำหนดหลักการทำงานและโครงสร้างพื้นฐานของระบบ M-ATN ไว้ว่าการประมวลผลคำในประโยคจะตรวจสอบจากซ้ายไปขวา กระบวนการตรวจสอบประกอบด้วย การตรวจสอบเงื่อนไข การย้อนกลับ และการจดจำผลของการตรวจสอบ ในระบบ M-ATN จะมีอาร์คทั้งหมด 3 ประเภทคือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1) Fixed Arc สัญลักษณ์  $n : m1$ 

ความหมาย  $m1$  เป็นเงื่อนไขที่ใช้ในการตรวจสอบประกอบด้วย 2 อาร์ค คือ JUMP Arc และ SEND Arc

หน้าที่ JUMP Arc ใช้ในการข้ามการตรวจสอบไปยังโนดต่อไปภายหลังจากเสร็จสิ้นการตรวจสอบเงื่อนไขของอาร์คที่มีลา

SEND Arc เป็นอาร์คที่บอกถึงการเสร็จสิ้นกระบวนการตรวจสอบในแต่ละโครงข่ายย่อยและการตรวจสอบจะย้อนกลับไปยังโครงข่ายก่อนหน้า

2) Network Arc สัญลักษณ์  $n : [m2]$ 

ความหมาย  $m2$  หมายถึงชื่อโครงข่ายซึ่งในโครงข่ายฉบับนี้ได้รวบรวมไว้ทั้งหมด 4 โครงข่าย สำหรับประโยคในภาษาไทยดังจะได้อธิบายต่อไป

หน้าที่ ทำการข้ามการตรวจสอบไปยังโครงข่ายที่ระบุด้วยค่า  $m2$  และเมื่อเสร็จสิ้นการตรวจสอบในโครงข่าย  $m2$  แล้วให้ย้อนกลับมาที่โครงข่ายเดิม

3) Category Arc สัญลักษณ์  $n : m3$ 

ความหมาย  $m3$  หมายถึงค่าที่บอกชนิดของคำ (Category) ซึ่งจะเป็นผลลัพธ์สุดท้ายของการตรวจสอบ

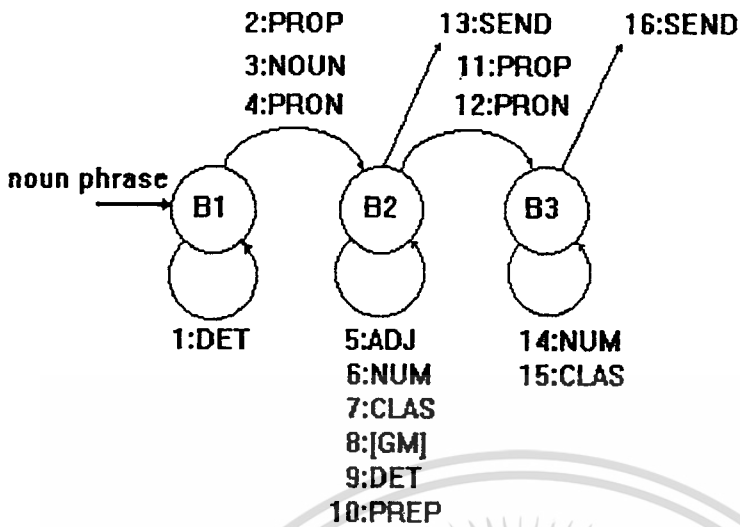
หน้าที่ ทำการตรวจสอบหาชนิดของคำตามหลักไวยากรณ์ภาษาไทยซึ่งได้กำหนดไว้ 4 โครงข่าย

## ก.) การกำหนดโครงข่าย M-ATN ของประโยคภาษาไทย

ในโครงข่ายนี้ได้รวบรวมรูปแบบโครงสร้างประโยคภาษาไทยเพื่อใช้ในระบบ M-ATN โดยอาศัยฐานข้อมูลทางภาษาศาสตร์ ที่ได้รับการวิจัยมาแล้ว และได้ทำการดัดแปลง ปรับปรุงเพื่อให้กระชับ และเหมาะสมกับการแสดงด้วยโครงข่าย และง่ายต่อการพัฒนาระบบซอฟต์แวร์คอมพิวเตอร์ ซึ่งผลสุดท้ายสามารถกำหนดได้ทั้งหมด 4 โครงข่าย โดยแต่ละโครงข่ายจะแสดงโครงสร้างของประโยคภาษาไทย ความหมายและชื่อของโครงข่ายทั้ง 4 มีดังต่อไปนี้

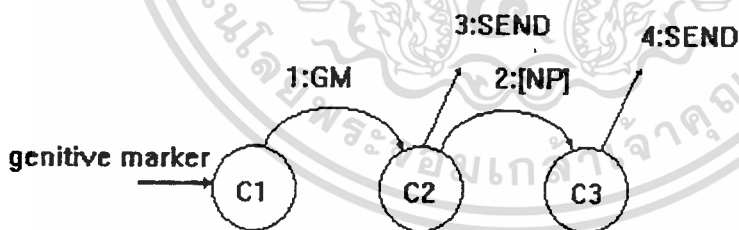
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้





รูปที่ 4.16 แสดงโครงขายนามวลี ( Noun Phrase : [NP] )

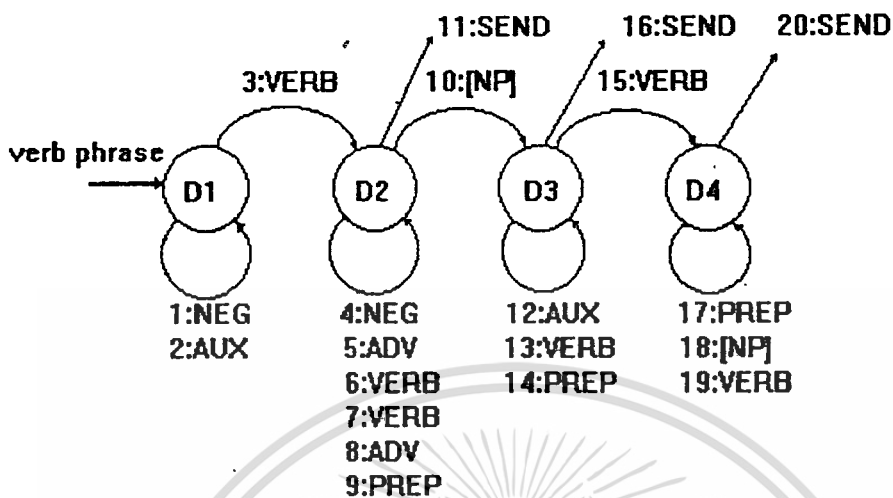
3) โครงข่ายของกลุ่มของ (Genitive Marker) [GM] เรียกว่า “กลุ่มของ” หมายถึง กลุ่มของคำว่า “ของ” และส่วนขยายคือ นามวลี ดังแสดงในรูปที่ 4.17



รูปที่ 4.17 แสดงโครงข่ายกลุ่มของ (Genitive Marker : [GM] )

4) โครงข่ายของกริยาวลี (Verb Phrase) VP เรียกว่า “กริยาวลี” หมายถึงกลุ่มของคำกริยา และส่วนขยายคำกริยาคด้วยซึ่งได้แก่ คำปฏิเสธ คำกริยานุเคราะห์(Auxiliary Verb) คำวิเศษณ์ (Adverb) คำกริยา (Verb) คำบุพบท และ นามวลี ดังแสดงในรูปที่ 4.18

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.18 แสดงโครงข่ายของกริยาวลี (Verb Phrase : [VP])

ข.) ข้อกำหนดและกระบวนการทำงานของโครงข่าย M-ATN

โครงข่ายที่กำหนดขึ้นในหัวข้อ ก.) นั้นจะมีข้อกำหนดในการทำงานดังนี้

- 1.) ต้องเริ่มต้นการตรวจสอบทั้งหมดที่โครงข่าย SN และจบสิ้นกระบวนการสุดท้ายที่โครงข่าย SN เช่นกัน
- 2.) มีลำดับการตรวจที่โนดใดๆ ตามลำดับตัวเลขที่กำหนดไว้ในเงื่อนไขของอาร์คนั้นๆ
- 3.) หากมีการเรียกใช้โครงข่ายย่อยและเงื่อนไขของอาร์คในโครงข่ายย่อยไม่ตรงตาม ชนิดของค่าที่ประมวลผลเลย ให้ถือว่าไม่ได้มีการเรียกใช้โครงข่ายย่อยนั้นๆ
- 4.) การจะข้ามโนดจากโนดหนึ่งไปยังอีกโนดหนึ่งซึ่งถูกเชื่อมต่อกับอาร์คใดๆจะต้องมีเงื่อนไขของอาร์คตรงกับชนิดของค่าที่ประมวลผล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากข้อกำหนดข้างต้นสามารถที่จะเขียนเป็นขั้นตอนการทำงานได้ในตารางที่ 4.6

ลำดับที่ 1 : เริ่มต้นจากประมวลผลที่ Entry โนคของ SN โดยเริ่มจากค่าแรก ( $W1 : I=1$ ) ของประโยค ซึ่งกำหนดให้เป็นค่าที่จะประมวลผลค่าแรกกับเงื่อนไขของอาร์คอันดับที่ 1 แล้วเริ่มลำดับที่ 2

ลำดับที่ 2 : ตรวจสอบเงื่อนไขของอาร์คทั้งหมดที่ออกจากโนคนี้ว่ามีเงื่อนไขใดที่ตรงกับชนิดของค่า ( $W1$ ) ที่จะประมวลผลหรือไม่ตามลำดับหมายเลขที่กำหนดไว้จากน้อยไปหามาก

2.1) ถ้าเงื่อนไขของอาร์คใดตรงตามชนิดของค่าที่ประมวลผล ( $W1$ ) ให้เลือกอาร์คนั้นไว้ให้กับค่านั้นไว้พร้อมกับเก็บบันทึกลำดับของโนคและลำดับของอาร์คที่เลือกไว้แล้วเลื่อนการประมวลผลไปที่คำต่อไปของประโยค แล้วเลื่อนการทำงานไปลำดับที่ 3

2.2) ถ้าไม่มีเงื่อนไขของอาร์คใดตรงกับชนิดของค่าที่ประมวลผล ( $W1$ ) ให้ย้อนกลับไปยังโนคก่อนหน้านั้นตามเงื่อนไขของอาร์คที่เข้ามา และให้นำคำศัพท์ก่อนหน้าคำศัพท์ที่ไม่สามารถประมวลผลได้ ( $W1-1$ ) มาประมวลผลเงื่อนไขของอาร์คใหม่ที่ตำแหน่งของอาร์คใหม่ แต่หากประมวลผลไม่ได้จนกระทั่งย้อนกลับไปถึง Entry โนคของ SN (ที่ตำแหน่งคำศัพท์ค่าแรก  $W1$  ของประโยค) แล้วไม่สามารถประมวลผลได้อีก ให้ถือว่าประโยคนั้นผิดไวยากรณ์ ให้หยุดการทำงาน (เก็บบันทึกลำดับของโนคและลำดับของอาร์คไว้)

2.3) ถ้าประมวลผลเงื่อนไขของอาร์คตรงกับชนิดของค่าที่ประมวลผล ( $W1$ ) ก่อนที่จะเลือกอาร์คนี้ไว้ให้ตรวจสอบว่าเส้นทางของอาร์คที่เลือกไว้นั้นขึ้นไปยังโนคที่เคยเก็บบันทึกไว้ว่า ไม่สามารถประมวลผลคำศัพท์คำต่อไป ( $W1+1$ ) กับเงื่อนไขของอาร์คที่ต่อออกจากโนคนั้นได้ให้ประมวลผลเงื่อนไขของอาร์คอื่น (ถ้ามี) และเก็บบันทึกลำดับของโนคและลำดับของอาร์คไว้

