สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

ระบบการวิเคราะห์ประโยคภาษาอังกฤษไปสู่ภาษากลาง เพื่อการแปลภาษาด้วยคอมพิวเตอร์

SYSTEM OF ANALYZING ENGLISH SENTENCES INTO INTERLINGUA FOR MACHINE TRANSLATION

หนังสืออ้างอิง
ห้ามนำออกนอกห้องสมุด
T 021206

นางสาวอรัญญา ปรีชาไว
MISS ARANYA PREECHAWAI

วิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิศวกรรมไฟฟ้า
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
พ.ศ. 2537

เลขหมู่
เลขทะเบียน 21206
วัน, เดือน, ปี - 3 ส.ค. 2537

ISBN 974-621-144-7

ลิขสิทธิ์ของบัณฑิตวิทยาลัย สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

SYSTEM OF ANALYZING ENGLISH SENTENCES INTO INTERLINGUA FOR MACHINE TRANSLATION

ARANYA PREECHAWAI

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIRMENTS FOR THE DEGREE OF MASTER OF ENGINEERING IN ELECTRICAL ENGINEERING GRADUATE SCHOOL

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

1994

ISBN 974-621-144-7

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อวิทยานิพนธ์	ระบบการวิเคราะห์ประโยคภาษาอังกฤษไปสู่ภาษากลาง เพื่อการแปลภาษาด้วยคอมพิวเตอร์
นักศึกษา	นางสาวอรัญญา ปรีชาไว
อาจารย์ผู้ควบคุมวิทยานิพนธ์	ผศ.ดร. รัตติกร วรากูลศิริพันธุ์
ระดับการศึกษา	วิศวกรรมศาสตรมหาบัณฑิต สาขาวิศวกรรมไฟฟ้า
ภาควิชา	อิเล็กทรอนิกส์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ปีการศึกษา	2537

บทคัดย่อ

ในการแปลภาษาด้วยระบบคอมพิวเตอร์นั้นมีหลายวิธี โดยใช้หลักการทางด้านการประมวลผลภาษาธรรมชาติและวิทยาการทางซอฟท์แวร์ ทั้งนี้วิธีการแปลโดยใช้ภาษากลางเป็นสื่อถ่ายทอดโครงสร้างทางไวยากรณ์ และความหมายของภาษาต้นแบบไปสู่การสร้างภาษาเป้าหมายนั้น เป็นวิธีการที่ให้ความถูกต้อง และมีประสิทธิภาพในการแปลสูง นอกจากนั้นยังเหมาะสำหรับการแปลหลายๆ ภาษาในระบบเดียวกัน เพราะจะลดจำนวนซอฟท์แวร์ จำนวนพจนานุกรม และขั้นตอนการถ่ายทอดให้น้อยลง

วิทยานิพนธ์นี้ได้เสนอแนวทางในการวิเคราะห์โครงสร้างประโยคภาษาอังกฤษเพื่อถ่ายทอดข้อมูลทางไวยากรณ์และความหมายไปสู่สื่อภาษากลางที่กำหนดโดยเริ่มต้นจากการวิเคราะห์หน่วยคำที่จะต้องมีขบวนการตัดส่วนเสริมของคำ และส่วนเสริมที่บอกกาลเวลา ซึ่งขั้นตอนเหล่านี้จะมีส่วนสัมพันธ์ไปถึงการออกแบบพจนานุกรมทางด้านรูปแบบ และข้อมูลภายในให้สมบูรณ์ ทั้งนี้เพื่อการถ่ายทอดทางความหมายของคำ ซึ่งเป็นส่วนสำคัญของงานวิจัยนี้ด้วย นอกจากนั้นเรายังได้ออกแบบระบบฐานข้อมูลทางไวยากรณ์ เพื่อแยกแยะโครงสร้างประโยคในการหาความสัมพันธ์ของคำในระดับแก่น โดยใช้หลักการประมวลผลภาษาธรรมชาติทางด้านทฤษฎีเคสของฟิลมอร์ และซอมสกีเป็นพื้นฐานการวิเคราะห์ จะมีการแก้ปัญหาความกำกวมทางภาษา เพื่อให้ได้ผลลัพธ์เป็นสื่อภาษากลางที่ต้องการ นอกจากนั้นจะกล่าวถึงการออกแบบและการใช้งานของซอฟท์แวร์ระบบวิเคราะห์โครงสร้างประโยคภาษาอังกฤษ ดังกล่าวด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Thesis Title	System of Analyzing English Sentences into Interlingua for Machine Translation
Student	Miss Aranya Preechawai
Thesis Adviser	Ass.Prof. Dr.Ruttikorn Varakulsiripunth
Level of Study	Master of Engineering in Electrical Engineering
Department	Electronics
	King Mongkut 's Institute of Technology Ladkrabang
Academic year	1994

Abstract

In the field of machine translation (MT) system, number of methods had been introduced so far. Most of methods are based on natural language processing and computer software techniques. Alternatively, the interlingua approach is generally used in most MT systems. The interlingua is used to be an intermediate expression in order to transfer the syntactic and semantic information of source language into target language generation. This methodology has been successfully recognized in both accuracy and efficiency. The interlingua approach is also suitable for multilingual MT system, because the amount of software and dictionaries, and also the transformation processes had been largely reduced.

In this thesis, we have proposed a new approach to analyze the English sentences structure. The goal of this approach is to represent the syntactic and semantic of English sentences in the well predefined interlingua expression. The analysis is started with morphological processing for parsing the prefixes, suffixes and tense's markers. This will relate to the designing of the dictionary format including conceptual meaning of each word. The structure of dictionary is one of the main objective of our research. Moreover, the knowledge base of syntactic information system is also created. This syntactic nowledge

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

base is designed based on the case theories introduced by Fillmore and Chomsky. Accordingly, an English sentence will be analyzed into interlingua in form of conceptual relation. The ambiguity that may occur is also considered and resolved in order to achieve a correct conceptual relation. Finally, we also describe the details of designing and implementation of this analysis system software.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้ สำเร็จลุล่วงได้ด้วยดี ผู้เขียนต้องขอขอบคุณทุกท่านดังนี้

- ผู้เขียนขอกราบขอบพระคุณ คุณพ่อ และคุณแม่ ผู้ซึ่งให้โอกาสในการเล่าเรียนรวมทั้งให้กำลังใจเสมอมา อันเป็นแรงบันดาลใจให้ผู้เขียนมีความพยายามและมานะในการทำการวิจัย และเขียนวิทยานิพนธ์ฉบับนี้

- ผศ.ดร. รัตติกร วรากูลศิริพันธุ์ อาจารย์ที่ปรึกษา ผู้เขียนต้องขอขอบพระคุณเป็นอย่างสูงที่กรุณาให้คำปรึกษา แนะแนวทางและให้ข้อคิดเห็นที่เป็นประโยชน์ต่อการทำการวิจัย ทั้งเอื้ออำนวยความสะดวกในการใช้อุปกรณ์ต่าง ๆ ที่ช่วยในการวิจัย รวมทั้งตรวจทานแก้ไขวิทยานิพนธ์

- ผศ. วรนุช เกิดสินธุ์ชัย อาจารย์ภาคคณิตศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้า ธนบุรี ผู้เขียนต้องขอขอบพระคุณเป็นอย่างสูงที่ได้ให้คำแนะนำและชี้แนวทางในการวิจัย รวมทั้งเจ้าหน้าที่วิจัยห้องปัญญาประดิษฐ์ สถาบันเทคโนโลยีพระจอมเกล้า ธนบุรี ทุกท่านที่ได้ให้คำปรึกษาในด้านภาษาศาสตร์ รวมทั้งได้ให้ข้อคิดและแนวทางในการทำวิจัย ผู้เขียนต้องขอขอบพระคุณเป็นอย่างสูง

ท้ายนี้ ผู้เขียนขอขอบพระคุณ

ผู้ช่วยศาสตราจารย์ ดร. รัตติกร วรากูลศิริพันธุ์

รองศาสตราจารย์ ดร. ชม กิ้มปาน

ดร. บุญธีร์ เครือตราชู

ดร. เอื้อน ปิ่นเงิน

ผู้ช่วยศาสตราจารย์ ดร. บุญวัฒน์ อัตชู

ที่ให้ความกรุณาเป็นกรรมการสอบวิทยานิพนธ์ในครั้งนี้

อรัญญา ปรีชาไว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

หน้าที่

บทคัดย่อภาษาไทย I

บทคัดย่อภาษาอังกฤษ II

กิตติกรรมประกาศ IV

สารบัญ V

สารบัญตาราง VIII

สารบัญรูป IX

บทที่ 1 บทนำ 1

1.1 ความเป็นมาของงานวิจัย 1

1.2 จุดประสงค์และขอบเขตของงานวิจัย 5

1.3 โครงร่างของวิทยานิพนธ์ 6

บทที่ 2 ฐานความรู้เกี่ยวกับไวยากรณ์ภาษาอังกฤษที่ใช้ในการวิเคราะห์ 7

2.1 บทนำ 7

2.2 การแบ่งประเภทของคำภาษาอังกฤษ 7

2.3 โครงสร้างทางไวยากรณ์ของประโยคภาษาอังกฤษ 12

บทที่ 3 การจัดเก็บฐานข้อมูลไวยากรณ์ในส่วนของพจนานุกรมคำศัพท์ 16

3.1 บทนำ 16

3.2 ส่วนข้อมูลทางโครงสร้างและความหมายของคำศัพท์ 17

3.2.1 ข้อมูลระดับหน่วยคำ 18

3.2.2 ข้อมูลระดับโครงสร้างของคำ 19

3.2.3 ข้อมูลระดับความหมายของคำ 35

3.3 ส่วนข้อมูลจำเพาะของคำศัพท์ประเภทคำกริยาที่มีการผันรูปไปตามกาลของประโยค 37

3.4 วิธีการค้นหาข้อมูลของคำศัพท์ในฐานข้อมูลพจนานุกรม 42

3.4.1 โครงสร้างการเก็บบันทึกข้อมูลคำศัพท์ในฐานข้อมูลของพจนานุกรม 43

3.4.2 ขบวนการค้นหาข้อมูลคำศัพท์ 47

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หน้าที่

บทที่ 4 โครงสร้างของภาษากลาง 54

4.1 บทนำ 54

4.2 หน่วยแสดงความหมายของคำ หรือ CP 55

4.3 การกำหนดส่วนแสดงความสัมพันธ์ของหน่วยความหมาย 56

4.3.1 ส่วนแสดงความสัมพันธ์ระหว่างหน่วยความหมายที่เกี่ยวกับวัตถุ 56

4.3.2 ส่วนแสดงความสัมพันธ์ระหว่างหน่วยความหมายภายในวัตถุ 59

4.4 ส่วนแสดงรูปแบบภาษากลางทางโครงสร้าง 61

4.5 ส่วนเสริมหน่วยความหมาย 62

4.5.1 ส่วนเสริมคำนาม 62

4.5.2 ส่วนเสริมคำกริยา 64

4.6 รูปแบบข้อมูลของภาษากลาง 67

4.7 รูปแบบภาษากลางจากตัวอย่างประโยคที่ผ่านขบวนการวิเคราะห์ 69

บทที่ 5 ขบวนการวิเคราะห์ประโยคภาษาอังกฤษ 71

5.1 บทนำ 71

5.2 ส่วนการวิเคราะห์ในระดับคำ 72

5.2.1 ข้อกำหนดของประโยคต้นแบบ 72

5.2.2 การประมวลผลของขบวนการวิเคราะห์ในระดับคำ 73

5.3 ส่วนการวิเคราะห์ในระดับวากยสัมพันธ์ 89

5.3.1 การออกแบบเครื่องมือของขบวนการวิเคราะห์ 89

5.3.2 การสร้างกฎไวยากรณ์ของขบวนการวิเคราะห์ 91

5.3.3 โครงสร้างเอาต์พุต 132

5.4 ส่วนการวิเคราะห์ในระดับอรรถศาสตร์ 132

5.4.1 โครงสร้างอินพุต 133

5.4.2 การประมวลผล 133

5.4.3 โครงสร้างเอาต์พุต 133

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หน้าที่

บทที่ 6 ซอฟท์แวร์ของระบบวิเคราะห์ประโยคภาษาอังกฤษไปสู่ภาษากลาง 135

6.1 บทนำ 135

6.2 การออกแบบและการเขียนซอฟท์แวร์ของระบบวิเคราะห์ประโยคภาษาอังกฤษ 135

6.2.1 ซอฟท์แวร์ของระบบส่วนการวิเคราะห์ในระดับคำ 137

6.2.2 ซอฟท์แวร์ของระบบส่วนการวิเคราะห์ในระดับวากยสัมพันธ์ 141

6.2.3 ซอฟท์แวร์ของระบบส่วนการวิเคราะห์ในระดับอรรถศาสตร์ 152

6.2.4 ซอฟท์แวร์ของระบบส่วนการแสดงผลแบบภาษากลาง 153

6.3 การใช้งานซอฟท์แวร์ของระบบการวิเคราะห์โครงสร้างประโยคภาษาอังกฤษ 154

6.4 สรุป 165

บทที่ 7 การประเมินประสิทธิภาพของระบบ 166

7.1 บทนำ 166

7.2 ผลการทดสอบ 167

7.3 ตัวอย่างประโยคอื่นๆ 174

7.4 การนำผลงานวิจัยไปประยุกต์ใช้กับงานอื่นๆ 174

7.4.1 การประยุกต์ใช้กับระบบการแปลหลายภาษาด้วยเครื่องคอมพิวเตอร์ 174

7.4.2 การประยุกต์ใช้กับงานตรวจสอบการสะกดคำในภาษาอังกฤษ 174

7.4.3 การประยุกต์ใช้กับงานตรวจสอบโครงสร้างประโยคภาษาอังกฤษ 175

7.5 แนวทางในการพัฒนาต่อของระบบ 175

ภาคผนวก ก 176

ภาคผนวก ข 178

เอกสารอ้างอิง 188

ประวัติผู้เขียน 189

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

หน้าที่

ตารางที่ 2.1 ประเภทและสัญญลักษณ์ของคำภาษาอังกฤษในกลุ่มต่าง ๆ 8
ตารางที่ 3.1 ตัวอย่างส่วนข้อมูลทางโครงสร้างและความหมายของคำศัพท์ในพจนานุกรม
ที่สร้างขึ้นเพื่อใช้สำหรับขบวนการวิเคราะห์ประโยคภาษาอังกฤษ 20
ตารางที่ 3.2 แสดงรูปแบบของคำกริยา 22
ตารางที่ 3.3 ตัวอย่างคำกริยาที่มีการผันรูปไปตามกาล 38
ตารางที่ 3.4 แสดงตัวอย่างการเก็บบันทึกข้อมูลจำเพาะของคำศัพท์ประเภทคำกริยา
ที่มีการผันรูปไปตามกาลของประโยคในพจนานุกรมคำศัพท์ 41
ตารางที่ 3.5 ตารางกำหนดตำแหน่งของกลุ่มคำศัพท์ 44
ตารางที่ 7.1 แสดงตัวอย่างประโยค 168

สารบัญรูป

หน้าที่

รูปที่ 1.1 โครงสร้างของระบบการแปลภาษาด้วยคอมพิวเตอร์ทั่วไป 2
รูปที่ 1.2 ระบบการแปลภาษาโดยใช้ภาษากลาง 3
รูปที่ 1.3 การวิเคราะห์ภาษาต้นแบบไปสู่ภาษากลาง 4
รูปที่ 1.4 แสดงโครงร่างวิทยานิพนธ์ 6
รูปที่ 3.1 แสดงการเชื่อมระหว่างข้อมูลทั้ง 2 ส่วนในพจนานุกรมคำศัพท์ 17
รูปที่ 3.2 แสดงข้อมูลทางโครงสร้างและความหมายของทั้ง 2 ส่วน 18
รูปที่ 3.3 แสดงสัญญลักษณ์ชนิดของคำทั้ง 2 กลุ่ม 31
รูปที่ 3.4 แสดงโครงสร้างการค้นหาข้อมูลคำศัพท์จากพจนานุกรม 43
รูปที่ 3.5 ตัวอย่างคำศัพท์ในกลุ่มที่ 1 หรือกลุ่มคำศัพท์ในกลุ่มอักขระ "A" 45
รูปที่ 3.6 โครงสร้างของการเก็บบันทึกข้อมูลคำกริยาที่ผันรูปตามกาล
และ m เป็นลำดับสุดท้าย 46
รูปที่ 3.7 โฟล์วชาร์ตขบวนการค้นหาคำศัพท์ 48
รูปที่ 3.8 แสดงโฟล์วชาร์ตของอัลกอริทึมแบบ Sequential Search 50
รูปที่ 3.9 แสดงโฟล์วชาร์ตของอัลกอริทึมแบบ Binary Search 53
รูปที่ 4.1 แสดงโครงข่ายความสัมพันธ์ของภาษากลางและตัวอย่างประโยค "He eats rice." 62
รูปที่ 4.2 แสดงรูปแบบข้อมูลของ node ใดๆ ในภาษากลาง 68
รูปที่ 4.3 แสดงโครงสร้างภาษากลางของตัวอย่างประโยค "He does not go to school" 69
รูปที่ 4.4 แสดงรูปแบบข้อมูลเป็นบล็อคของตัวอย่างประโยค "He does not go to school" 70
รูปที่ 5.1 แสดงแผนภูมิการวิเคราะห์ในระดับคำ 74
รูปที่ 5.2 แสดงโฟล์วชาร์ตของอัลกอริทึมในกลุ่มที่ 1 82
รูปที่ 5.3 แสดงโฟล์วชาร์ตของอัลกอริทึมในกลุ่มที่ 2 84
รูปที่ 5.4 แสดงโฟล์วชาร์ตของอัลกอริทึมในกลุ่มที่ 3 86
รูปที่ 5.5 แสดงโฟล์วชาร์ตของอัลกอริทึมในกลุ่มที่ 4 88
รูปที่ 5.6 แสดงตำแหน่ง window ณ จุดเริ่มต้นของขบวนการวิเคราะห์ใน
ระดับวากยสัมพันธ์ที่คำแรกสุดของประโยค 90

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 5.7 แสดงตำแหน่งของ window ต่างๆ เมื่อ monitor window เปลี่ยนตำแหน่งของคำบ่งชี้ 91
รูปที่ 5.8 แสดงรูปแบบภาษากลางจากตัวอย่างประโยค "My brother wants to play football in field." 132
รูปที่ 5.9 แสดงโครงสร้างภาษากลางของตัวอย่างประโยค "My brother wants to play football in field." 133
รูปที่ 5.10 แสดงรูปแบบข้อมูลเป็นบล็อคของตัวอย่างประโยค "My brother wants to play football in field." 134
รูปที่ 6.1 แสดงขั้นตอนการทำงานของซอฟท์แวร์ระบบวิเคราะห์ 136
รูปที่ 6.2 แสดงโฟล์วชาร์ตของซอฟท์แวร์ระบบส่วนการวิเคราะห์ในระดับคำ 138
รูปที่ 6.3 แสดงลักษณะการเรียงลำดับของอักขระในประโยคตัวอย่าง "I have a house.". 139
รูปที่ 6.4 แสดงตัวอย่างการเข้าหาคำศัพท์ในพจนานุกรมคำศัพท์ของประโยคตัวอย่าง "I have a house." 140
รูปที่ 6.5 แสดงข้อมูลที่ติดมากับหน่วยคำของประโยคตัวอย่าง "I have a house." 140
รูปที่ 6.6 แสดงโฟล์วชาร์ตของซอฟท์แวร์ของระบบส่วนการวิเคราะห์ในระดับวากยสัมพันธ์ 142
รูปที่ 6.7 แสดงโฟล์วชาร์ตการประมวลผลส่วนขยายหน้าคำนามหลัก 143
รูปที่ 6.8 แสดงโฟล์วชาร์ตการประมวลผลส่วนขยายหลังคำนามหลัก 144
รูปที่ 6.9 แสดงโฟล์วชาร์ตการประมวลผลคำปฏิเสธกับคำกริยาหลัก 145
รูปที่ 6.10 แสดงโฟล์วชาร์ตการประมวลผลคำกริยาช่วยกับคำกริยาหลัก 146
รูปที่ 6.11 แสดงโฟล์วชาร์ตการประมวลผลส่วนขยายประโยคกับคำกริยาหลัก 147
รูปที่ 6.12 แสดงโฟล์วชาร์ตการประมวลผลคำกริยาวิเศษณ์กับคำกริยาหลัก 148
รูปที่ 6.13 แสดงโฟล์วชาร์ตการประมวลผลคำกริยาหลักกับคำบุพบทและคำนามที่อยู่หลังคำบุพบท 149
รูปที่ 6.14 แสดงโฟล์วชาร์ตการประมวลผลหารูปแบบของคำกริยา 150
รูปที่ 6.15 แสดงตัวอย่างผลลัพธ์ของหน่วยคำที่ผ่านการประมวลผลทางด้านประเภทของหน่วยคำที่มีมากกว่าหนึ่งของประโยคตัวอย่าง "I have a house." 151

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หน้าที่

รูปที่ 6.16 แสดงผลลัพธ์ในรูปแผนภาพต้นไม้ของส่วนการวิเคราะห์ทางด้านวากยสัมพันธ์ของประโยคตัวอย่าง "I have a house." 151
รูปที่ 6.17 แสดงโฟล์วชาร์ตของซอฟท์แวร์ระบบส่วนการวิเคราะห์ในระดับอรรถศาสตร์ 152
รูปที่ 6.18 แสดงผลลัพธ์ในรูปแบบแผนภาพต้นไม้ของการวิเคราะห์ระดับอรรถศาสตร์ของประโยคตัวอย่าง "I have a house." 153
รูปที่ 6.19 รูปแบบภาษากลางของประโยคตัวอย่าง ของประโยคตัวอย่าง "I have a house." 154
รูปที่ 6.20 แสดงหน้าจอแรกของระบบ 155
รูปที่ 6.21 แสดงหน้าจอการรับข้อมูลทางแป้นพิมพ์ 156
รูปที่ 6.22 แสดงผลการตรวจสอบอักขระ 157
รูปที่ 6.23 แสดงการค้นหาคำศัพท์ในพจนานุกรมคำศัพท์ 158
รูปที่ 6.24 แสดงประเภทของคำในประโยค 159
รูปที่ 6.25 แสดงผลลัพธ์ของกฎโครงสร้างวลี 160
รูปที่ 6.26 แสดงผลลัพธ์ทางด้านวากยสัมพันธ์ 161
รูปที่ 6.27 แสดงผลลัพธ์ทางด้านอรรถศาสตร์ 162
รูปที่ 6.28 แสดงรูปแบบภาษากลาง 163
รูปที่ 6.29 แสดงหน้าจอต้องการทำงานต่อหรือไม่ 164

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

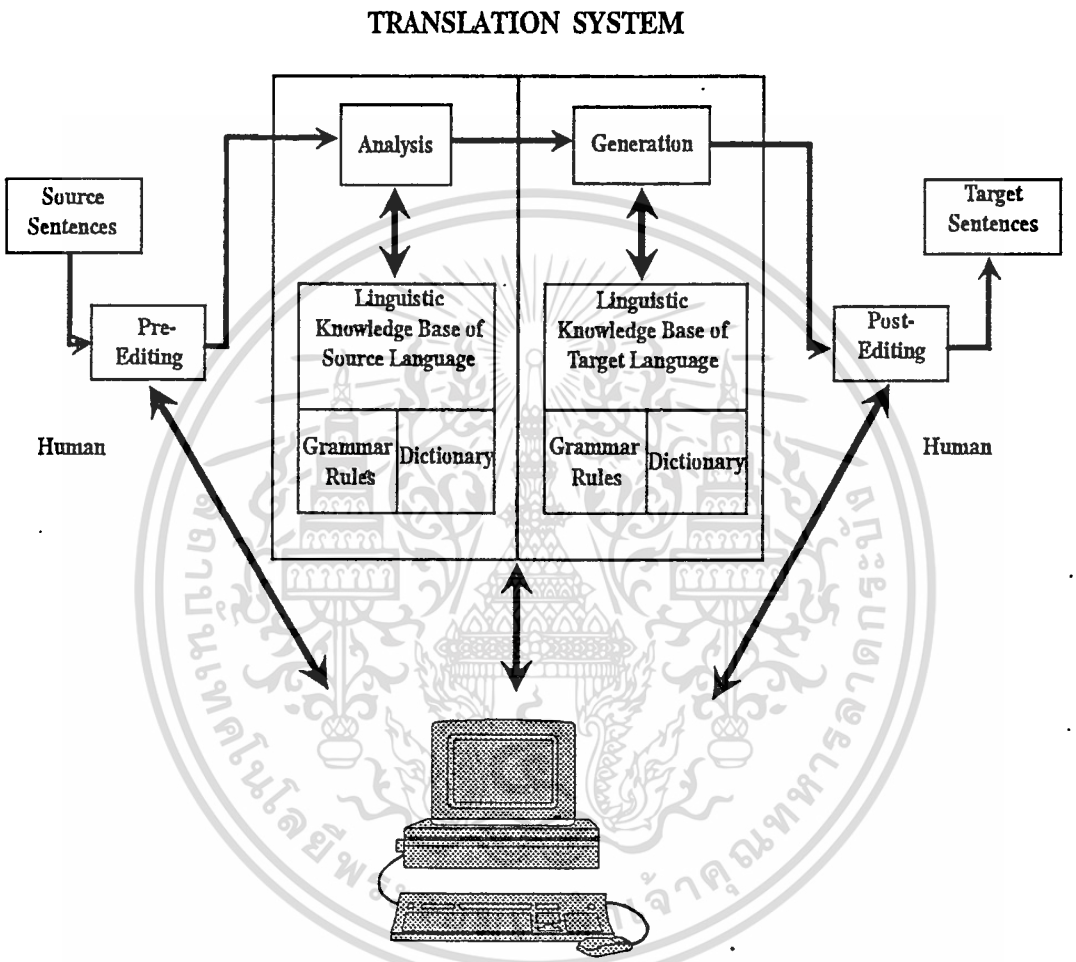
1.1 ความเป็นมาของงานวิจัย

การแปลภาษาด้วยระบบคอมพิวเตอร์เป็นวิทยาการที่ได้จากการประยุกต์ทฤษฎีทางด้านการประมวลผลภาษาธรรมชาติ (Natural Language Processing) มาใช้ในการวิเคราะห์ (Analysis) ภาษาต้นแบบและในการสังเคราะห์หรือสร้าง (Synthesis or Generation) ภาษาเป้าหมาย ในการแปลภาษาหนึ่งไปเป็นภาษาอื่นด้วยระบบคอมพิวเตอร์นั้นจะมีกระบวนการที่คล้ายกับวิธีการแปลภาษาด้วยมนุษย์ คือ มนุษย์จะต้องอ่านภาษาต้นแบบ แล้วทำความเข้าใจกับภาษาต้นแบบนั้น แล้วจึงแปลมาเป็นภาษาเป้าหมายตามความคิดที่มนุษย์เข้าใจ ซึ่งผู้ที่จะแปลภาษานั้นจะต้องเข้าใจหรือมีความรู้ ทั้งภาษาต้นแบบและภาษาเป้าหมาย ในทำนองเดียวกันการแปลภาษาด้วยคอมพิวเตอร์ก็ต้องใช้การวิเคราะห์ภาษาต้นแบบ การสังเคราะห์ประโยคและความรู้เหมือนกับมนุษย์ ซึ่งมนุษย์จะใช้ความรู้ที่จดจำไว้ในสมองเป็นเครื่องมือในการแปล ส่วนระบบคอมพิวเตอร์จะแปลภาษาได้ต้องอาศัยกระบวนการประมวลผลภาษาธรรมชาติกับฐานความรู้ทางภาษาศาสตร์ (Linguisitic Knowledge Base) ปัญหาของการแปลภาษาด้วยระบบคอมพิวเตอร์จึงอยู่ที่การทำให้ระบบคอมพิวเตอร์เข้าใจโครงสร้างและความหมายของประโยคหรือข้อความที่จะแปลได้เหมือนมนุษย์

ดังนั้นการแปลภาษาด้วยระบบคอมพิวเตอร์จึงต้องเริ่มจากขบวนการวิเคราะห์ประโยคของภาษาต้นแบบ (Source Sentence) เพื่อหาข้อมูลทางด้านโครงสร้าง (Syntax) และความหมาย (Semantic) โดยอาศัยฐานความรู้ทางภาษาศาสตร์ของภาษาต้นแบบนั้น ได้แก่พจนานุกรมและกฎไวยากรณ์ต่างๆ ดังแสดงในรูปที่ 1.1 ตัวอย่างข้อมูลทางด้านโครงสร้างประโยค ได้แก่ หน้าที่ของคำ (Category) เช่น การเป็นคำนาม (Noun), กริยา (Verb), คำสรรพนาม (Pronoun), คำวิเศษณ์ (Adjective), คำบุพบท (Preposition) หรือ ความสัมพันธ์ของคำในประโยคเช่นการเป็นประธาน (Subject) การเป็นส่วนการกระทำ (Action Verb) การเป็นกรรม (Object) เป็นต้น ส่วนตัวอย่างข้อมูลทางด้านความหมาย ได้แก่ จุดประสงค์ของประโยคนั้น (Objective) เวลาที่เกิดเหตุการณ์ในประโยคนั้น (Time and Tense) เป็นต้น ข้อมูลเหล่านี้จะถูกส่งต่อให้ขบวนการสร้างประโยคของภาษาเป้าหมาย (Target Sentence) ทำการ สังเคราะห์บนฐานความรู้ของภาษาเป้าหมายนั้นซึ่งประกอบด้วยพจนานุกรมและกฎไวยากรณ์เช่นกัน ในการแปลภาษาด้วยระบบคอมพิวเตอร์ทั่วไปดังแสดงในรูปที่ 1.1 นั้น ในบางครั้งอาจจะมีความจำเป็น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ต้องจัดการกับประโยค ต้นแบบ (Pre-editing) ก่อนที่จะป้อนเข้าสู่ขบวนการวิเคราะห์ เช่น การตรวจสอบตัวสะกด



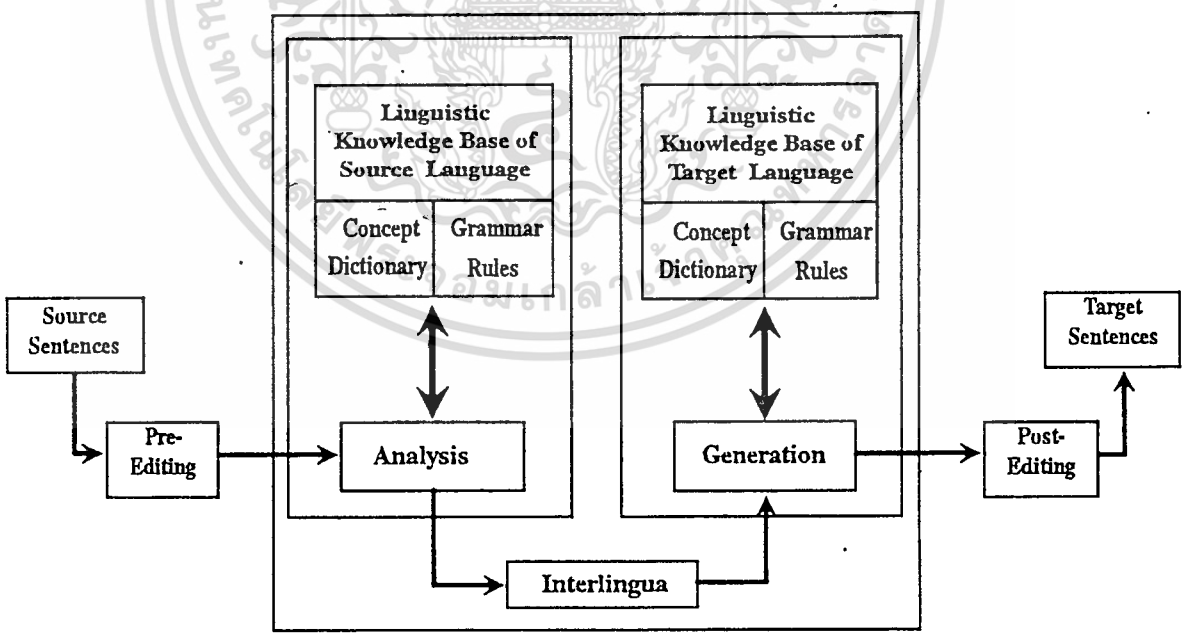
รูปที่ ง.1 โครงสร้างของระบบการแปลภาษาด้วยคอมพิวเตอร์ทั่วไป

(Spelling Check) หรือการตัดประโยค เป็นต้น ในทำนองเดียวกัน อาจจะมีความจำเป็นต้องจัดการกับผลลัพธ์ประโยคเป้าหมายที่สังเคราะห์ได้ (Post-editing) ให้มีความถูกต้องยิ่งขึ้น เช่นมีความเป็นธรรมชาติในทางความหมายของภาษามากขึ้น เป็นต้น

จากการศึกษาวิจัยวิธีการแปลภาษาด้วยระบบคอมพิวเตอร์จากอดีตจนถึงปัจจุบัน (ดังได้อธิบายไว้ในภาคผนวก ก) พบว่า วิธีการแปลโดยใช้สื่อภาษากลาง (Intermediate Language หรือ Interlingua) เป็นตัวกลางในการถ่ายทอด (Transfer) ข้อมูลโครงสร้างทางไวยากรณ์และความหมายของ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ในเชิงพาณิชย์ ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประโยคภาษาต้นแบบไปสู่การสร้างประโยคภาษาเป้าหมายนั้น เป็นวิธีการที่ได้รับการยอมรับว่าให้ความถูกต้องและมีประสิทธิภาพในการแปลสูง เหมาะสำหรับการแปลหลายๆ ภาษาในระบบเดียวกัน เพราะลดความยุ่งยากและลดเวลาในการวิเคราะห์ประโยคภาษาต้นแบบและการสร้างประโยคภาษาเป้าหมาย เมื่อเปรียบเทียบกับกระบวนการแปลภาษาแบบถ่ายทอดโดยตรง (Direct Machine Translation Strategy) ทั้งนี้เนื่องจากไม่จำเป็นต้องคำนึงถึงคู่ภาษาที่จะแปลเพราะการวิเคราะห์ประโยค ภาษาต้นแบบใดๆ ก็ตามให้อยู่ในรูปของภาษากลางแล้วก็จะสามารถแปลเป็นภาษาเป้าหมายอะไรก็ได้ โดยที่ไม่ต้องทำการวิเคราะห์ใหม่และในทำนองเดียวกัน ทางด้านการสร้างประโยคภาษาเป้าหมายใดๆ นั้น ก็เป็นการสังเคราะห์จากภาษากลางซึ่งเป็นผลของการวิเคราะห์ภาษาต้นแบบใดๆ เพียงอย่างเดียว เมื่อได้มีการสังเคราะห์แล้วก็ไม่ต้องทำซ้ำอีกต่อไป นอกจากนั้นการแปลภาษาโดยใช้ภาษากลางนี้จะใช้พจนานุกรมทางด้านมโนทัศน์ของภาษานั้นกับภาษากลาง (Concept Dictionary) เป็นฐานข้อมูลในการวิเคราะห์ และสังเคราะห์เพียงชุดเดียวต่อหนึ่งภาษา ซึ่งต่างจากการแปลภาษาแบบถ่ายทอดโดยตรงที่จะต้องใช้พจนานุกรมคู่ภาษาถึง n ชุด ในการแปลจากภาษาต้นแบบ 1 ภาษาไปสู่ภาษาเป้าหมาย n ภาษา รูปที่ 1.2 ได้แสดงให้เห็นถึงโครงสร้างโดยทั่วไปของระบบการแปลภาษาด้วยคอมพิวเตอร์โดยการใช้ภาษากลางเป็นสื่อ



รูปที่ 1.2 ระบบการแปลภาษาโดยใช้ภาษากลาง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประสิทธิภาพของการแปลภาษาด้วยระบบคอมพิวเตอร์โดยใช้สื่อกลางเป็นหลักดังกล่าวข้างต้น จึงขึ้นอยู่กับความสมบูรณ์และความถูกต้องของส่วนประกอบ 4 ส่วนได้แก่

ก) รูปแบบภาษากลาง

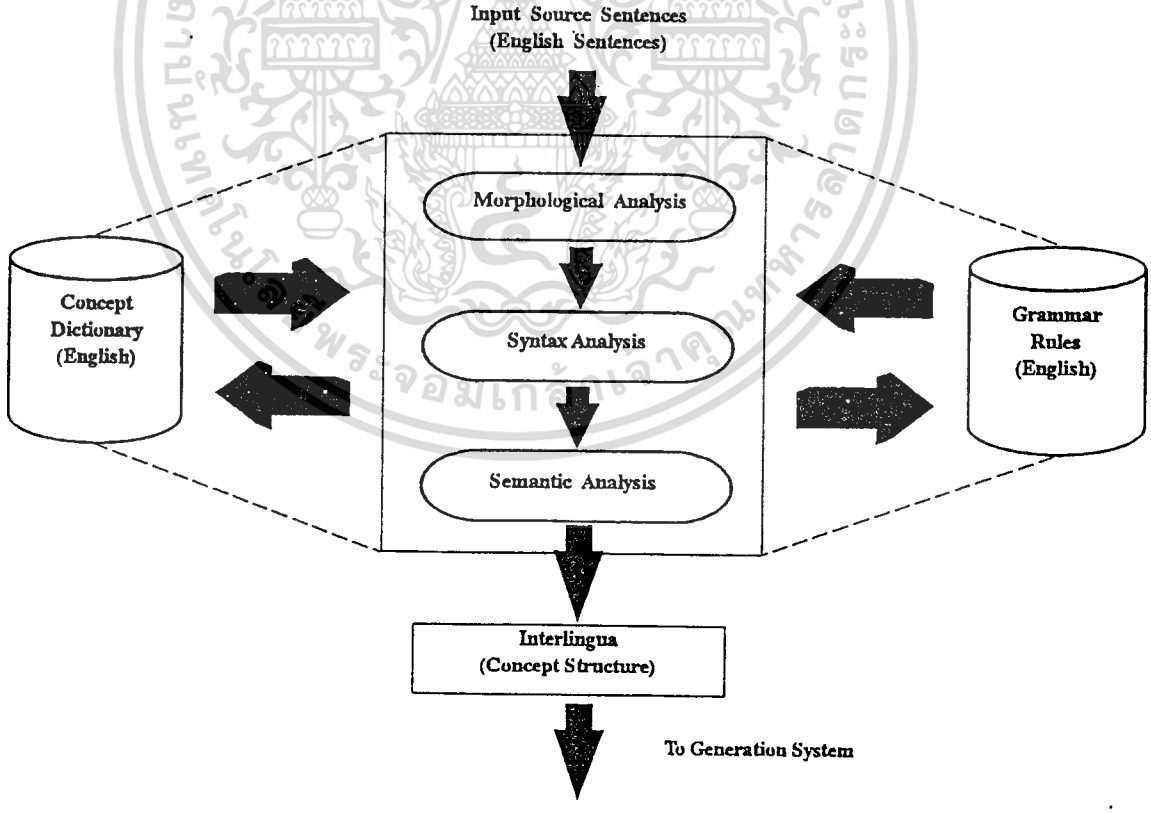
ข) ฐานความรู้ทางภาษาศาสตร์ที่รวมไปถึงกฎไวยากรณ์

ค) ระบบการวิเคราะห์ภาษาต้นแบบไปสู่ภาษากลาง

ง) ระบบการสังเคราะห์ภาษาเป้าหมายจากภาษากลาง

ส่วนประกอบทั้งสี่ต่างมีความสำคัญเท่าเทียมกันและแต่ละส่วนเป็นงานศึกษาวิจัยที่มีขอบเขตกว้างขวางที่ยังไม่มีข้อสรุปที่สมบูรณ์ในปัจจุบัน

ดังนั้นวิทยานิพนธ์ฉบับนี้ จึงขอเสนอผลงานวิจัยและพัฒนาขั้นพื้นฐานในส่วนของการวิเคราะห์ภาษาต้นแบบไปสู่ภาษากลาง รูปแบบของภาษากลางและฐานความรู้ทางภาษาศาสตร์ได้แก่ โครงสร้างของพจนานุกรมคำศัพท์ตลอดจนกฎไวยากรณ์ต่างๆ โดยใช้ภาษาอังกฤษเป็นกรณีศึกษาวิจัยดังแสดงในรูปที่ 1.3



รูปที่ 1.3 การวิเคราะห์ภาษาต้นแบบไปสู่ภาษากลาง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.2 จุดประสงค์และขอบเขตของงานวิจัย

งานวิจัยนี้ได้เสนอขบวนการวิเคราะห์ประโยคภาษาอังกฤษไปสู่ภาษากลางภายใต้หลักการประมวลผลภาษาธรรมชาติตลอดจนการพัฒนาซอฟท์แวร์ของระบบการวิเคราะห์นี้ เพื่อการวิเคราะห์ประมวลผลด้วยระบบคอมพิวเตอร์ อันจะเป็นพื้นฐานและผลการวิจัยที่จะนำไปสู่การพัฒนาระบบการแปลหลายๆ ภาษาด้วยระบบคอมพิวเตอร์เดียวกัน โดยใช้สื่อภาษากลางในอนาคต

ขบวนการวิเคราะห์ประโยคภาษาอังกฤษไปสู่ภาษากลางของงานวิจัยนี้ แบ่งเป็น 3 ขั้นตอนใหญ่ๆ ตามลำดับดังนี้

ขั้นตอนที่ 1 : ส่วนการวิเคราะห์ในระดับคำ (Morphological Analysis)

ขั้นตอนที่ 2 : ส่วนการวิเคราะห์ในระดับโครงสร้างทางวากยสัมพันธ์หรือไวยากรณ์ (Syntax Analysis)

ขั้นตอนที่ 3 : ส่วนการวิเคราะห์ในระดับอรรถศาสตร์หรือความหมาย (Semantic Analysis)

แต่ละขั้นตอนจะมีการประมวลผลต่อเนื่องสัมพันธ์กันดังแสดงในรูปที่ 1.3

ในการวิเคราะห์ของแต่ละขั้นตอน จะใช้กฎไวยากรณ์ (Grammar Rules) ภาษาอังกฤษประกอบกับพจนานุกรมทางความหมายและโครงสร้างของคำศัพท์ (Concept Dictionary) ภาษาอังกฤษประกอบ เพื่อการหาความสัมพันธ์ระดับแก่นของคำแต่ละคำในประโยค โดยไม่ให้มีความกำกวมเกิดขึ้น ความสัมพันธ์ระดับแก่นของคำในประโยคนี้จะใช้หลักการทางการกสัมพันธ์ (Case Relation) ของฟิลมอร์ [1] เป็นพื้นฐานและงานวิจัยนี้ได้เสนอให้ผลลัพธ์โครงสร้างระดับแก่น (Conceptual Structure) ของประโยคที่วิเคราะห์นี้เป็นภาษากลาง

นอกจากนั้นในงานวิจัยนี้ได้เสนอรูปแบบของพจนานุกรมทางความหมายและโครงสร้างของคำศัพท์ภาษาอังกฤษ ภายใต้หลักการทางภาษาศาสตร์และคอมพิวเตอร์ซอฟท์แวร์ เพื่อให้มีข้อมูลที่สมบูรณ์เพียงพอ และสะดวกต่อการใช้งานร่วมกับขบวนการวิเคราะห์ในขั้นตอนต่างๆ โดยมีการจัดหมวดหมู่ของคำศัพท์ การกำหนดรูปแบบของคำกริยา (Verb Pattern) เป็นต้น คำศัพท์ทั้งหมดนี้จะนำข้อมูลมาจากพจนานุกรมมาตรฐานของลองแมน [2]

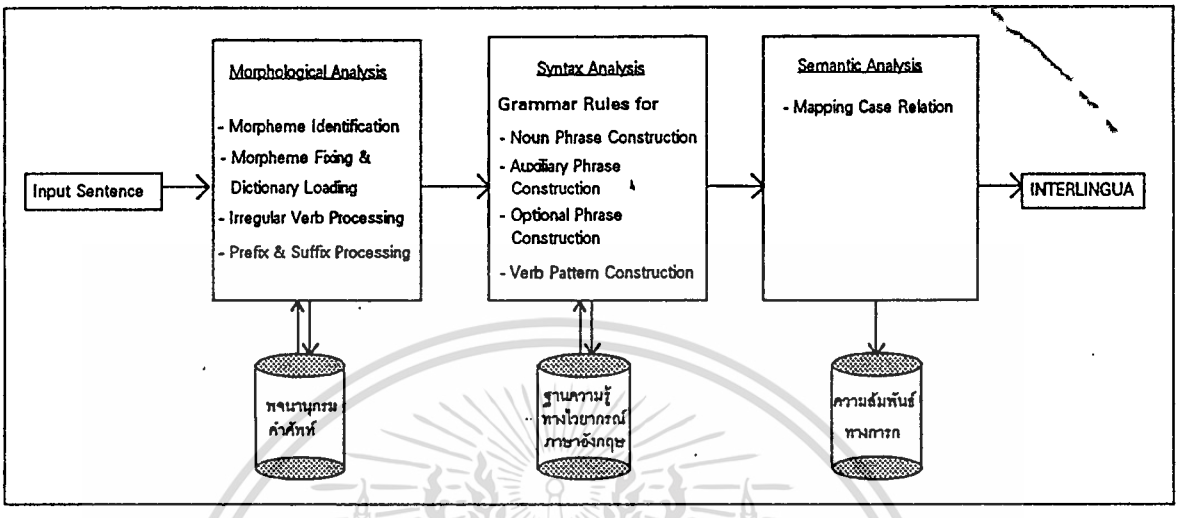
สำหรับการสร้างกฎไวยากรณ์ จะใช้หลักการของกฎโครงสร้างวลี (Phrase Structure Rule) เป็นพื้นฐานไวยากรณ์ของภาษาอังกฤษจะรวบรวมมาจากตำราไวยากรณ์ภาษาอังกฤษที่ได้มาตรฐาน[4],[5]

และงานวิจัยยังได้เสนอรูปแบบของภาษากลางที่แสดงโครงสร้างระดับแก่นของประโยคที่สามารถนำไปประยุกต์ใช้กับภาษาใดๆ ก็ได้ โดยคำนึงถึงความสมบูรณ์ทางภาษาที่จะนำไปเป็นสื่อกลางในการแปลระหว่างหลายๆ ภาษาด้วยระบบคอมพิวเตอร์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.3 โครงร่างของวิทยานิพนธ์

ในส่วนโครงร่างของวิทยานิพนธ์ฉบับนี้ดังแสดงในรูปที่ 1.4



รูปที่ 1.4 แสดงโครงร่างวิทยานิพนธ์

จากรูปที่ 1.4 เมื่อเรานำประโยคมาวิเคราะห์ผ่านขั้นตอนการวิเคราะห์ระดับคำ (Morphological Analysis) เพื่อให้ได้รากคำศัพท์โดยอ้างอิงจากพจนานุกรมคำศัพท์และมีการดึงข้อมูลจากพจนานุกรมคำศัพท์เพื่อใช้ในการวิเคราะห์ระดับวากยสัมพันธ์ต่อไป ซึ่งในระดับนี้จะต้องนำฐานความรู้ทางไวยากรณ์ภาษาอังกฤษมาสร้างเป็นกฎไวยากรณ์ที่ใช้ในการวิเคราะห์ประโยคเพื่อให้ได้ความสัมพันธ์ของคำแต่ละคำของประโยคในหน้าที่ของคำแต่ละคำนั้น ๆ เมื่อได้ความสัมพันธ์ระดับไวยากรณ์แล้วก็จะเข้าสู่การวิเคราะห์ระดับอรรถศาสตร์เพื่อให้ได้โครงสร้างของประโยคที่ไม่ขึ้นกับภาษาใดโดยใช้ความสัมพันธ์ทางการกเข้ามาแสดงความสัมพันธ์ในประโยค และท้ายที่สุดเราก็จะได้รูปแบบภาษากลางที่ไม่ขึ้นกับภาษาใด ๆ

โดยที่วิทยานิพนธ์นี้ประกอบด้วยบทต่างๆ ดังนี้

บทที่ 1 ว่าด้วยเรื่องบทนำ

บทที่ 2 ว่าด้วยเรื่องฐานความรู้เกี่ยวกับไวยากรณ์ภาษาอังกฤษที่ใช้ในการวิเคราะห์

บทที่ 3 ว่าด้วยเรื่องการจัดเก็บฐานข้อมูลไวยากรณ์ในส่วนของพจนานุกรมคำศัพท์

บทที่ 4 ว่าด้วยเรื่องโครงสร้างของภาษากลาง

บทที่ 5 ว่าด้วยเรื่องขบวนการวิเคราะห์ประโยคภาษาอังกฤษ

บทที่ 6 ว่าด้วยเรื่องซอฟท์แวร์ของระบบวิเคราะห์ประโยคภาษาอังกฤษไปสู่ภาษากลาง

บทที่ 7 ว่าด้วยเรื่องการประเมินประสิทธิภาพของระบบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2
ฐานความรู้เกี่ยวกับไวยากรณ์ภาษาอังกฤษที่ใช้ในการวิเคราะห์

2.1 บทนำ

จากการศึกษาและค้นคว้าเกี่ยวกับหลักไวยากรณ์ภาษาอังกฤษที่ได้รับการพิมพ์เผยแพร่ในตำราที่ได้มาตรฐานต่างๆ [3],[4] ผู้วิจัยได้ทำการปรับปรุงและเรียบเรียงประเภทของคำ โครงสร้างทางไวยากรณ์ของประโยคและรูปแบบของคำกริยา ให้เหมาะสมกับการสร้างกฎไวยากรณ์เพื่อการวิเคราะห์โครงสร้างประโยคภาษาอังกฤษในการหาความสัมพันธ์ของคำแต่ละคำในประโยค ที่จะให้เกิดความกำกวมน้อยที่สุด และแปลงเป็นภาษากลางได้อย่างถูกต้องทั้งด้านโครงสร้างและความหมาย ทั้งนี้โดยคำนึงถึงประโยชน์และความเหมาะสมที่นำไปใช้กับระบบการแปลภาษาด้วยคอมพิวเตอร์

2.2 การแบ่งประเภทของคำภาษาอังกฤษ

ในการวิเคราะห์ประโยคภาษาอังกฤษระดับคำ (Morphological Analysis) ข้อมูลฐานความรู้จากพจนานุกรมคำศัพท์ภาษาอังกฤษที่จำเป็นเริ่มแรกได้แก่ ประเภทของคำ (Category) ที่จะนำไปสู่การกำหนดหน้าที่ของคำนั้นๆ เมื่อไปปรากฏอยู่ในประโยค (Part of Speech) ดังนั้น งานวิจัยจึงได้จัดแบ่งประเภทของคำออกเป็น 10 กลุ่ม ตามคุณสมบัติของคำนั้นๆ และในบางกลุ่มยังได้กระจายเป็นกลุ่มย่อย (Sub-category) ตามประเภทของคำที่มีคุณสมบัติส่วนรายละเอียดเหมือนกัน ทั้งนี้เพื่อความละเอียดและถูกต้องในการเป็นฐานข้อมูลสำหรับการวิเคราะห์โครงสร้างประโยคภาษาอังกฤษ ประเภทของคำในกลุ่มต่างๆ ที่งานวิจัยได้จัดแบ่งไว้ ตลอดจนสัญญลักษณ์นี้ที่ใช้แทนประเภทของคำนั้นแสดงไว้ในตารางที่ 2.1 รายละเอียดประเภทของคำมีดังต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 2.1 ประเภทและสัญญลักษณ์ของคำภาษาอังกฤษในกลุ่มต่าง ๆ

Category		Sub-category		Example
Name	Symbol	Name	Symbol	
Noun	N	Proper Noun	NPRP	England,Thailand
		Cardinal Number	NCNM	one,two,zero,half
		Ordinal Number	NONM	first,second
		Nominal Possesive	NPSS	Somsak's
		Verbial Noun	NVBN	process,attemp
		Adjectival Noun	NAJN	absence,most
		Common Noun	NCMN	house,table,dog
Pronoun	PRON	Personal Pronoun	PPRS	I,you,him,my
		Demonstrative Pronoun	PDMN	this,that,those
		Definite Pronoun	PDEF	all,both
		Indefinite Pronoun	PIND	other,much
		Number Pronoun	PONE	one,ones
Normal Verb	V	Transitive Verb	TV	borrow,eat,give
		Intransitive Verb	IV	walk,sleep
Auxiliary Verb	AUX	Modallity	XMDL	will,can,could
		Verb to be	XXBE	is,am,are,was,were
		Verb to do	XXDO	do,does,did,done
		Verb to have	XHAV	has,have,had
Determiner	DET	Definite Article	DTHE	the
		Indefinite Article	DDAN	a,an
		Personal Determiner	DPRS	our,you,your,her,
		Demonstrative Determiner	DDMN	this,that,those
		Indefinite Determiner	DIND	some,few,little
		Cardinal Number Determiner	DCNM	one,two,three,
Adjective	ADJ	Adjective before Noun	ALRP	good
		Adjective after Noun	ANRN	elect
Adverb	ADV	Adverb berfore Verb	ADVF	hardly,usually
		Adverb after Verb	ADVV	well,quickly
Negation	NEG	———	——	not,no,none
Conjunction	CONJ	———	——	and
Preposition	PREP	———	——	at,in,on,up

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.1 กลุ่มคำนาม (Noun ใช้สัญญลักษณ์ N) ได้แก่คำที่ใช้เรียก คน สัตว์ และสิ่งของ โดยงานวิจัยนี้ได้จัดแบ่งเป็น 7 กลุ่มย่อย ดังนี้

1) วิสามัญนาม (Proper Noun ใช้สัญญลักษณ์ NPRP) เป็นคำนามชี้เฉพาะ เช่น ชื่อเฉพาะของบุคคล สัตว์ สิ่งของ และ สถานที่ ตัวอย่างเช่น Thailand, Japan, England, John, Aranya, KMITL เป็นต้น

2) คำนามที่แสดงจำนวนนับ (Cardinal Number ใช้สัญญลักษณ์ NCNM) ตัวอย่างเช่น one, two, zero, half เป็นต้น

3) คำนามที่บอกลำดับที่ก่อน-หลัง (Ordinal Number ใช้สัญญลักษณ์ NONM) ตัวอย่างเช่น first, second เป็นต้น

4) คำนามที่แสดงความเป็นเจ้าของของคำนามที่ตามมา (Nominal Possesive ใช้ สัญญลักษณ์ NPSS) จะอยู่ในรูปของวิสามัญนาม หรือสามัญนามกับอะพอสโทฟี่ s ตัวอย่างเช่น Somsak's หมายถึง — ของสมศักดิ์ , dog's หมายถึง — ของสุนัข เป็นต้น

5) คำที่เป็นอาการนาม (Verbial Noun ใช้สัญญลักษณ์ NVBN) ตัวอย่างเช่น process, attempt เป็นต้น

6) คำนามที่แสดงการขยายคำอื่น ๆได้ (Adjectival Noun ใช้สัญญลักษณ์ NAJN) ทำนองเดียวกับคำวิเศษณ์ ตัวอย่างเช่น sentitiveness, absence, most เป็นต้น

7) สามัญนาม (Common Noun ใช้สัญญลักษณ์ NCMN) ซึ่งเป็นคำนามโดยทั่วๆไป ที่ใช้เรียก คน สัตว์ สิ่งของ ตัวอย่างเช่น house, table เป็นต้น

2.2.2 กลุ่มคำสรรพนาม (Pronoun ใช้สัญญลักษณ์ PRON) ได้แก่ คำที่ใช้แทนคำนาม และอาจแทนด้วยคำบ่งชี้ที่มีความหมายรวมถึงคำนามที่ละไว้ด้วย โดยงานวิจัยนี้ได้แบ่งเป็น 5 กลุ่มย่อย ดังนี้

1) บุรุษสรรพนามซึ่งเป็นคำสรรพนามที่แทนตัวผู้พูด ผู้ที่พูดด้วย หรือผู้ที่กล่าวถึง (Personal Pronoun ใช้สัญญลักษณ์ PPRS) ตัวอย่างเช่น I, you, we, him, my เป็นต้น

2) สรรพนามบ่งชี้ที่ใช้แทนสิ่งที่ผู้พูดชี้เฉพาะเจาะจงถึง (Demonstrative Pronoun ใช้สัญญลักษณ์ PDMN) ตัวอย่างเช่น this, that, those, these เป็นต้น

3) สรรพนามชี้เฉพาะที่ใช้แทนเพื่อบอกขอบเขตของสิ่งที่พูดถึง (Definite Pronoun ใช้สัญญลักษณ์ PDEF) ตัวอย่างเช่น all, both เป็นต้น

4) สรรพนามไม่ชี้เฉพาะที่ใช้แทนคนหรือสิ่งทั่ว ๆ ไปไม่มีขอบเขตแน่นอน (Indefinite Pronoun ใช้สัญญลักษณ์ PIND) ตัวอย่างเช่น other, others, much, many, some เป็นต้น

5) คำยกเว้นที่แสดงหน้าที่เป็นคำนามได้โดยที่คำเดิมเป็นคำแสดงจำนวนนับ (Number Pronoun ใช้สัญญลักษณ์ PONE) ตัวอย่างเช่น ones, one เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.3 กลุ่มคำกริยาปกติ (Normal Verb ใช้สัญญลักษณ์ V) ได้แก่ คำที่แสดงกริยาอาการหรือสภาพของประธานในประโยค และถือเป็นคำหลักของประโยคได้จัดแบ่งเป็น 2 กลุ่มย่อย ดังนี้

1) สกรรมกริยา (Transitive Verb ใช้สัญญลักษณ์ TV) ได้แก่คำกริยาที่ต้องมีกรรมมารองรับ ตัวอย่างเช่น borrow, eat, give เป็นต้น

2) อกรรมกริยา (Intransitive Verb ใช้สัญญลักษณ์ IV) ได้แก่คำกริยาที่ไม่ต้องมีกรรมมารองรับ ตัวอย่างเช่น walk, sleep เป็นต้น

2.2.4 กลุ่มคำกริยาช่วย (Auxiliary Verb ใช้สัญญลักษณ์ AUX) ได้แก่ คำที่เกิดร่วมกับคำกริยาเพื่อแสดงกาล มาลา วาจก ลักษณาการ และเจตนาของประโยค ได้จัดแบ่งเป็น 4 กลุ่มย่อย ดังนี้

1) คำที่แสดงมาลาของประโยค (Modalityใช้สัญญลักษณ์ XMDL) ตัวอย่างเช่น will, can, could, should, shall เป็นต้น

2) คำที่มีความหมายในภาษาไทยว่า "เป็น"หรือ" อยู่ " หรือ" คือ" อย่างใดอย่างหนึ่งในประโยค คือ Verb to be นั่นเอง (Verb to beใช้สัญญลักษณ์ XXBE) ตัวอย่าง เช่น is, am, are, were, was เป็นต้น

3) คำที่เป็นกริยาช่วย (Verb to do ใช้สัญญลักษณ์ XXDO) เพื่อให้ประโยคสมบูรณ์ ตัวอย่างเช่น do, does, did เป็นต้น

4) คำที่เป็นกริยาช่วย (Verb to have ใช้สัญญลักษณ์ XHAV) เพื่อให้ประโยคสมบูรณ์ ตัวอย่างเช่น has, have, had เป็นต้น

2.2.5 กลุ่มคำบ่งชี้หรือคำกำหนด (Determiner ใช้สัญญลักษณ์ DET) ได้แก่ คำที่ใช้แสดงความเฉพาะเจาะจงของคำนาม โดยปรากฏร่วมกับคำนาม ได้จัดแบ่งเป็น 6 กลุ่มย่อย ดังนี้

1) คำที่บ่งชี้เฉพาะเมื่อกล่าวถึงเป็นครั้งที่สองหรือต้องการกล่าวเน้นโดยเฉพาะ (Definite Article ใช้สัญญลักษณ์ DTHE) คือ the

2) คำบ่งชี้ที่ไม่ชี้เฉพาะเจาะจง และมีความหมายเป็นเอกพจน์ (Indefinite Article ใช้สัญญลักษณ์ DDAN) คือ a, an

3) คำบ่งชี้ซึ่งเป็นคำที่แทนตัวผู้พูด ผู้ที่พูดด้วยหรือผู้ที่กล่าวถึง (Personal Determiner ใช้สัญญลักษณ์ DPRS) ตัวอย่างเช่น our, your, his, her, its เป็นต้น

4) คำบ่งชี้ที่ใช้แทนสิ่งที่ผู้พูดชี้เฉพาะเจาะจงถึง (Demonstrative Determiner ใช้สัญญลักษณ์ DDMN) ตัวอย่างเช่น this, that, those เป็นต้น

5) คำบ่งชี้ที่ไม่ชี้เฉพาะที่ใช้แทนคน หรือสิ่งทั่วๆ ไป ไม่มีขอบเขตที่แน่นอน (Indefinite Determiner ใช้สัญญลักษณ์ DIND) ตัวอย่างเช่น some, few, fewer, fewest, little, less, least, many,

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

more เป็นต้น

6) คำบ่งชี้ที่ใช้แสดงจำนวนนับ (Cardinal Number Determiner ใช้สัญญลักษณ์ DCNM) ตัวอย่างเช่น one, two, three เป็นต้น

2.2.6 กลุ่มคำคุณศัพท์ (Adjective ใช้สัญญลักษณ์ ADJ) ได้แก่ คำขยายคำนามเพื่อบอกลักษณะของคำนาม ว่ามีลักษณะอย่างไร ได้จัดแบ่งเป็น 2 กลุ่มย่อย ดังนี้

1) คำคุณศัพท์ที่วางอยู่หน้าคำนาม (Adjective before Noun ใช้สัญญลักษณ์ ALRP) และจะรวมทั้งคำคุณศัพท์ที่วางอยู่หลังกริยาช่วย(verb to be) ตัวอย่างเช่น good, back เป็นต้น

2) คำคุณศัพท์ที่วางอยู่หลังคำนาม (Adjective after Noun ใช้สัญญลักษณ์ ANRN) ตัวอย่างเช่น elect, alight เป็นต้น

2.2.7 กลุ่มคำกริยาวิเศษณ์ (Adverb ใช้สัญญลักษณ์ ADV) ได้แก่ คำขยายกริยาเพื่อบอกว่าเหตุการณ์หรือสิ่งที่เกิดขึ้นนั้น เกิดเมื่อไหร่ ที่ไหน หรืออย่างไร ได้จัดแบ่งเป็น 2 กลุ่มย่อย ดังนี้

1) คำกริยาวิเศษณ์ที่วางอยู่หลังกริยาทั่วไป (Adverb after Verb ใช้สัญญลักษณ์ ADVV) ตัวอย่างเช่น well, quickly, slowly เป็นต้น

2) คำกริยาวิเศษณ์ที่วางอยู่หน้ากริยา (Adverb before Verb ใช้สัญญลักษณ์ ADVF) และหลังคำกริยาช่วย ตัวอย่างเช่น always เป็นต้น

2.2.8 กลุ่มคำปฏิเสธ (Negation ใช้สัญญลักษณ์ NEG) คือ คำที่แสดงการปฏิเสธหรือเพื่อให้ความหมายกลับเป็นตรงข้าม ตัวอย่างเช่น not, never เป็นต้น

2.2.9 กลุ่มคำสันธาน (Conjunction ใช้สัญญลักษณ์ CONJ) คือ คำที่ใช้เชื่อมวลี หรือประโยคเข้าด้วยกัน เพื่อแสดงความสัมพันธ์ในแง่ใดแง่หนึ่ง ตัวอย่างเช่น with, but, and, or เป็นต้น

2.2.10 กลุ่มคำบุพบท (Preposition ใช้สัญญลักษณ์ PREP) คือคำที่นำหน้าคำนามหรือคำสรรพนามเพื่อบอกตำแหน่งสถานที่ และความสัมพันธ์ระหว่างคำกริยาและคำนามหรือสรรพนามที่ตามมา ตัวอย่างเช่น in, at, on, under เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3 โครงสร้างทางไวยากรณ์ของประโยคภาษาอังกฤษ

จากการศึกษาโครงสร้างทางไวยากรณ์พื้นฐานของประโยคภาษาอังกฤษพบว่า ประโยคประกอบไปด้วยกลุ่มคำที่เรียงลำดับก่อนหลังจากซ้ายไปขวาอย่างมีกฎเกณฑ์ ตามหน้าที่ของคำนั้นๆ ที่สัมพันธ์กับคำข้างเคียง โดยแต่ละคำจะมีความหมาย (Meaning) ในตัวเอง เรียกคำเหล่านี้ว่าเป็นคำที่มีความหมาย (Meaning Words) หรือคำศัพท์ (Vocaburaries) ถ้าแบ่งกลุ่มคำที่ถูกเขียนเรียงกันในประโยคภาษาอังกฤษออกเป็นส่วนย่อย ๆ ตามคุณสมบัติและหน้าที่ของกลุ่มคำเหล่านั้นจะได้วลี (Phrase) และวลีเหล่านี้นั่นเองที่จะเป็นตัวกำหนดความหมาย (Semantic) ของการที่ผู้เขียนหรือผู้พูดประโยคนั้น ๆ ต้องการแสดงออกให้ผู้อ่านหรือผู้ฟังได้รับทราบ

งานวิทยานิพนธ์นี้ได้ทำการแบ่งกลุ่มคำที่จะปรากฏในประโยคภาษาอังกฤษออกเป็น 5 วลี ตามหลักไวยากรณ์ภาษาอังกฤษ โดยเน้นในรายละเอียดเพื่อประโยชน์ต่อการสร้างกฎไวยากรณ์ (Grammar Rule) ที่จะใช้ในการสร้างซอฟท์แวร์ของขบวนการวิเคราะห์โครงสร้างประโยคภาษาอังกฤษที่ให้ความถูกต้องมากที่สุดต่อไป วลีทั้ง 5 ประเภท ได้แก่

ก) นามวลี (Noun Phrase จะใช้สัญญลักษณ์ NP)
ข) กริยาวลี (Verb Phrase จะใช้สัญญลักษณ์ VP)
ค) กริยาวิเศษณ์วลี (Adverb Phrase จะใช้สัญญลักษณ์ ADVP)
ง) คุณศัพท์วลี (Adjective Phrase จะใช้สัญญลักษณ์ ADJP)
จ) บุพบทวลี (Preposition Phrase จะใช้สัญญลักษณ์ PP)