2.4) หากระบบทำการประมวลผลเงื่อนไขของอาร์คที่จะต้องอ้างอิงถึงโครงข่ายอื่น (NPN, PPN) ให้ข้ามไปประมวลผลตามโครงข่ายนั้นๆ และประมวลผลเงื่อนไขตามลำดับที่ 2 เมื่อประมวลผลเสร็จสิ้นให้ย้อนกลับไปยังโครงข่ายที่ออกมา

ลำดับที่ 3 : เริ่มต้นการประมวลผลใหม่กับคำต่อไป ( $W_{i+1}$ ) ที่โนดซึ่งอยู่ต่อจากอาร์คที่เลือกไว้ โดยย้อนการทำงานไปที่ลำดับที่ 2 จนกว่าจะหมดคำที่จะประมวลผล (เมื่อ  $i=n$ ) หรือโนดที่จะเริ่มการประมวลผลจึงจะเลื่อนการทำงานไปลำดับที่ 4

ลำดับที่ 4 : ให้นำเงื่อนไขของอาร์ค (ชนิดของคำ) ที่เลือกไว้ทั้งหมด มาเรียงกันตามลำดับเพื่อแสดงโครงสร้างทางไวยากรณ์ที่ถูกต้องของประโยคนั้น

#### ตารางที่ 4.6 แสดงกระบวนการการทำงานของระบบ M-ATN

ด้วยระบบ M-ATN ทำให้สามารถลดขั้นตอนการประมวลผลให้น้อยลงและรัดกุมขึ้นเมื่อเปรียบเทียบกับระบบ ATN ซึ่งจะทำให้ความถูกต้องและความรวดเร็วในการคัดเลือกโครงสร้างของประโยค เพราะเราได้ลดความซับซ้อนให้น้อยลง เมื่อเปรียบเทียบกับระบบอื่นๆ ที่กล่าวมา การออกแบบและการสร้าง ระบบซอฟต์แวร์ให้กับกระบวนการ M-ATN ก็ทำได้สะดวก ซึ่งในโครงการนี้ได้ใช้ภาษาซี (Borland c ++ Version 4.5) ในการเขียนซอฟต์แวร์ทั้งหมด โดยการแบ่งเป็นส่วนต่างๆ ไปตามขั้นตอนการประมวลผลของระบบ M-ATN ซึ่งจะกล่าวในบทที่ 5

#### ค.) ตัวอย่างการประมวลผลประโยคภาษาไทยด้วยโครงข่าย M-ATN

##### ประโยคตัวอย่างที่ 1 “นักเรียนไปโรงเรียน”

เมื่อผ่านกระบวนการแยกคำศัพท์แล้วจะได้ผลลัพธ์คือ

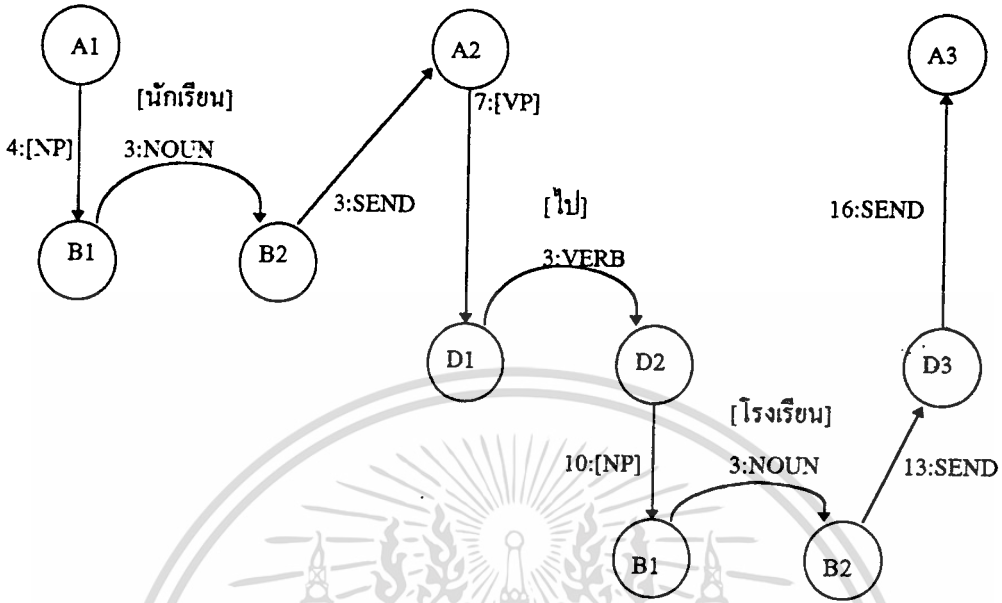
นักเรียน/ ไป/ โรงเรียน

ซึ่งแต่ละคำจะมีชนิดของคำกำกับมาด้วยดังนี้

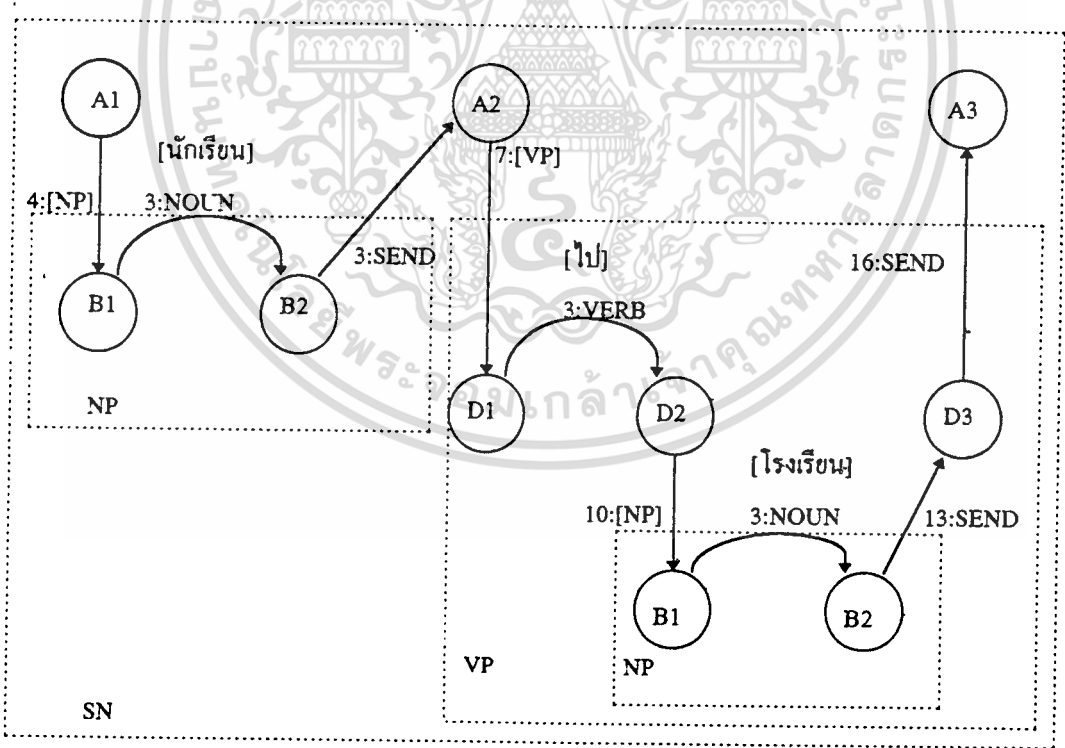
นักเรียน	ไป	โรงเรียน
NOUN	VERB	NOUN

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อนำคำศัพท์เหล่านี้มาประมวลผลด้วยโครงข่าย M-ATN จะได้ผลดังรูปที่ 4.19



(ก) แสดงการวิเคราะห์โครงสร้างประโยค “นักเรียนไปโรงเรียน”



(ข) แสดงการวิเคราะห์โครงสร้างประโยคด้วยโครงข่ายต่างๆ ( [SN], [NP], [VP] )

รูปที่ 4.19 แสดงตัวอย่างการประมวลผลประโยค “นักเรียนไปโรงเรียน”

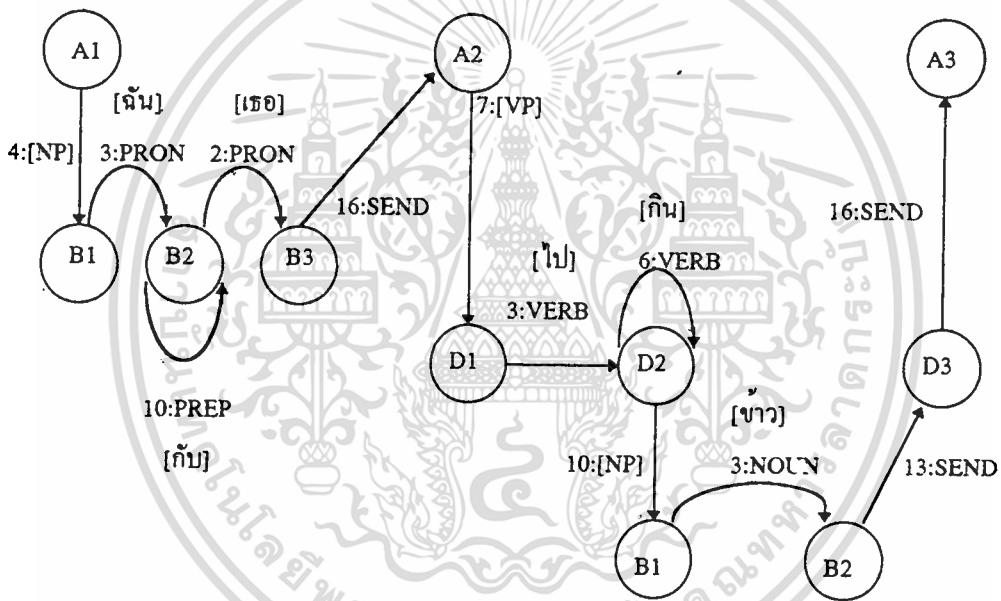
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประโยคตัวอย่างที่ 2 “ฉันกับเธอไปกินข้าว”

เมื่อผ่านกระบวนการแยกคำศัพท์แล้วจะได้ผลลัพธ์คือ

ฉัน	กับ	เธอ	ไป	กิน	ข้าว
PRON	PREP	PRON	VERB	VERB	NOUN

เมื่อนำคำศัพท์เหล่านี้มาประมวลผลด้วยโครงข่าย M-ATN จะได้ผลดังรูปที่ 4.20



รูปที่ 4.20 แสดงตัวอย่างการประมวลผลประโยค “ฉันกับเธอไปกินข้าว”

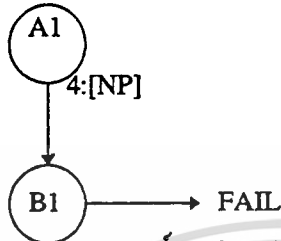
ประโยคตัวอย่างที่ 3 “โคลงเรือจนเรือโคลง”

เมื่อผ่านกระบวนการแยกคำศัพท์แล้วจะได้ผลลัพธ์ 4 รูปแบบคือ

รูปแบบที่ 1	โคลง/	เรือ/	จน/	เรือ/	โคลง
รูปแบบที่ 2	โค/	ลง/	เรือ/	จน/	เรือ/ โคลง
รูปแบบที่ 3	โคลง/	เรือ/	จน/	เรือ/	โค/ ลง
รูปแบบที่ 4	โค/	ลง/	เรือ/	จน/	เรือ/ โคลง/ ลง

ดังนั้นเราจะใช้กระบวนการประมวลผลโครงข่าย M-ATN เพื่อเลือกประโยคที่ต้องการ  
รูปแบบประโยคที่ 1