2.3.1 นามวลี (Noun Phrase :NP) ได้แก่วลีที่มีคำนามเป็นแกนหลักของกลุ่มคำ นามวลีนี้ อาจจะมีโครงสร้างเป็นเพียงคำนามหรือคำสรรพนามเพียงคำเดียว หรืออาจจะมีส่วนขยายมาประกอบในโครงสร้างสำหรับงานวิจัยนี้ได้แบ่ง นามวลีไว้ทั้งหมด 5 โครงสร้างดังนี้

ก) โครงสร้างนามวลีที่ 1 ได้แก่ นามวลีที่ประกอบไปด้วยคำบ่งชี้ (Determiner:DET) ตามด้วยคำนาม

NP ——> [DET] + N ตัวอย่างเช่น

DET	N
A	book

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข) โครงสร้างนามวลีที่ 2 ได้แก่ นามวลีที่ประกอบด้วยคำบ่งชี้ (Determiner:DET) ตามด้วยคำคุณศัพท์ (Adjective) และคำนาม

NP ——> [DET] + [ADJ] + N ตัวอย่างเช่น

DET	ADJ	N
A	red	book

ค) โครงสร้างนามวลีที่ 3 ได้แก่ นามวลีที่ประกอบไปด้วยคำบ่งชี้ (Determiner:DET) ตามด้วยคำนาม และบุพบทวลี

NP ——> [DET] + N + [PP] ตัวอย่างเช่น

DET	N	PP
A	book	of new bookstore

ง) โครงสร้างนามวลีที่ 4 ได้แก่ นามวลีที่ประกอบไปด้วยคำบ่งชี้ (Determiner:DET) ตามด้วยคำคุณศัพท์ คำนามและบุพบทวลี

NP ——> [DET] + [ADJ] + N + [PP] ตัวอย่างเช่น

DET	ADJ	N	PP
A	red	book	of new bookstore

จ) โครงสร้างนามวลีที่ 5 ได้แก่ นามวลีที่ประกอบไปด้วยนามวลี และนามวลี

NP ——> NP + NP ตัวอย่างเช่น

DET	ADJ	N	PP	CONJ	DET	ADJ	N
A	red	book	of new bookstore	and	a	blue	book

2.3.2 **กริยาวลี** (Verb Phrase :VP) ได้แก่ วลีที่มีคำกริยาเป็นแกนหลักของกลุ่มคำ กริยาวลีอาจจะมีโครงสร้างเป็นคำกริยาเพียงคำเดียวหรืออาจจะมีส่วนขยายประกอบในโครงสร้าง สำหรับงานวิจัยนี้ได้จัดแบ่งกริยาวลีออกเป็น 2 โครงสร้างดังนี้

ก) โครงสร้างกริยาวลีที่ 1 ได้แก่ คำกริยาวิเศษณ์ และคำกริยา

VP ——> [ADV] + V ตัวอย่างเช่น

ADV	V
always	give

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข) โครงสร้างกริยาวลีที่ 2 ได้แก่ คำกริยา และ คำกริยาวิเศษณ์

VP ——> V + [ADV] ตัวอย่างเช่น

V	ADV
go	slowly

2.3.3 **กริยาวิเศษณ์วลี** (Adverb Phrase : ADVP) ได้แก่วลีที่มีคำกริยาวิเศษณ์เป็นแกนหลักของกลุ่มคำ ซึ่งอาจจะประกอบด้วยคำกริยาวิเศษณ์เพียงคำเดียวหรืออาจจะมีคำกริยาวิศษณ์ มากกว่า 1 คำ มาเขียนเรียงต่อกันก็จะได้รูปแบบต่าง ๆ ที่ปรากฏในประโยค เช่น

กริยาวิเศษณ์วลีที่ประกอบด้วย คำกริยาวิเศษณ์ มากกว่า 1 คำ

ADVP ——> [ADV] + ADV ตัวอย่างเช่น

ADV	ADV	ADV
as	too	briefly

2.3.4 **คุณศัพท์วลี** (Adjective Phrase : ADJP) ได้แก่วลีที่มีคำคุณศัพท์เป็นแกนหลักของกลุ่มคำ อาจประกอบด้วยคำคุณศัพท์เพียงคำเดียว หรือมีส่วนขยายและเมื่อรวมส่วนขยายจะทำให้คุณศัพท์วลีมีโครงสร้างที่แตกต่างออกไป เช่น

คุณศัพท์วลีที่ประกอบด้วย คำกริยาวิเศษณ์ และ คำคุณศัพท์

ADJP ——> [ADV] + ADJ ตัวอย่างเช่น

ADV	ADJ
all	right

2.3.5 **บุพบทวลี** (Preposition Phrase : PP) เป็นวลีที่มีคำบุพบทเป็นแกนหลักของกลุ่มคำ ซึ่งมีโครงสร้างประกอบด้วยคำบุพบทและส่วนขยาย โดยงานวิจัยได้แบ่งออกเป็น 2 โครงสร้างดังนี้

ก) โครงสร้างบุพบทวลีที่ 1 ได้แก่ บุพบทวลีที่ประกอบด้วยคำบุพบทตามด้วยคำบ่งชี้ (Determiner:DET) และคำนาม

PP ——> PREP + [DET] + N ตัวอย่างเช่น

PREP	DET	N
at	my	house

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข) โครงสร้างบุพบทวลีที่ 2 ได้แก่ บุพบทวลีที่ประกอบด้วยคำบุพบทตามด้วยคำบ่งชี้ (Determiner:DET) และคำคุณศัพท์ (Adjective) และคำนาม

PP ——> PREP + [DET] + [ADJ] + N **ตัวอย่างเช่น**

PREP	DET	ADJ	N
with	my	red	bag

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3
การจัดเก็บฐานข้อมูลไวยากรณ์ในส่วนของพจนานุกรมคำศัพท์

3.1 บทนำ

ในการวิเคราะห์ประโยคภาษาอังกฤษทางด้านโครงสร้างและความหมายให้อยู่ในรูปแบบภาษากลางเพื่อประโยชน์ในการถ่ายทอดข้อมูลในการแปลไปสู่ภาษาเป้าหมายได้อย่างถูกต้องสมบูรณ์นั้น จะต้องอาศัยฐานความรู้เกี่ยวกับไวยากรณ์ภาษาอังกฤษเป็นหลักดังที่กล่าวมาแล้วในบทที่ 2 ในทางปฏิบัตินั้นงานวิจัยได้นำฐานความรู้ทางไวยากรณ์นี้แทรกเข้าในขบวนการวิเคราะห์ในรูปของกฎไวยากรณ์ (Grammar Rules) ดังจะได้กล่าวโดยละเอียดในบทที่ 5 ตลอดจนการแยกส่วนข้อมูลพื้นฐานทางไวยากรณ์ของคำที่มีการกำหนดไว้อย่างแน่นอนโดยจัดเก็บบันทึกข้อมูลเหล่านั้นลงในหน่วยความจำถาวรของระบบการวิเคราะห์ในรูปแบบของฐานข้อมูลทางด้านพจนานุกรมคำศัพท์ เพื่อใช้ประกอบกับกฎไวยากรณ์ในระหว่างขบวนการวิเคราะห์

ข้อมูลพื้นฐานทางไวยากรณ์ของคำศัพท์ภาษาอังกฤษที่จะนำมาประยุกต์ใช้กับขบวนการวิเคราะห์ประโยคในงานวิจัยนี้จะต้องให้รายละเอียดในส่วนรากของคำ (Root of Words) ประเภทของคำ (Category) ชนิดของคำ (A Kind of Words) ตัวแทนของคำนั้นในภาษากลาง (Conceptual Primitive) และสำหรับคำศัพท์ที่มีประเภทของคำเป็นคำกริยาจะต้องมีข้อมูลทางด้านรูปแบบคำกริยา (Verb Patterns) ที่แสดงความสัมพันธ์ทางด้านหน้าที่ของคำที่อยู่ข้างเคียงกับคำกริยานั้น ๆ ตลอดจนข้อมูลทางโครงสร้างและความหมายของคำกริยา (Syntax and Semantic Information) ที่แสดงความสัมพันธ์กับโครงสร้างลึก (Deep Structure) ในรูปแบบการจับคู่ (Mapping) หน้าที่ของคำในประโยคที่วิเคราะห์ไปสู่หน้าที่ของคำนั้นในประโยคภาษากลาง

จากการวิจัยของวิทยานิพนธ์ฉบับนี้พบว่าเพื่อให้ระบบการจัดเก็บพจนานุกรมคำศัพท์ประหยัดเนื้อที่หน่วยความจำถาวรมากที่สุดจะต้องลดความซ้ำซ้อนของข้อมูลคำศัพท์ วิเคราะห์เอาเฉพาะข้อมูลที่จำเป็นตลอดจนได้พิจารณาถึงความคล่องตัวในการเรียกข้อมูลคำศัพท์มาใช้ระหว่างขบวนการวิเคราะห์ประโยคร่วมกับกฎไวยากรณ์ดังกล่าว และคำนึงถึงความไม่ซับซ้อนในการออกแบบและพัฒนาระบบซอฟท์แวร์ ด้วยเหตุผลดังกล่าวมานี้ งานวิจัยจึงได้จัดแบ่งข้อมูลของพจนานุกรมคำศัพท์ออกเป็น 2 ส่วน ได้แก่

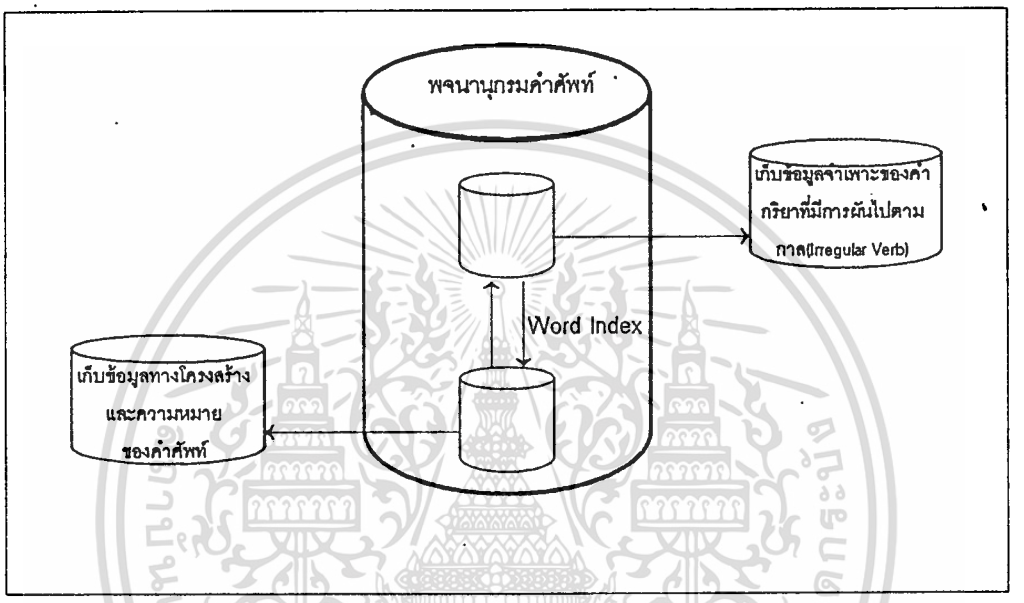
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

(ก) ส่วนข้อมูลทางโครงสร้างและความหมายของคำศัพท์

(ข) ส่วนข้อมูลจำเพาะของคำกริยาที่มีการผันรูปไปตามกาลของประโยค (Irregular Verb)

โดยทั้ง 2 ส่วนนี้จะมีดัชนีของหน่วยคำศัพท์ (Word Index) เป็นตัวเชื่อมดังแสดงในรูปที่ 3.1 รายละเอียดของข้อมูลแต่ละส่วนจะได้อธิบายต่อไป



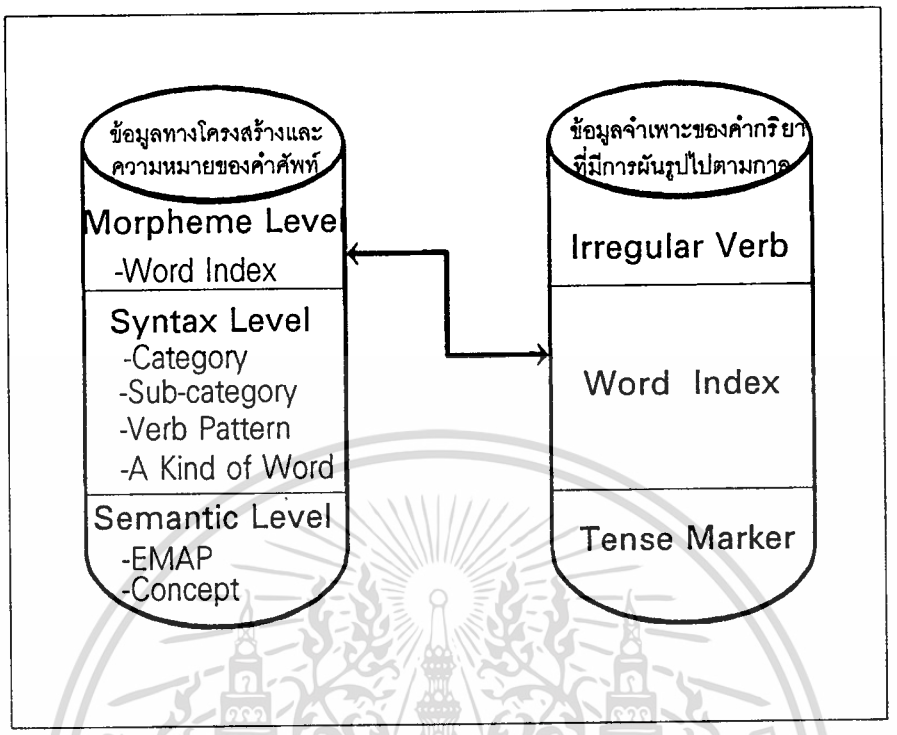
รูปที่ 3.1 แสดงการเชื่อมระหว่างข้อมูลทั้ง 2 ส่วนในพจนานุกรมคำศัพท์

3.2 ส่วนข้อมูลทางโครงสร้างและความหมายของคำศัพท์

ในส่วนนี้จะประกอบด้วย ข้อมูลระดับหน่วยคำ ข้อมูลระดับโครงสร้างของคำ และข้อมูลระดับความหมายของคำ ดังแสดงในรูปที่ 3.2 และตัวอย่างในตาราง 3.1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

21206



รูปที่ 3.2 แสดงข้อมูลทางโครงสร้างและความหมายของทั้ง 2 ส่วน

3.2.1 ข้อมูลระดับหน่วยคำ (Morpheme Level Information)

ข้อมูลส่วนนี้จะเป็นรายละเอียดเกี่ยวกับรากของคำหรือในบางครั้งเรียกว่าเปลือกของคำ (Surface of Words) ซึ่งงานวิจัยนี้ได้กำหนดให้เป็นดัชนีของหน่วยคำ (Word Index) ในการแยกแยะหน่วยคำศัพท์ของประโยคที่วิเคราะห์ และเป็นดัชนีในการค้นหาข้อมูลของคำศัพท์นั้น ๆ ในพจนานุกรม ในวิทยานิพนธ์นี้จะใช้สัญญลักษณ์ย่อเรียกข้อมูลระดับหน่วยคำว่า ENTRY คำเหล่านี้ถ้าเป็นคำนามที่มีการเติม s หรือ es ในรูปของพหูพจน์ จะมีรากคำเป็นคำเอกพจน์ เช่น boxes, cars, companies จะมีรากคำเป็น box, car, company ตามลำดับเป็นต้น

ถ้าเป็นคำกริยาก็จะมีรากคำเป็นต้นคำเดิมที่ยังมิได้มีการผันรูปไปตามกาลหรือตามประธานที่เป็นเอกพจน์หรือพหูพจน์ เช่น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

do, does, did, doing, done	จะมีรากคำเป็น do
be, is, am, are, being, was, were, been	จะมีรากคำเป็น be
has, have, had	จะมีรากคำเป็น have
play, plays, played, playing	จะมีรากคำเป็น play
wash, washes, washed, washing	จะมีรากคำเป็น wash เป็นต้น

ถ้าเป็นคำคุณศัพท์ที่เป็นคำเปรียบเทียบ (Comparison) จะมีรากคำเป็นคำเดิมที่ยังไม่มีการเปรียบเทียบระหว่างสิ่งใด ๆ ที่จะบ่งบอกให้ทราบว่าคนหรือสิ่งใดมีลักษณะที่แตกต่างกันในแง่ของการเปรียบเทียบแต่ยังคงมีรูปเดิมแฝงอยู่ เช่น

big, bigger, bigest	จะมีรากคำเป็น big
small, smaller, smallest	จะมีรากคำเป็น small
thick, thicker, thickest	จะมีรากคำเป็น thick
cold, colder, coldest	จะมีรากคำเป็น cold เป็นต้น

3.2.2 ข้อมูลระดับโครงสร้างของคำ (Syntax Level Information)

ข้อมูลในระดับนี้ ประกอบด้วย ประเภทของคำ ชนิดของคำ และรูปแบบของคำกริยา

3.2.2.1 ประเภทของคำ (Category ใช้สัญญลักษณ์ย่อว่า CAT)

เพื่อการวิเคราะห์ประโยคภาษาอังกฤษได้อย่างถูกต้องสมบูรณ์ จึงได้แบ่งประเภทของคำศัพท์ออกเป็น 10 หมวดคือ หมวดคำนาม หมวดคำสรรพนาม หมวดคำกริยา หมวดคำกริยาช่วย หมวดคำบ่งชี้เฉพาะ หมวดคำคุณศัพท์ หมวดคำกริยาวิเศษณ์ หมวดคำปฏิเสธ หมวดคำบุพบท และหมวดคำเชื่อม ดังที่ได้กล่าวมาแล้วในบทที่ 2 โดยในแต่ละหมวดจะประกอบด้วยหมวดคำย่อย (Sub-category ใช้สัญญลักษณ์ย่อว่า SUBCAT) ต่อไปอีก เช่น หมวดคำนาม ประกอบด้วยหมวดคำย่อย 7 ชนิด ได้แก่ วิสามัญนาม คำนามที่แสดงจำนวนนับ คำนามที่บอกลำดับที่ก่อน-หลัง คำอาการนาม คำนามที่แสดงการขยายคำอื่น ๆ ได้เช่นเดียวกับคำวิเศษณ์ คำนามที่แสดงความเป็นเจ้าของ สามัญนาม โดยวิทยานิพนธ์ได้กำหนดสัญญลักษณ์ย่อกำกับประเภทของคำเหล่านี้ เพื่อสะดวกในการเขียนดังได้กล่าวรายละเอียดแล้วในบทที่ 2 ตัวอย่างของข้อมูลในส่วนประเภทของคำนี้ได้แสดงไว้ในตารางที่ 3.1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.1 ตัวอย่างตารางส่วนข้อมูลทางโครงสร้างและความหมายของคำศัพท์ในพจนานุกรมที่สร้างขึ้นเพื่อใช้สำหรับขบวนการวิเคราะห์ประโยคภาษาอังกฤษ

Morpheme Level	Syntax Level				Semantic Level	
Word Index (ENTRY)	Category (CAT)	Sub-category (SUBCAT)	Verb Pattern (vp)	A Kind of Word (AKO)	EMAP	Concept (CP)
's	n	npss		111		A
a	det	ddan				A
accompany	vt	v	2	2111	sub=agt,dob=obj	ACCOMPANY
accord	vi	v	1	2111	sub=agt	ACCORD
accord	n	ncmn		2211		ACCORD
be	vi,vt	v	1,2	233	sub=agt,dob=obj	BE
be	aux	xxbe				BE
call	n	nvbn		2211		CALL
call	vi,vt	v	1,2	2111	sub=agt,dob=obj	CALL
car	n	ncmn		1322		CAR
her	det	dprs		111		SHE
her	pron	pprs		111		SHE
him	pron	pprs		111		HE
in	adj	alrp		231		IN
in	adv	advv		231		IN
in	prep	prep				IN

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2.2.2 รูปแบบของคำกริยา (Verb Pattern ใช้สัญญลักษณ์ย่อว่า vp)

วิทยานิพนธ์ฉบับนี้ได้รวบรวมรูปแบบกริยาของประโยคใจความเดียวในภาษาอังกฤษ[4],[5] เพื่อใช้เป็นข้อมูลพื้นฐานในการวิเคราะห์ประโยคส่วนที่เกี่ยวกับคำกริยา จากการวิจัยพบว่ารูปแบบของคำกริยาที่ใช้กันมากในประโยคใจความเดียว และสามารถให้ข้อมูลในการวิเคราะห์ครอบคลุมไปถึงประโยคอื่น ๆ ได้ และจากการพิจารณาถึงความซ้ำซ้อนของรูปแบบกริยา ดังกล่าวทั้งหมด จึงได้ทำการแบ่งออกเป็น 4 กลุ่ม ดังแสดงในตารางที่ 3.2 ทั้งนี้เพื่อให้การวิเคราะห์ในส่วนคำกริยาของประโยคกระทำได้รวดเร็วขึ้นดังต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.2 แสดงรูปแบบของคำกริยา

รูปแบบคำกริยา	ตัวอย่างประโยค
vp1: SUB+IV+COMP1	His behaviour and his principle do not accord. The weather clear. The president can not condone killing. My dream may come true one day. That school answers for the child's safety.
vp2: SUB+IV+to+INF+COMP2	Your brother aches to *be* home. My mother want to *play* football in the field. I really want to *talk* to my friend. Her brother likes to *read* in bed. My mother loathes to *travel* by bus.
vp3: SUB+TV+DOB+COMP3	He sometimes accompany a convoy of ships Your friend hate anyone listening. I have been hearing him in a concert. He presented the hospital with an ambulance. That film interests him. He finally gives me some gifts.
vp4: SUB+TV+DOB+to+INF+COMP4	My grandmother has allowed the children to *play* in our garden. He has heard him to *sing* in a concert. His impudence provoked her to *slap* his face. I understood him to *say.* Nobody do not like people *to tell* the truth.

ความหมายและนิยามของสัญญลักษณ์ในตารางที่ 3.2

SUB : Subject หมายถึง ประธานของประโยค

vp : Verb Pattern หมายถึง รูปแบบคำกริยา

IV : Intransitive Verb หมายถึง กริยาที่ไม่ต้องมีกรรมมารองรับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

TV : Transitive Verb หมายถึง กริยาที่ต้องมีกรรมมารองรับ

COMP : Complement หมายถึง ส่วนเติมแต่งประโยคให้มีใจความสมบูรณ์ยิ่งขึ้น

DOB : Direct Object หมายถึง กรรมตรงของประโยค

INF : Infinitive หมายถึง คำกริยาที่อยู่ในรูปกริยาช่องที่ 1 ซึ่งได้แก่รากคำหรือเปลือกคำของคำกริยาที่เป็น ENTRY นั่นเอง

คำหรือวลีที่มีเส้นใต้ หมายถึงคำหรือวลีที่ทำหน้าที่เป็นประธานในประโยค (SUB) เช่น

His impudence, He, I, My brother เป็นต้น

คำหรือวลีที่เป็นตัวปกติหมายถึงคำหรือวลีที่ทำหน้าที่เป็นกริยาประโยค เช่น answers, likes, has heard เป็นต้น

คำหรือวลีที่เป็นตัวเข้ม หมายถึงคำหรือวลีที่เป็นส่วนเติมแต่ง (COMP) ของประโยค เช่น

for the child's safety, in a concert. to my friend เป็นต้น

คำหรือวลีที่ขีดเส้นใต้สองเส้น หมายถึงคำหรือวลีที่ทำหน้าที่เป็นกรรมของประโยค เช่น anyone , him , the hospital เป็นต้น

คำที่เป็นตัวเอียง ขีดเส้นใต้และเป็นตัวเข้ม หมายถึงคำที่เป็นกริยาช่องที่ 1 (INF) เช่น *play ,read, tell* เป็นต้น

<u>รูปแบบคำกริยากลุ่มที่ 1 (Verb Pattern 1 :vp1)</u>

เป็นกลุ่มของรูปแบบคำกริยาที่ไม่ต้องมีกรรมมารับ (IV หรือ Intransitive Verb) ก็ทำให้ประโยคมีใจความสมบูรณ์ได้ทางความหมายในระดับที่ให้ข้อมูลต่อผู้อ่านหรือผู้ฟังว่าประโยคนั้นมีประธานเป็นใครหรืออะไร และประธานของประโยคกำลังทำอะไรอยู่ ตัวอย่างคำกริยาที่ไม่ต้องมีกรรมมารับ เช่น run, go, come เป็นต้น หลังคำกริยาเหล่านี้ ซึ่งเป็นส่วนท้ายของประโยค อาจจะมีคำหรือกลุ่มคำหรือวลีเป็นส่วนเติมแต่งให้ประโยคมีความหมายสมบูรณ์ชัดเจนขึ้นว่า ประธานทำอะไร กับใคร ที่ไหน เมื่อไร เป็นต้น

เช่น He went to school yesterday.

SUB- IV COMP1

แสดงให้เห็นว่าประธาน He(เขา) ได้ go(ไป) school yesterday (โรงเรียนเมื่อวานนี้) แต่ส่วนเติม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แต่งประโยคนี้จะละไว้ (มีหรือไม่มี) ก็ไม่ทำให้ประโยคผิดหลักไวยากรณ์ภาษาอังกฤษ ดังนั้นจึงสรุปรูปแบบคำกริยากลุ่มที่ 1 ได้เป็น

vp1: SUB+IV+COMP1

ตัวอย่างประโยคอื่น ๆ

His behaviour and his principle do not accord.
SUB IV

The weather clear.
SUB IV

The president can not condone killing.
SUB IV COMP1

My dream may come true one day.
SUB IV COMP1

That school answers for the child's safety.
SUB IV COMP1

รูปแบบคำกริยากลุ่มที่ 2 (Verb Pattern 2 :vp2)

ได้แก่รูปแบบของคำกริยาต่อเนื่อง (serial Verb) ที่มีกริยาหลักหรือกริยาตัวแรกของประโยคเป็นกริยาที่ไม่ต้องมีกรรมมารับ (IV)และมี "to" เชื่อมหรือนำหน้าคำกริยาที่ต่อเนื่อง ซึ่งคำกริยาตัวที่สองนี้จะเป็น IV หรือ TV ก็ได้ แต่ต้องอยู่ในรูปกริยาช่องที่ 1 (INFหรือInfinitive) หรือรากของคำกริยานั้น ๆ คำกริยาต่อเนื่องในกลุ่มนี้ จะให้ความหมายของการกระทำของประธานประโยค 2 อย่างต่อเนื่องกัน และอาจจะมีส่วนเติมแต่งประโยคหลังคำกริยาที่ 2 ได้เช่นกัน ตัวอย่างเช่น

I go to buy a book in that shop.
SUB IV INF COMP2

จากการวิจัยของวิทยานิพนธ์ฉบับนี้พบว่า คำกริยาต่อเนื่องมีผลต่อการแปลไปเป็นภาษาอื่น ๆ ในด้านการหาคู่คำกริยาที่ถูกต้องทั้งด้านโครงสร้างและความหมาย ดังนั้นจึงได้จัดให้เป็นรูปแบบคำกริยากลุ่มที่ 2 แยกออกจากคำกริยากลุ่มที่ 1 ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

vp2: | SUB+IV+to+INF+COMP2 |

ตัวอย่างประโยคอื่น ๆ

Your brother aches to be home.
SUB IV INF COMP2

My mother want to play football in the field.
SUB IV INF COMP2

I really want to talk to my friend.
SUB IV INF COMP2

Her brother likes to read in bed.
SUB IV INF COMP2

My mother loathes to travel by bus.
SUB IV INF COMP2

เช่นเดียวกับรูปแบบคำกริยากลุ่มที่ 1 ที่ส่วนเติมแต่งประโยคนั้นอาจจะไม่มีประโยคยังคงความหมายได้เช่น จากตัวอย่างประโยค "Her brother likes to read in bed." ถ้าตัดส่วนเติมแต่งออกจะได้ประโยค "Her brother likes to read" ซึ่งก็มีใจความเช่นเดียวกัน

รูปแบบคำกริยากลุ่มที่ 3 (Verb Pattern 3 :vp3)

เป็นกลุ่มของรูปแบบคำกริยาที่ต้องมีกรรมมารับ (TV หรือ Transitive Verb) จึงจะได้ความสมบูรณ์ขั้นพื้นฐานตามหลักไวยากรณ์ภาษาอังกฤษ ได้แก่ประโยคมีประธานเป็นอะไร ทำอะไร และกระทำต่ออะไร คำที่แสดงถึงสิ่งของที่ถูกกระทำจะเป็นคำนามที่ทำหน้าที่ในประโยคเป็นกรรมตรง (DOB หรือ Direct Object) ตัวอย่างคำกริยา เช่น buy, eat, understand เป็นต้น หลังคำกริยาเหล่านี้จะมีกรรมมารับ และส่วนท้ายของประโยคอาจจะมีกลุ่มคำหรือวลี เป็นส่วนเติมแต่งให้ประโยคมีความหมายสมบูรณ์ชัดเจนขึ้นว่า ประธานทำอะไร กับใคร ที่ไหน เมื่อไร เป็นต้น เช่น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

She buys a pen from this shop.

SUB TV DOB COMP3

แสดงให้เห็นว่าประธาน She(หล่อน) buys(ซื้อ) a pen(ปากกา) from this shop (จากร้านนี้) ซึ่งส่วนเติมแต่งนี้จะละไว้ ก็ไม่ทำให้ประโยคมีความหมายที่ผิดไปจากหลักไวยากรณ์ ดังนั้นจึงสรุปรูปแบบคำกริยากลุ่มที่ 3 เป็นดังนี้

vp3: SUB+TV+DOB+COMP3

ตัวอย่างประโยคอื่น ๆ

He sometimes accompany a convoy of ships.

SUB TV DOB

Your friend hate anyone listening.

SUB TV DOB COMP3

I have been hearing him in a concert.

SUB TV DOB COMP3

He presented the hospital with an ambulance.

SUB TV DOB COMP3

That film interests him.

SUB TV DOB

He finally gives me some gifts.

SUB TV DOB COMP3

จากตัวอย่างประโยค "He finally gives me some gifts." ถ้าตัดส่วนเติมแต่งออกจะได้ประโยคที่เหลือคือ "He finally gives me" ซึ่งจะเห็นว่าก็เป็นประโยคที่มีใจความได้เหมือนกันคือทำให้ทราบว่าประธานของประโยคเป็นใคร และทำอะไร ต่อใคร ซึ่งประโยคก็ยังถูกตามหลักไวยากรณ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปแบบคำกริยากลุ่มที่ 4 (Verb Pattern 4 :vp4)

ได้แก่รูปแบบคำกริยาต่อเนื่อง (Serial Verb) ที่มีกริยาหลักหรือกริยาตัวแรกของประโยคเป็นกริยาที่ต้องมีกรรมมารับ และหลังคำที่เป็นกรรมของประโยคจะมีคำว่า "to" เชื่อมหรือนำหน้าคำกริยาที่ต่อเนื่องซึ่งคำกริยาตัวที่สองนี้จะเป็น IV หรือ TV ก็ได้ แต่ต้องอยู่ในรูปกริยาช่องที่ 1 คำกริยาต่อเนื่องในกลุ่มนี้จะให้ความหมายของการกระทำของประธานประโยค 2 อย่างที่ต่อเนื่องกันและอาจมีส่วนเติมแต่งประโยคหลังคำกริยาที่ 2 ได้ เช่นกัน ตัวอย่างเช่น

I allowed that boy to play in my field.
SUB TV DOB INF COMP4

แสดงให้เห็นประธาน I(ฉัน) ได้ allow(อนุญาต) that boy(เด็กชาย) คนนั้นให้ to play(เล่น) in my field(ในสนามของฉัน) แต่ส่วนเติมแต่งประโยคนี้จะละไว้ก็ได้ จึงได้จัดรูปแบบคำกริยากลุ่มที่ 4 แยกออกจากคำกริยากลุ่มที่ 3 เป็นรูปแบบคำกริยาดังนี้

vp4: SUB+TV+DOB+to+INF+COMP4

ตัวอย่างประโยคอื่น ๆ

My grandmother has allowed the children to play in our garden.
SUB TV DOB INF COMP4

He has heard him to sing in a concert.
SUB TV DOB INF COMP4

His impudence provoked her to slap his face.
SUB TV DOB INF COMP4

I understood him to say.
SUB TV DOB INF

Nobody do not like people to tell the truth.
SUB TV DOB INF COMP4

จากตัวอย่างประโยค "He has heard him to sing in a concert." ถ้าประโยคถูกตัดส่วนเติมแต่งเหลือเป็นประโยค "He has heard him to sing." ประโยคก็ได้ใจความเช่นกัน

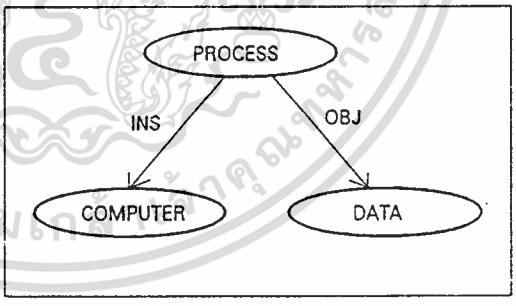
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2.2.3 ชนิดของคำ (A Kind of Words ใช้สัญญลักษณ์ย่อว่า AKO)

ชนิดของคำจะเป็นข้อมูลในพจนานุกรมคำศัพท์ระดับโครงสร้างของคำส่วนที่สามดังแสดงในตารางที่ 3.1 ช่องที่ 5 โดยที่วิทยานิพนธ์นี้ได้ทำการวิเคราะห์จัดแบ่งหมวดหมู่ของชนิดของคำตามความหมายแก่นแท้ (Conceptual Meaning) หรือความหมายลึก (Deep Meaning) ของคำนั้น ๆ ความหมายที่กล่าวนี้จะแตกต่างกันไปตามวัฒนธรรม อารยธรรมหรือประเพณีของการใช้ภาษานั้น ๆ การแบ่งชนิดของคำนี้จะมีประโยชน์ต่อการสร้างความสัมพันธ์ (Relation) ระหว่างคำ การเลือกประธานของประโยคหรือผู้กระทำ การเลือกกรรมของประโยคหรือผู้ถูกกระทำ ตลอดจนการเลือกคำกริยาหรือรูปแบบกริยาที่เหมาะสมได้อย่างถูกต้องยิ่งขึ้นตามวัฒนธรรมของภาษานั้น ๆ ตัวอย่างเช่น

ก) ตัวอย่างการสร้างความสัมพันธ์เพื่อหาประธานและกรรมของประโยค

ตัวอย่างที่ ก.1) Computer processes data (คอมพิวเตอร์ประมวลผลข้อมูล) ในประโยคนี้ถ้าเรามีข้อมูลชนิดของคำ "computer" ว่าเป็น วัตถุที่มีตัวตนแต่เป็นสิ่งไม่มีชีวิต เราก็จะได้ข้อมูลเพิ่มขึ้นว่า สิ่งไม่มีชีวิตแสดงการกระทำไม่ได้ ดังนั้นในประโยคนี้ "computer" เป็นเพียงเครื่องมือของการประมวลผลจะต้องมีบุคคล (Human) ใช้ computer ในการประมวลผลและข้อมูล (data) ก็จะเป็นกรรมของการกระทำ (process) นั่นคือ ภาษากลางของประโยคนี้จะเป็น *

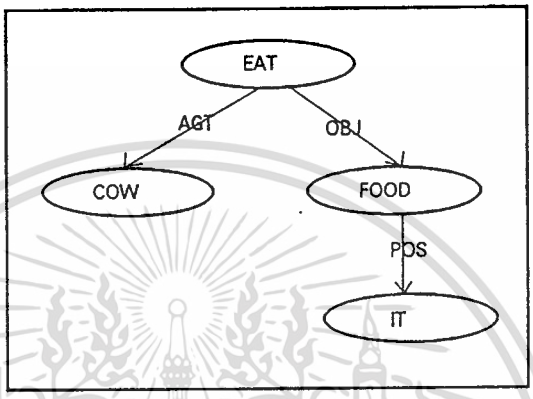


ตัวอย่างที่ ก.2) A cow is eating her food (วัวตัวเมียกำลังกินอาหารของมันอยู่) ในประโยคนี้ ถ้าเรามีข้อมูลชนิดของคำ "cow(วัวตัวเมีย)" ว่าเป็นสัตว์ เราก็จะวิเคราะห์ได้ว่า "her(ของหล่อน)" ในที่นี้มิใช่มนุษย์ ดังนั้นเมื่อทำการแปลเป็นภาษาอื่นเช่น ภาษาไทยตามวัฒนธรรมไทยจะไม่ใช้

* เกี่ยวกับโครงสร้างของภาษากลางที่กำหนดขึ้นในวิทยานิพนธ์นี้ จะได้อธิบายโดยละเอียดในบทที่ 4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สรรพนาม "หล่อน" แทนสัตว์ ทำให้การวิเคราะห์เชิงการแปลงข้อมูลสำหรับคำว่า "her" จะต้องมีข้อมูลบ่งว่า "her" เป็นสรรพนามแทนสัตว์ ซึ่งจากข้อมูลนี้จะทำให้การสร้างภาษาไทย (คือแปลเป็นภาษาไทย) สำหรับคำว่า "her" ในประโยคนี้จะได้เป็น "ของมัน" โดยที่ผู้กระทำกริยา "eat(กิน)" ได้แก่ "cow(วัวตัวเมีย)" และผู้ถูกกระทำคือ "food(อาหาร)" นั่นเอง ทำให้ภาษากลางของประโยคนี้เป็น*



ข) ตัวอย่างการเลือกกริยาหรือรูปแบบกริยา

วัฒนธรรมของบางภาษามีการเลือกคำกริยาไปตามชนิดของคำที่เป็นประธานหรือผู้กระทำ เช่นในภาษาไทยใช้ "กิน" กับสัตว์ ใช้ "รับประทาน" กับมนุษย์ทั่วไป ใช้ "ฉัน" กับพระภิกษุ แต่ในภาษาอังกฤษจะใช้ "eat" กับทั้งสัตว์ และมนุษย์ และใช้ "have หรือ has" กับมนุษย์เท่านั้น เป็นต้น

จากตัวอย่างประโยคที่ได้วิเคราะห์ให้เห็นถึงประโยชน์ของการกำหนดข้อมูลชนิดของคำดังกล่าวข้างต้น วิทยานิพนธ์ได้เล็งเห็นความสำคัญในข้อนี้ จึงได้ทำการวิจัยวิเคราะห์แบ่งชนิดของคำในภาษาอังกฤษให้ได้ข้อมูลที่ละเอียดเพียงพอ โดยจัดแบ่งเป็นหมวดหมู่ให้ลดหลั่นกันไปตามความสัมพันธ์ของคำ เพื่อให้สะดวกต่อการวิเคราะห์ด้วยกฎไวยกรณ์ต่อไป

จากการวิจัยของวิทยานิพนธ์นี้ได้แบ่งชนิดของคำในภาษาอังกฤษออกเป็น 2 กลุ่มใหญ่ ได้แก่ กลุ่มคำรูปธรรม และกลุ่มคำนามธรรม

---

* เกี่ยวกับโครงสร้างของภาษากลางที่กำหนดขึ้นในวิทยานิพนธ์นี้ จะได้อธิบายโดยละเอียดในบทที่ 4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กลุ่มคำรูปธรรม ได้แก่คำที่มีความหมายเกี่ยวกับสิ่งที่มีตัวตนหรืออยู่ในรูปของวัตถุ เช่น man, boy, girl , cat, dog, table, chair เป็นต้น

กลุ่มคำนามธรรม ได้แก่คำที่มีความหมายเกี่ยวกับสภาพของวัตถุ เหตุการณ์ หรือการกระทำ เช่น walk, play, run, work เป็นต้น

ชนิดของคำทั้ง 2 กลุ่ม จะถูกจัดให้มีลักษณะลดหลั่นกัน (Hierarchy) จากสิ่งที่มีความหมายมากกว่าไปสู่สิ่งที่มีความหมายน้อยกว่าตัวอย่างเช่น คำว่า "animal(สัตว์)" กับคำว่า "dog (สุนัข)" และ "cat (แมว)" เมื่อพิจารณาแล้วจะพบว่า "dog หรือ cat" ต่างก็มีความหมายว่าเป็นสัตว์ชนิดหนึ่งแต่จะมีลักษณะเฉพาะลงไปว่าเป็นสุนัขหรือแมว ส่วนคำว่า "animal(สัตว์)" จะมีความหมายถึงสัตว์ทุกชนิดไม่เฉพาะเจาะจง จึงกำหนดได้ว่า "dog หรือ cat" จะเป็นสมาชิกย่อย(sub-set) ของคำว่า "animal" นั่นก็จะแสดงให้ทราบว่า คำว่า "animal" จะมีความหมายมากกว่าหรือกว้างกว่าคำว่า "dog หรือ cat" เป็นต้น โดยที่สิ่งที่กินความน้อยกว่าจะแทนด้วยจำนวนตัวเลขที่มากขึ้นเช่นคำว่า dog, cat เป็นต้น ซึ่งการจัดความหมายของคำที่มีอยู่ในภาษาให้เป็นหมวดหมู่ เพื่อเป็นประโยชน์ในการวิเคราะห์ความหมายของประโยค หรือเป็นการช่วยลดความกำกวม ที่อาจเกิดขึ้นได้ในประโยคดังที่ได้กล่าวมาในตัวอย่างข้างต้น ดังนั้นจึงจำเป็นต้องมีการแบ่งกลุ่มทางความหมายให้กับคำแต่ละคำ โดยจัดหมวดหมู่การแทนชนิดของคำด้วยกลุ่มรหัสตัวเลขทำให้สะดวกในทางซอฟท์แวร์ของการเก็บบันทึกในพจนานุกรม และการอ้างถึงในกฎไวยากรณ์ระหว่างการวิเคราะห์ โดยจัดเป็นลำดับขั้นจากกลุ่มใหญ่จนถึงกลุ่มย่อยเราเรียกว่า AKO (A Kind Of) ดังแสดงในรูปที่ 3.3 ซึ่งมีรายละเอียดดังต่อไปนี้

ชนิดคำกลุ่มที่ 1 กลุ่มคำรูปธรรม (Concrete Words ใช้สัญญลักษณ์ CCRT) แบ่งเป็น

กลุ่มย่อยที่ 1.1 บุคคลและกลุ่มที่ประกอบด้วยคน (Concrete Nominal ใช้สัญญลักษณ์ CNOM) แบ่งเป็น

กลุ่มรหัส 1.1.1 บุคคล (Person) ตัวอย่างเช่น He, Expert, Consumer เป็นต้น

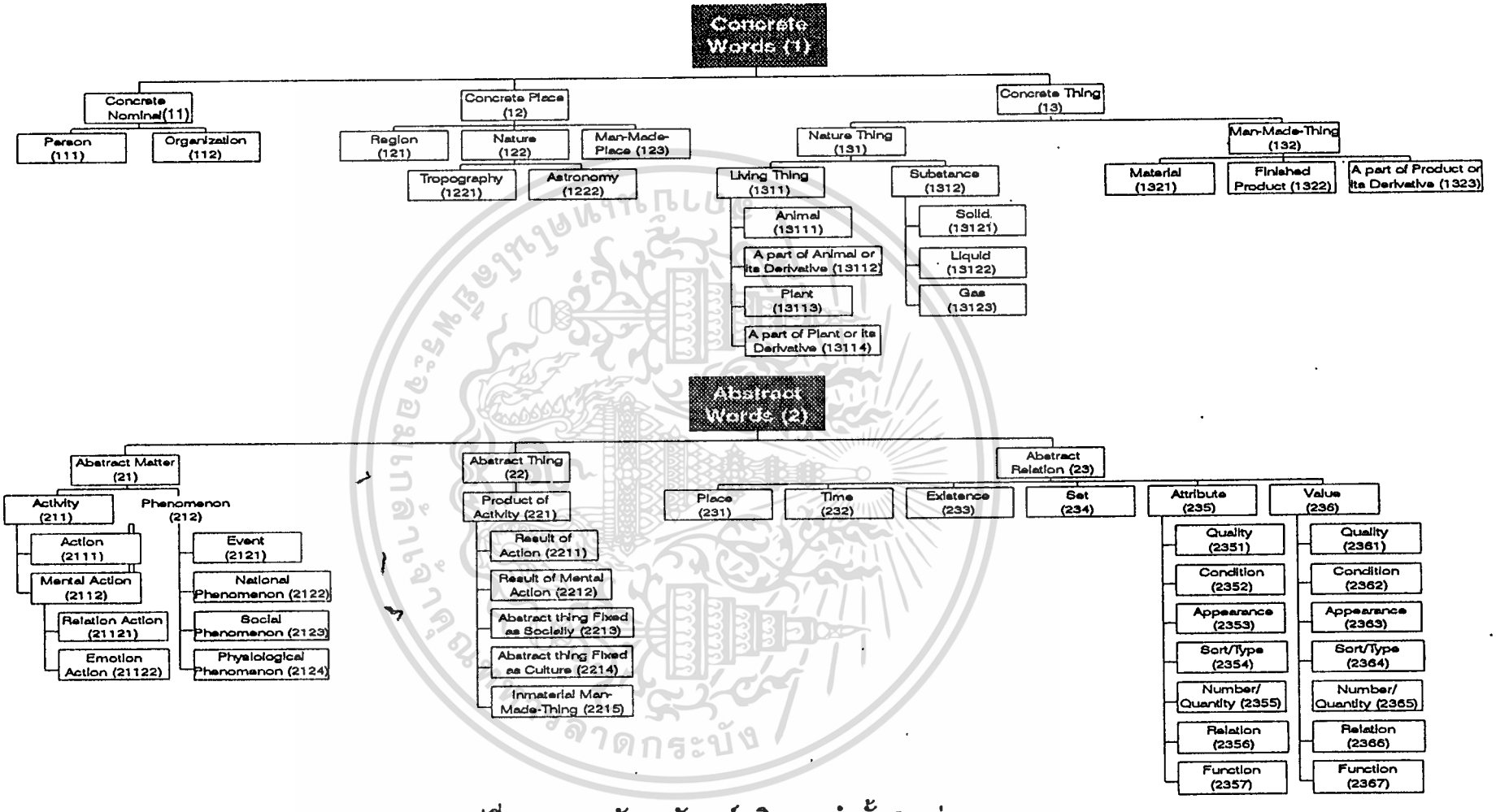
กลุ่มรหัส 1.1.2 กลุ่มหรือองค์กรต่าง ๆ (Organization) ตัวอย่างเช่น NEC, THAI, Malaysia เป็นต้น

กลุ่มย่อยที่ 1.2 สถานที่ (Concrete Place ใช้สัญญลักษณ์ CPLC) แบ่งเป็น

กลุ่มรหัส 1.2.1 สถานที่เป็นดินแดนหรือมีขอบเขตที่แน่นอน (Region) ตัวอย่างเช่น Province, Amphur, Bangkok เป็นต้น

กลุ่มรหัส 1.2.2 สถานที่เป็นธรรมชาติ (Nature) แบ่งเป็น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.3 แสดงสัญญลักษณ์ชนิดของคำทั้ง 2 กลุ่ม

กลุ่มรหัส 1.2.2.1 สถานที่ที่เป็นภูมิประเทศ (Topography) ตัวอย่างเช่น field, mountain, garden, seaboard เป็นต้น

กลุ่มรหัส 1.2.2.2 สถานที่ที่เกี่ยวข้องกับดาราศาสตร์ (Astronomy) ตัวอย่างเช่น world,sun, moon เป็นต้น

กลุ่มรหัส 1.2.3 สถานที่ที่มนุษย์สร้างขึ้น (Man-Made-Place) ตัวอย่างเช่น Post-office, public library, NEC-building เป็นต้น

กลุ่มย่อยที่ 1.3 สิ่งของ (Concrete Thing ใช้สัญญลักษณ์ CTNG) แบ่งเป็น

กลุ่มรหัส 1.3.1 สิ่งของที่เกิดตามธรรมชาติ (Nature Thing) แบ่งเป็น

กลุ่มรหัส 1.3.1.1 สิ่งมีชีวิต (Living Thing) แบ่งเป็น

กลุ่มรหัส 1.3.1.1.1 สัตว์ (Animal) ตัวอย่างเช่น dog, cat, fish เป็นต้น

กลุ่มรหัส 1.3.1.1.2 สัตว์ หรือผลิตผลของสัตว์ (A Part of Animal or its Derivative) ตัวอย่างเช่น hand, leg,bone เป็นต้น

กลุ่มรหัส 1.3.1.1.3 พืช (Plant) ตัวอย่างเช่น flower, rose, rice เป็นต้น

กลุ่มรหัส 1.3.1.1.4 ส่วนต่าง ๆ ของพืช หรือผลิตผลของพืช (A Part of Plant or its Derivative) ตัวอย่าง เช่น root, branch, leaf เป็นต้น

กลุ่มรหัส 1.3.1.2 สสาร (Substance) แบ่งเป็น

กลุ่มรหัส 1.3.1.2.1 ของแข็ง (Solid) ตัวอย่างเช่น ice, gold, silver, gem เป็นต้น

กลุ่มรหัส 1.3.1.2.2 ของเหลว (Liquid) ตัวอย่างเช่น water, oil เป็นต้น

กลุ่มรหัส 1.3.1.2.3 ก๊าซ (Gas) ตัวอย่างเช่น air, nitrogen, carbon-dioxide เป็นต้น

กลุ่มรหัส 1.3.2 สิ่งของที่มนุษย์สร้างขึ้น (Man-Made-Thing) แบ่งเป็น

กลุ่มรหัส 1.3.2.1 วัตถุดิบ (Material) ตัวอย่างเช่น plastic, cloth เป็นต้น

กลุ่มรหัส 1.3.2.2 ผลิตภัณฑ์ (Finished Product) ตัวอย่างเช่น table, car, computer, train เป็นต้น

กลุ่มรหัส 1.3.2.3 ส่วนของผลิตภัณฑ์หรือผลที่ได้จากผลิตภัณฑ์ (A Part of Product or its derivative) ตัวอย่างเช่น lid, razorblade เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชนิดคำกลุ่มที่ 2 กลุ่มคำนามธรรม (Abstract Words ใช้สัญญลักษณ์ ABCT) แบ่งเป็น

กลุ่มย่อยที่ 2.1 นามธรรมที่ครอบคลุมเหตุการณ์และกริยาอาการต่าง ๆ (Abstract Matter ใช้สัญญลักษณ์ AMAT) แบ่งเป็น

กลุ่มรหัส 2.1.1 กริยาและอาการต่าง ๆ (Activity) แบ่งเป็น

กลุ่มรหัส 2.1.1.1 การกระทำ (Action) ตัวอย่างเช่น activity, reaction, play, decide เป็นต้น

กลุ่มรหัส 2.1.1.2 การกระทำที่เกี่ยวกับความรู้สึกนึกคิด (Mental Action) แบ่งเป็น

กลุ่มรหัส 2.1.1.2.1 การกระทำที่ต้องใช้เหตุผลและความคิด (Rational Action) ตัวอย่างเช่น think, control เป็นต้น

กลุ่มรหัส 2.1.1.2.2 การกระทำที่เกี่ยวข้องกับความรู้สึก (Emotion Action) ตัวอย่างเช่น expect, try, like, sorry เป็นต้น

กลุ่มรหัส 2.1.2 ปรากฏการณ์ (Phenomenon) แบ่งเป็น

กลุ่มรหัส 2.1.2.1 เหตุการณ์ (Event) ตัวอย่างเช่น result, social, danger, successful เป็นต้น

กลุ่มรหัส 2.1.2.2 ปรากฏการณ์ธรรมชาติ (National Phenomenon) ตัวอย่างเช่น temperature, pollution เป็นต้น

กลุ่มรหัส 2.1.2.3 ปรากฏการณ์ทางสังคม (Social Phenomenon) ตัวอย่างเช่น second world war เป็นต้น

กลุ่มรหัส 2.1.2.4 ปรากฏการณ์ทางสรีระหรือทางร่างกาย (Physiological Phenomenon) ตัวอย่างเช่น born, digestion เป็นต้น

กลุ่มย่อยที่ 2.2 นามธรรมที่มีลักษณะคล้ายสิ่งที่สามารถจับต้องได้ (Abstract Thing ใช้สัญญลักษณ์ ATHG) แบ่งเป็น

กลุ่มรหัส 2.2.1 ผลของกริยาและอาการต่าง ๆ (Product of Activity) แบ่งเป็น

กลุ่มรหัส 2.2.1.1 ผลของการกระทำ (Result of Action) ตัวอย่างเช่น plan, description, plan, conclude, concept เป็นต้น

กลุ่มรหัส 2.2.1.2 ผลของการกระทำที่เกี่ยวข้องกับความรู้สึกนึกคิด (Result of Mental Action) ตัวอย่างเช่น theory, remembrance, love เป็นต้น

กลุ่มรหัส 2.2.1.3 สิ่งที่เป็นนามธรรมที่เกี่ยวข้องกับสังคม (Abstract Thing Fixed as Socially) ตัวอย่างเช่น songkarn, loikratong เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กลุ่มรหัส 2.2.1.4 สิ่งที่เป็นนามธรรมที่เกี่ยวข้องกับวัฒนธรรม (Abstract Thing Fixed as Culture) ตัวอย่างเช่น olimpic, boxing, subject, cenema เป็นต้น

กลุ่มรหัส 2.2.1.5 สิ่งที่ไม่ใช่วัตถุที่มนุษย์สร้างขึ้น (Inmaterial Man-Made-Thing) ตัวอย่างเช่น program, technology เป็นต้น

กลุ่มย่อยที่ 2.3 ความสัมพันธ์เชิงนามธรรม (Abstract Relation ใช้สัญญลักษณ์ AREL) แบ่งเป็น

กลุ่มรหัส 2.3.1 สถานที่ (Place) ตัวอย่างเช่น upon, righ, left, bottom เป็นต้น

กลุ่มรหัส 2.3.2 เวลา (Time) ตัวอย่างเช่น yesterday, tomorow, today, january เป็นต้น

กลุ่มรหัส 2.3.3 สถานะ (Existence) ตัวอย่างเช่น exist, have, be เป็นต้น

กลุ่มรหัส 2.3.4 กลุ่ม (Set) ตัวอย่างเช่น group, set เป็นต้น

กลุ่มรหัส 2.3.5 คุณสมบัติของวัตถุหรือสิ่งของ (Attribute) แบ่งเป็น

กลุ่มรหัส 2.3.5.1 คุณลักษณะ (Quality) ตัวอย่างเช่น character, specification เป็นต้น

กลุ่มรหัส 2.3.5.2 สภาวะ (Condition) ตัวอย่างเช่น condition, situation เป็นต้น

กลุ่มรหัส 2.3.5.3 ลักษณะที่ปรากฏให้เห็น (Appearance) ตัวอย่างเช่น colour, figure เป็นต้น

กลุ่มรหัส 2.3.5.4 ประเภทหรือชนิด (Sort/Type) ตัวอย่างเช่น type, kind เป็นต้น

กลุ่มรหัส 2.3.5.5 จำนวน (Number/Quantity) ตัวอย่างเช่น highness, volume เป็นต้น

กลุ่มรหัส 2.3.5.6 ความสัมพันธ์ (Relation) ตัวอย่างเช่น relation เป็นต้น

กลุ่มรหัส 2.3.5.7 หน้าที่และประสิทธิภาพ (Function) ตัวอย่างเช่น ability, efficiency เป็นต้น

กลุ่มรหัส 2.3.6 ค่าของคุณสมบัติหรือสิ่งของ (Value) แบ่งเป็น

กลุ่มรหัส 2.3.6.1 ค่าที่แสดงลักษณะ (Quality) ตัวอย่างเช่น hard, high, easy, basic, difficult เป็นต้น

กลุ่มรหัส 2.3.6.2 ค่าที่แสดงสภาวะ (Condition) ตัวอย่างเช่น dengerous, security เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กลุ่มรหัส 2:3.6.3 ค่าที่แสดงลักษณะที่ปรากฏ (Appearance) ตัวอย่างเช่น circle, big, little เป็นต้น

กลุ่มรหัส 2.3.6.4 ค่าที่แสดงชนิดหรือประเภท (Sort/Type) ตัวอย่างเช่น digital-type เป็นต้น

กลุ่มรหัส 2.3.6.5 ค่าที่แสดงจำนวนหรือปริมาตร (Number /Quantity) ตัวอย่างเช่น 25 years, a little of, long time เป็นต้น

กลุ่มรหัส 2.3.6.6 ค่าที่แสดงความสัมพันธ์ (Relation) ตัวอย่างเช่น similar, look like, contact เป็นต้น

กลุ่มรหัส 2.3.6.7 ค่าที่แสดงหน้าที่และประสิทธิภาพ (Function) ตัวอย่างเช่น excellent (quality), high (efficiency) เป็นต้น

3.2.3 **ข้อมูลระดับความหมายของคำ** (Semantic Level Information)

ข้อมูลระดับความหมายของคำศัพท์ในที่นี้มิได้หมายถึงคำแปลของคำจากภาษาหนึ่งไปเป็นอีกภาษาหนึ่ง เช่น "hen" แปลว่า "ไก่" เป็นต้น แต่เป็นการให้ข้อมูลทางด้านกรอบทางความหมายของการใช้คำศัพท์นั้นๆ ร่วมกับคำศัพท์อื่น ๆ ในประโยค เป็นการแสดงความสัมพันธ์ที่อาจจะเกิดขึ้นระหว่างคำที่กำหนดกับคำข้างเคียงอื่นที่ปรากฏในประโยคเดียวกัน ซึ่งงานวิจัยนี้ได้แสดงข้อมูลของคำศัพท์ส่วนนี้ในรูปของ การเชื่อมความสัมพันธ์ทางโครงสร้างที่แสดงหน้าที่ของคำทางความหมายระดับผิว (Surface Meaning หรือ Part of speech) กับหน้าที่ของคำนั้นๆ ทางความหมายระดับลึก หรือความหมายระดับแก่น (Deep Meaning หรือ Conceptual Meaning) และจากการวิจัยพบว่าในการวิเคราะห์ประโยคภาษาอังกฤษ เพื่อนำไปสู่ภาษากลางในการแปลเป็นภาษาอื่น ๆ ด้วยคอมพิวเตอร์นั้น ข้อมูลในส่วนนี้วิเคราะห์สร้างขึ้นได้เฉพาะคำศัพท์ที่เป็นคำกริยาเท่านั้น

นอกจากนั้นจะให้ข้อมูลทางด้านคำที่จะกำหนดเป็นตัวแทนของคำศัพท์ต่าง ๆ ในระดับหน่วยคำที่มีความหมายระดับแก่นในกรอบเดียวกัน เรียกตัวแทนของคำนี้ว่า "มโนภาพของคำ Conceptual Primitive" รายละเอียดของข้อมูลทั้ง 2 ส่วน ที่กำหนดในวิทยานิพนธ์นี้ มีดังต่อไปนี้

<u>3.2.3.1 ข้อมูลส่วนแสดงกรอบความสัมพันธ์ของคำ (Relation Frame)</u>

ดังที่ได้กล่าวมาข้างต้นแล้วว่าในทางภาษาศาสตร์นั้น ข้อมูลในส่วนนี้วิเคราะห์สร้างขึ้นได้เฉพาะคำศัพท์ที่เป็นคำกริยาเท่านั้นซึ่งตรงกับจุดประสงค์และแนวทางของการวิจัยและพัฒนาระบบวิเคราะห์โครงสร้างประโยคภาษาอังกฤษไปสู่ภาษากลางของวิทยานิพนธ์นี้ที่ใช้คำกริยาหลัก (Main Verb) ของประโยคเป็นจุดเริ่มต้นในการวิเคราะห์โครงสร้างของประโยค หาหน้าที่ของคำต่าง ๆ ในประโยคนั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูลส่วนนี้จะแสดงกรอบการเชื่อมความสัมพันธ์ของคำกริยากับคำอื่นๆ ในประโยคทางโครงสร้างที่แสดงหน้าที่ของคำทางความหมายระดับผิวกับหน้าที่ของคำนั้นๆ ทางความหมายลึก ตัวอย่างเช่น

ประโยค : He eats rice. กริยาหลักในที่นี้คือ "eat" เมื่อพิจารณาจะพบว่าในทางความหมายระดับผิว

- คำที่ทำหน้าที่เป็นประธานของประโยคคือ "He" มีค่าหน้าที่ทางความหมายระดับผิวต่อคำกริยาหลัก "eat" = SUB
- คำที่ทำหน้าที่เป็นกรรมตรงของประโยคคือ "rice" มีค่าหน้าที่ทางความหมายระดับผิวต่อคำกริยาหลัก "eat" = DOB

ในทำนองเดียวกันเมื่อพิจารณาทางความหมายระดับแก่นจะได้

- คำที่เป็นผู้กระทำกริยาหลักคือ "He" มีค่าหน้าที่ทางความหมายระดับแก่น = AGT
- คำที่เป็นสิ่งถูกกระทำโดยกริยาหลักคือ "rice" มีค่าหน้าที่ทางความหมายระดับแก่น=OBJ เป็นต้น

การแสดงข้อมูลในส่วนนี้ได้กำหนดรูปแบบข้อมูล (data format) ที่แสดงการเชื่อมความสัมพันธ์ดังกล่าวในรูปของความเท่ากันของหน้าที่ เราเรียกวิธีการนี้ว่า "English Surface and Deep Meaning Mapping" กำหนดชื่อฟังก์ชันสำหรับแสดงค่านี้ว่า EMAP (ย่อจาก English Mapping) ดังนั้นจากตัวอย่างประโยคข้างต้น คำกริยา eat จะมีค่า EMAP =(SUB=AGT, DOB=OBJ) เป็นต้น

ในการวิเคราะห์สร้างข้อมูลส่วนนี้ให้กับคำกริยาต่าง ๆ นั้น จะต้องพิจารณาถึงความเป็นไปได้ที่ถูกต้องตามหลักไวยากรณ์ของภาษา จะต้องรวบรวมรูปแบบประโยคต่าง ๆ ที่เกี่ยวข้องกับคำกริยานั้นมาวิเคราะห์พิจารณาอย่างถี่ถ้วนจึงจะสามารถกำหนดค่า EMAP ให้กับคำกริยานั้น ๆ ได้ วิทยานิพนธ์นั้นได้ทุ่มเทเวลาให้กับการวิจัยในส่วนนี้มากเพื่อให้ได้ข้อมูลที่มากพอต่อการวิเคราะห์ประโยคภาษาอังกฤษต่อไป

อนึ่งค่าหน้าที่ทางความหมายระดับแก่นของคำที่มีต่อคำกริยาหลัก เช่น AGT, OBJ เป็นต้นนั้น วิทยานิพนธ์นี้ ได้อาศัยหลักการวิเคราะห์และกำหนดขึ้นตามวิธีของไวยากรณ์การก (Case Grammar) ซึ่งจะได้กล่าวโดยรายละเอียดในบทที่ 4 ว่าด้วยการกำหนดโครงสร้างของภาษากลางต่อไป ตารางที่ 3.1 ได้แสดงตัวอย่างของการเก็บบันทึกข้อมูลส่วนนี้ในพจนานุกรมคำศัพท์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2.3.2 ข้อมูลส่วนแสดงมโนภาพของคำ (Conceptual Primitive ใช้สัญญลักษณ์ย่อว่า CP)

เป็นส่วนกำหนดคำที่จะเป็นตัวแทนทางความหมายระดับแก่นให้กับกลุ่มคำศัพท์ที่มีความหมายระดับแก่นในกรอบเดียวกัน เพื่อให้เป็นหน่วยคำใหม่สำหรับคำเหล่านั้นในโครงสร้างภาษากลาง เช่น

- his, him, he ต่างมีเปลือกของคำต่างกัน แต่มีความหมายระดับแก่นที่เกี่ยวกับ "เขา" งานวิทยานิพนธ์นี้จะกำหนดให้กลุ่มคำนี้มี CP = HE
- play, playing, played ต่างมีเปลือกของคำต่างกัน แต่มีความหมายระดับแก่นที่เกี่ยวกับ "การเล่น" งานวิทยานิพนธ์นี้จะกำหนดให้กลุ่มคำนี้มี CP =PLAY
- big, bigger , biggest ต่างมีเปลือกของคำต่างกัน แต่มีความหมายระดับแก่นที่เกี่ยวกับ "ใหญ่" งานวิทยานิพนธ์นี้จะกำหนดให้กลุ่มคำนี้มี CP = BIG
- man, men, gentleman ต่างมีเปลือกของคำต่างกัน แต่มีความหมายระดับแก่นที่เกี่ยวกับ "ผู้ชาย" งานวิทยานิพนธ์นี้จะกำหนดให้กลุ่มคำนี้มี CP = MAN เป็นต้น

กล่าวอีกนัยหนึ่งเสมือนกับการกำหนดสัญญลักษณ์ให้เป็นตัวแทนของกลุ่มคำเหล่านั้น เพื่อให้สะดวกต่อการทำความเข้าใจ วิทยานิพนธ์นี้ได้กำหนดสัญญลักษณ์ที่เป็น CP โดยใช้รากศัพท์ที่มีรูปแบบง่ายที่สุด (Simple Form) ของกลุ่มคำเหล่านั้น และแสดงในรูปของตัวอักษรพิมพ์ใหญ่

ดังแสดงในช่องสุดท้ายของตารางที่ 3.1 เกี่ยวกับตัวอย่างการเก็บบันทึกข้อมูลส่วนนี้ในพจนานุกรมคำศัพท์

3.3 ส่วนข้อมูลจำเพาะของคำศัพท์ประเภทคำกริยาที่มีการผันรูปไปตามกาลของประโยค (Irregular Verb Information)