โคลง	เรือ	จน	เรือ	โคลง
VERB	NOUN	LINK	NOUN	VERB

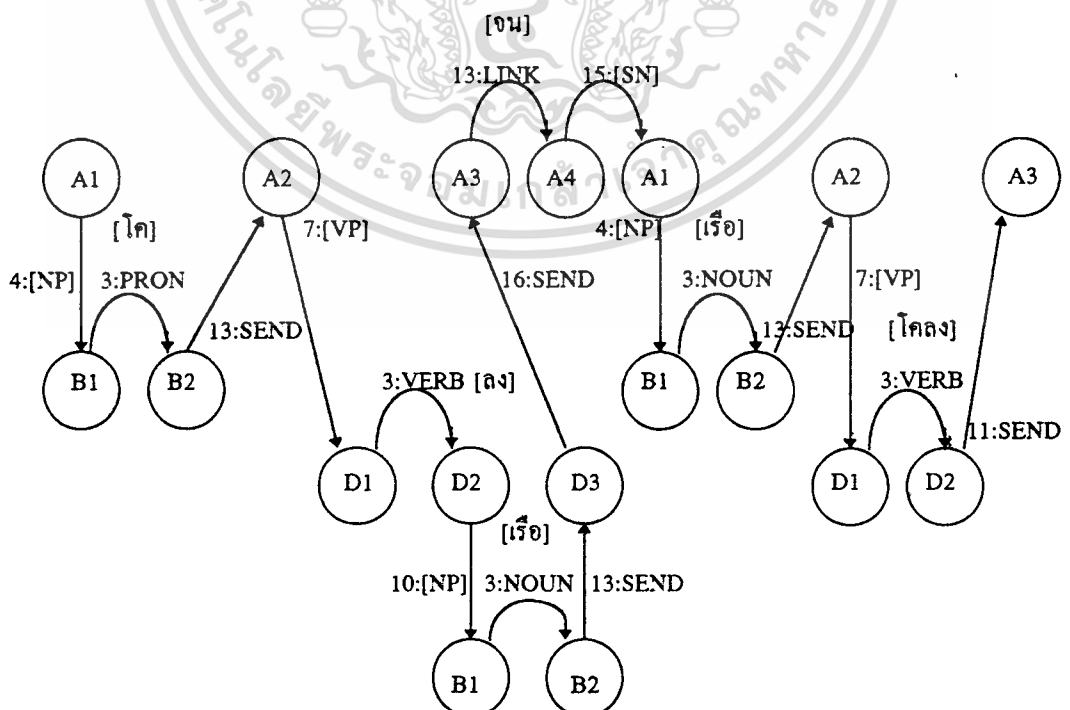


รูปที่ 4.21 แสดงการวิเคราะห์ประโยค “โคลง/เรือ/จน/เรือ/โคลง”

เพราะคำศัพท์คำแรกเป็นชนิดกริยา เมื่อประมวลผลโครงข่าย SN (A1) และโครงข่าย NP (B1) ปรากฏว่าไม่สามารถประมวลผลได้ จึงถือว่าประโยคในรูปแบบที่ 1 นี้ผิดไวยากรณ์

รูปแบบประโยคที่ 2

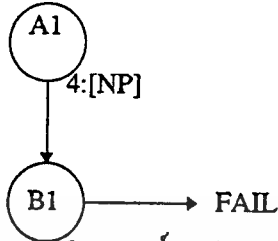
โค	ลง	เรือ	จน	เรือ	โคลง
NOUN	VERB	NOUN	LINK	NOUN	VERB



รูปที่ 4.22 แสดงการวิเคราะห์ประโยค “โค/ลง/เรือ/จน/เรือ/โคลง”

รูปแบบประโยคที่ 3

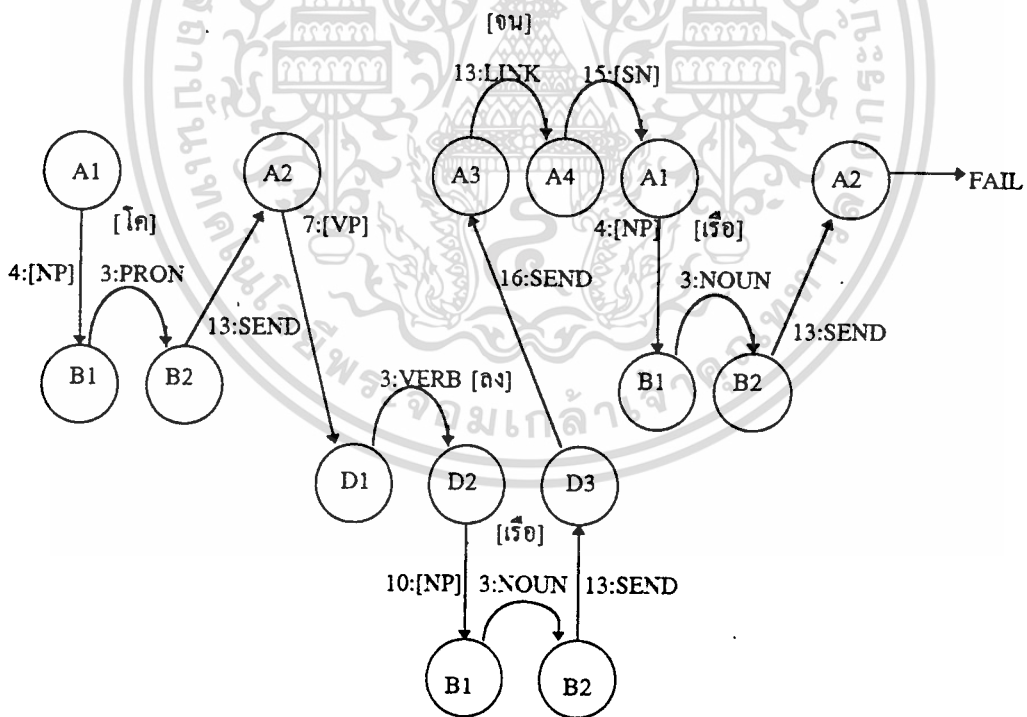
โคลง	เรือ	จน	เรือ	โค	ลง
VERB	NOUN	LINK	NOUN	NOUN	VERB



รูปที่ 4.23 แสดงการวิเคราะห์ประโยค “โคลง/เรือ/จน/เรือ/โค/ลง”

รูปแบบประโยคที่ 4

โค	ลง	เรือ	จน	เรือ	โค	ลง
NOUN	VERB	NOUN	LINK	NOUN	NOUN	VERB



รูปที่ 4.24 แสดงการวิเคราะห์ประโยค “โค/ลง/เรือ/จน/เรือ/โค/ลง”

จากการประมวลผลประโยคทั้ง 4 รูปแบบ จะเห็นว่าประโยคที่ถูกต้องตามหลักไวยากรณ์ที่กำหนดโดยโครงข่าย M-ATN นั้นคือประโยครูปแบบที่ 2 ( โค/ลง/เรือ/จน/เรือ/โคลง) เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

#### 4.4 การวิเคราะห์และแทนรูปภาษากลาง

เมื่อผ่านการประมวลผลจากกระบวนการวิเคราะห์โครงสร้างประโยคภาษาไทยแล้ว เอกลักษณ์ที่ได้ก็คือ คำศัพท์ และข้อมูลของคำที่จำเป็นเช่น ชนิดของคำ ประเภทของคำศัพท์ที่แบ่งตามความหมาย (AKO) เป็นต้น ซึ่งขั้นตอนต่อไปก็คือ นำเอาคำศัพท์และข้อมูลของคำเหล่านั้นมาวิเคราะห์เพื่อกำหนดความสัมพันธ์ ว่าคำศัพท์แต่ละคำมีความสัมพันธ์ระหว่างกันอย่างไร แล้วสร้างเป็นโครงข่ายทางความหมาย เพื่อแทนในรูปภาษากลางต่อไป

##### 4.4.1 ขั้นตอนการวิเคราะห์และสร้างโครงข่ายทางความหมาย

###### 1.) แยกประธาน / กริยา / กรรม ของประโยค

ให้ SUB = ประธาน  
VERB = กริยา  
DOB = กรรม

ดังนั้น ประโยค = SUB / VERB / DOB

โดย SUB = ทุกคำที่อยู่หน้า VERB ตัวแรก  
DOB = ทุกคำที่อยู่หลัง VERB ตัวสุดท้าย  
VERB = ทุกคำที่อยู่ระหว่าง SUB และ DOB

###### 2.) หากกริยาหลักของประโยค

ในที่นี้เราจะถือเอาว่า กริยาตัวแรกของประโยคเป็นกริยาหลักของประโยค แต่มีข้อยกเว้นคือ ถ้าคำศัพท์ “ไป” หรือ “มา” เป็นกริยาตัวแรกและมีกริยาคำอื่นอีก ให้ถือเป็นการบอกทิศทาง และให้เอาคำถัดไปเป็นกริยาหลัก และถ้ามีคำศัพท์ “ไป” หรือ “มา” เพียงคำเดียว ให้ถือเอาเป็นกริยาหลักได้เลย

###### 3.) วิเคราะห์ประธานของประโยค (SUB)

ประโยคของประโยคนั้นจะเป็นนามวลี ที่มีคุณลักษณะเฉพาะ (Attribute) จำนวน (Num) และการแสดงการเป็นเจ้าของ (Gm)

คุณลักษณะ (ATTRIBUTE) จะต้องประกอบไปด้วยคำที่มีชนิดต่อไปนี้

Attrib = [ ADJ, DET, (CLAS+DET) ]

จำนวน (NUMBER) จะต้องประกอบไปด้วยคำที่มีชนิดต่อไปนี้

Num = [ NUM+CLAS ]

การเป็นเจ้าของ (GENITIVE MARKER) โดยการใช้กฎเกณฑ์ดังนี้

ก. คำศัพท์ที่อยู่หลังคำศัพท์ “ของ” เป็น นามเฉพาะ (PROP)

นาม (NOUN)

สรรพนาม (PRON) เท่านั้น

ข. คำศัพท์ที่อยู่หน้าคำศัพท์ “ของ” เป็น สรรพนามไม่ได้ เช่น

“ฉัน ของ ....”

“เธอ ของ ....”

ค. คำศัพท์ที่อยู่หลังคำศัพท์ “ของ” ที่มีประเภทตามความหมาย (A Kind Of) เป็น  
องค์กร, บุคคล ดังนั้น คำศัพท์ที่อยู่หน้าคำศัพท์ “ของ” เป็นได้ทั้งคำ นาม, สรรพ  
นาม และนามเฉพาะ

ง. คำศัพท์ที่อยู่หลังคำศัพท์ “ของ” ที่มีประเภทตามความหมาย (A Kind Of) เป็น  
สัตว์ ดังนั้น คำศัพท์ที่อยู่หน้าคำศัพท์ “ของ” จะต้องมี A Kind Of เป็นส่วน  
ประกอบของสัตว์ (Part of animal) เช่น “ขา ของ ช้าง”

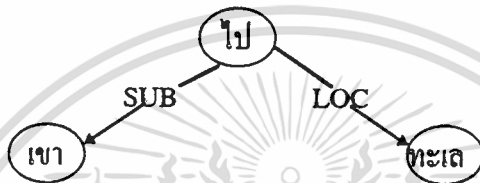
“ขน ของ เม่น”

จ. คำศัพท์ที่อยู่หลังคำศัพท์ “ของ” ที่มีประเภทตามความหมาย (A Kind Of) เป็น  
พืช ดังนั้น คำศัพท์ที่อยู่หน้าคำศัพท์ “ของ” จะต้องมี A Kind Of เป็นส่วน  
ประกอบของพืช (Part of plant) เช่น “ใบ ของ ผัก”

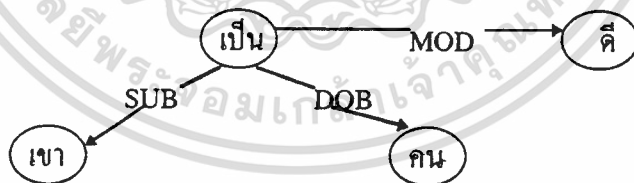
“กิ่ง ของ ต้นไม้”