ข้อมูลของคำศัพท์ในพจนานุกรมที่กล่าวมาในข้อ 3.2 ทั้งหมดนั้น เป็นส่วนข้อมูลทางโครงสร้างและความหมายของคำศัพท์ วิทยานิพนธ์นี้ได้จัดเก็บบันทึกข้อมูลเหล่านี้ในรูปของฐานข้อมูลดังแสดงในตารางที่ 3.1 โดยกำหนดให้เปลือกของคำ (ที่กล่าวโดยรายละเอียดในข้อ 3.2.1 แล้ว) เป็นดัชนีหน่วยคำหรือ ENTRY ในการค้นหาและอ่านข้อมูลในส่วนต่าง ๆ ของคำศัพท์นั้น ๆ เมื่อมีคำสั่งให้ค้นหาจากซอฟท์แวร์ของขบวนการวิเคราะห์ประโยค(ที่จะอธิบายในบทที่ 6) ในทางโปรแกรมซอฟท์แวร์แล้วดัชนีหน่วยคำหรือ ENTRY จะเปรียบเสมือนหน่วยชี้หรือพอยน์เตอร์ (Pointer) ที่จะบ่งตำแหน่งของหน่วยบันทึกข้อมูล (Data Records) ในฐานข้อมูลของหน่วยความจำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากการศึกษาค้นคว้าวิจัยเกี่ยวกับคำกริยาของภาษาอังกฤษพบว่า มีคำกริยาจำนวนหนึ่งที่มีการผันรูปคำไปตามกาล (Tense) ที่ไม่เป็นไปตามกฎของการเติม "ed" ต่อท้ายคำปกติ แต่จะมีการแปลงรูปคำเป็นลักษณะเฉพาะของคำกริยานั้น ๆ เช่น ตัวอย่างในตารางที่ 3.3

ตารางที่ 3.3 ตัวอย่างคำกริยาที่มีการผันรูปไปตามกาล

รูปคำกริยาปกติ Original	รูปคำกริยาแสดงอดีตกาล Past Tense	รูปคำกริยาแสดง(อดีต)กาลสมบูรณ์ (Past) Perfect Tense
arise	arose	arisen
be	was,were	been
choose	chose	chosen
do	dis	done
eat	ate	eaten
tall	fell	fallen
get	got	got,gotten
have	had	had
inlay	inlaid	inlaid
know	knew	known
lie	lay	lain
make	made	made
outdo	outdid	outdone
pay	paid	paid
run	ran	run
say	said	said
take	took	taken
underlie	underlay	underlain
wake	woke	woken

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากผลการวิจัยของวิทยานิพนธ์พบว่า การใช้กฎไวยากรณ์เพียงอย่างเดียวในการวิเคราะห์กาลของประโยคที่มีคำกริยาผันรูปไปตามกาลที่ไม่ปกติหรือไม่เป็นไปตามกฎของการเติม "ed" ดังกล่าว จะยุ่งยากและเกิดความผิดพลาดได้ง่าย ตลอดจนการเก็บบันทึกในรูปแบบเดียวกับส่วนข้อมูลทางโครงสร้างและความหมาย จะทำให้เกิดความซ้ำซ้อนของข้อมูลของคำศัพท์กริยาเดียวกันเนื่องจากมีเปลือกของคำต่างกันตามรูปกาลของประโยค เช่น เปลือกคำกริยา do หรือ did หรือ done ต่างมีข้อมูลส่วนโครงสร้างและความหมายเหมือนกันคือมี

CAT	= vi,vt
SUBCAT	= v
vp	= 1,2
AKO	= 2111
EMAP	= sub=agt,dob=obj
CP	= DO

ดังนั้น จึงกำหนดให้รูปแบบคำกริยาปกติ (Original หรือ Infinitive Verb) เป็นตัวแทนของคำกริยาที่ผันรูปไปตามกาล กล่าวอีกนัยหนึ่งก็คือ ให้คำกริยาปกติเป็นดัชนีหน่วยคำหรือ ENTRY ของคำกริยาที่มีการผันรูปไปตามกาล เพื่อเชื่อมโยงไปสู่การค้นหาข้อมูลในส่วนโครงสร้างและความหมาย โดยกำหนดเป็นตารางบันทึกเฉพาะขึ้นเป็นอีกส่วนหนึ่งในฐานข้อมูลพจนานุกรมคำศัพท์เรียกว่า **"ส่วนข้อมูลจำเพาะของคำศัพท์ประเภทคำกริยาที่มีการผันรูปไปตามกาลของประโยค"** และเพื่อให้เป็นประโยชน์ต่อการวิเคราะห์โครงสร้างประโยควิทยานิพนธ์นี้ได้วิเคราะห์แบ่งการเก็บบันทึกข้อมูลในส่วนนี้ออกเป็น 3 หน่วย ได้แก่

ก) หน่วยคำกริยาที่ผันรูปตามกาล (Irregular Verb ใช้สัญญลักษณ์ย่อ IRRV)

ข) หน่วยคำกริยาปกติ ที่จะกำหนดให้เป็น ENTRY ของคำกริยาข้อ ก)

ค) หน่วยแสดงสถานภาพของกาล (Tense Marker) ของหน่วยคำกริยาข้อ ก)

วิทยานิพนธ์นี้ได้รวบรวมหน่วยคำกริยาที่ผันรูปไปตามกาลจากหนังสืออ้างอิง[7] ได้จำนวน 288 คำ ซึ่งจะเป็นส่วนข้อมูลต้นแบบของงานวิจัยนี้

การกำหนดหน่วยแสดงสถานภาพของกาลจะเป็นข้อมูลที่เป็นประโยชน์ต่อขบวนการวิเคราะห์โครงสร้างประโยคในการตัดสิน "กาล (Tense)" ของประโยคได้อย่างถูกต้อง จากตัวอย่างในตารางที่ 3.3 ซึ่งเป็นข้อมูลทางภาษาศาสตร์ วิทยานิพนธ์นี้จึงได้กำหนดรูปแบบของค่าหน่วยแสดงสถานภาพของกาล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ของคำกริยาในทางซอฟท์แวร์ด้วยฟังก์ชัน TENSE ที่มีค่าเพียง 2 กรณี คือ Past Tense หรือ (Past) Perfect Tense และกำหนดให้

PAST แสดงค่าสถานภาพของคำกริยาที่แสดงกาลของประโยคเป็น Past Tense

PPER แสดงค่าสถานภาพของคำกริยาที่แสดงกาลของประโยคเป็น (Past) Perfect Tense

ตัวอย่างเช่น "did" และ "done" จะมี ENTRY เป็น "do"

และ did มีค่าสถานภาพของกาลเป็น TENSE.{PAST}

ส่วน done มีค่าสถานภาพของกาลเป็น TENSE.{PPER} เป็นต้น

ตารางที่ 3.4 ได้แสดงตัวอย่างส่วนการเก็บบันทึกข้อมูลจำเพาะของคำศัพท์ประเภทคำกริยาที่มีการผันรูปไปตามกาลของประโยค ซึ่งจะเป็นข้อมูลอีกส่วนหนึ่งในพจนานุกรมคำศัพท์ ดังที่ได้กล่าวมาแล้วในบทที่ 3.1 (กรุณาดูรูปที่ 3.1 และรูปที่ 3.2)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.4 แสดงตัวอย่างการเก็บบันทึกข้อมูลจำเพาะของคำศัพท์ประเภทคำกริยาที่มีการผันรูปไปตามกาลของประโยคในพจนานุกรมคำศัพท์

Irregular Verb (IRRV)	Word Index (ENTRY)	Tense Marker (TENSE)
abode	abide	PAST,PPER
arose	arise	PAST
arisen	arise	PPER
awoke	awake	PAST
awoken	awake	PPER
been	be	PPER
bore	bear	PAST
borne	bear	PPER
beaten	beat	PPER
became	become	PAST
befell	befall	PAST
befallen	befall	PPER
begot	beget	PAST
begotten	beget	PPER
began	begin	PAST
begun	begin	PPER
beheld	behold	PAST,PPER
bent	bend	PAST,PPER
could	can	PAST,PPER
caught	catch	PAST,PPER
dealt	deal	PAST,PPER
fed	feed	PAST,PPER
gainsaid	gainsay	PAST

จากตารางที่ 3.4 ในช่อง Tense Marker จะเห็นว่าในหนึ่งคำมีโอกาสที่จะแสดงรูปของกาล (Tense) ได้มากกว่า 1 กาล โดยจะมีรูปของกริยาที่ผันรูปไปตามกาลเหมือนกัน เช่น คำว่า "abide" จะมีรูปกริยาที่ผันไปตามกาลเพียงรูปเดียวคือ "abode" ไม่ว่าจะอยู่ใน Past Tense หรือ (Past) Perfect Tense ก็ตาม ซึ่งวิทยานิพนธ์นี้จะมีกฎไวยากรณ์เข้าไปตัดสินความกำกวมได้คือ ถ้าหน้าคำกริยาที่ผันรูปไปเป็น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Verb to have ก็แสดงว่าประโยคนั้นเป็น (Past) Perfect Tense แต่ถ้าไม่มีก็แสดงว่าเป็นประโยค Past Tense เป็นต้น

3.4 วิธีการค้นหาข้อมูลของคำศัพท์ในฐานข้อมูลพจนานุกรม

ในระหว่างการวิเคราะห์ประโยคภาษาอังกฤษด้วยกฎไวยากรณ์ตามขบวนการวิเคราะห์ทางภาษา (ที่จะกล่าวโดยรายละเอียดในบทที่ 5) เพื่อ

การแยกแยะหน่วยคำและค้นหาข้อมูลของหน่วยคำจากพจนานุกรม

↓

หาความสัมพันธ์ระหว่างคำในระดับผิว

↓

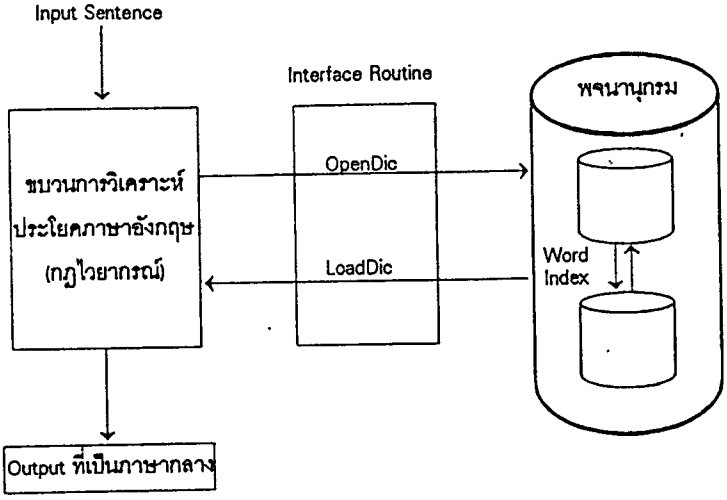
หาความสัมพันธ์ระหว่างคำในระดับแก่น

↓

สร้างภาษากลางให้กับประโยคต้นแบบที่ให้ข้อมูลทางภาษาระดับแก่นของประโยคต้นแบบนั้น

ตามลำดับแต่ละขั้นตอนจำเป็นจะต้องอาศัยข้อมูลทางภาษาคำศัพท์แต่ละคำจากฐานข้อมูลพจนานุกรมที่มีการวิเคราะห์ทางภาษาและเก็บบันทึกล่วงหน้าไว้แล้ว (ดังที่ได้กล่าวมาแล้วจากบทที่ 3 หัวข้อที่ 3.1 ถึง 3.3) วิทยานิพนธ์นี้ได้กำหนดหน่วยโปรแกรมซอฟท์แวร์ที่เป็นตัวเชื่อมหรืออินเตอร์เฟสรูทีน (Interface Routine) ระหว่างโปรแกรมซอฟท์แวร์ของขบวนการวิเคราะห์กับฐานข้อมูลพจนานุกรม เมื่อไรก็ตามที่ขบวนการวิเคราะห์ (Analyser) ต้องการค้นหาข้อมูลของคำศัพท์ในพจนานุกรมจะใช้คำสั่ง OpenDic {คำศัพท์} อินเตอร์เฟสรูทีนก็จะทำการค้นหาข้อมูลของคำศัพท์ในวงเล็บ {คำศัพท์} โดยใช้คำศัพท์นั้น เป็นดัชนีของหน่วยคำ (หมายเหตุ: อัลกอริทึมของขบวนการค้นหาคำศัพท์จะได้กล่าวต่อไป) เมื่อพบคำศัพท์นั้นในพจนานุกรม ก็จะทำการนำข้อมูลของคำศัพท์นั้นทั้งหมดไปให้กับขบวนการวิเคราะห์ด้วยคำสั่ง LoadDic {คำศัพท์} ดังแสดงในรูปที่ 3.4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.4 แสดงโครงสร้างการค้นหาข้อมูลคำศัพท์จากพจนานุกรม

3.4.1 โครงสร้างการเก็บบันทึกข้อมูลคำศัพท์ในฐานข้อมูลของพจนานุกรม

ก่อนที่จะได้อธิบายถึงขบวนการค้นหาคำศัพท์ในฐานข้อมูลของพจนานุกรมนั้น ความจำเป็นที่จะต้องทราบถึงโครงสร้างของการจัดเก็บบันทึกข้อมูลคำศัพท์ในหน่วยความจำหรือฐานข้อมูลของพจนานุกรม

ดังที่ได้กล่าวมาแล้วในหัวข้อ 3.2 และ 3.3 ว่าวิทยานิพนธ์นี้ได้แบ่งข้อมูลคำศัพท์ออกเป็น 2 ส่วนใหญ่ ๆ ได้แก่ ส่วนข้อมูลคำศัพท์ทางโครงสร้างและความหมาย และส่วนข้อมูลจำเพาะของคำศัพท์ประเภทคำกริยาที่มีการผันรูปไปตามกาล เพื่อความสะดวก ความรวดเร็วและความแม่นยำในการค้นหาข้อมูลคำศัพท์ในพจนานุกรมดังกล่าว จึงได้ทำการกำหนดโครงสร้างการจัดเก็บบันทึกข้อมูลดังนี้

3.4.1.1 โครงสร้างของส่วนข้อมูลคำศัพท์ทางโครงสร้างและความหมาย

ข้อมูลคำศัพท์ในส่วนนี้จะถูกจัดเก็บแยกเป็นกลุ่ม ๆ ตามอักขระตัวแรกของคำศัพท์นั้น ๆ เรียงลำดับตัวเลขแสดงตำแหน่งของกลุ่ม (Block Number หรือ Block No.) ตามลำดับของอักขระภาษาอังกฤษ (Alphabet Ordering) ซึ่งมีทั้งหมด 26 ตัวอักษร ทำให้มีกลุ่มคำศัพท์ทั้งหมด 26 กลุ่ม เช่นกลุ่มคำศัพท์ที่มีอักขระตัวแรกเป็น "A" จะอยู่ในกลุ่มที่ 1 มี Block No. = 1 กลุ่มคำศัพท์ที่ขึ้นต้นด้วยอักขระ "Z" จะอยู่ในกลุ่มที่ 26 มี Block No. = 26 เป็นต้น ในทางระบบซอฟท์แวร์แล้วจะทำการจัดเก็บตารางเปรียบเทียบตำแหน่งของกลุ่มคำศัพท์กับอักขระตัวเลขดังตารางที่ 3.5 จากการค้นหาตัวเลขของ Block No. ซึ่ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.5 ตารางกำหนดตำแหน่งของกลุ่มคำศัพท์

First Character of Word's Entry	Block Number of Words
A	1
B	2
C	3
D	4
E	5
F	6
G	7
H	8
I	9
J	10
K	11
L	12
M	13
N	14
O	15
P	16
Q	17
R	18
S	19
T	20
U	21
V	22
W	23
X	24
Y	25
Z	26

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จะเป็นรหัสกุญแจ (Key Pointer) ในการเปิดแฟ้มข้อมูลของกลุ่มคำศัพท์นั้น ๆ ต่อไป โดยที่คำศัพท์ภายในกลุ่มต่าง ๆ จะถูกจัดเก็บเรียงตามลำดับค่ารหัสแอสกี้ (ASCII Code) ของอักขระแต่ละตัวในแต่ละคำศัพท์ตามวิธีการของการเรียงดัชนี (Indexing Order) เช่น ในกลุ่มที่ 1 AA จะมาก่อน AB หรือ ABX จะมาก่อน ABZ เป็นต้น นั่นคือดัชนีของคำ (Word's Entry) จะถูกจัดเก็บเรียงภายในกลุ่มคำศัพท์ตามลำดับของค่าแอสกี้ของอักขระดังตัวอย่างในรูปที่ 3.5

Block No.	Indexing of. Word's Entry	Word's Syntax and Semantic Data	
1	A	CAT=det,SUBCAT=ddan,CP=A	First Word's record in this block
	Accompany	CAT=vi,SUBCAT=v,vp=2, AKO=2111,EMAP=sub=agt, dob=obj,CP=ACCOMPANY	2nd Word's record in this block
	Accord	CAT=n,SUBCAT=ncmn, AKO=2211,CP=ACCORD	
	Alone	CAT=adj,SUBCAT=anrn, AKO=2355,CP=ALONE	
	Arrive	CAT=vi,SUBCAT=v,vp=1, AKO=2111,EMAP=sub=agt, CP=ARRIVE	
	At	CAT=prep,SUBCAT=prep,CP=AT	Last Word's record in this block

รูปที่ 3.5 ตัวอย่างคำศัพท์ในกลุ่มที่ 1 หรือกลุ่มคำศัพท์ในกลุ่มอักขระ "A"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.4.1.2 โครงสร้างของส่วนข้อมูลจำเพาะของคำกริยาที่ผันรูปไปตามกาล

เนื่องจากข้อมูลในส่วนนี้จะมีจำนวนน้อยเมื่อเปรียบเทียบกับส่วนแรกที่กล่าวมา ดังนั้น วิทยานิพนธ์นี้ได้ทำการจัดเก็บข้อมูลส่วนนี้ในรูปของตารางดังเช่นตารางที่ 3.4 ซึ่งกล่าวมาแล้วในข้อ 3.3 โดยจะเรียงข้อมูลส่วนนี้ตามลำดับค่าแอสกี้ของอักขระของกริยาผันรูปในรูปแบบการเรียงดัชนี (Indexing Order) เช่นกัน ดังตัวอย่างในรูปที่ 3.6 และอักขระตัวแรกของ Word's Entry

Word Order	Word Indexing of Irregular Verbs	Word's Entry	Tense Marker	
1	abode	abide	PAST,PPER	First record of Irregular Verb
2	arose	arise	PAST	2nd record of Irregular Verb
3	been	be	PPER	
4	borne	bear	PPER	
5	beaten	beat	PPER	
6	caught	catch	PAST,PPER	
7	dealt	deal	PAST,PPER	
8	fed	feed	PAST,PPER	
m	——	——	——	Last record of Irregular Verb

รูปที่ 3.6 โครงสร้างของการเก็บบันทึกข้อมูลคำกริยาที่ผันรูปตามกาล และ m เป็นลำดับสุดท้าย

ของคำกริยาที่ผันรูปตามกาลนี้จะเป็นดัชนีในการค้นหาข้อมูลทางโครงสร้างและความหมายของคำกริยานี้โดยเปรียบเทียบจากตาราง 3.5 เพื่อไปเปิดแฟ้มข้อมูลของกลุ่มคำศัพท์นั้นต่อไปข้อมูลต้นแบบ ในส่วนนี้ของวิทยานิพนธ์ได้แสดงไว้ในภาคผนวก ข

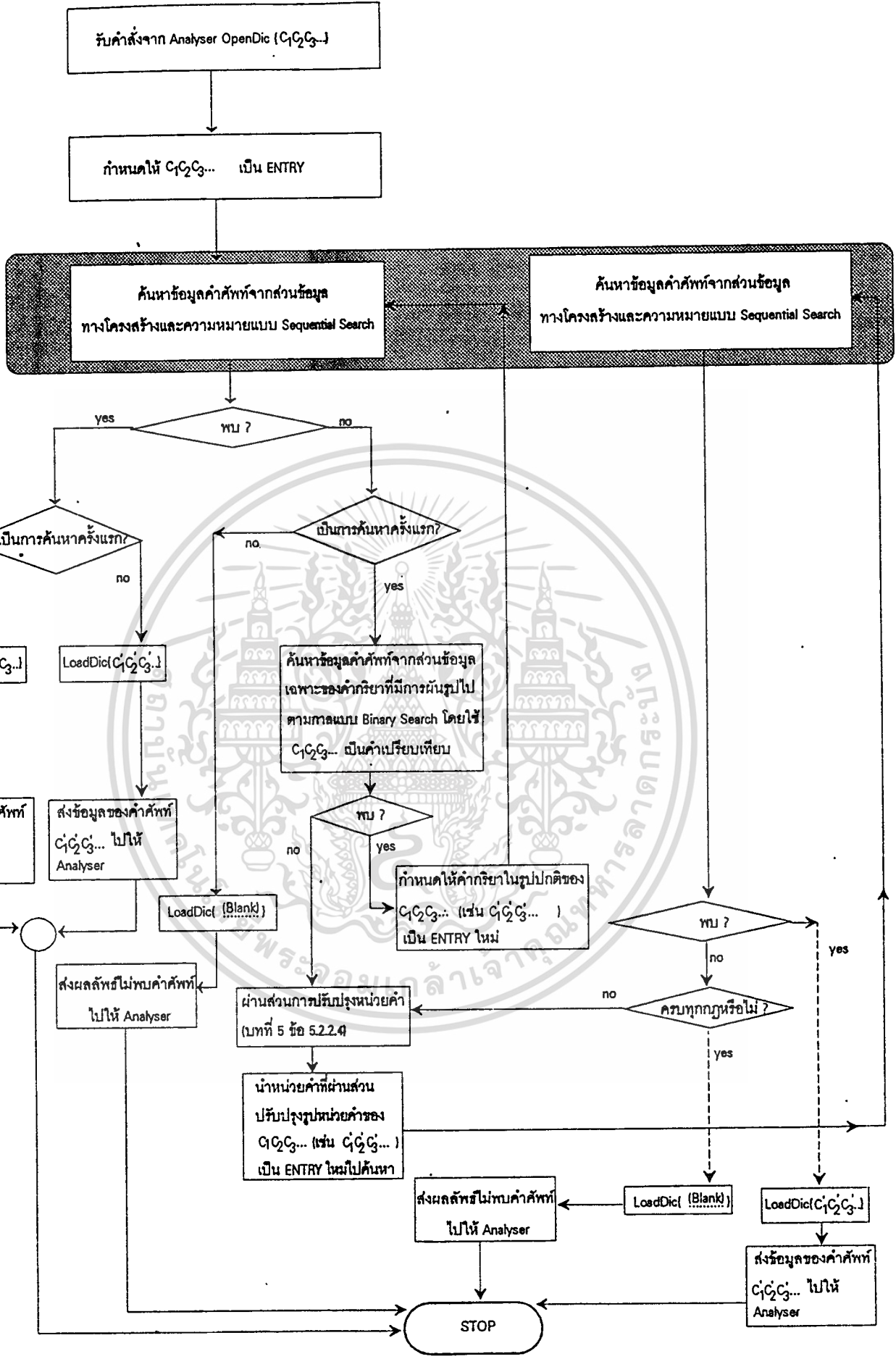
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.4.2 ขบวนการค้นหาข้อมูลคำศัพท์

สมมุติให้ $C_1C_2C_3...C_n$ เป็นลำดับของอักขระของคำที่ต้องการค้นหาข้อมูล (n=1,2,....)

เมื่อมีคำสั่ง OpenDic {$C_1C_2C_3...C_n$} จากขบวนการวิเคราะห์ (Analyser) อิตเตอร์เฟสรูทีนก็จะทำการค้นหาข้อมูลของคำศัพท์นั้นในพจนานุกรมโดยใช้ $C_1C_2C_3...C_n$ เป็น Entry เริ่มต้นโดยจะทำการค้นหาจากส่วนข้อมูลทางโครงสร้างและความหมายก่อนเป็นลำดับแรก เมื่อไม่พบจะตั้งสมมติฐานว่าเป็นคำกริยาผันรูปตามกาล อินเตอร์เฟสรูทีนก็จะทำการค้นหาจากตารางของข้อมูลส่วนคำกริยาผันรูปตามกาล เมื่อพบคำศัพท์นั้นจะนำ Word's Entry ของคำนั้นกลับไปค้นหาข้อมูลทางโครงสร้างและความหมายใหม่ ขบวนการทั้งหมดได้แสดงไว้ในรูปที่ 3.7

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.7 โฟล์วชาร์ตขบวนการค้นหาคำศัพท์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การค้นหาจากส่วนข้อมูลคำศัพท์ทางโครงสร้างและความหมายในวิทยานิพนธ์นี้จะใช้วิธี Sequential Search แบบการเปรียบเทียบค่า ASCII Code ของอักขระแต่ละตัวของคำศัพท์ตามขบวนการในอัลกอริทึมที่ 1

อัลกอริทึมที่ 1 Sequential Search

Step 1.1: หากลุ่มคำศัพท์ในพจนานุกรมที่เริ่มต้นด้วยอักขระ C_1 (อักขระตัวแรก) โดยเปรียบเทียบจากตารางที่ 3.5 เพื่อหารหัสกุญแจเปิดแฟ้ม

Step 1.2: เมื่อพบกลุ่มคำศัพท์ให้ทำการเปรียบเทียบอักขระแต่ละตัวตามลำดับระหว่าง ENTRYกับคำศัพท์ในกลุ่มนี้ตั้งแต่คำแรกสุดโดยการเปรียบเทียบความเท่ากันของอักขระในตำแหน่งเดียวกัน

Step 1.3: จากการเปรียบเทียบถ้า

1.3.1 พบความเท่ากันของทุกอักขระในตำแหน่งเดียวกันให้ทำการส่งข้อมูลทั้งหมดของคำศัพท์นั้นไปให้ Analyser ด้วยคำสั่ง LoadDic {$C_1C_2C_3...C_n$} แล้วไป Step 1.4 หรือ

1.3.2 ถ้าไม่พบความเท่ากันของทุกตัวอักขระในตำแหน่งเดียวกันให้พิจารณาว่า

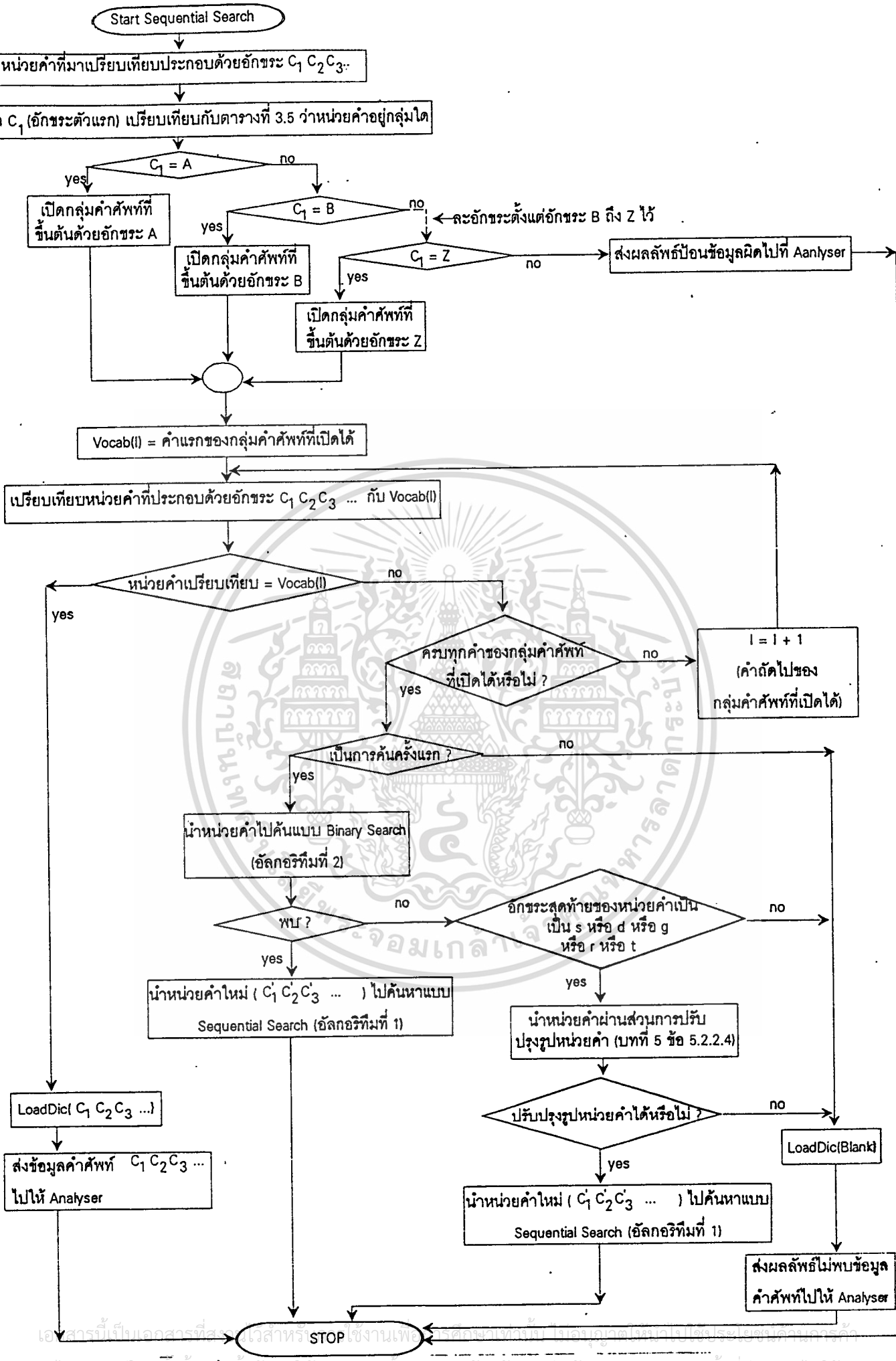
1.3.2.1 ถ้าเป็นการค้นหาครั้งแรกให้ตั้งสมมติฐานว่า $C_1C_2C_3...C_n$ เป็นคำกริยาที่ผันรูปตามกาล กำหนดให้ $C_1C_2C_3...C_n$ เป็นคำเปรียบเทียบแล้วข้ามไปที่ Step 2.1 ของอัลกอริทึมที่ 2 หรือ

1.3.2.2 ถ้าเป็นการค้นหาครั้งที่ 2 ให้ตั้งสมมติฐานว่าไม่มีการบันทึกข้อมูลของคำศัพท์นั้นหรือการสะกดคำศัพท์ $C_1C_2C_3...C_n$ ไม่ถูกต้องให้ส่งผลลัพธ์ไม่พบคำศัพท์ไปให้ Analyser ด้วยคำสั่ง LoadDic {(Blank)} แล้วไป Step 1.4

Step 1.4: หยุดการค้นหาย้อนการทำงานของระบบไปที่ Analyser

สำหรับโฟล์วชาร์ตการทำงานของอัลกอริทึมนี้แสดงในรูปที่ 3.8

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.8 แสดงโฟล์วชาร์ตของอัลกอริทึมแบบ Sequential Search

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น และต้องอ้างอิงถึงเจ้าของผลงานทุกครั้งที่มีการนำไปใช้

ในกรณีที่การค้นหาครั้งแรกไม่พบคำศัพท์ในส่วนข้อมูลทางโครงสร้างและความหมายให้ทำการตั้งสมมติฐานใหม่ว่า $C_1C_2C_3...C_n$ อาจจะเป็นคำกริยาที่ผันรูปไปตามกาล ดังนั้นขบวนการค้นหา คำศัพท์จะข้ามไป (Jumpping) ค้นหาจากส่วนข้อมูลจำเพาะของคำศัพท์ประเภทคำกริยาที่ผันรูปตามกาลในวิทยานิพนธ์นี้จะใช้วิธีค้นหาในส่วนนี้ด้วยวิธีแบบ Binary Search เพื่อกำหนดคำที่จะใช้เปรียบเทียบกับ $C_1C_2C_3...C_n$ โดยการเปรียบเทียบค่าอักขระแต่ละตัวในตำแหน่งเดียวกันตามลำดับ ทั้งนี้เพื่อค้นหา ENTRY ของคำนั้นจากตารางในรูปที่ 3.6 ตามขบวนการที่แสดงในอัลกอริทึมที่ 2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

อัลกอริทึมที่ 2 Binary Search

กำหนดให้ F : เป็นลำดับคำ (Word Order) ของคำแรกในกลุ่มคำกริยาที่ผันรูปตามกาลจากตารางในรูปที่ 3.6 ที่จะใช้เปรียบเทียบ

L : เป็นลำดับคำ (Word Order) ของคำสุดท้ายในกลุ่มคำกริยาที่ผันรูปตามกาลจากตารางในรูปที่ 3.6 ที่จะใช้เปรียบเทียบ

MID : เป็นลำดับของคำกริยาในตารางรูปที่ 3.6 ที่เป็นตัวเปรียบเทียบกับ $C_1C_2C_3...C_n$ ซึ่งจะเป็นคำกึ่งกลางอยู่ระหว่างคำที่ F และคำที่ L นั่นคือ MID = INT[(F+L)/2] *

Step 2.1: เริ่มต้นกำหนดให้ F=1 และ L=m

Step 2.2: คำนวณหาค่า MID = INT [(F+L)/2]

Step 2.3: นำคำกริยาผันรูปที่มีลำดับคำ =MID มาเปรียบเทียบค่าแอสกี้ของอักขระแต่ละตัวตามลำดับกับ $C_1C_2C_3...C_n$ ถ้า

2.3.1 พบความเท่ากันของทุกอักขระในตำแหน่งเดียวกันให้ทำการนำ ENTRY ของคำไปค้นหาข้อมูลใน Step 1.1 ของอัลกอริทึมที่ 1 โดยใช้ ENTRY นี้เป็นคำเปรียบเทียบใหม่ แล้วไป Step 2.6 หรือ