#### 4.) วิเคราะห์กรรมของประโยค

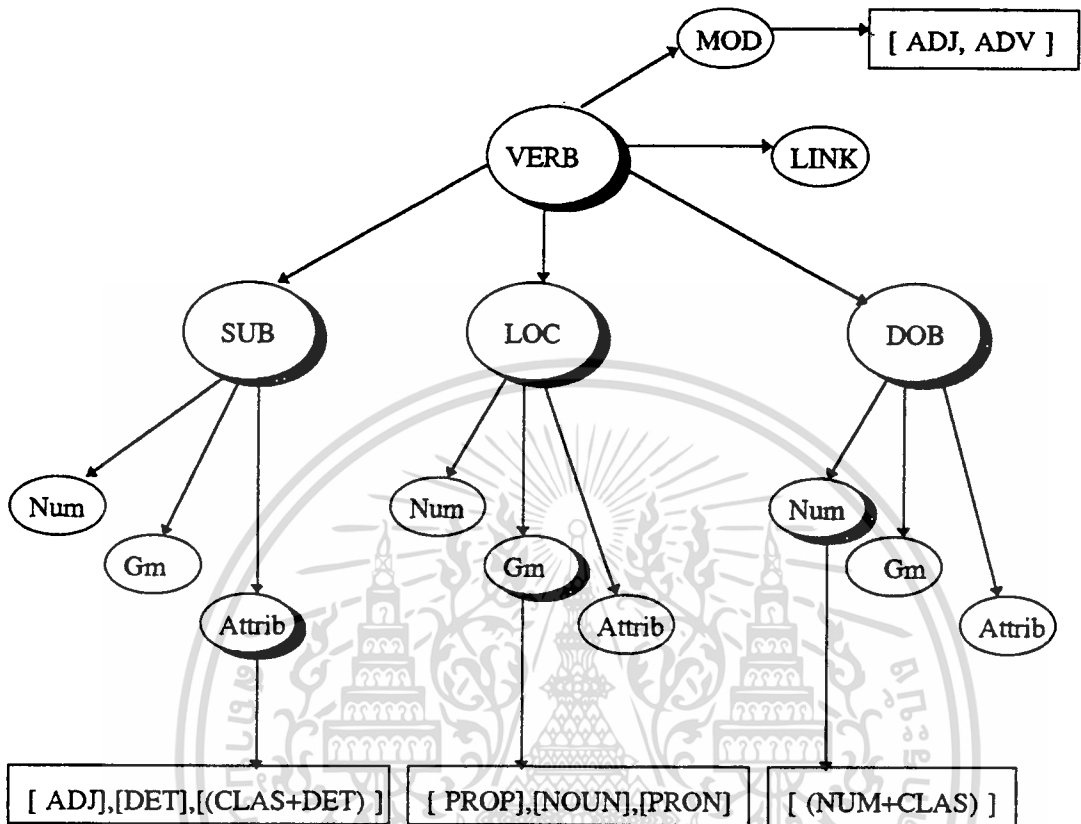
ให้หาคำศัพท์ที่มีชนิดเป็น PREP ในส่วนของ DOB เช่นคำศัพท์ “ที่” โดยถ้าคำศัพท์ที่ตามหลัง PREP นั้นเป็น PROP หรือ NOUN ให้หมายถึงการบอกสถานที่ เช่นประโยค “เขาไปที่ทะเล”



แต่ถ้าคำศัพท์ที่อยู่หลัง PREP นั้นไม่ใช่ PROP หรือ NOUN ให้ถือว่าเป็นส่วนขยาย (MOD) ของ DOB เช่นประโยค “เขาเป็นคนที่ดี”



#### 4.4.2 โครงข่ายความหมายและการแทนรูปภาษากลางของประโยค



รูปที่ 4.25 แสดงรูปแบบโครงข่ายทางความหมายของประโยค

สำหรับการแทนโครงข่ายทางความหมายด้วยภาษากลางนั้น ทำโดยการการนำข้อมูลของคำศัพท์มาแทนคำศัพท์ภาษาไทย ซึ่งในโครงการนี้ได้เสนอวิธีการง่ายๆดังต่อไปนี้

VERB(value) := { SUB, DOB, LOC, MOD, LINK }

SUB(value) := { Num, Gm, Attrib }

SUB.Num := { ( NUM(value) + CLAS(value) ) }

SUB.Gm := { [PROP(value)], [NOUN(value)], [PRON(value)] }

SUB.Attrib := { [ADJ(value)], [DET(value)], [(CLAS(value)+DET(value))] }

DOB(value) := { Num, Gm, Attrib }

DOB.Num := { ( NUM(value) + CLAS(value) ) }

DOB.Gm := { [PROP(value)], [NOUN(value)], [PRON(value)] }

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

DOB.Attrib := {[ADJ(value)],[DET(value)],[(CLAS(value)+DET(value))]}  
 LOC(value) := {Num, Gm, Attrib}  
 LOC.Num := { ( NUM(value) + CLAS(value) ) }  
 LOC.Gm := { [PROP(value)],[NOUN(value)],[PRON(value)] }  
 LOC.Attrib := {[ADJ(value)],[DET(value)],[(CLAS(value)+DET(value))]}  
 MOD(value) := { [ADJ(value)], [ADV(value)] }  
 LINK(value) := {ประโยชน์}



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
 ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 5

### ซอฟต์แวร์และการออกแบบระบบวิเคราะห์ภาษา

#### 5.1 บทนำ

การออกแบบและการเขียนโปรแกรมของระบบใดก็ตาม เราจะต้องมีการกำหนดสิ่งต่างๆที่ซอฟต์แวร์หรือระบบควรมี (specification) เช่น ข้อมูลนำเข้า (input) ข้อมูลที่เป็นผลลัพธ์ของระบบ (output) ว่าควรมีรูปแบบอย่างไร และกระบวนการทำงานแต่ละส่วนทำหน้าที่อะไร ทำอย่างไร ติดต่อกันระหว่างกระบวนการอย่างไร เป็นต้น

ระบบวิเคราะห์ภาษาไทยก็เช่นเดียวกัน เราจะได้้นำความรู้และทฤษฎีต่างๆที่ได้กล่าวมาแล้วในส่วนประกอบแต่ละส่วนในบทที่ 4 เหล่านั้นมาออกแบบและสร้างเป็นโปรแกรมคอมพิวเตอร์ขึ้น โดยมีลำดับการทำงานจากกระบวนการแยกหน่วยคำศัพท์จากประโยคอินพุต และจะได้ผลลัพธ์เป็นคำศัพท์ที่แยกเป็นคำๆ จากนั้นจึงนำคำศัพท์ที่แยกแล้วนั้นผ่านเข้าเป็นอินพุตของกระบวนการตรวจสอบโครงสร้างภาษาไทยโดยใช้โครงข่าย M-ATN ว่าเป็นประโยคที่ถูกต้องตามหลักไวยากรณ์หรือไม่ และทำการกำหนดความสัมพันธ์ของคำศัพท์เหล่านั้น และขั้นตอนสุดท้ายเราจะนำคำศัพท์และความสัมพันธ์ของคำศัพท์เหล่านั้นมาแปลหรือแทนในรูปของภาษากลาง ซึ่งในบทนี้จะได้อธิบายถึงการออกแบบซอฟต์แวร์ส่วนต่างๆของระบบ

#### 5.2 การออกแบบระบบวิเคราะห์ภาษาไทย

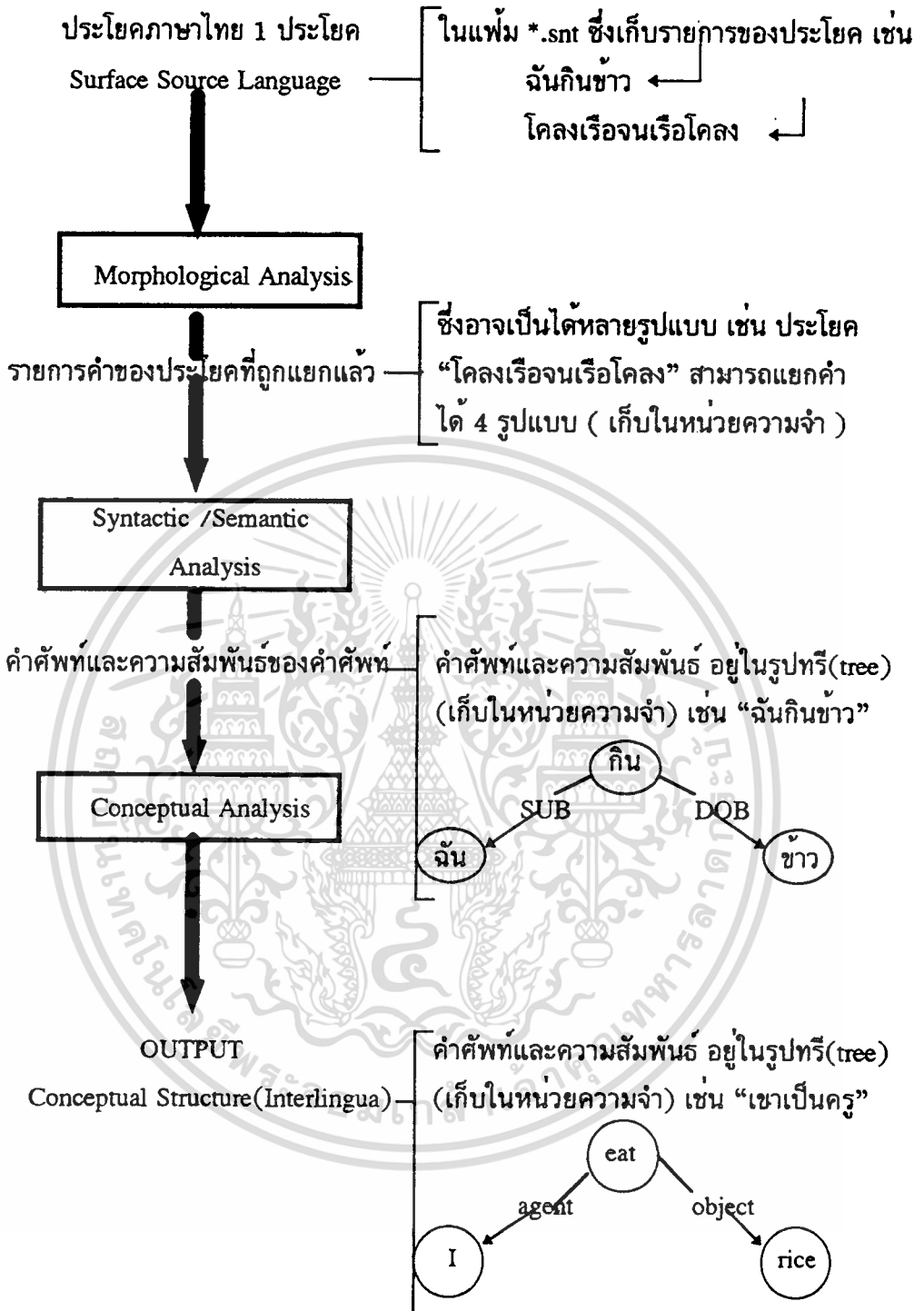
จากที่ได้กล่าวมาแล้ว ระบบวิเคราะห์ภาษานั้นได้แบ่งการทำงานออกเป็น 3 ส่วนใหญ่ๆดังนี้

- 1.) กระบวนการแยกคำศัพท์
- 2.) การวิเคราะห์โครงสร้างและความหมาย
- 3.) การแทนรูปภาษากลาง และ

นอกจากนี้ยังมีอีกส่วนหนึ่งที่สำคัญกับทุกๆส่วนนั่นคือ พจนานุกรม ที่จะให้ข้อมูลแก่ระบบทั้งหมด

#### โครงสร้างและส่วนประกอบของระบบแสดงผังรูปที่ 5.1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะวิธีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 5.1 แสดงส่วนประกอบต่างๆและข้อมูลเข้าออกของแต่ละส่วน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 5.8 การออกแบบพจนานุกรม

พจนานุกรมจะเป็นฐานข้อมูลที่เกี่ยวข้องเกี่ยวกับคำศัพท์ภาษาไทยทั้งหมด ที่จำเป็นอย่างอื่นในการวิเคราะห์ภาษาไทย ข้อมูลของคำศัพท์เหล่านี้จะถูกใช้ในตั้งแต่กระบวนการแยกคำ การวิเคราะห์โครงสร้าง และการวิเคราะห์เพื่อกำหนดความสัมพันธ์ทางความหมาย จนถึงการแทนความสัมพันธ์ของคำศัพท์ด้วยภาษากลาง เช่น การแยกคำ จะใช้คำศัพท์ต่างๆในการทำงาน การวิเคราะห์โครงสร้าง จะใช้คำศัพท์ ชนิด และข้อมูลอื่นๆที่เกี่ยวข้อง เช่น คำศัพท์ที่ใช้แทนภาษากลางซึ่งในที่นี้เราจะใช้เป็นภาษาอังกฤษ (Conceptual Word) เช่น “ฉัน” จะใช้คำว่า “I” นอกจากนี้มีข้อมูลอีกที่จะใช้ในการวิเคราะห์ภาษาคือ A Kind Of (AKO) ซึ่งเป็นข้อมูลที่จะบอกว่าคำศัพท์คำนั้น โดยความหมายแล้วมีลักษณะอย่างไร เช่น “คน” คือสิ่งมีชีวิต เป็นต้น

#### 5.8.1 ข้อมูลต่างๆของคำศัพท์ในพจนานุกรม

- 1.) TKEY : คำศัพท์ภาษาไทย
  - 2.) CP (Conceptual Word) : คำศัพท์ภาษาอังกฤษที่ใช้แทนภาษากลาง
  - 3.) CAT (Cat) : ชนิดของคำศัพท์ ซึ่งมี 16 ชนิดดังที่ได้อธิบายมาแล้ว เช่น NOUN VERB PRON PREP LINK เป็นต้น
  - 4.) A Kind Of (AKO) : บอกว่าโดยความหมายแล้วคำศัพท์นั้นมีลักษณะอย่างไร เช่น 111 Person : เป็นบุคคล เช่น เขา เพื่อน คู่รัก ผู้ปกครอง
  - 5.) VP (Verb pattern) : รูปแบบของคำกริยา เช่นคำว่า “กิน” VPคือ “SUB VERB DOB” และได้แสดงไว้ในตารางที่ 5.1
  - 6.) CMPTYPE : คือ ตำแหน่งของคำบุรพบทที่ต้องเกิดกับคำกริยาในประโยค โดยปรากฏหลังคำกริยาตามลำดับในตำแหน่งที่ 2 และ 3 เช่น คำว่า “ทะเลาะ” ต้องมีคำบุรพบท “กับ” เกิดขึ้นด้วยเสมอเพื่อบอกผู้ร่วมทะเลาะ และปรากฏในตำแหน่งที่ 2 หลังคำกริยา จึงเป็น CMP2 เช่นประโยค “เขาทะเลาะกับน้อง”
 

1	2
---	---
- แต่กริยา “ซื้อ” จะต้องมีคำบุรพบท “ให้” เกิดหลังกรรมในตำแหน่งที่ 3 จึงเป็น CMP3 เช่นประโยค “แม่ ซื้อ ขนม ให้ ลูก”
- |   |   |   |
|---|---|---|
| 1 | 2 | 3 |
|---|---|---|
- 7.) VALUE OF CMP : คือค่าของ CMP ซึ่งเป็นคำศัพท์ภาษาไทยชนิดบุรพบท เช่น “กับ”

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 5.3.2 ระบบแฟ้มข้อมูลของพจนานุกรม

แฟ้มสำหรับเก็บข้อมูลคำศัพท์ในพจนานุกรมนั้นมีอยู่ด้วยกัน 2 แฟ้มคือ

#### 1.) แฟ้มดัชนี ( DATA.NDX )

เก็บข้อมูลชนิดตัวเลขขนาด 4 ไบท์ ( Byte ) ในแต่ละเรคคอร์ดเพื่อเป็น

ตัวชี้ไปยังคำศัพท์ตามดัชนีทั้ง 3 โดยที่ ดัชนีที่ 1 มี 44 คำ

ดัชนีที่ 2 มี 68 คำ

ดัชนีที่ 3 มี 48 คำ ( 2 - 49 )

และพื้นที่สงวนไว้สำหรับระบบ 19 ไบท์

ดังนั้นแฟ้ม DATA . NDX จึงมีขนาดเท่ากับ  $44*68*49*4+19 = 574,483$

#### 2.) แฟ้มคำศัพท์ ( DATA.DIC )

เก็บคำศัพท์ และข้อมูลต่างๆของแต่ละคำ ดังที่ได้อธิบายไว้แล้วในหัวข้อ

##### 5.3.1 ดังนี้

ลำดับ	ชื่อ Filed	ความหมาย	ชนิด	ขนาด(ไบท์)	ตัวอย่าง
1	TKEY	คำศัพท์	อักษร	50	“ฉัน”, “วิ่ง”
2	CP	Conceptual Word	อักษร	50	“1”, “RUN”
3	CAT	ชนิดของคำศัพท์	อักษร	5	NOUN,PRON
4	AKO	A Kind Of	อักษร	5	13121:ทอง,หิน,เงิน
5	VP	Verb Pattern	อักษร	25	“SUB VERB DOB”
6	CMP_TYPE	Type of CMP	ตัวเลข	1	0: ไม่มี, 1: CMP2, 2: CMP3
7	VALUE_CMP	Value of CMP	อักษร	50	“กับ”, “ให้”
8	REPEAT	ชี้คำที่ TKEY เหมือนกัน	ตัวเลข	4	3, 78, 45879
9	NEXT	ชี้คำต่อไปในกลุ่ม เดียวกัน	ตัวเลข	4	3, 78, 45879

ตารางที่ 5.1 แสดงข้อมูลของคำศัพท์ในพจนานุกรม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ดัชนีที่ 1	ดัชนีที่ 2	ดัชนีที่ 3	คำศัพท์&ข้อมูล	REPEAT	NEXT
ฉ	1	49			
		48			
		3			
		2			
		49			
	๒	48			
			ฉน/PRON/I		
		3			
		2			
		49			
		48	ฉน/VERB/EAT		

รูปที่ 5.3 แสดงตัวอย่างเพิ่มข้อมูลคำศัพท์ “ฉน”

#### 5.4 การออกแบบกระบวนการแยกคำศัพท์

##### 5.4.1 อินพุทและเอาต์พุทของกระบวนการแยกคำศัพท์

อินพุทของกระบวนการทำงานนี้อยู่ในรูปแบบของประโยค เช่น

“ฉนกินข้าว”, “นายแดงชกนายดำ” เป็นต้น

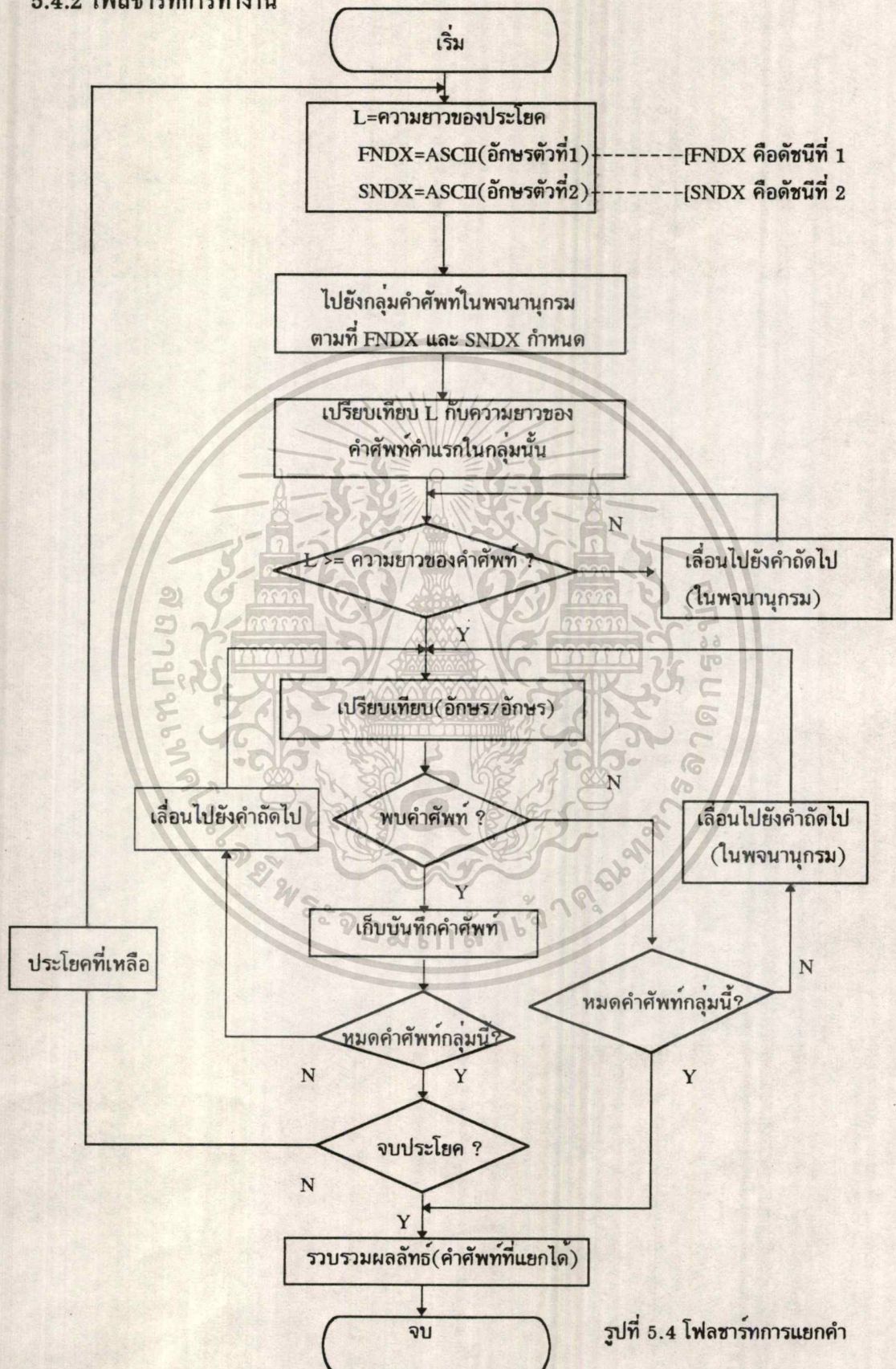
และประโยคเหล่านี้อาจจะถูกเก็บในแฟ้มข้อมูลที่มีส่วนขยายเป็น SNT (Sentence) เช่น Sent1.snt หรืออาจจะป้อนประโยคภาษาไทยจากคีย์บอร์ดในขณะที่ทำงาน

ส่วนเอาต์พุท ซึ่งถูกเก็บในหน่วยความจำ (Memory) เก็บเป็นเอเรย์ของรูปแบบของการแยกคำศัพท์ เช่น

Mark[1]	{Word1/Word2/Word3/.....Wordn}
Mark[2]	{Word1/Word2/Word3/.....Wordn}
Mark[n]	{Word1/Word2/Word3/.....Wordn}

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่ออนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