2.3.2 พบความไม่เท่ากันของอักขระตำแหน่งแรกสุดให้ไป Step 2.4

Step 2.4: ทำการเปรียบเทียบค่าแอสกี้ของอักขระในตำแหน่งแรกสุดที่มีค่าไม่เท่ากัน ถ้า

2.4.1 ค่าแอสกี้ของอักขระของคำ ENTRY เริ่มต้นมีค่ามากกว่าอักขระของคำเปรียบเทียบจากตารางในรูปที่ 3.6 ในลำดับ MID ให้ทำการกำหนดค่าเริ่มต้นใหม่โดยให้ F = MID และ L = L แล้วย้อนกลับไป Step 2.2 หรือ

2.4.2 ค่าแอสกี้ของอักขระของคำ ENTRY เริ่มต้น มีค่าน้อยกว่าอักขระของคำเปรียบเทียบจากตารางในรูปที่ 3.6 ในลำดับ MID ให้ทำการกำหนดค่าเริ่มต้นใหม่โดยให้ F = F และ L = MID แล้วย้อนกลับไป Step 2.2

Step 2.5: ถ้าหมดคำเปรียบเทียบให้ไป Step 2.6

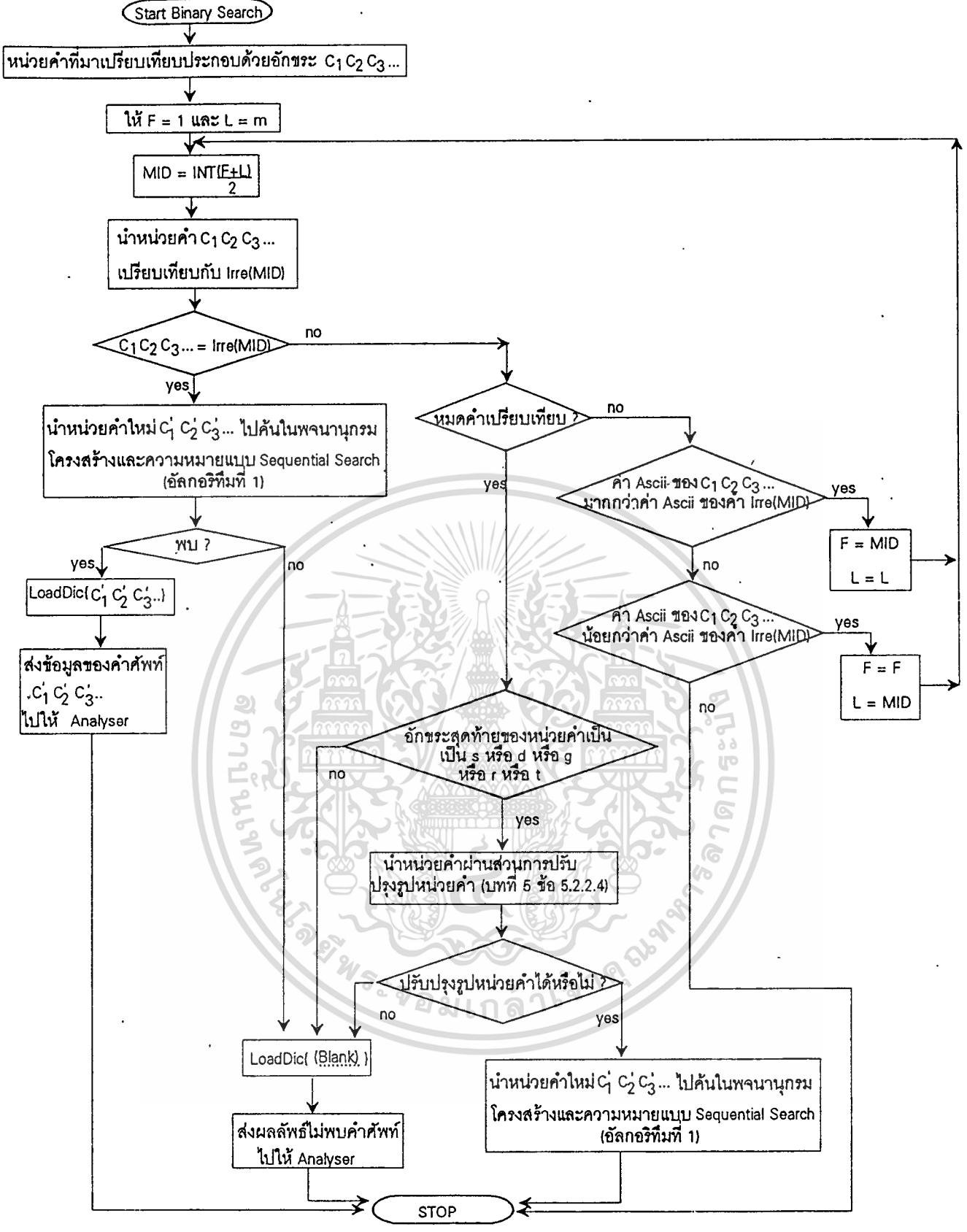
Step 2.6: ถ้าไม่พบคำศัพท์ของคำกริยาที่ผันรูปตามกาลให้นำหน่วยคำไปเข้าส่วนการปรับปรุงรูปหน่วยคำ (บทที่ 5 ข้อ 5.2.2.4)

Step 2.7: หยุดการคำนวณ

สำหรับโฟล์วชาร์ตการทำงานของอัลกอริทึมนี้แสดงในรูปที่ 3.9

* หมายเหตุ INT[(F+L)/2] หมายถึงเลขจำนวนเต็มบวกที่มีค่าน้อยกว่าหรือเท่ากับ (F+L)/2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



<u>หมายเหตุ</u>

Irre(MID) : เป็นคำกริยาที่ผันรูปตามกาลลำดับที่ MID

รูปที่ 3.9 แสดงโฟล์วชาร์ตของอัลกอริทึมแบบ Binary Search

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4
โครงสร้างของภาษากลาง

4.1 บทนำ

ในหลักการของภาษาศาสตร์ทางการประมวลผลภาษาธรรมชาติ (Natural Language Processing) แขนงการแปลภาษาด้วยเครื่อง (Machine Translation) นั้น ได้มีนักวิจัยหลายท่านสรุปผลการวิจัยของตนเองว่าการใช้ภาษากลาง (Intermediate Language หรือ Interlingua ใช้สัญญลักษณ์ย่อว่า IL) เป็นตัวแทนทางภาษาที่จะแสดงความหมายและโครงสร้างของหลาย ๆ ภาษาได้พร้อมกันในรูปแบบเดียว (เรียกว่า Multilanguage Representation)[1] จะเป็นวิธีการแปลหลายภาษาในระบบเดียวที่ให้ผลการแปลรวดเร็วสำหรับการแปลปริมาณมาก ๆ ทั้งยังประหยัดปริมาณซอฟท์แวร์และให้ผลการแปลที่ถูกต้องมากขึ้น จึงนับได้ว่าภาษากลางเป็นรูปแบบสัญญลักษณ์ภาษาทางปัญญาประดิษฐ์ที่ถูกพัฒนาและกำหนดขึ้นให้เป็นตัวแทนของหลาย ๆ ภาษา ที่มีลักษณะความเป็นกลางไม่ขึ้นโดยตรงกับภาษาใดเรียกว่าเป็น "ภาษาอิสระ (Language Independent)" ดังนั้นการใช้ภาษากลางเป็นสื่อของการถ่ายทอดข้อมูลทางภาษาของประโยคจากภาษาหนึ่งไปสู่ภาษาอื่น ๆ โดยเฉพาะในระบบการแปลภาษาด้วยเครื่องคอมพิวเตอร์จึงได้รับความนิยมและเป็นที่ยอมรับของนักวิจัยทั่วไป

การกำหนดส่วนประกอบหลักของภาษากลาง เพื่อแทนความหมายของภาษาต่างๆ นั้นได้มีการวิจัยและพัฒนาและนำมาใช้ในระบบแปลหลายภาษาต่าง ๆ[1] ซึ่งส่วนใหญ่มีการกำหนดโครงสร้างภาษากลางโดยแบ่งเป็น 2 ส่วนได้แก่ ส่วนของหน่วยแสดงความหมาย (Unit of meaning) ของคำในภาษาต้นฉบับ เรียกว่า "Conceptual Primitive (CP)"[6]และส่วนแสดงความสัมพันธ์ของหน่วยความหมาย เพื่อแทนความหมายของเนื้อหาในภาษาต้นฉบับ เรียกส่วนนี้ว่า ไวยากรณ์ภาษากลาง (Interlingual Grammar)

การกำหนดรูปแบบภาษากลางของวิทยานิพนธ์นี้ได้อาศัยหลักเกณฑ์ข้างต้น และได้วิจัยพัฒนาให้เหมาะสมยิ่งขึ้น เพื่อสะดวกต่อการศึกษาทำความเข้าใจและสามารถใช้เป็นตัวกลางในการถ่ายทอดข้อมูลระหว่างภาษาต่าง ๆ ได้ละเอียดและถูกต้องมากที่สุด ตลอดจนสะดวกต่อการนำไปประยุกต์ใช้กับภาษาต่าง ๆ และง่ายต่อการพัฒนาทางซอฟท์แวร์ โดยแบ่งรายละเอียด ออกเป็น 4 ส่วน ได้แก่ ส่วนของหน่วยแสดงความหมายของคำหรือ CP ส่วนแสดงความสัมพันธ์ของหน่วยคำ ส่วนแสดงรูปแบบภาษากลาง และส่วนเสริมหน่วยความหมาย ดังที่จะได้อธิบายในลำดับต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2 หน่วยแสดงความหมายของคำหรือ CP

วิทยานิพนธ์นี้ได้กำหนด CP เพื่อเป็นตัวแทนของคำในภาษาอังกฤษและภาษาอื่น ๆ โดยพิจารณาจากความหมายลึก (Deep Meaning หรือ Concept) ของคำนั้น ๆ ประกอบกับคุณสมบัติ (Property) ของคำนั้น ๆ และจัดหมวดหมู่คำที่มีความหมายลึกเดียวกันไว้ในกลุ่มเดียวกัน แล้วกำหนดให้มี CP ที่เป็นตัวแทนเดียวกัน ในการพิจารณาความหมายลึกเพื่อจัดกลุ่มคำดังกล่าวนี้ จะต้องอาศัยความรู้ทางวัฒนธรรมของการใช้ภาษานั้น ๆ วิทยานิพนธ์นี้ได้ใช้ภาษาอังกฤษเป็นกรณีศึกษา และความรู้ทางภาษาไทยเป็นส่วนอ้างอิงในการพิจารณากำหนด CP อนึ่งการกำหนดสัญญลักษณ์ให้กับ CP นั้น จะเป็นรูปแบบใดก็ได้ที่เป็นความเข้าใจของผู้ใช้ เพื่อเป็นความสะดวกและความเข้าใจต่อการศึกษาวิจัย วิทยานิพนธ์ฉบับนี้ได้กำหนด CP โดยใช้คำภาษาอังกฤษและเขียนด้วยตัวอักษรพิมพ์ใหญ่

ตัวอย่างการกำหนดหน่วยแสดงความหมายของคำ

1) คำว่า eat, ate, eats, eaten, eating ต่างมีรูปแบบส่วนผิวต่างกันไปตามไวยากรณ์ที่ถูกใช้ในประโยคแต่ก็มีความหมายลึกเดียวกันที่แสดงอาการ "กิน หรือรับประทานสิ่งของเข้าสู่ร่างกาย" ในทางภาษากลางจึงจัดให้อยู่ในกลุ่มเดียวกัน และใช้สัญญลักษณ์ CP เป็น "EAT"

2) คำว่า zero, zeroes ต่างมีรูปแบบส่วนผิวที่ต่างกันไปตามไวยากรณ์ที่ถูกใช้ในประโยค แต่ก็มีความหมายลึกอย่างเดียวกันที่มีความหมายว่า "ศูนย์ หรือเลขศูนย์ " ในทางภาษากลางจึงจัดให้อยู่กลุ่มเดียวกัน และใช้สัญญลักษณ์ CP เป็น "ZERO"

3) คำว่า his, he, him ต่างมีรูปแบบส่วนผิวที่ต่างกันไปตามไวยากรณ์ที่ถูกใช้ในประโยค แต่ก็มีความหมายลึกอย่างเดียวกันที่มีความหมายว่า "เขา(ผู้ชาย)" ในทางภาษากลางจึงจัดให้อยู่กลุ่มเดียวกัน และใช้สัญญลักษณ์ CP เป็น "HE"

4) คำว่า big, bigger, biggest ต่างมีรูปแบบส่วนผิวที่ต่างกันไปตามไวยากรณ์ที่ถูกใช้ในประโยค แต่ก็มีความหมายลึกอย่างเดียวกันที่มีความหมายว่า "ใหญ่" ในทางภาษากลางจึงจัดให้อยู่กลุ่มเดียวกัน และใช้สัญญลักษณ์ CP เป็น "ZERO"

5) คำว่า woman, women ต่างมีรูปแบบส่วนผิวที่ต่างกันไปตามไวยากรณ์ที่ถูกใช้ในประโยคแต่ก็มีความหมายลึกอย่างเดียวกันที่มีความหมายว่า "ผู้หญิง" ในทางภาษากลางจึงจัดให้อยู่กลุ่มเดียวกัน และใช้สัญญลักษณ์ CP เป็น "WOMAN"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3 การกำหนดส่วนแสดงความสัมพันธ์ของหน่วยความหมาย

วิทยานิพนธ์นี้ได้ทำการกำหนดส่วนแสดงความสัมพันธ์ของหน่วยความหมาย โดยอาศัยทฤษฎีทางการกสัมพันธ์ (Case Relation) เป็นแม่แบบ[6],[7] และได้ทำการแบ่งรายละเอียดของการพิจารณากำหนดเพื่อให้เหมาะสมออกเป็น 2 ส่วน ได้แก่ ส่วนแสดงความสัมพันธ์ระหว่างหน่วยความหมายที่เกี่ยวกับวัตถุ และส่วนแสดงความสัมพันธ์ระหว่างหน่วยความหมายภายในวัตถุ ในวิทยานิพนธ์นี้จะใช้ CASE เป็นสัญญลักษณ์แทนหน่วยการกสัมพันธ์

4.3.1 ส่วนแสดงความสัมพันธ์ระหว่างหน่วยความหมายที่เกี่ยวกับวัตถุ

เป็นการแสดงความสัมพันธ์ระหว่างเหตุการณ์หรือสถานะกับวัตถุ (Deep Case Relation) หรือความสัมพันธ์ระหว่างหน่วยความหมายที่เป็นคำนามกับคำกริยา ความสัมพันธ์ในส่วนนี้เป็นความสัมพันธ์ทางการก (Case Relation) ซึ่งมี 6 แบบดังนี้

1) **ผู้กระทำ (AGT=AGenTive)** หมายถึงความสัมพันธ์ระหว่างคำนามกับกริยาในลักษณะที่คำนามเป็นผู้ก่อการกระทำ หรือทำให้เกิดเหตุการณ์ขึ้น จะรวมทั้งสิ่งมีชีวิตและสิ่งไม่มีชีวิต แต่เป็นลักษณะทำให้เกิดเหตุการณ์นั้นขึ้น ซึ่งจะมีลักษณะดังต่อไปนี้

ก. AGT ที่เป็นสิ่งมีชีวิตและมีความหมายถึงว่าเป็นคน หรือสัตว์ก็ได้ เช่น

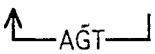
ผู้กระทำอาการ ; He entered the house.[*]

หมายถึง He เป็น AGT ของกริยา enter

ผู้กระทำอาการรู้สึก ; I understood him to say.

ข. AGT ที่เป็นสิ่งไม่มีชีวิตและทำให้เกิดเหตุการณ์ขึ้น เช่น

The wind fills the sail.



---

[*] หมายเหตุ เครื่องหมาย A B หมายถึงหน่วยคำ A สัมพันธ์กับหน่วยคำ B โดยที่ A เป็น AGT ของ B สำหรับ

ความสัมพันธ์อื่น ๆ จะอธิบายได้ในทำนองเดียวกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2) ผู้ถูกกระทำ (OBJ=OBJect) หมายถึงความสัมพันธ์ระหว่างคำนามกับคำกริยาในลักษณะที่คำนามเป็นผู้ถูกกระทำหรือผู้ที่รับสภาพหรือได้รับผลกระทบกระเทือนจากการกระทำรวมทั้งสิ่งไม่มีชีวิตที่ถูกกระทำโดยกริยา

ก. OBJ สิ่งมีชีวิตที่ได้รับผลจากการกระทำโดยกริยานั้นๆ เช่น

My mother hits him violently.

OBJ

หมายถึง him เป็น OBJ ของกริยา hit

ข. OBJ สิ่งไม่มีชีวิตที่ได้รับผลกระทบจากกริยานั้นๆ เช่น

Your brother plays the football in the field.

OBJ

3) เครื่องมือ (INS=INStrument) หมายถึงความสัมพันธ์ระหว่างคำนามกับกริยาในลักษณะที่คำนามเป็นเครื่องมือที่ใช้ทำกริยา หรือวิธีการที่ทำให้เกิดเหตุการณ์

ก. INS ที่มีลักษณะเป็นเครื่องมือที่ใช้ทำกริยา เช่น

He has assisted her with my car.

INS

หมายถึง car เป็น INS ของกริยา assist

ข. INS ที่เป็นวิธีการให้เกิดเหตุการณ์ เช่น

She solves this problem by matric theory.

INS

4) จุดหมาย (TAR=TARget) หมายถึงความสัมพันธ์ระหว่างคำนามกับกริยาในลักษณะบ่งบอกถึงจุดมุ่งหมาย จุดปลายทาง จุดสิ้นสุด ของกริยา หรือเป้าหมายของกริยาที่มิได้เป็นกรรมแท้ของกริยาได้แก่

ก. TAR ที่เป็นจุดมุ่งหมายหรือจุดหมายปลายทางของการกระทำ เช่น

He presented an ambulance to the hospital.

TAR

หมายถึง hospital เป็น TAR ของกริยา present

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข. TAR ที่เป็นสถานะสุดท้ายของการเปลี่ยนแปลง เช่น

The apple changes from green to red.

ค. TAR ที่เป็นจุดหมายเกี่ยวกับสถานที่ เช่น

I ache to go home.

5) **ต้นกำเนิด** (SOR=SOuRce) หมายถึงความสัมพันธ์ระหว่างคำนามกับกริยาในลักษณะบอกจุดตั้งต้นของการกระทำ ซึ่งแสดงโดยกริยาหรือที่มาของสิ่งของที่ส่งจากผู้ให้ไปยังผู้รับ หรือที่มาของเหตุการณ์ เช่น

That old man has alighted from a bus.

หมายถึง bus เป็น SOR ของกริยา alight

6) **เวลา** (TIM=TIMe) หมายถึงความสัมพันธ์ระหว่างคำนามกับกริยาในลักษณะที่กล่าวถึงเวลาที่เริ่มต้น เวลาจบสิ้น ช่วงเวลาที่ใช้ในกริยานั้น ๆ โดยลักษณะของเวลาจะกำหนดโดย Subfeature หรือ สภาพที่แสดงโดยกริยา ได้แก่

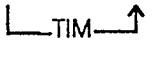
ก. TIM ที่เป็นเวลาเริ่มต้น เช่น

She get up at 7 o'clock.

หมายถึง 7 o'clock เป็น TIM ของกริยา get up

ข. TIM ที่เป็นเวลาจบสิ้น เช่น

The meeting adjourned at five o'clock.



ค. TIM ที่เป็นช่วงเวลา เช่น

She took a bus for 2 hours.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ง. TIM ที่เป็นสภาพของการแสดงโดยกริยา เช่น

She go to school every day.

└TIM┘

4.3.2 **ส่วนแสดงความสัมพันธ์ระหว่างหน่วยความหมายภายในวัตถุ** (Object Inner Concept Relation) หรือความสัมพันธ์ระหว่างหน่วยความหมายที่ไม่ใช่คำนามกับคำกริยา ซึ่งมี 9 แบบดังนี้

1) **ตัวเชื่อมขนาน** (PAR=PARallel) หมายถึง ความสัมพันธ์ระหว่างคำนามมากกว่าหนึ่งกับกริยาในลักษณะที่กระทำอาการร่วมกัน หรือคำนามที่สามารถกระทำอาการมากกว่าหนึ่งอาการในเวลาเดียวกัน หรือเหตุการณ์ที่เกิดต่อเนื่องกัน ได้แก่

ก. PAR ของการกระทำที่ร่วมกันของคำนามมากกว่าหนึ่งเช่น

John and Merry heard the news.

└PAR┘

หมายถึง Merry เป็น PAR ของคำนาม John

ข. PAR ของการกระทำอาการที่มากกว่าหนึ่งในเวลาเดียวกัน เช่น

I really want to talk to my friend.

└PAR┘

ค. PAR ของเหตุการณ์ที่เกิดต่อเนื่องกัน เช่น

He walk and then he crawled.

└PAR┘

2) **ส่วนขยาย** (MOD=MODifier) หมายถึงความสัมพันธ์ระหว่างคำกริยากับหน่วยคำที่เป็นส่วนขยายกริยา เช่น

I really want to talk to my friend.

└MOD┘

หมายถึง really เป็น MOD ของกริยา want

3) **ส่วนขยายคำนาม** (NMOD=Noun MODifier) หมายถึงความสัมพันธ์ระหว่างคำที่อยู่หน้านามซึ่งทำหน้าที่เป็นส่วนขยายคำนาม เช่น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

The pupils as sembled in the school hall.

↑NMOD┘

หมายถึง school เป็น NMOD ของคำนาม hall

4) **จำนวน** (QUANT=QUANTity) หมายถึงความสัมพันธ์ระหว่างคำนามกับคำที่บ่งบอกถึงปริมาณหรือ จำนวนของคำนามนั้น ๆ เช่น

The two houses adjoin.

↑QUANT┘

หมายถึง two เป็น QUANT ของคำนาม house

5) **ตัวบ่งชี้** (REF=REFerence) หมายถึงความสัมพันธ์ระหว่างคำนามกับคำคุณศัพท์ที่ชี้เฉพาะ เช่น

That school answers for the child's safety.

↑REF┘

หมายถึง That เป็น REF ของคำนาม school

6) **ลักษณะ** (ATT=ATTribute) หมายถึงความสัมพันธ์ระหว่างคำที่บ่งบอกลักษณะหรือคุณสมบัติ กับคำนามนั้น ๆ เช่น

His sister drives slowly the red car .

↑ATT┘

หมายถึง red เป็น ATT ของคำนาม car .

7) **ผู้เป็นเจ้าของ** (POSS=POSsessor) หมายถึงความสัมพันธ์ระหว่างคำนามกับคำคุณศัพท์แสดงความเป็นเจ้าของ เช่น

My mother hits him violently.

↑POSS┘

หมายถึง My เป็น POSS ของคำนาม mother

8) **ค่า** (VAL=Value) หมายถึงความสัมพันธ์ระหว่างคำนามกับคำที่บ่งบอกปริมาณของคำนั้นๆ ที่ไม่ได้บอกจำนวนที่แน่นอน เช่น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

He finally gives me some gifts.

↑VAL┘

หมายถึง some เป็น VAL ของคำนาม gift

9) ชื่อ (NAM=NAMing) หมายถึงความสัมพันธ์ระหว่างคำนามกับคำนามในลักษณะที่คำนามตัวหนึ่งทำให้ความหมายของคำนามทั่วไปชัดเจนยิ่งขึ้น เช่น

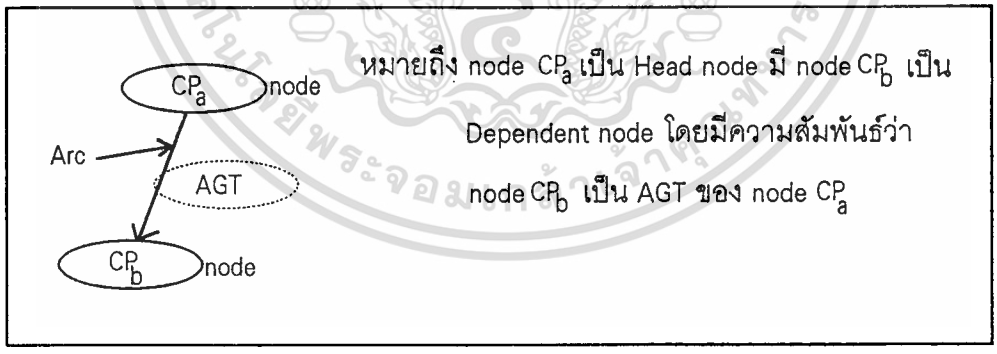
Today,I will see the American football .

↑NAM┘

หมายถึง American เป็น NAM ของคำนาม football

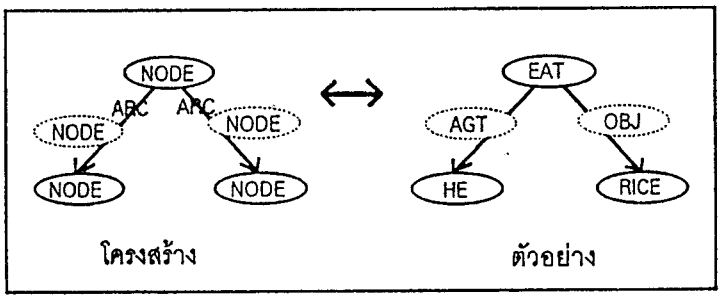
4.4 ส่วนแสดงรูปแบบภาษากลางทางโครงสร้าง

ในส่วนนี้จะเป็นลักษณะโครงข่ายแสดงความสัมพันธ์ โดยแสดงโครงสร้างของภาษากลางในรูปแบบของต้นไม้[9] ประกอบด้วย "โนด (NODE)" และ "อาร์ค (ARC)" โดยโนดจะแทน CP และ CASE ส่วนอาร์คจะแสดงความสัมพันธ์ของโนดในลักษณะของคู่ความสัมพันธ์ที่ขึ้นต่อกัน (Head-Dependent Relation) ตัวอย่างเช่น



วิทยานิพนธ์ได้ใช้วิธีนี้แสดงภาษากลางของผลการวิเคราะห์โครงสร้างประโยคภาษาอังกฤษ ทำให้เข้าใจง่ายและสะดวกต่อการศึกษา ตลอดจนสะดวกต่อการพัฒนาเป็นรูปแบบของซอฟท์แวร์ดังเช่น ตัวอย่างประโยค "He eats rice." จะมีคำว่า "HE, EAT, RICE, AGT และOBJ"เป็นโนดของโครงข่ายความสัมพันธ์ ส่วนอาร์คของโครงข่ายความสัมพันธ์ในวิทยานิพนธ์นี้ใช้เครื่องหมายลูกศรเป็นตัวบอกความสัมพันธ์ระหว่างโนดโดยมีหัวลูกศรเป็นการบ่งบอกทิศทางความสัมพันธ์ดังแสดงในรูปที่ 4.1 ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.1 แสดงโครงข่ายความสัมพันธ์ของภาษากลาง และตัวอย่างประโยค "He eats rice."

4.5 ส่วนเสริมหน่วยความหมาย

เป็นส่วนที่ทำให้ประโยคมีความชัดเจนมากยิ่งขึ้นเช่นทำให้ทราบว่าประโยคนั้นเกิดขึ้นที่ไหน เวลาใด โดยในส่วนเสริมที่บอกเหตุการณ์ต่างๆ ว่าเกิดขึ้นที่ไหน เวลาใดนั้นเรียกว่า ส่วนเสริมคำนามเช่น คำบุพบท "of, from, at" เป็นต้น และส่วนเสริมที่ ทำให้ทราบว่าประโยคนั้นผู้กระทำอาการมีลักษณะการกระทำอย่างไร เช่นกำลังทำอยู่(Present) หรือได้ทำไปแล้ว(Past) หรือ สามารถทำได้(can) เป็นต้น ในในส่วนนี้เรียกว่า ส่วนเสริมคำกริยาดังนั้นจึงได้แบ่งส่วนเสริมหน่วยความหมาย แบ่งเป็น 2 ชนิดได้แก่ ส่วนเสริมคำนาม (Noun Attribute) และ ส่วนเสริมคำกริยา (Verb Attribute) ดังมีรายละเอียดดังนี้

4.5.1 ส่วนเสริมคำนาม

ส่วนเสริมคำนามนี้ได้แก่ Article คำบุพบท และคำเชื่อม โดย Article จะบอกรายละเอียดเกี่ยวกับคำนามคือ คำนามเอกพจน์ คำนามพหูพจน์ คำนามชี้เฉพาะ คำนามไม่ชี้เฉพาะ คำบุพบทจะเป็นตัวชี้การก (Case Marker) และคำเชื่อมจะเป็นตัวบอกความสัมพันธ์ระหว่างมโนภาพภายในวัตถุ (Object Inner Concept Relation) ดังนี้

1) คำบุพบท "of" เป็นตัวชี้การก POSS ตัวอย่างเช่น

He sometimes accompanys a convoy of ships.

└———POSS———┘

2) คำบุพบท "from" เป็นตัวชี้การก SOR ตัวอย่างเช่น

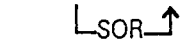
The old man has alighted from a bus.

└SOR┘

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

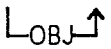
3) คำบุพบท "in" เป็นตัวชี้การก SOR ตัวอย่างเช่น

The pupils assembled in the school hall.



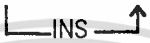
4) คำบุพบท "with" เป็นตัวชี้การก OBJ ตัวอย่างเช่น

You agree with him.



5) คำบุพบท "with" เป็นตัวชี้การก INS ตัวอย่างเช่น

He has assisted her with my car.



6) คำบุพบท "by" เป็นตัวชี้การก INS ตัวอย่างเช่น

My mother loathes to travel by bus.

7) คำบุพบท "to" เป็นตัวชี้การก TAR ตัวอย่างเช่น

He presented an ambulance to the hospital.

8) คำบุพบท "for" เป็นตัวชี้การก TAR ตัวอย่างเช่น

That school answers for the child's safety.

9) คำบุพบท "at" เป็นตัวชี้การก TAR ตัวอย่างเช่น

We have arrived at a village.

10) คำบุพบท "at" เป็นตัวชี้การก TIM ตัวอย่างเช่น

The meeting adjourned at five o'clock.