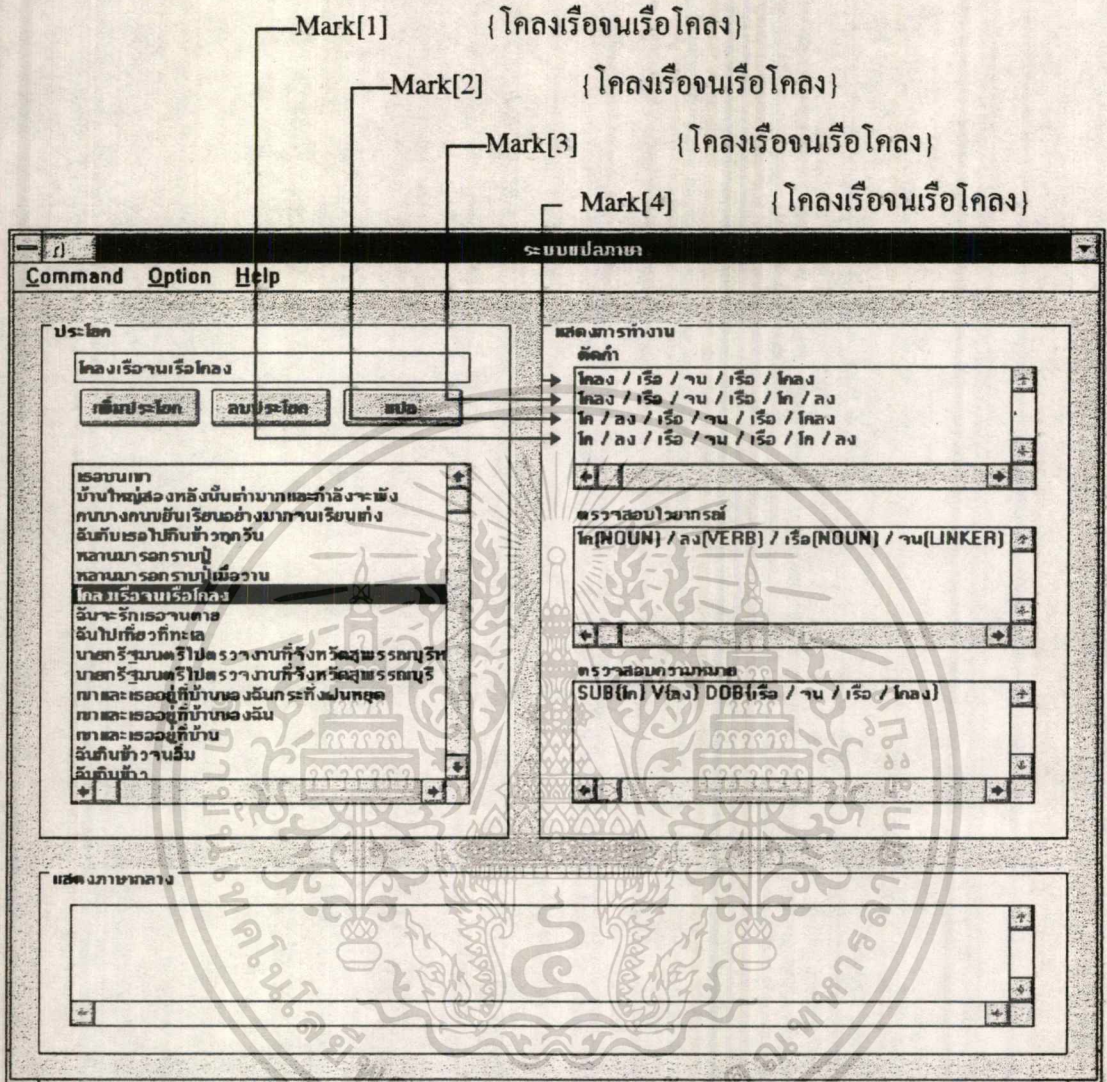
## 5.4.2 โพลีชาร์ทการทำงาน



รูปที่ 5.4 โพลีชาร์ทการแยกคำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.4.3 ตัวอย่างการทำงานของกระบวนกรแยกคำ



รูปที่ 5.5 แสดงตัวอย่างการทำงานของกระบวนกรแยกคำ

5.5 การออกแบบกระบวนกรวิเคราะห์โครงสร้างประโยค

5.5.1 อินพุตและเอาต์พุตของกระบวนกรวิเคราะห์โครงสร้างประโยค

ข้อมูลอินพุตของกระบวนกรนี้คือ เอาต์พุตของกระบวนกรแยกคำศัพท์ดังกล่าวแล้ว เช่น

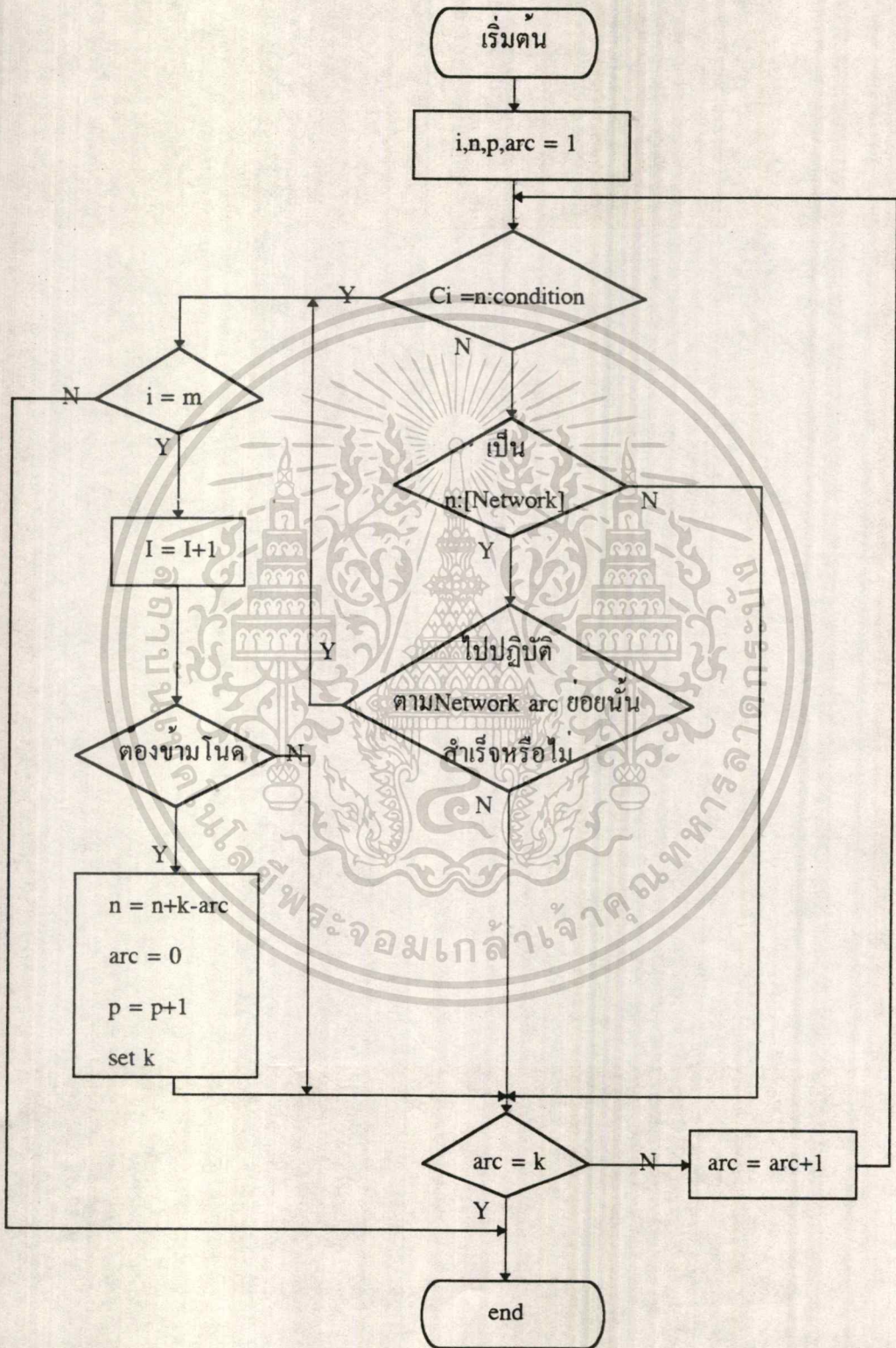
Mark[n] {Word1/ Word2/ Word3/ .....Word n}

ส่วนเอาต์พุตของระบบนี้คือ ประโยคที่ถูกตัดตามหลักไวยากรณ์ ซึ่งจะถูกเก็บอยู่ในหน่วยความจำของระบบ เช่น

โค[NOUN]/ ลง[VERB]/ เรือ[NOUN]/ วน[LINK]/ เรือ[NOUN]/ โคลง[VERB]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 5.5.2 โฟลว์ชาร์ทการทำงานของระบบวิเคราะห์โครงสร้างประโยค

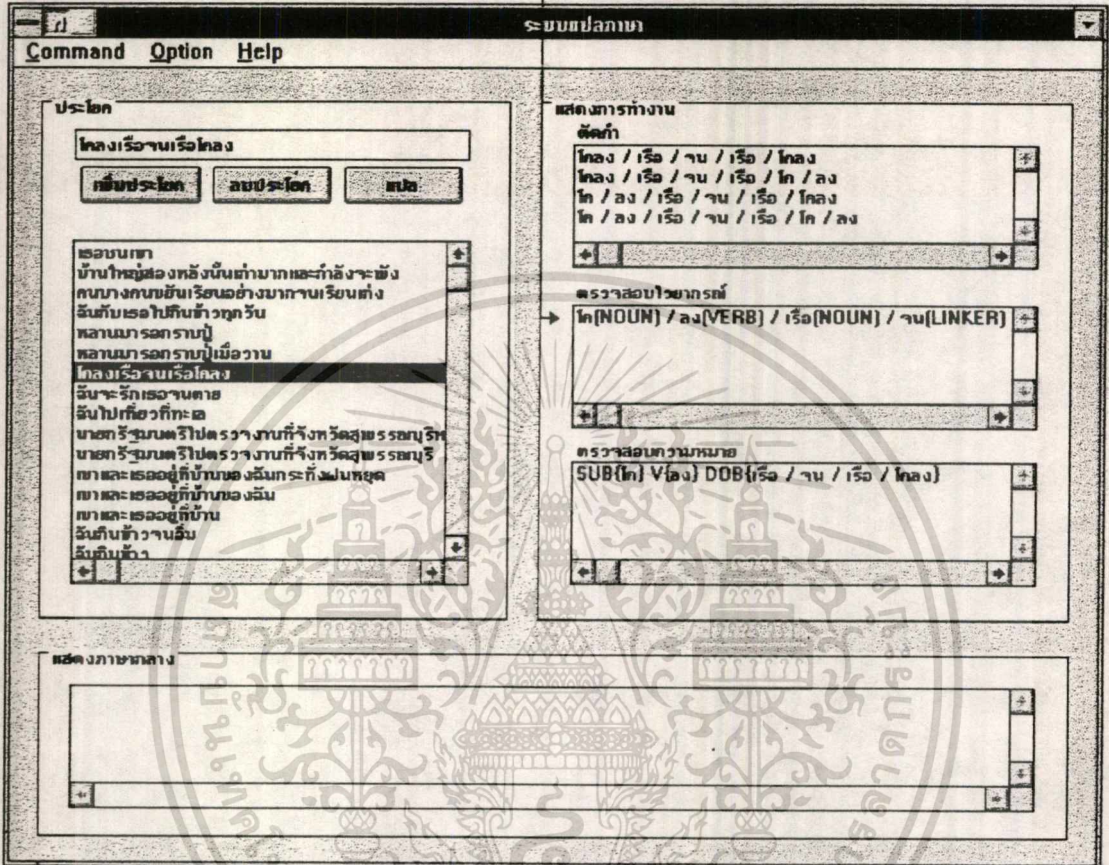


รูปที่ 5.6 แสดงโฟลว์ชาร์ทของกระบวนการวิเคราะห์โครงสร้างประโยค

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 5.5.3 ตัวอย่างการทำงานของกระบวนการวิเคราะห์ประโยค

#### ผลลัพธ์ของการวิเคราะห์ประโยค (โคลงเรือจนเรือโคลง)



รูปที่ 5.7 ตัวอย่างการทำงานของกระบวนการวิเคราะห์ประโยค

### 5.6 เพิ่มข้อมูลต่างๆในระบบ

- 1.) DATA.DIC : เก็บข้อมูลของคำศัพท์ภาษาไทยในพจนานุกรม
- 2.) DATA.NDX : เก็บข้อมูลตำแหน่งของคำศัพท์ เพื่อการเข้าถึงอย่างรวดเร็ว
- 3.) DATA.KEY : เก็บคำศัพท์ เพื่อช่วยการคำนวณกับพจนานุกรมให้ดีขึ้น
- 4.) \*.SNT : เก็บรายการประโยคอินพุทของระบบ
- 5.) MT.EXE : เป็นโปรแกรมระบบวิเคราะห์ภาษา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 6

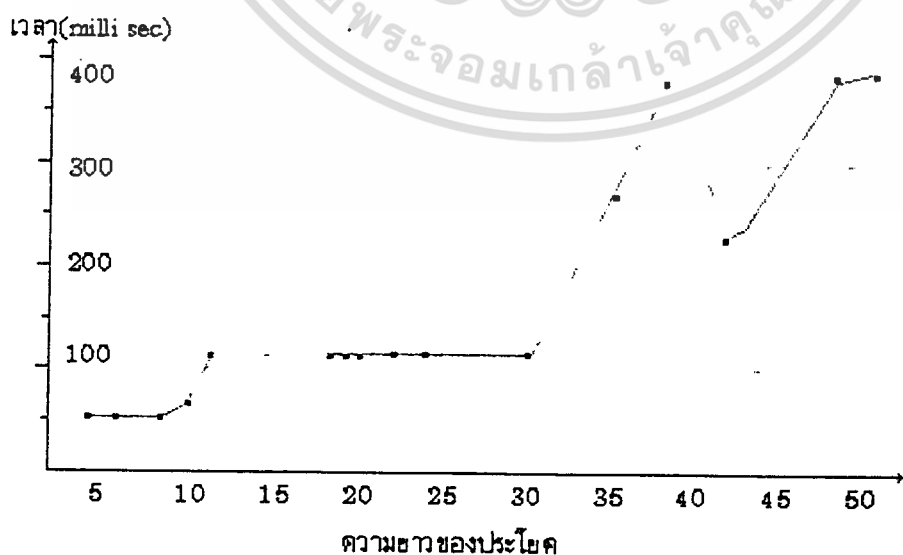
### บทสรุปและการประเมินผล

#### 6.1 บทนำ

ระบบการวิเคราะห์ภาษาไทย ได้แยกการทำงานออกเป็นส่วนๆ ดังที่ได้กล่าวมาแล้วนั้น แต่ส่วนรับอินพุต และให้ผลลัพธ์ที่แตกต่างกัน เช่น กระบวนการแยกคำศัพท์รับอินพุตเป็นประโยคภาษาไทย และให้อำนาจในรูปแบบของคำศัพท์ที่แยกเป็นคำๆ เป็นต้น

#### 6.2 การทดลองแยกคำ

จากประโยคภาษาไทยประมาณ 100 ประโยค คำศัพท์ในพจนานุกรมประมาณ 1000 คำ และได้ทดลองแยกคำศัพท์จากประโยค โดยใช้เครื่องไมโครคอมพิวเตอร์รุ่น 486DX4 ความเร็ว 100 เมกเฮิร์ต (Mhz) ซึ่งได้ผลสรุปดังกราฟในรูปที่ 6.1 โดยใช้ประโยคที่มีความยาวตั้งแต่ 2-50 และกำหนดให้แกน X (แนวนอน) แทนความยาวของประโยค ส่วนแกน Y (แนวตั้ง) แทนเวลาที่ใช้ในการประมวลผล



รูปที่ 6.1 แสดงกราฟผลลัพธ์การแยกคำศัพท์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากข้อมูลในรูปที่ 6.1 จะเห็นว่าเมื่อความยาวของประโยคเพิ่มขึ้น เวลาที่ใช้ในการแยก คำศัพท์ก็จะมากขึ้นตามไปด้วย แต่จะมีบางช่วงที่มีความยาวของประโยคมากขึ้น แต่เวลาที่ใช้กลับ น้อยลง เป็นเพราะคำศัพท์ในช่วงนั้นมีน้อยกว่า หรือคำศัพท์อาจจะอยู่ที่ช่วงต้น ๆ ของกลุ่มคำศัพท์ เหล่านั้นก็ได้

### 6.3 การทดลองส่วนวิเคราะห์โครงสร้างประโยค

จากการทดลองทำการวิเคราะห์โครงสร้างประโยคพบว่า การประมวลผลด้วยโครงข่าย M-ATN นั้นยังไม่สามารถควบคุมประโยคในภาษาไทยได้ทั้งหมด รูปแบบประโยคที่สามารถ ประมวลผลได้นั้น มีลักษณะดังนี้

ประธาน                      กริยา                      กรรม

ประธาน และกรรมในที่นี้สามารถเป็น นามวลี ส่วน กริยา ก็สามารเป็น กริยาวลีได้เช่น กัน คือมีส่วนขยายได้ เช่น

กริยา	กิน,	เดิน
ขยายกริยา	จะ,	กำลัง
กริยาวลี	จะกิน, กำลังเดิน	เดินเร็ว เป็นต้น

ประโยคตัวอย่าง เช่น

ฉัน	กิน	ข้าว
ฉันกับเธอ	กิน	ข้าว
เขา	กำลังกิน	ข้าว
นายแดง	วิ่งเร็ว	กว่า นายดำ
ภรรยาของเขา	เป็น	ครู
คอมพิวเตอร์	ทำ	งานเร็ว แต่ คน ทำ งานช้า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 6.4 การประยุกต์ใช้งานระบบวิเคราะห์ภาษาไทย

### 6.4.1 การแปลภาษาด้วยคอมพิวเตอร์ ( Machine Translation )

ในการแปลภาษาไทยด้วยคอมพิวเตอร์นั้น เราจะต้องมีระบบการแยกคำศัพท์จากประโยค ระบบการวิเคราะห์โครงสร้าง และความหมาย และระบบการแทนรูปภาพกลาง รวมกันเป็นส่วนประกอบส่วนแรกคือส่วนวิเคราะห์ภาษา (Analysis) ซึ่งสามารถใช้ความรู้ และทฤษฎีต่างๆ ในโครงการนี้ได้

### 6.4.2 การตรวจสอบตัวสะกด หรือคำผิดในงานจดจำอักษร

ในงานจดจำรูปแบบตัวอักษร จะใช้การแปลงข้อมูลตัวอักษรในรูปของกราฟฟิก เป็นข้อมูลของตัวอักษรในรูปของไฟล์ข้อความ ซึ่งหากการแปลงข้อมูลดังกล่าวเกิดการผิดพลาด ข้อมูลผลลัพธ์ไม่สมบูรณ์ ไม่ว่าจะเกิดจากสาเหตุใดก็ตาม ข้อมูลในรูปของไฟล์ข้อความก็จะเป็นคำศัพท์ที่ผิดด้วย ดังนั้นเราจึงสามารถนำระบบแยกคำศัพท์มาใช้ในการตรวจสอบคำผิดได้

## 6.5 แนวทางการพัฒนาระบบวิเคราะห์ภาษาไทยด้วยคอมพิวเตอร์

ระบบวิเคราะห์ภาษาในโครงการนี้นั้น สามารถวิเคราะห์ได้ทีละประโยค และรับข้อมูล (ประโยค) ได้โดยการป้อนจากคีย์บอร์ด และอ่านจากไฟล์ \*.SNT ซึ่งสามารถบันทึกประโยคเหล่านั้นได้ที่โปรแกรมหลักของระบบ และไฟล์ดังกล่าวยังเป็นรูปแบบของตัวเองเท่านั้น

ในการพัฒนาระบบวิเคราะห์ภาษาไทยนี้ จึงควรจะให้สามารถรับประโยคอิพุทได้จากหลายแหล่ง เช่น ไฟล์ \*.DOC, \*.CW เป็นต้น และยังสามารถประมวลผลได้ทีละหลายๆ ประโยคอีกด้วย โดยการโหลดข้อมูลประโยคอินพุทจากไฟล์ข้อมูลดังกล่าวเข้ามายังหน่วยความจำ ทีละหลายๆ ประโยค โดยเราอาจจะโหลดทีละหนึ่งย่อหน้า แล้วทำการวิเคราะห์ทีละประโยค ซึ่งสามารถอาศัยข้อมูลจากประโยคข้างเคียงเพื่อช่วยในการวิเคราะห์ด้วย

## 6.6 บทสรุป

ถึงแม้โครงการนี้จะไม่สามารถแก้ปัญหาของการประมวลผลภาษาไทยในงานด้านการแปลภาษาด้วยคอมพิวเตอร์ได้ทั้งหมด แต่ก็สามารถที่จะแก้ปัญหาได้บางส่วน เพราะการทำโครงการนี้ต้องการความรู้ทางด้านภาษาศาสตร์อย่างมากในการหาวิธีการที่จะประมวลผลให้ถูกต้องมากที่สุด อีกทั้งการประมวลผลไวยากรณ์และความหมายของภาษาโดยโปรแกรมคอมพิวเตอร์นั้นมีข้อจำกัดอยู่มาก โดยเฉพาะภาษาไทย ซึ่งเป็นภาษาที่หากฎเกณฑ์หรือสูตรที่แน่นอนสำหรับการโปรแกรมภาษาคอมพิวเตอร์ในการประมวลผลได้ยากมาก แต่อย่างน้อยที่สุด แนวทางในการทำโครงการนี้น่าจะให้แนวความคิดที่ดี และความรู้สำหรับการพัฒนาวิธีการประมวลผลภาษาไทยในงานด้านการแปลภาษา หรือด้านอื่น ๆต่อไปอีกในอนาคต



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## กิตติกรรมประกาศ

ความสำเร็จของการทำโครงการนี้ เนื่องจากการได้รับความช่วยเหลือและสนับสนุนจากหลายๆฝ่ายด้วยกัน ขอขอบพระคุณอาจารย์ประสาร ตั้งติสานนท์ อาจารย์ที่ปรึกษา ผู้ซึ่งให้ความอนุเคราะห์ทั้งความรู้ คำปรึกษา และตำราต่างๆ ขอขอบคุณคุณสมศักดิ์ จันวัน เจ้าหน้าที่สำนักวิจัยฯ สำหรับแนวความคิด ตำรา ตลอดจนความรู้ต่างๆเกี่ยวกับระบบวิเคราะห์ภาษาไทย ขอขอบคุณศักดิ์คาและฝั่งที่ให้อืมเครื่องพิมพ์ ขอขอบคุณพงที่ช่วยพิมพ์ Abstract และขอบคุณเพื่อนๆทุกคนสำหรับความช่วยเหลือและกำลังใจที่มีให้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## หนังสืออ้างอิง