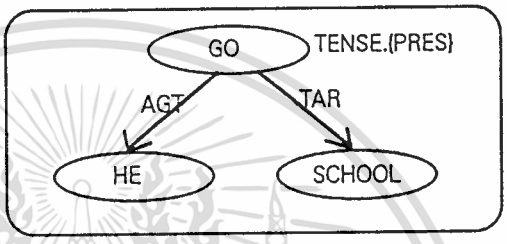
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.5.2 ส่วนเสริมคำกริยา

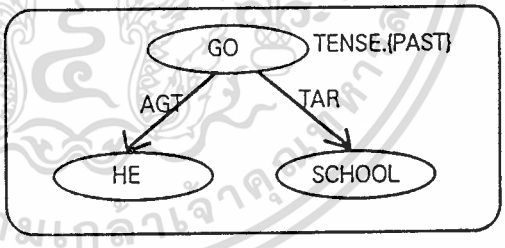
ส่วนเสริมคำกริยานี้ได้แก่คำช่วยกริยา (Auxiliary Verb) คำช่วยกริยานี้จะบอกส่วนเสริม (Modality) ของหน่วยความหมายที่เป็นเหตุการณ์หรือสถานะ ซึ่งในงานวิจัยนี้ส่วนเสริมหน่วยความหมายที่เป็นเหตุการณ์หรือสถานะ แบ่งเป็น 4 ชนิดได้แก่

1) กาล (Tense) แบ่งย่อยเป็น

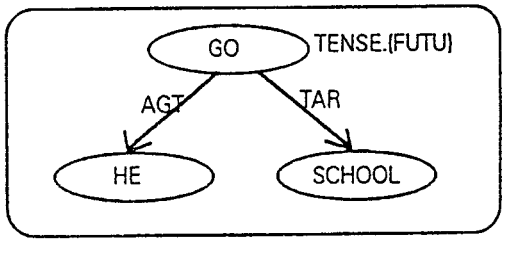
1.1) Present Tense ในระบบซอฟท์แวร์ของขบวนการวิเคราะห์ กำหนดค่าแสดงสถานะภาพนี้เป็น TENSE.(PRES) ตัวอย่างเช่นประโยค "He goes to school." จะมีการเก็บข้อมูลดังภาพข้างล่างนี้



1.2) Past Tense ในระบบซอฟท์แวร์ของขบวนการวิเคราะห์ กำหนดค่าแสดงสถานะภาพนี้เป็น TENSE.(PAST) ตัวอย่างเช่นประโยค "He went to school." จะมีการเก็บข้อมูลดังภาพข้างล่างนี้

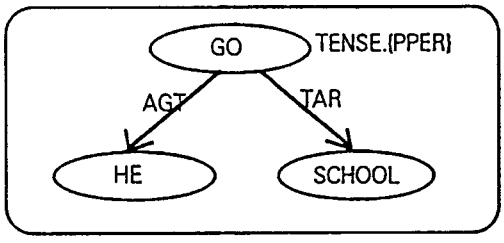


1.3) Future Tense ในระบบซอฟท์แวร์ของขบวนการวิเคราะห์ กำหนดค่าแสดงสถานะภาพนี้เป็น TENSE.(FUTU) ตัวอย่างเช่น ประโยค "He will go to school." จะมีการเก็บข้อมูลดังภาพข้างล่างนี้



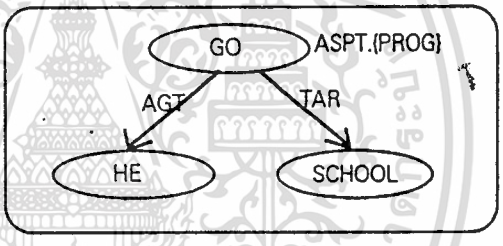
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.4) Present Perfect Tense ในระบบซอฟท์แวร์ของขบวนการวิเคราะห์ กำหนดค่าแสดงสถานะภาพนี้เป็น TENSE.(PPER) ตัวอย่างเช่น ประโยค "He has gone to school." จะมีการเก็บข้อมูลดังภาพข้างล่างนี้

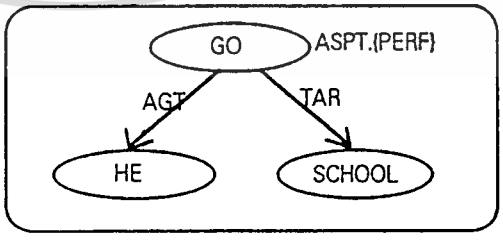


2) ลักษณาการ (Aspect) แบ่งย่อยเป็น

2.1) ลักษณาการกำลังดำเนินอยู่ (Progressive) ในระบบซอฟท์แวร์ของขบวนการวิเคราะห์ กำหนดค่าแสดงสถานะภาพนี้เป็น ASPT.(PROG) ตัวอย่างเช่น ประโยค "He is going to school." จะมีการเก็บข้อมูลดังภาพข้างล่างนี้



2.2) ลักษณาการสมบูรณ์ (Perfect) ในระบบซอฟท์แวร์ของขบวนการวิเคราะห์ กำหนดค่าแสดงสถานะภาพนี้เป็น ASPT.(PERF) ตัวอย่างเช่น ประโยค "He had gone to school." จะมีการเก็บข้อมูลดังภาพข้างล่างนี้

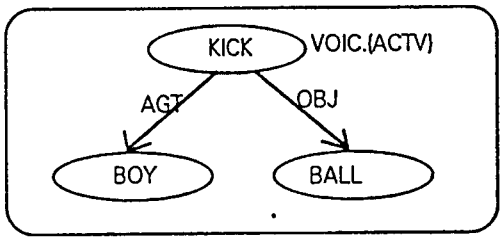


3) วาจก (Voice) แบ่งย่อยเป็น

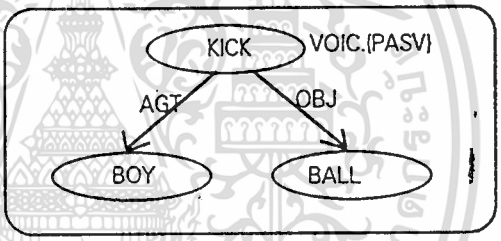
3.1) กรรตุวาจก (Active Voice) ในระบบซอฟท์แวร์ของขบวนการวิเคราะห์ กำหนดค่าแสดงสถานะภาพนี้เป็น VOIC.(ACTV) ตัวอย่างเช่น ประโยค

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

"The boy kicked the ball." จะมีการเก็บข้อมูลดังภาพข้างล่างนี้

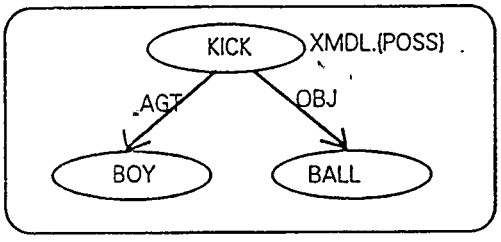


3.2) กรรมวาจก (Passive Voice) ในระบบซอฟท์แวร์ของขบวนการวิเคราะห์กำหนดค่าแสดงสถานะภาพนี้เป็น VOIC.(PASV) ตัวอย่าง เช่น ประโยค "The ball was kicked by the boy." จะมีการเก็บข้อมูลดังภาพข้างล่างนี้



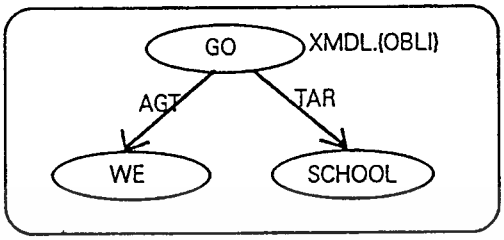
4) มาลา (Modal) แบ่งย่อยเป็น

4.1) ความเป็นไปได้ (Possibility) ได้แก่ can เป็นต้น ในระบบซอฟท์แวร์ของขบวนการวิเคราะห์กำหนดค่าแสดงสถานะภาพนี้เป็น XMDL.(POSS) ตัวอย่างเช่น ประโยค "The boy can kick the ball." จะมีการเก็บข้อมูลดังภาพข้างล่างนี้

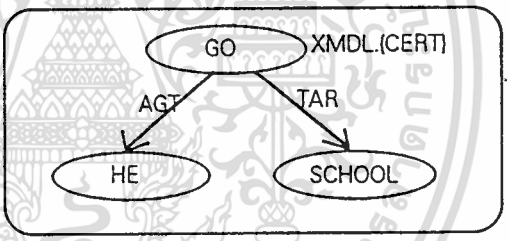


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

4.2) การมีพันธะหน้าที่ (Obligation) ได้แก่ should เป็นต้น ในระบบซอฟท์แวร์ของขบวนการวิเคราะห์กำหนดค่าแสดงสถานะภาพนี้เป็น XMDL.{OBLI} ตัวอย่างเช่น ประโยค "We should go to school." จะมีการเก็บข้อมูลดังภาพข้างล่างนี้



4.3) ความแน่ใจ (Certainty) ได้แก่ must เป็นต้น ในระบบซอฟท์แวร์ของขบวนการวิเคราะห์กำหนดค่าแสดงสถานะภาพนี้เป็น XMDL.{CERT} ตัวอย่างเช่น ประโยค "He must go to school." จะมีการเก็บข้อมูลดังภาพข้างล่างนี้