1. นิตยา กาญจนะวรรณ, รศ ดร., “การแปลภาษาด้วยเครื่อง ทฤษฎีและวิธีการ”, ภาควิชาภาษาไทย และภาษาตะวันออก คณะมนุษยศาสตร์ มหาวิทยาลัยรามคำแหง, 488 หน้า, 2534
2. นิตยา กาญจนะวรรณ, รศ ดร., “การศึกษาโครงสร้างภาษาไทยเพื่อใช้ในระบบการแปลภาษา ด้วยเครื่องคอมพิวเตอร์”, คณะมนุษยศาสตร์ มหาวิทยาลัยรามคำแหง, 205 หน้า, 2529
3. พรทิพย์ พุกผาสุก, “การศึกษาการละกริยาในภาษาพูดในกรุงเทพฯตามแนววรรณคดี”, ภาควิชาภาษาไทย บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 231 หน้า, 2524
4. ภาณุ สังขวร, “ความสัมพันธ์ทางอรรถศาสตร์ระหว่างคำนามกับคำกริยาในประโยคภาษาไทย”, ภาควิชาภาษาไทย จุฬาลงกรณ์มหาวิทยาลัย, 128 หน้า, 2536
5. ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ, “การแปลภาษาด้วยคอมพิวเตอร์”, สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ กระทรวงวิทยาศาสตร์เทคโนโลยี และสิ่งแวดล้อม, 124 หน้า, 2536
6. สมศักดิ์ จันวัน, “ระบบวิเคราะห์ภาษาไทยด้วยคอมพิวเตอร์”, บัณฑิตวิทยาลัย, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 196 หน้า, 2534

## ภาคผนวก ก

## แบบโครงสร้างของประโยคโดยใช้คำกริยา (Verb Pattern)

เป็นการกำหนดแบบโครงสร้างของประโยคโดยใช้คำกริยาเป็นหลัก แบบโครงสร้างนี้จะแสดงว่าประโยคแต่ละแบบประกอบด้วยคำซึ่งมีหน้าที่ไวยากรณ์ และมีความสัมพันธ์ระหว่าง concept เช่นใด ซึ่งรวมทั้ง concept ที่เป็นคำนามกับคำนาม กริยากับคำนาม และอื่นๆและสำหรับแบบโครงสร้างของประโยคนั้น คำกริยาจะเป็นคำหลัก(head concept) และคำที่นำหน้าและตามหลังกริยา จะมีความสัมพันธ์อย่างไรกับคำกริยา โดยใช้การอธิบายความสัมพันธ์การก (case relation) เทียบไว้กับความสัมพันธ์ทางไวยากรณ์ ซึ่งได้กำหนดไว้ดังนี้

แบบที่	Verb Pattern	ตัวอย่าง	Tmaps
1	SUB V	เขาเดิน	SUB=OBJ
2	V DOB	มีดอกไม้บนโต๊ะ	DOB=OBJ
3	SUB V AUX	ลูกหมาเกิดแล้ว	SUB=OBJ
4	SUB V ADP	เขาลูกขึ้นแล้ว	SUB=OBJ
5	SUB V CMP2	นำประกอบควยไฮโครเจน และออกซิเจน	SUB=OBJ, CMP2=ELM
6	SUB V DOB CMP3	เขาซื้อขนมไหลูก	SUB=AGT, DOB=OBJ, CMP3=TAR
7	SUB V DOB	เขากินข้าวเหนียว	SUB=AGT, DOB=OBJ
8	SUB V DOB ADP	เขาเก็บอาหารไว้ในตู้	SUB=AGT, DOB=OBJ
9	SUB V DOB IOB	เขาแจกขนมเด็ก	SUB=AGT, DOB=OBJ, IOB=TAR

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## ภาคผนวก ข

## AKO (A Kind Of)

AKO (A Kind Of) เป็นข้อมูลที่ใช้จัดความหมายของคำที่มีอยู่ในภาษาให้เป็นหมวดหมู่ เพื่อเป็นประโยชน์ในการวิเคราะห์ความหมายของประโยค หรือเป็นการช่วยลดความกำกวมที่อาจเกิดขึ้นได้ในประโยค ดังตัวอย่างเช่น "ตากลับบ้าน" ประโยคนี้จะเกิดความกำกวมขึ้นที่คำนาม "ตา" ซึ่งเครื่องคอมพิวเตอร์ไม่สามารถบอกได้ว่า คำว่า "ตา" จะมีความหมายว่า "grandfather" หรือ "eye" จึงต้องอาศัยคำที่อยู่ใกล้เคียงมาช่วยในการตัดสินใจ ซึ่งก็คือคำกริยา "กลับ" จะพบว่า ข้อมูลคำกริยาตัวนี้ ต้องการประธานที่เป็นสิ่งมีชีวิต ซึ่งในที่นี้ก็คือ "ตา" ที่มีความหมายว่า "grandfather" ดังนั้น จึงจำเป็นต้องมีการแบ่งกลุ่มทางความหมายให้กับคำโดยจัดหมวดหมู่ของความหมายเป็นลำดับขั้น จากกลุ่มที่ใหญ่จนถึงกลุ่มย่อยซึ่งมีดังต่อไปนี้

## 1 Concrete : รูปธรรม แบ่งเป็น

11 Subject : บุคคลและกลุ่มที่ประกอบด้วยคน แบ่งเป็น

111 Person : บุคคล เช่น ผู้ปกครอง เขา ผู้ดูแล คู่รัก เพื่อน

112 Organization : กลุ่มหรือองค์กรต่างๆ เช่น บริษัท กรรมการ กลุ่ม คณะ

12 Concrete place : สถานที่ แบ่งเป็น

121 Region : สถานที่ที่เป็นดินแดนหรือมีขอบเขตที่แน่นอน เช่น จังหวัด อำเภอ หมู่บ้าน

122 Nature : สถานที่ที่เป็นธรรมชาติ แบ่งเป็น

1221 Topography : สถานที่ที่เป็นภูมิประเทศ เช่น ทุ่ง เทือกเขา ชายทะเล

1222 Astronomy : สถานที่ที่เกี่ยวข้องกับดาราศาสตร์ เช่น โลก พระอาทิตย์

123 Man-made place : สถานที่ที่มนุษย์สร้างขึ้น เช่น ทางด่วน ดิถ บ้าน

13 Concrete thing : สิ่งของ แบ่งเป็น

131 Nature thing : สิ่งของที่เกิดตามธรรมชาติ แบ่งเป็น

1311 Living thing : สิ่งมีชีวิต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

13111 Animal : สัตว์ เช่น หนอน หมา แมว

13112 A part of animal or its derivative : ส่วนของสิ่งมีชีวิต หรือ สิ่งที่ได้จากสิ่งมีชีวิต เช่น มือ เท้า แขน ไข่น้ำตา

13113 Plant : พืช เช่น ดอกไม้ มะนาว กุหลาบ

13114 A part of plant or its derivative : ส่วนของพืชหรือ สิ่งที่ได้จากพืช เช่น กิ่งไม้ ราก เปลือก

1312 Substance : สสาร แบ่งเป็น

13121 Solid : ของแข็ง เช่น ทอง เงิน เพชร

13122 Liquid : ของเหลว เช่น น้ำ ปรัชมน้ำมัน

13123 Gas : ก๊าซ เช่น อากาศ ไอน้ำ ก๊าซธรรมชาติ

132 Man-made thing : สิ่งของที่มนุษย์สร้างขึ้น แบ่งเป็น

1321 Material : วัตถุดิบ เช่น พลาสติก ผ้า

1322 Finished product : ผลิตภัณฑ์ เช่น โต๊ะ ดวงไฟ รถยนต์

1323 A part of product or its derivative : ส่วนของผลิตภัณฑ์หรือ ผลที่ได้จากผลิตภัณฑ์ เช่น ขาโต๊ะ ขี้อยางลบ ใบมีด

2 Abstract : นามธรรม แบ่งออกเป็น

21 Abstract matter : นามธรรมที่ครอบคลุมเหตุการณ์และกริยาอาการต่างๆ แบ่งเป็น

211 Activity : กริยาและอาการต่างๆ แบ่งออกเป็น

2111 Action : การกระทำ เช่น ปฏิกริยา เล่น เดินทาง ไข้

2112 Mental action : การกระทำที่เกี่ยวกับความรู้สึกนึกคิด แบ่งเป็น

21121 Rational action : การกระทำที่ต้องใช้เหตุผลและความคิด เช่น คิด ตัดสินใจ ควบคุม

21122 Emotional action : การกระทำที่เกี่ยวข้องกับความรู้สึก เช่น รัก หวัง โกรธ เสียใจ

212 Phenomenon : ปรากฏการณ์ แบ่งออกเป็น

2121 Event : เหตุการณ์ เช่น ดวง โชค ความสำเร็จ

2122 Natural phenomenon : ปรากฏการณ์ธรรมชาติ เช่น พายุฝน

สุริยุปราคา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2123 Social phenomenon : ปรากฏการณ์ทางสังคม เช่น การประท้วง สงคราม การจลาจล

2124 Physiological phenomenon : ปรากฏการณ์ทางสรีระหรือทางร่างกาย เช่น อาการป่วย การเกิด การตาย การย่อยอาหาร

22 Abstract thing : นามธรรมที่มีลักษณะคล้ายสิ่งที่สามารถจับต้องได้ แบ่งเป็น

221 Product of activity : ผลของกริยาและอาการต่างๆ แบ่งออกเป็น

2211 Result of action : ผลของการกระทำ เช่น คำอธิบาย แผนการณ้ขอเสนอ

2212 Result of mental action : ผลของการกระทำที่เกี่ยวข้องกับความรูสึกนึกคิด เช่น ความจำ ความเป็นห่วง ความรัก

2213 Abstract thing fixed as socially : สิ่งที่เป็นนามธรรมที่เกี่ยวข้องกับสังคม เช่น งานปีใหม่ งานลอยกระทง

2214 Abstract thing fixed as culture : สิ่งที่เป็นนามธรรมที่เกี่ยวข้องกับวัฒนธรรม เช่น ภาษาไทย รำวง มวยไทย

2215 Immaterial man-made thing : สิ่งที่ไม่ใช่วัตถุที่มนุษย์สร้างขึ้น เช่น วรรณคดี เพลง บทความ โปรแกรมคอมพิวเตอร์

23 Abstract relation : ความสัมพันธ์เชิงนามธรรม แบ่งออกเป็น

231 Place : สถานที่ เช่น ขวา ซ้าย บน แดงๆ

232 Time : เวลา เช่น เมื่อวานนี้ สัปดาห์นี้ สมัยธนบุรี

234 Set : กลุ่ม เช่น เป็นแถว เรียงราย เป็นกลุ่ม

235 Attribute : คุณสมบัติของวัตถุหรือสิ่งของ แบ่งเป็น

2351 Quality : คุณลักษณะ เช่น ลักษณะ อุปนิสัย

2352 Condition : สภาวะ เช่น สภาพร่างกาย สถานการณ์

2353 Appearance : ลักษณะที่ปรากฏให้เห็น เช่น รูปร่าง สี ลาย

2354 Sort/Type : ประเภทหรือชนิด เช่น แบบ ชนิด ประเภท เผ่า

2355 Number/Quantity : จำนวน เช่น ปริมาตร ความสูง

2356 Relation : ความสัมพันธ์ เช่น ความเกี่ยวข้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2357 Function : หน้าที่และประสิทธิภาพ เช่น ความสามารถ ประสิทธิภาพ  
ภาพ คุณภาพ

236 Value : ค่าของคุณสมบัติหรือสิ่งของ แบ่งเป็น

2361 Quality : ค่าที่แสดงลักษณะ เช่น ใจดี สูง ง่าย

2362 Condition : ค่าที่แสดงสภาวะ เช่น เป็นอันคราย ปลอดภัย

2363 Appearance : ค่าที่แสดงลักษณะที่ปรากฏ เช่น กลุ่ม สีเหลือง  
สว่าง แดง

2364 Sort/Type : ค่าที่แสดงชนิดหรือประเภท เช่น รุ่น 2503 แบบดิจิทัล

2365 Number/Quantity : ค่าที่แสดงจำนวนหรือปริมาณ เช่น 10 วัน  
มาก นาน

2366 Relation : ค่าที่แสดงความสัมพันธ์ เช่น โกล้เคียง คล้าย ติดกัน

2367 Function : ค่าที่แสดงหน้าที่และประสิทธิภาพ เช่น (ประสิทธิภาพ)  
ภาพสูง (คุณภาพ)เยี่ยม