4.6 รูปแบบข้อมูลของภาษากลาง

การเก็บบันทึกข้อมูลทางโครงสร้างของภาษากลางให้อยู่ในรูปแบบทางซอฟท์แวร์นั้น จะทำการเก็บบันทึกตามรูปแบบของโครงสร้างในลักษณะเป็นบล็อคของข้อมูลของแต่ละ CP node และ CASE node โดยมีพอยต์เตอร์บอกลำดับของการต่อเชื่อมข้อมูลและความสัมพันธ์ของแต่ละ CP node หรือ CASE node ข้อมูลของแต่ละ CP node หรือ CASE node จะเป็นข้อมูลทางภาษาที่เป็นผลจากการวิเคราะห์ประโยคต้นแบบ นั่นคือผลของขบวนการวิเคราะห์ประโยคต้นแบบไปสู่ภาษากลางนั่นเอง โครงสร้างการเก็บข้อมูล (Data Structure) ของภาษากลางของ CP node หรือ CASE node ใด ๆ ได้แสดงไว้ในรูปที่ 4.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
***** node N *****          ----------> ส่วนที่(1)
CP.{AAA...},               ----------> ส่วนที่ (2)
CAT.{BBB...},              ----------> ส่วนที่(3)
TENSE.{CCC...},            ----------> ส่วนที่(4)
NEG.{DDD...},              ----------> ส่วนที่(5)
XMDL.{eee...},             ----------> ส่วนที่(6)
DET.{FFF..}                ----------> ส่วนที่(7)
%
NO. DOWN-NODE              ----------> ส่วนที่(8)
NO. UP-NODE                ----------> ส่วนที่(9)
%%
```

รูปที่ 4.2 แสดงรูปแบบข้อมูลของ node ใด ๆ ในภาษากลาง

จากรูปที่ 4.2 ในแต่ละส่วนมีรายละเอียดดังนี้

ส่วนที่(1) หมายถึง ลำดับของโนดที่มาเชื่อมกันเป็นโครงข่าย โดยที่เราจะเริ่มตั้งแต่คำที่เป็นกริยาหลักของประโยค ซึ่ง N จะมีค่าตั้งแต่ 0-n

ส่วนที่(2) หมายถึง หน่วยความหมาย เช่น CP.{GO}

ส่วนที่(3) หมายถึง ประเภทของคำ เช่น CAT.{V}

ส่วนที่(4) หมายถึง กาลของประโยคอินพุต เช่น TENSE.{PRES}

ส่วนที่(5) หมายถึง ประโยคอินพุตเป็นประโยคบอกเล่าหรือประโยคปฏิเสธจะมี 2 ค่าคือ T หรือ F เช่นถ้าประโยคเป็นประโยคปฏิเสธ แล้วส่วนนี้ จะมีค่าเป็น NEG.{T} เป็นต้น

ส่วนที่(6) หมายถึง มาลาของประโยค เช่น XMDL.{poss}

ส่วนที่(7) หมายถึง ในกรณีที่คำคำนั้นมีคำชี้เฉพาะ (Determiner) คือ a,an,the นำหน้า เช่น DET.{A}

ส่วนที่(8) หมายถึง หมายเลขโนดที่ต่อเชื่อมอยู่ข้างล่าง เช่น 1,2

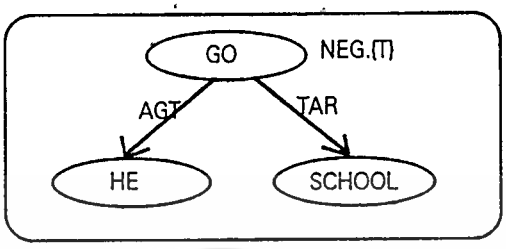
ส่วนที่(9) หมายถึง หมายเลขโนดที่ต่อเชื่อมอยู่ข้างบน เช่น 0

<u>หมายเหตุ</u> ส่วนที่(4),(5) และส่วนที่(6) จะปรากฏที่หน่วยคำที่เป็นกริยาหลักของประโยคเท่านั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

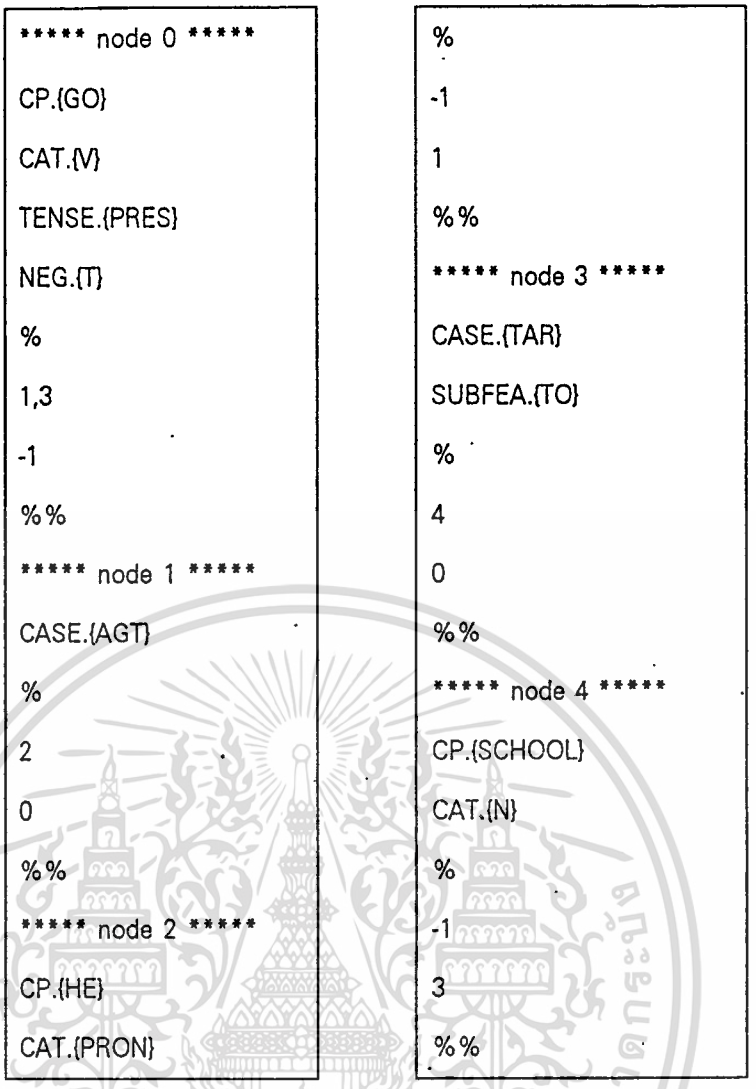
4.7 รูปแบบภาษากลางจากตัวอย่างประโยคที่ผ่านขบวนการวิเคราะห์

ประโยคตัวอย่างเช่น "He does not go to school." เมื่อผ่านขบวนการวิเคราะห์จะแสดงในรูปแบบของต้นไม้หรือรูปแบบทางโครงสร้างดังในรูปที่ 4.3 และ แสดงรูปแบบข้อมูลเป็นบล็อคดังรูปที่ 4.4



รูปที่ 4.3 แสดงโครงสร้างภาษากลาง
ของตัวอย่างประโยค "He does not go to school."

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.4 แสดงรูปแบบข้อมูลเป็นบล็อคของ
ตัวอย่างประโยค "He does not go to school."

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
ขบวนการวิเคราะห์ประโยคภาษาอังกฤษ

5.1 บทนำ

ดังที่ได้กล่าวมาแล้วในบทที่ 1 ว่า ขบวนการวิเคราะห์ประโยคภาษาอังกฤษไปสู่ภาษากลางของวิทยานิพนธ์นี้ ได้แบ่งออกเป็น 3 ขั้นตอนใหญ่ ๆ ที่ทำงานเป็นลำดับต่อเนื่องกัน เพื่อให้สะดวกต่อการพัฒนาและให้ความละเอียดและความถูกต้องในการวิเคราะห์มากขึ้น ตลอดจนสะดวกต่อการแก้ไขในภายหลังเพราะซอฟท์แวร์ของขบวนการวิเคราะห์ที่พัฒนาจะมีลักษณะเป็นสามรูทีน (Routine) ใหญ่ ๆ ตามขบวนการทั้ง 3 ขั้นตอนดังกล่าว นอกจากนี้การออกแบบเช่นนี้จะสะดวกต่อการแก้ไขกฎไวยากรณ์ภายในขบวนการวิเคราะห์ระหว่างการประมวลผล (Processing) ในรูปของการย้อนสำรวจโปรแกรม (Program Debugging)

การแบ่งขบวนการวิเคราะห์ประโยคภาษาอังกฤษไปสู่ภาษากลางออกเป็น 3 ขั้นตอนใหญ่ ๆ ของวิทยานิพนธ์นี้ได้คำนึงถึงหลักภาษาศาสตร์ของข้อมูลทางภาษาที่ควรจะมีการประมวลผลเป็นลำดับ เพื่อให้เกิดความสมบูรณ์ในการถ่ายทอดข้อมูลทั้งหมดไปสู่ภาษากลาง โดยการเริ่มต้นจากประโยคต้นแบบที่เป็นประโยคแสดงด้วยตัวอักษรที่นำมาเรียงกันเป็นคำ และเป็นประโยคในรูปแบบผิวนอก (Surface Form) ที่มนุษย์สามารถอ่านทำความเข้าใจได้ ขบวนการวิเคราะห์จะทำการประมวลผลประโยคดังกล่าวด้วยการใช้ฐานข้อมูลทางภาษาในรูปของกฎไวยากรณ์ (Grammar Rules) (ดังที่กล่าวในบทที่ 2) ประกอบกับข้อมูลจากพจนานุกรมคำศัพท์ (Dictionary) (ดังที่กล่าวมาในบทที่ 3) เพื่อหาโครงสร้างและความหมายลึกของประโยคดังกล่าว และทำการถ่ายทอดข้อมูลทางภาษาดังกล่าวทั้งหมดของประโยคนั้นให้อยู่ในรูปแบบของภาษากลาง (ดังที่กล่าวมาในบทที่ 4) อันเป็นเป้าหมายของงานวิทยานิพนธ์นี้

ขบวนการวิเคราะห์ประโยคภาษาอังกฤษทั้ง 3 ขั้นตอนที่มีการประมวลผลตามลำดับได้แก่

ก) ส่วนการวิเคราะห์ในระดับคำ (Morphological Analysis)

ข) ส่วนการวิเคราะห์ในระดับวากยสัมพันธ์ หรือไวยากรณ์ (Syntax Analysis)

ค) ส่วนการวิเคราะห์ในระดับอรรถศาสตร์ หรือความหมาย (Semantic Analysis)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.2 ส่วนการวิเคราะห์ในระดับคำ (Morphological Analysis)

เป็นขบวนการวิเคราะห์ลำดับแรกสุดเพื่อการแยกแยะโครงสร้างของประโยคต้นแบบ (Source Sentence) ที่อยู่ในรูปแบบผิวนอก (Surface form) จุดประสงค์ของขั้นตอนนี้คือ การค้นหาหน่วยคำ (Morpheme) และข้อมูลของหน่วยคำนั้น หลังจากนั้นก็จะนำหน่วยคำประกอบกับข้อมูลของหน่วยคำเหล่านั้นมาวิเคราะห์หาความสัมพันธ์ระหว่างหน่วยคำทางด้านวากยสัมพันธ์ (Syntax Analysis) และหาความหมายทางด้านอรรถศาสตร์ (Semantic Analysis) ให้กับหน่วยคำโดยการอ้างอิงข้อมูลจากพจนานุกรมคำศัพท์ (Dictionary)

ในการค้นหาหน่วยคำที่ปรากฏในประโยคของภาษาต่าง ๆ นั้น มีความยากง่ายแตกต่างกันไปแล้วแต่คุณสมบัติของภาษานั้นๆ ในภาษาไทย ภาษาญี่ปุ่น ภาษาจีน หรือภาษาเกาหลี นั้นจะไม่มีการเว้นช่องว่างระหว่างคำ ดังนั้นในการค้นหาหน่วยคำด้วยระบบคอมพิวเตอร์ จึงต้องมีการกำหนดกระบวนการค้นหา (Searching Algorithm) หน่วยคำ และสร้างขึ้นเป็นระบบซอฟท์แวร์ กระบวนการนี้เรียกว่า "การตัดหน่วยคำ (Word Segmentation)" หรือ "การแยกแยะหน่วยคำ (Word Parsing)" สำหรับในภาษาอังกฤษหรือภาษาของกลุ่มประเทศยุโรปนั้น มีการเว้นช่องว่าง (Space) ระหว่างคำ จึงเป็นการง่าย ในการที่จะกำหนดหน่วยคำให้กับระบบคอมพิวเตอร์ วิทยานิพนธ์ได้อาศัยช่องว่างระหว่างคำในประโยคภาษาอังกฤษเป็นเครื่องหมายในการกำหนด 1 หน่วยคำ เพื่อนำไปเปรียบเทียบกับคำในพจนานุกรมคำศัพท์เพื่อค้นหาข้อมูลทางภาษาต่อไป รายละเอียดของขบวนการวิเคราะห์ในระดับคำมีดังต่อไปนี้

5.2.1 ข้อกำหนดของประโยคต้นแบบ

ประโยคภาษาอังกฤษต้นแบบหรือประโยคเริ่มต้น ซึ่งจะเป็นประโยคอินพุตเริ่มต้นของขบวนการวิเคราะห์ในระดับคำจะต้องประกอบด้วย

ก) ตัวอักษรหรืออักขระ (Characters) ที่เป็นอักษรพิมพ์ใหญ่ ได้แก่

A, B, C, ..., Z (ทั้งหมด 26 ตัวอักษร) และอักษรพิมพ์เล็ก

a, b, c, ..., z (ทั้งหมด 26 ตัวอักษร)

ข) เครื่องหมายพิเศษ ได้แก่ ,(comma), ;(semi-colon), :(colon), .(full stop), ?(question mark)

ค) ช่องว่างระหว่างหน่วยคำ

ง) เครื่องหมายกำหนดการสิ้นสุดของ 1 ประโยค ได้แก่ .(full stop), ?(question mark)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.2.2 การประมวลผลของขบวนการวิเคราะห์ในระดับคำ

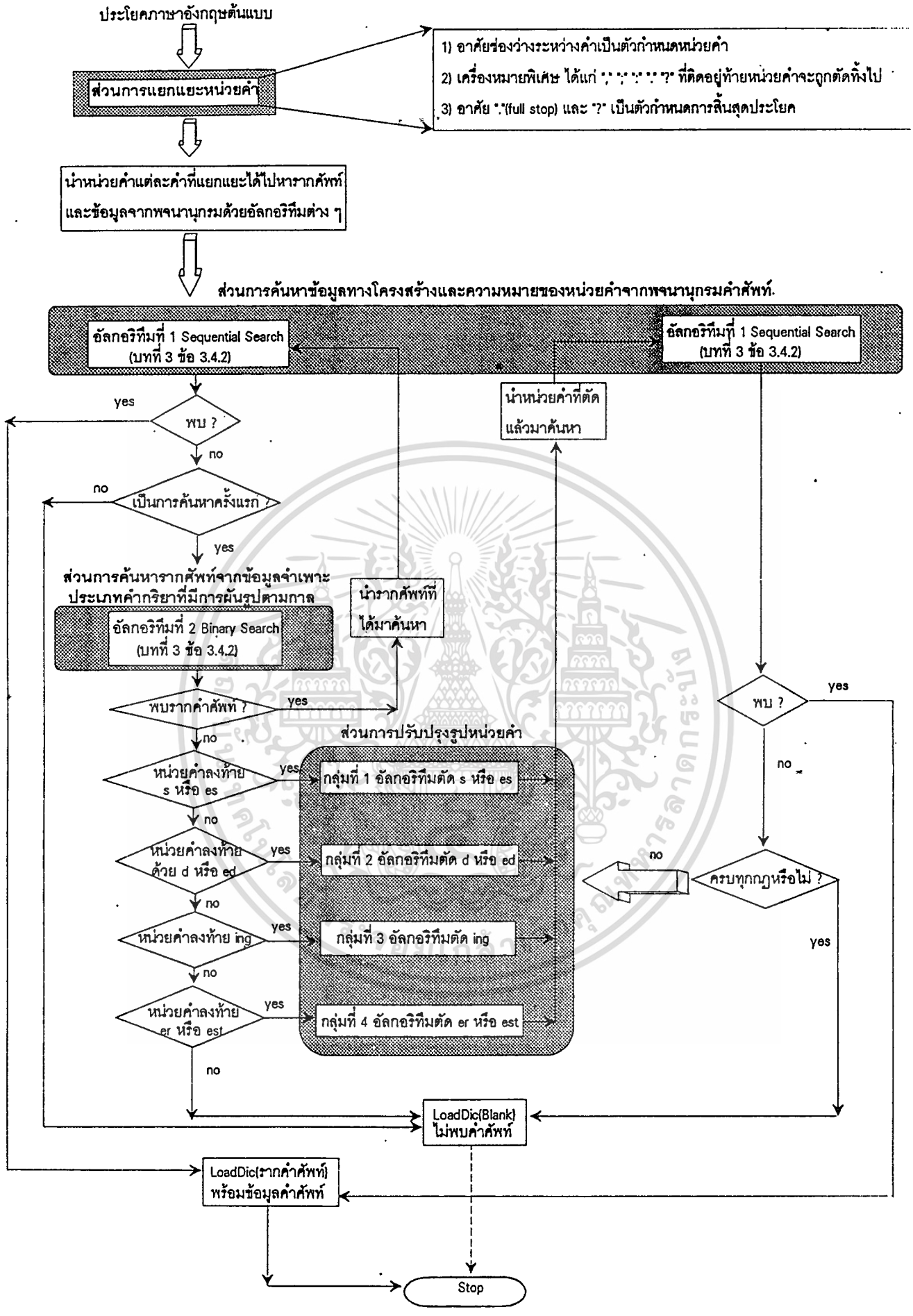
การประมวลผลของขบวนการวิเคราะห์ในระดับคำจะเป็นขบวนการแรกสุดของขบวนการวิเคราะห์โครงสร้างประโยคภาษาอังกฤษต้นแบบที่ถูกป้อนเข้าสู่คอมพิวเตอร์และถูกส่งเข้าสู่ซอฟท์แวร์ของระบบ วิทยานิพนธ์ได้วิจัยและพัฒนาฐานความรู้ของระบบซอฟท์แวร์ในส่วนนี้ให้สามารถแยกแยะหน่วยคำเหมือนกับความสามารถของมนุษย์ที่แยกแยะหน่วยคำได้ด้วยการดูและอ่านประโยค (เรียกคุณสมบัตินี้ว่า "Humanlike Computer's Vision") และเมื่อได้หน่วยคำแล้ว จะนำแต่ละหน่วยคำที่แยกแยะได้ไปค้นหาข้อมูลทางภาษาจากฐานข้อมูลภายในพจนานุกรมคำศัพท์ โดยการเปรียบเทียบกับหน่วยคำที่ถูกบันทึกอยู่ภายในพจนานุกรม ตรวจสอบหาหน่วยคำที่มีการเขียนด้วยอักขระเหมือนกันตามลำดับทุกตัวอักษรเฉพาะขั้นตอนนี้จะเหมือนกับการตรวจสอบความถูกต้องของการสะกดของหน่วยคำที่ปรากฏอยู่ในประโยคต้นแบบ (เรียกคุณสมบัตินี้ว่า การตรวจสอบตัวสะกดของคำ หรือ Spelling Check) แต่ในบางครั้งจะต้องใช้ฐานความรู้ทางภาษาในการพิจารณาดัดแปลงหน่วยคำที่แยกแยะได้จากประโยคต้นแบบที่มีการผันรูปไปตาม พจน์ หรือ กาล หรือ การเปรียบเทียบ ด้วยการพิจารณาตัดส่วนเสริมท้ายคำ หรือ "วิภัตติ (Suffix)" ออกและปรับปรุงหน่วยคำเพื่อนำไปค้นหาข้อมูลจากพจนานุกรมคำศัพท์ต่อไป

จากที่กล่าวมาข้างต้น วิทยานิพนธ์นี้ได้ทำการศึกษาวิจัยบนพื้นฐานของคุณสมบัติทางไวยากรณ์ของภาษาอังกฤษ และได้ข้อสรุปของการแบ่งหน่วยประมวลผลของขบวนการวิเคราะห์ในระดับคำออกเป็น 4 ส่วนย่อยได้แก่

ก) ส่วนการแยกแยะหน่วยคำ
ข) ส่วนการค้นหาข้อมูลทางโครงสร้างและความหมายของหน่วยคำจากพจนานุกรมคำศัพท์
ค) ส่วนการค้นหารากศัพท์จากข้อมูลจำเพาะของคำศัพท์ประเภทคำกริยาที่มีการผันรูปไปตามกาลของประโยค
ง) ส่วนการปรับปรุงรูปหน่วยคำ

โดยแต่ละส่วนจะมีรายละเอียดในการประมวลผลทางภาษาที่ต่อเนื่องสัมพันธ์กันตามลำดับดังแสดงในแผนภูมิในรูปที่ 5.1 รายละเอียดของการประมวลผลในแต่ละส่วนมีดังต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



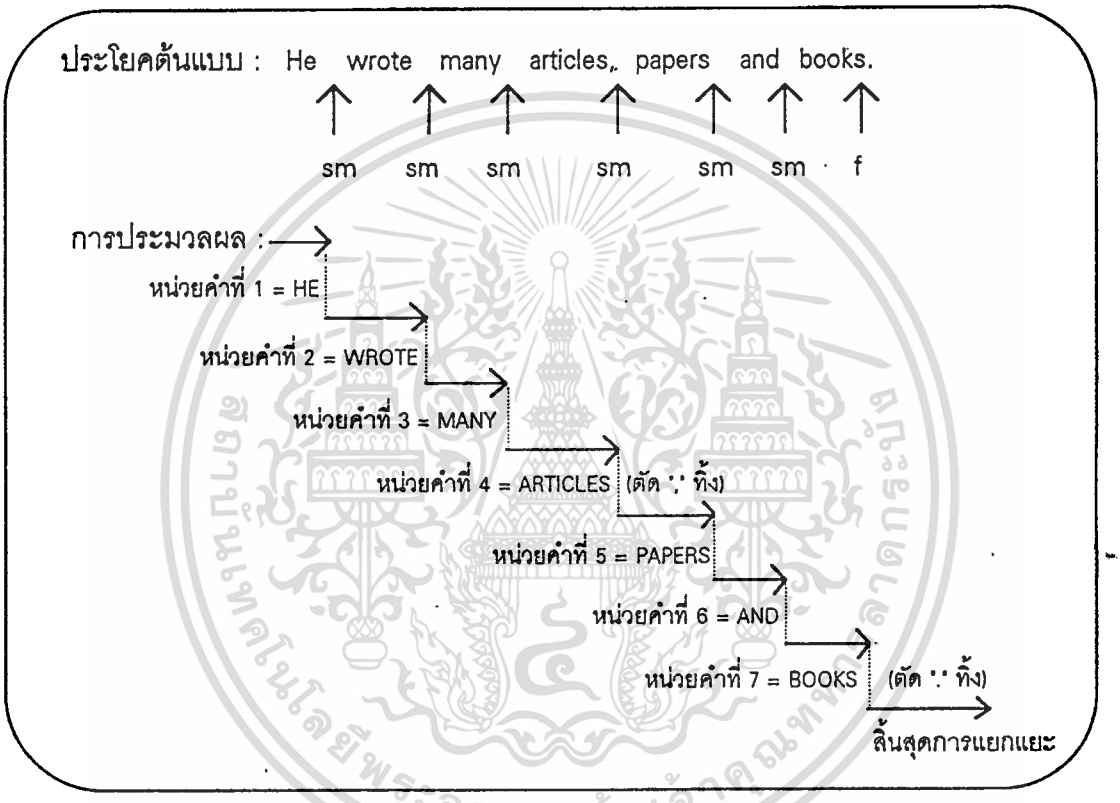
รูปที่ 5.1 โฟล์วชาร์ตของการวิเคราะห์ในระดับคำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.2.2.1 ส่วนการแยกแยะหน่วยคำ (Morpheme Segmentation)

เมื่อประโยคภาษาอังกฤษต้นแบบถูกป้อนเข้าสู่ระบบซอฟท์แวร์ในส่วนนี้จะทำการตรวจหาเครื่องหมายช่องว่างระหว่างคำ (Space Marker) โดยเริ่มค้นหาจากซ้ายไปขวา เมื่อพบเครื่องหมายช่องว่างจะทำการแยกกลุ่มอักขระระหว่างช่องว่างทั้งหมดเป็นแต่ละหน่วยคำเริ่มต้น และเมื่อพบจุด "." (fullstop) หรือ "?" (Question Mark) ถือว่าเป็นการสิ้นสุดประโยค จะทำการส่งหน่วยคำให้กับการประมวลผลในส่วนต่อไป

ตัวอย่างเช่น (sm : ย่อจาก Space Marker, f : fullstop)



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.2.2.2 ส่วนการค้นหาข้อมูลทางโครงสร้างและความหมายของคำศัพท์จากพจนานุกรม

หน่วยคำทั้งหมดที่ได้จากการแยกแยะในข้อ 5.2.2.1 จะเป็นหน่วยคำเริ่มต้นที่กำหนดให้เป็น ENTRY ด้วยคำสั่ง LoadDic {ENTRY} เพื่อการค้นหารากศัพท์และข้อมูลทางภาษาจากพจนานุกรมในส่วนโครงสร้างและความหมายของคำศัพท์ด้วยซอฟท์แวร์ของอัลกอริทึมที่ 1 (การค้นแบบ Sequential Search) ดังที่ได้กล่าวมาแล้วในบทที่ 3 ข้อ 3.4.2

การประมวลผลในส่วนนี้จะพบรากศัพท์และข้อมูลของหน่วยคำเริ่มต้นทันที สำหรับหน่วยคำที่ไม่มีการผันรูป (เช่น He และ many และ and ในประโยคตัวอย่างข้างต้น) แต่สำหรับหน่วยคำที่มีการผันรูปหรือเปลี่ยนแปลงรูปคำ (เช่น wrote, articles, papers และ books ในประโยคตัวอย่างข้างต้น) จะถูกส่งไปค้นหาหรือปรับปรุงในการประมวลผลส่วนอื่นเสียก่อน แล้วจึงย้อนกลับมาหารากศัพท์และข้อมูลจากการประมวลผลในส่วนนี้ใหม่ (ดังแสดงในรูปที่ 5.1)

ตัวอย่างเช่น ผลลัพธ์จาก 5.2.2.1 ได้ : หน่วยคำที่ 1 = HE

หน่วยคำที่ 2 = WROTE

หน่วยคำที่ 3 = MANY

หน่วยคำที่ 4 = ARTICLES

หน่วยคำที่ 5 = PAPERS

หน่วยคำที่ 6 = AND

หน่วยคำที่ 7 = BOOKS

ผลลัพธ์ส่วนแรกที่จะได้จาก 5.2.2.2ได้แก่ : 1) LoadDic{HE} พร้อมข้อมูล

CAT = PRON

SUBCAT = PPRS

AKO = 111

CP = HE

2) LoadDic{MANY} พร้อมข้อมูล

CAT = ADJ

SUBCAT = ALRP

AKO = 2363

CP = MANY

3) LoadDic{AND} พร้อมข้อมูล

CAT = CONJ

SUBCAT = CCRG

CP = AND

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.2.2.3 ส่วนการค้นหารากศัพท์จากข้อมูลจำเพาะของคำศัพท์ประเภทคำกริยาที่มีการผันรูปไปตามกาลของประโยค

ในกรณีของหน่วยคำที่ผ่านการค้นหาข้อมูลจากส่วนข้อ 5.2.2.2 ซึ่งเป็นการค้นหาครั้งแรก แล้วไม่พบรากศัพท์ จะถูกตั้งสมมติฐานให้เป็นคำกริยาที่ผันรูปไปตามกาล หน่วยคำนั้นจะถูกนำไปค้นหารากศัพท์จากส่วนข้อมูลจำเพาะของคำศัพท์ประเภทคำกริยาที่มีการผันรูปไปตามกาลของประโยค ด้วยซอฟท์แวร์ของอัลกอริทึมที่ 2 (การค้นหาแบบ Binary Search) ดังที่ได้กล่าวมาแล้วในบทที่ 3 ข้อ 3.4.2

การประมวลผลในส่วนนี้จะเป็นการค้นหารากศัพท์ของหน่วยคำที่พิจารณาเมื่อพบรากศัพท์แล้วจะทำการกำหนดให้รากศัพท์ของหน่วยคำเป็น ENTRY ใหม่ เพื่อย้อนกลับไปค้นหาข้อมูลของหน่วยคำจากส่วนการค้นหาข้อมูลทางโครงสร้างและความหมายของคำศัพท์จากพจนานุกรมต่อไป พร้อมกันนี้จะได้ข้อมูลทางภาษาแสดงสถานะกาล (Tense) ของคำกริยานั้น

ถ้าในขั้นตอนนี้ไม่พบรากศัพท์ของหน่วยคำ ระบบก็จะตั้งสมมติฐานใหม่ว่า หน่วยคำที่พิจารณาเป็นคำที่มีการเปลี่ยนรูปแบบไปตามหน้าที่ที่ถูกใช้ในประโยค หน่วยคำนั้นจะถูกนำไปประมวลผลในส่วนการปรับปรุงรูปหน่วยคำต่อไปในข้อ 5.2.2.4

ตัวอย่างเช่น

คำว่า "WROTE" ในประโยคตัวอย่างข้างต้นเมื่อผ่านขบวนการประมวลผลในส่วนนี้ จะพบรากศัพท์คำว่า "WRITE" พร้อมกับค่า TENSE(PAST) และเมื่อนำรากศัพท์ "WRITE" เป็น ENTRY ไปค้นหาข้อมูลทางโครงสร้างและความหมายก็จะได้ผลลัพธ์ดังนี้

LoadDic(WRITE) พร้อมข้อมูล :

CAT = TV

SUBCAT = V

AKO = 2111

CP = WRITE

TENSE = PAST

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.2.2.4 ส่วนการปรับปรุงรูปหน่วยคำ (Morpheme Modification)

สำหรับหน่วยคำที่มีการเปลี่ยนรูปแบบไปตามหน้าที่ที่ถูกใช้ในประโยคด้วยการถูกเติมที่ท้ายคำด้วยส่วนเสริมหรือวิภัตติ (Suffix) * ได้แก่

ก) การเปลี่ยนตามพจน์ของคำนามด้วยการเติม "___s" หรือ "___es" เช่น

เอกพจน์(รากศัพท์) ——————>	พหูพจน์
book	books
box	boxes
duty	duties
knife	knives
mango	mangoes
article	articles

ข) หน่วยคำกริยาที่เปลี่ยนรูปตามประธานในประโยคปัจจุบันกาล (Present Tense) ด้วยการเติม s หรือ es เช่น

รูปปกติ(รากศัพท์) ——————>	เมื่อประธานเป็นบุรุษที่ 3 หรือคำนามเอกพจน์
go	goes
play	plays
cry	cries
apply	applies

ค) คำกริยาเปลี่ยนรูปตามกาลของประโยค ด้วยการเติม "___d", "___ed" หรือ "___ing" เช่น

รูปปกติ(รากศัพท์) ——————>	รูปที่เปลี่ยนไป
cry	cried (Past tense)
walk	walked (Past tense)
stop	stopped (Past tense)

---

*หมายเหตุ สำหรับหน่วยคำที่มีการเปลี่ยนรูปโดยการเติมส่วนหน้าคำด้วยอุปสรรค(Prefix) เช่น disconnect, impossible เป็นต้น ในทางภาษาแล้วถือว่าเป็นการสร้างหน่วยคำใหม่ ดังนั้นงานวิจัยนี้ได้กำหนดให้หน่วยคำประเภทนี้เป็นรากศัพท์อิสระและมีข้อมูลบันทึกอยู่ในพจนานุกรมคำศัพท์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปปกติ(รากศัพท์) ---------->	รูปที่เปลี่ยนไป
dance	danced (Past tense)
cry	crying (Continuous tense)
walk	walking (Continuous tense)
stop	stopping (Continuous tense)
lie	lying (Continuous tense)
dance	dancing (Continuous tense)
fit	fitting (Continuous tense)
travel	travelling(Continuous tense)

ง) คำคุณศัพท์ที่มีการเปลี่ยนรูปเมื่อใช้ในการเปรียบเทียบขนาดด้วยการเติม "___er", "__est" เช่น

รูปปกติ(รากศัพท์) ---------->	ในความหมาย "__กว่า" ---------->	ในความหมาย "__ที่สุด"
dry	drier	driest
big	bigger	biggest
brave	braver	bravest
tall	taller	tallest
small	smaller	smallest

หน่วยคำเริ่มต้นใดที่ผ่านการค้นหารากศัพท์ครั้งแรกจากข้อ 5.2.2.2 และ 5.2.2.3 แล้ว ไม่พบในพจนานุกรมจะถูกนำมาพิจารณาตรวจสอบทางด้านส่วนเสริมท้ายคำดังที่กล่าวมาในข้อ ก, ข, ค และ ง ด้วยซอฟท์แวร์ของอัลกอริทึมการตัดส่วนเสริมท้ายคำของการประมวลผลในส่วนการปรับปรุงรูปหน่วยคำ วิทยานิพนธ์นี้ได้แบ่งอัลกอริทึมของการตัดส่วนเสริมท้ายหน่วยคำออกเป็น 4 กลุ่ม ได้แก่

กลุ่มที่ 1 อัลกอริทึมของการตัดส่วนเสริมท้ายหน่วยคำที่เป็น "___s" หรือ "___es"
กลุ่มที่ 2 อัลกอริทึมของการตัดส่วนเสริมท้ายหน่วยคำที่เป็น "___d" หรือ "___ed"
กลุ่มที่ 3 อัลกอริทึมของการตัดส่วนเสริมท้ายหน่วยคำที่เป็น "___ing"
กลุ่มที่ 4 อัลกอริทึมของการตัดส่วนเสริมท้ายหน่วยคำที่เป็น "___er" หรือ "___est"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

<<การพิจารณาเบื้องต้น>>

(1) ถ้าคำศัพท์นั้นลงท้ายด้วย s หรือ es จะผ่านอัลกอริทึมกลุ่มที่ 1 หรือ

(2) ถ้าคำศัพท์นั้นลงท้ายด้วย d หรือ ed จะผ่านอัลกอริทึมกลุ่มที่ 2 หรือ

(3) ถ้าคำศัพท์นั้นลงท้ายด้วย ing จะผ่านอัลกอริทึมกลุ่มที่ 3 หรือ

(4) ถ้าคำศัพท์นั้นลงท้ายด้วย er หรือ est จะผ่านอัลกอริทึมกลุ่มที่ 4 หรือ

(5) ถ้าไม่เป็นไปตามข้อ (1), (2), (3) และ (4) ให้ส่งผลไม่พบคำศัพท์ด้วยคำสั่ง LoadDic{Blank} กลับไปยังขบวนการวิเคราะห์และหยุดการค้นหา

<<ข้อกำหนดเบื้องต้น>>

เมื่อหน่วยคำที่พิจารณาเป็นไปตามกรณีที่ (1) หรือ (2) หรือ (3) หรือ (4) ดังกล่าวข้างต้น หน่วยคำนั้นจะถูกประมวลผลตามขั้นตอนของอัลกอริทึมที่เกี่ยวข้อง โดยมีข้อกำหนดเบื้องต้นของการประมวลผลดังนี้

<u>กำหนดให้</u>

(1) C_X : เป็นอักขระตัวสุดท้ายของหน่วยคำที่พิจารณา

(2) x : เป็นตำแหน่งของอักขระตัวสุดท้ายของหน่วยคำที่พิจารณา โดยที่ x เป็นเลขจำนวนเต็มบวก

<ตัวอย่างเช่น> หน่วยคำ knives จะมี suffix เป็น "es" ในกรณีนี้จะได้

C_X = s, C_{X-1} = e, C_{X-2} = v,...

และ x = 6 (เป็นลำดับตำแหน่งของอักขระ "s" ในหน่วยคำนับจากขวาไปซ้าย)

(3) SetVowel : เป็นเซ็ตของกลุ่มสระ a, e, i, o, u

โดยในอัลกอริทึมของแต่ละกลุ่มมีรายละเอียดดังต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กลุ่มที่ 1 อัลกอริทึมของการตัดส่วนเสริมท้ายหน่วยคำที่เป็น "___s" หรือ "___es"

ณ ขณะนี้อักขระ C_X คือ "s" เริ่มการประมวลผลตามลำดับดังนี้

Step 1 : ตัดอักขระ C_X (นั่นคือตัด "s") แล้วไป Step 7

(ตัวอย่างเช่น tables—> table เป็นต้น)

Step 2 : -ถ้าอักขระ C_{X-1} คือ "e" ให้ตัดอักขระ C_{X-1} แล้วไป Step 7 และ

-ถ้าไม่ใช่ให้ไป Step 6

(ตัวอย่างเช่น boxes –> box, brushes –> brush เป็นต้น)

Step 3 : -ถ้าอักขระ C_{X-2} คือ "i" ให้เปลี่ยน "i" เป็น "y" แล้วไป Step 8 และ

-ถ้าไม่ใช่ให้ไป Step 4

(ตัวอย่างเช่น duties –> duty, armies –> army เป็นต้น)

Step 4 : -ถ้าอักขระ C_{X-2} คือ "v" ให้เปลี่ยน "v" เป็น "f" แล้วไป Step 7 และ

-ถ้าไม่ใช่ให้ไป Step 6

(ตัวอย่างเช่น thieves –> thief เป็นต้น)

Step 5 : ขณะนี้อักขระ C_{X-2} คือ "v" ให้เปลี่ยน "v" เป็น "fe" แล้วไป Step 8

(ตัวอย่างเช่น knives–> knife, wives –> wife เป็นต้น)

Step 6 : ให้ส่งผลไม่พบหน่วยคำนั้นในพจนานุกรมคำศัพท์ด้วยคำสั่ง LoadDic(Blank) แล้วหยุดการค้นหา

Step 7 : นำหน่วยคำที่เหลือไปค้นหาข้อมูลจากส่วนการค้นหาข้อมูลทางโครงสร้างและความหมายของคำศัพท์จากพจนานุกรม

-ถ้าพบไปที่ Step 9 และ

-ถ้าไม่พบให้ย้อนกลับไปที่ Step ถัดจาก Step ที่เรียกมา

Step 8 : นำหน่วยคำที่ปรับปรุงแล้วไปค้นหาข้อมูลจากส่วนการค้นหาข้อมูลทางโครงสร้างและความหมายของคำศัพท์จากพจนานุกรม

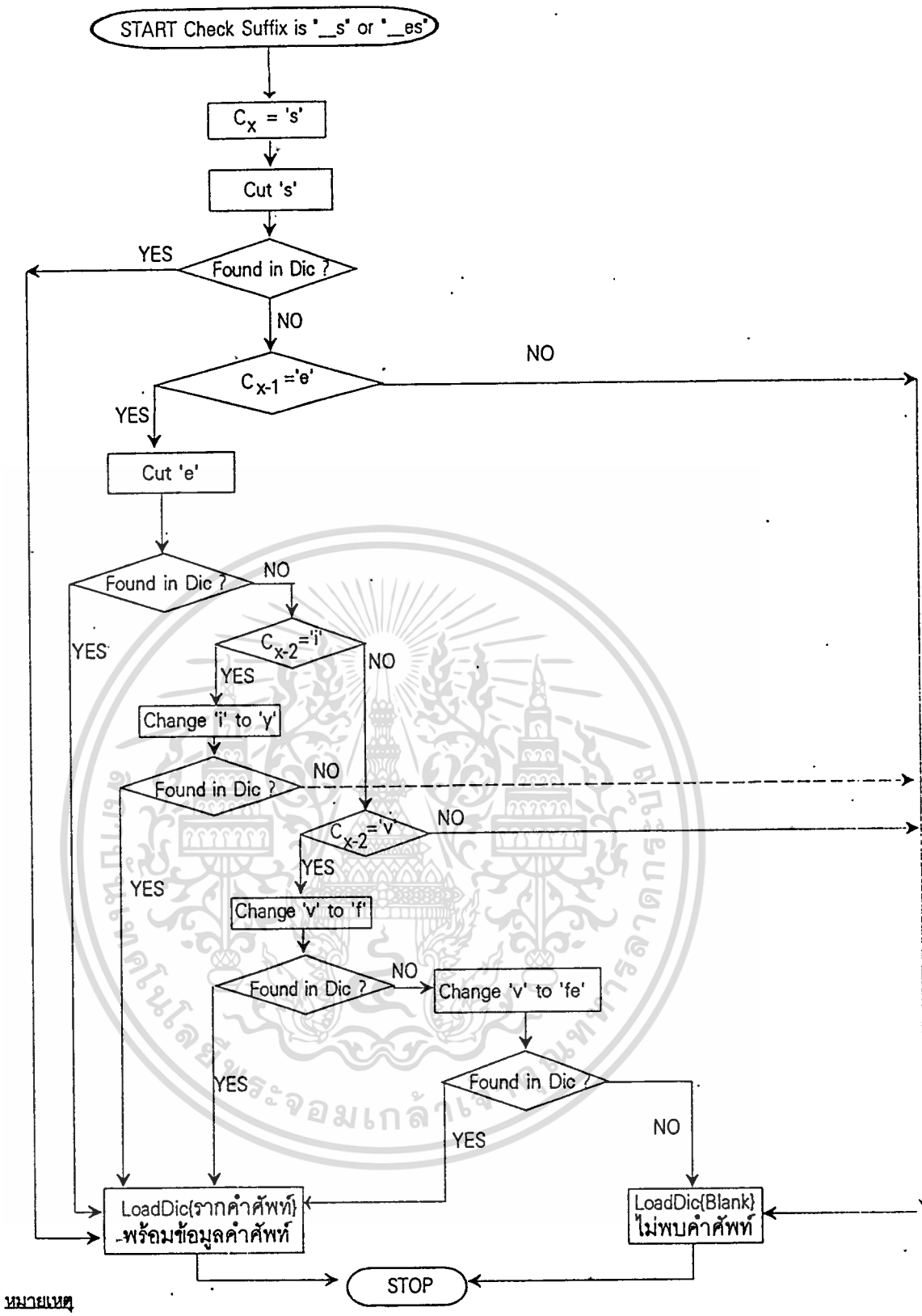
-ถ้าพบไป Step 9 และ

-ถ้าไม่พบไป Step 6

Step 9 : ทำการโหลดข้อมูลของหน่วยคำจากคำศัพท์ที่พบด้วยคำสั่ง LoadDic(รากศัพท์ที่พบ) กลับไปให้ขบวนการวิเคราะห์แล้วหยุดการค้นหา

อัลกอริทึมนี้แสดงในโฟล์วชาร์ตรูปที่ 5.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



หมายเหตุ

Dic : หมายถึงพจนานุกรมคำศัพท์ส่วนโครงสร้างและความหมาย

C_x : เป็นอักขระตัวสุดท้ายของหน่วยคำที่พิจารณา

x : เป็นตำแหน่งอักขระตัวสุดท้ายของหน่วยคำที่พิจารณา

โดยที่ x เป็นเลขจำนวนเต็มบวก

รูปที่ 5.2 แสดงโฟล์วชาร์ตของอัลกอริทึมในกลุ่มที่ 1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กลุ่มที่ 2 อัลกอริทึมของการตัดส่วนเสริมท้ายหน่วยคำที่เป็น "___d" หรือ "___ed"

ณ ขณะนี้อักขระ C_X คือ "d" เริ่มการประมวลผลตามลำดับดังนี้

Step 1 : ตัดอักขระ C_X (นั่นคือตัด "d") แล้วไป Step 6

(ตัวอย่างเช่น danced ——> dance เป็นต้น)

Step 2 : -ถ้าอักขระ C_{X-1} คือ "e" ให้ตัดอักขระ C_{X-1} แล้วไป Step 6 และ

-ถ้าไม่ใช่ให้ไป Step 5

(ตัวอย่างเช่น played ——> play เป็นต้น)

Step 3 : -ถ้าอักขระ C_{X-2} คือ "i" ให้เปลี่ยน "i" เป็น "y" แล้วไป Step 7 และ

-ถ้าไม่ใช่ให้ไป Step 4

(ตัวอย่างเช่น maried --> mary , cried --> cry เป็นต้น)

Step 4 : -ถ้าอักขระ C_{X-2} และอักขระ C_{X-3} เป็นอักขระที่ซ้ำกัน และอักขระ(C_{X-4})

เป็นสมาชิกใน SetVowel ให้ตัดอักขระ C_{X-2} แล้วไป Step 7 และ

-ถ้าไม่ใช่ให้ไป Step 5

(ตัวอย่างเช่น fitted --> fit, stopped --> stop ,controlled -->control เป็นต้น)

Step 5 : ให้ส่งผลไม่พบหน่วยคำนั้นในพจนานุกรมคำศัพท์ด้วยคำสั่ง LoadDic(Blank) แล้วหยุดการค้นหา

Step 6 : นำหน่วยคำที่เหลือไปค้นหาข้อมูลจากส่วนการค้นหาข้อมูลทางโครงสร้างและความหมายของคำศัพท์จากพจนานุกรม

-ถ้าพบไปที่ Step 8 และ

-ถ้าไม่พบให้ย้อนกลับไปที่ Step ถัดจาก Step ที่เรียกมา

Step 7 : นำหน่วยคำที่ปรับปรุงแล้วไปค้นหาข้อมูลจากส่วนการค้นหาข้อมูลทางโครงสร้างและความหมายของคำศัพท์จากพจนานุกรม

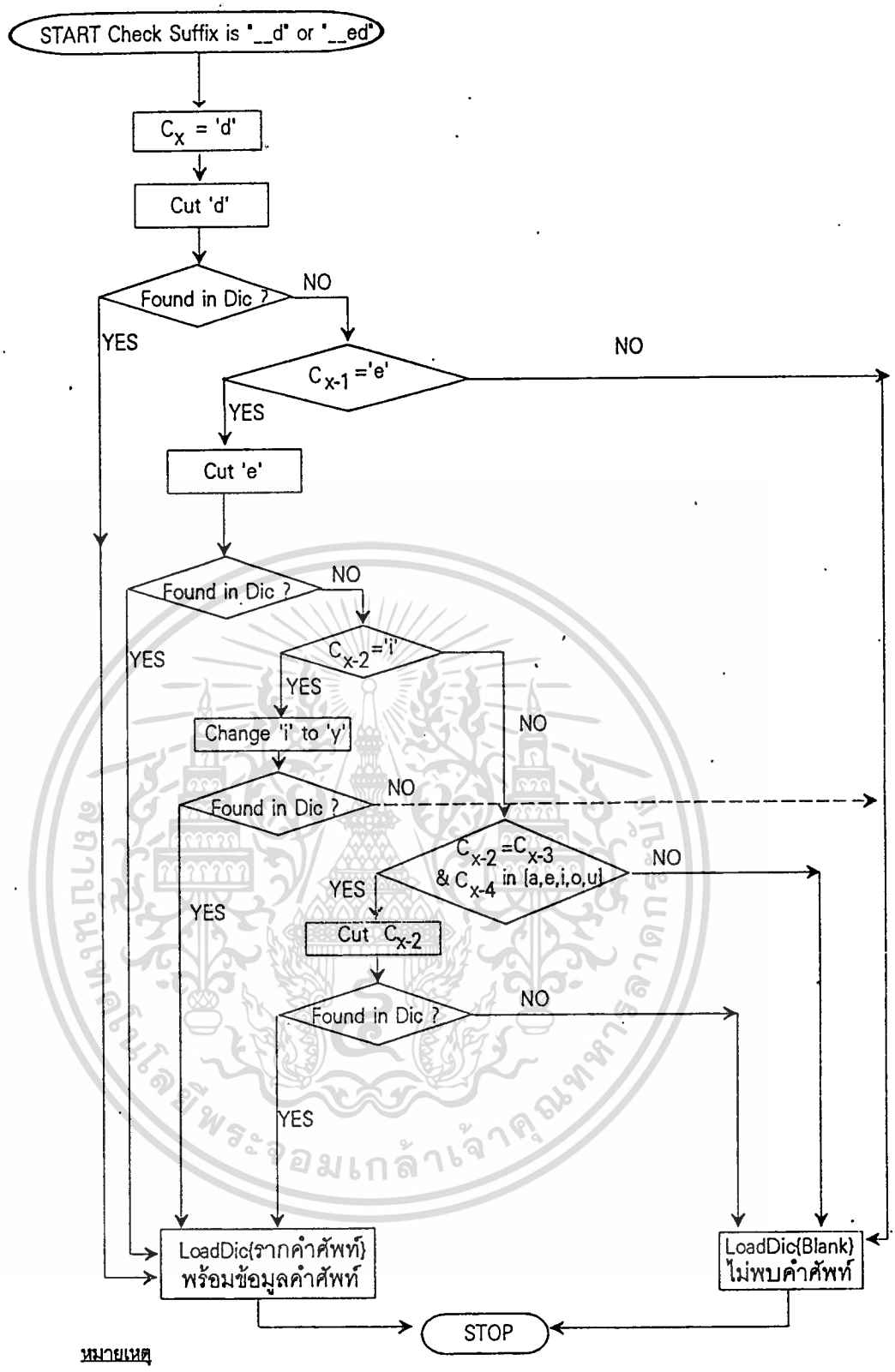
-ถ้าพบไป Step 8 และ

-ถ้าไม่พบไป Step 5

Step 8 : ทำการโหลดข้อมูลของหน่วยคำจากคำศัพท์ที่พบด้วยคำสั่ง LoadDic(รากศัพท์ที่พบ) กลับไปให้ขบวนการวิเคราะห์แล้วหยุดการค้นหา

อัลกอริทึมนี้แสดงในโฟล์วชาร์ตรูปที่ 5.3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



หมายเหตุ

Dic : หมายถึงพจนานุกรมคำศัพท์ส่วนโครงสร้างและความหมาย

C_X : เป็นอักขระตัวสุดท้ายของหน่วยคำที่พิจารณา

x : เป็นตำแหน่งอักขระตัวสุดท้ายของหน่วยคำที่พิจารณา

โดยที่ x เป็นเลขจำนวนเต็มบวก

รูปที่ 5.3 แสดงโฟล์วชาร์ตของอัลกอริทึมในกลุ่มที่ 2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กลุ่มที่ 3 อัลกอริทึมของการตัดส่วนเสริมท้ายหน่วยคำที่เป็น "___ing"

ณ ขณะนี้อักขระ C_X คือ "g" เริ่มการประมวลผลตามลำดับดังนี้

Step 1 : -ถ้าอักขระ C_X คือ "g" อักขระ C_{X-1} คือ "n" และอักขระ C_{X-2} คือ "i" ให้ตัดอักขระ C_X , C_{X-1} และ อักขระ C_{X-2}(นั่นคือตัด "ing") แล้วไป Step 6 และ

-ถ้าไม่ใช่ให้ไป Step 5

(ตัวอย่างเช่น playing ----> play เป็นต้น)

Step 2 : -ถ้าอักขระ C_{X-3} คือ "y" ให้เปลี่ยน "y" เป็น "ie" แล้วไป Step 7 และ

-ถ้าไม่ใช่ให้ไป Step 3

(ตัวอย่างเช่น lying –> lie , dying –> die เป็นต้น)

Step 3 : -ถ้าอักขระ C_{X-3} และอักขระ C_{X-4} เป็นอักขระที่ซ้ำกัน และอักขระ(C_{X-5}) เป็นสมาชิกใน SetVowel ให้ตัดอักขระ C_{X-3} แล้วไป Step 7 และ

-ถ้าไม่ใช่ให้ไป Step 4

(ตัวอย่างเช่น fitting –> fit, stopping–> stop ,controlling –>control เป็นต้น)

Step 4 : ให้เพิ่ม "e" ที่ท้ายหน่วยคำแล้วไป Step 7

(ตัวอย่างเช่น dancing ---> dance, moving ----> move เป็นต้น)

Step 5 : ให้ส่งผลไม่พบหน่วยคำนั้นในพจนานุกรมคำศัพท์ด้วยคำสั่ง LoadDic(Blank) แล้วหยุดการค้นหา

Step 6 : นำหน่วยคำที่เหลือไปค้นหาข้อมูลจากส่วนการค้นหาข้อมูลทางโครงสร้างและความหมายของคำศัพท์จากพจนานุกรม

-ถ้าพบไปที่ Step 8 และ

-ถ้าไม่พบให้ย้อนกลับไปที่ Step ถัดจาก Step ที่เรียกมา

Step 7 : นำหน่วยคำที่ปรับปรุงแล้วไปค้นหาข้อมูลจากส่วนการค้นหาข้อมูลทางโครงสร้างและความหมายของคำศัพท์จากพจนานุกรม

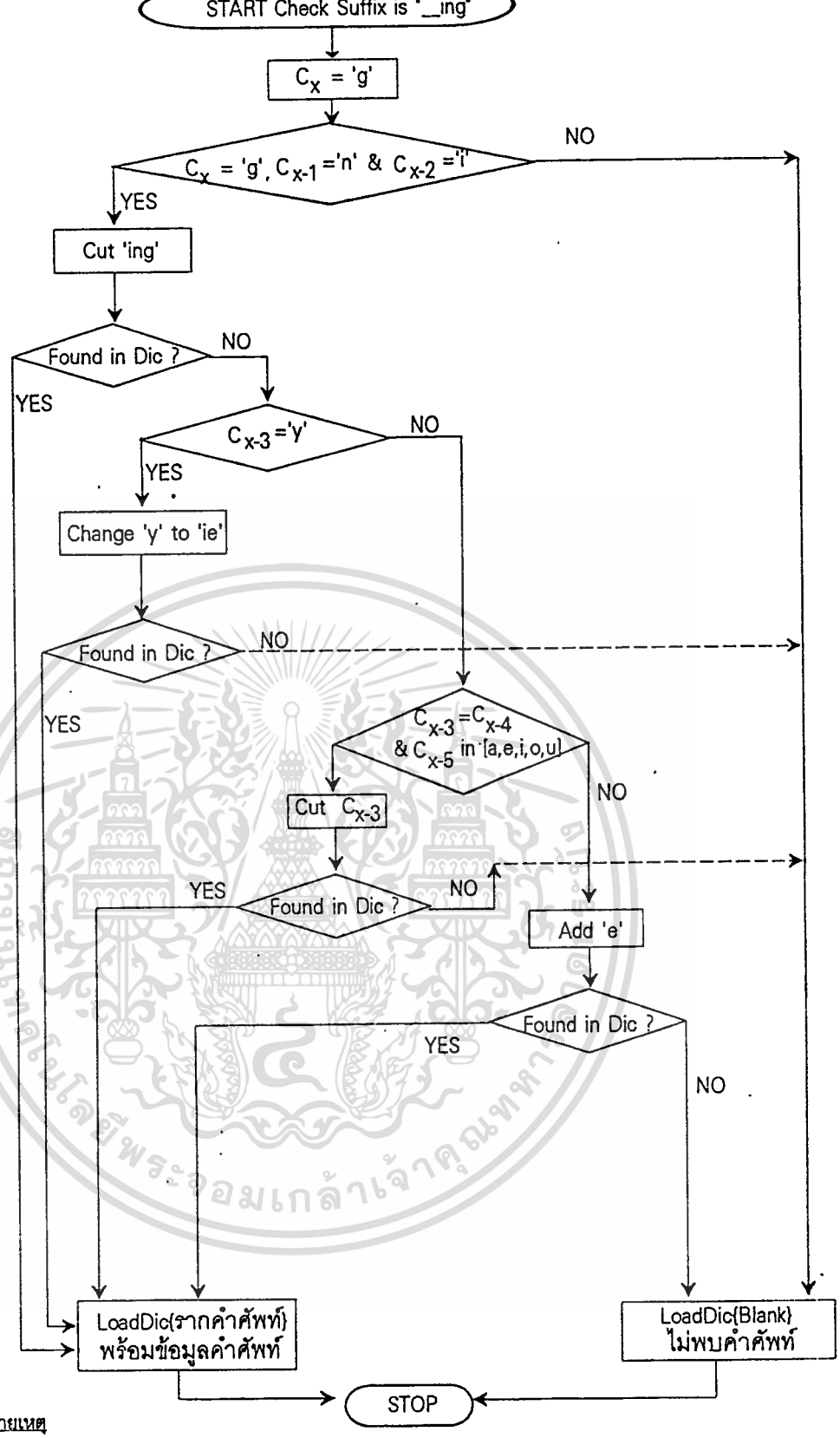
-ถ้าพบไป Step 8 และ

-ถ้าไม่พบไป Step 5

Step 8 : ทำการโหลดข้อมูลของหน่วยคำจากคำศัพท์ที่พบด้วยคำสั่ง LoadDic(รากศัพท์ที่พบ) กลับไปให้ขบวนการวิเคราะห์แล้วหยุดการค้นหา

อัลกอริทึมนี้แสดงในโฟล์วชาร์ตรูปที่ 5.4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



หมายเหตุ

Dic : หมายถึงพจนานุกรมคำศัพท์ส่วนโครงสร้างและความหมาย

C_x : เป็นอักขระตัวสุดท้ายของหน่วยคำที่พิจารณา

x : เป็นตำแหน่งอักขระตัวสุดท้ายของหน่วยคำที่พิจารณา

โดยที่ x เป็นเลขจำนวนเต็มบวก

รูปที่ 5.4 แสดงโฟล์วชาร์ตของอัลกอริทึมในกลุ่มที่ 3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กลุ่มที่ 4 อัลกอริทึมของการตัดส่วนเสริมท้ายหน่วยคำที่เป็น "___er" หรือ "___est"

ณ ขณะนี้อักขระ C_X คือ "r" หรือ อักขระ C_X คือ "t" เริ่มการประมวลผลตามลำดับดังนี้

Step 1 : -ถ้าอักขระ C_X คือ "r" อักขระ C_{X-1} คือ "e" ให้ตัดอักขระ C_X และอักขระ C_{X-1} (นั่นคือตัด "er") และให้ n = 2* แล้วไป Step 3 และ

-ถ้าไม่ใช่ให้ไปที่ Step 2

(ตัวอย่างเช่น taller ---> tall เป็นต้น)

Step 2 : -ถ้าอักขระ C_X คือ "t" อักขระ C_{X-1} คือ "s" และอักขระ C_{X-2} คือ "e" ให้ตัดอักขระ C_X ,C_{X-1} และอักขระ C_{X-2} (นั่นคือตัด "est") และให้ n=3* แล้วไป Step 3 และ

-ถ้าไม่ใช่ให้ไปที่ Step 7

(ตัวอย่างเช่น highest ---> high เป็นต้น)

Step 3 : นำหน่วยคำที่เหลือไปค้นหาข้อมูลจากส่วนการค้นหาข้อมูลทางโครงสร้างและความหมายของคำศัพท์จากพจนานุกรม

-ถ้าพบไปที่ Step 9

-ถ้าไม่พบให้ไปที่ Step 4

Step 4 : -ถ้าอักขระที่อยู่ก่อน "e" (C_{X-n}) คือ "i" และอักขระ C_{X-n-1} ไม่เป็นสมาชิกใน SetVowel ให้เปลี่ยน "i" เป็น "y" แล้วไป Step 8 และ

-ถ้าไม่ใช่ให้ไป Step 5 (ตัวอย่างเช่น drier,driest --> dry เป็นต้น)

Step 5 : -ถ้าอักขระ C_{X-n} และอักขระ C_{X-n-1} เป็นอักขระที่ซ้ำกัน และอักขระ C_{X-n-2} เป็นสมาชิกใน SetVowel ให้ตัดอักขระ C_{X-n} แล้วไป Step 8 และ

-ถ้าไม่ใช่ให้ไป Step 6 (ตัวอย่างเช่น bigger, biggest ---> big เป็นต้น)

Step 6 : ให้เพิ่ม "e" ที่ท้ายหน่วยคำแล้วไป Step 8

(ตัวอย่างเช่น braver, bravest ---> brave, truer,truest ---> true เป็นต้น)

Step 7 : ให้ส่งผลไม่พบหน่วยคำนั้นในพจนานุกรมคำศัพท์ด้วยคำสั่ง LoadDic{Blank} แล้วหยุดการค้นหา

Step 8 : นำหน่วยคำที่ปรับปรุงแล้วไปค้นหาข้อมูลจากส่วนการค้นหาข้อมูลทางโครงสร้างและความหมายของคำศัพท์จากพจนานุกรม -ถ้าพบไป Step 9 และ -ถ้าไม่พบไป Step 7

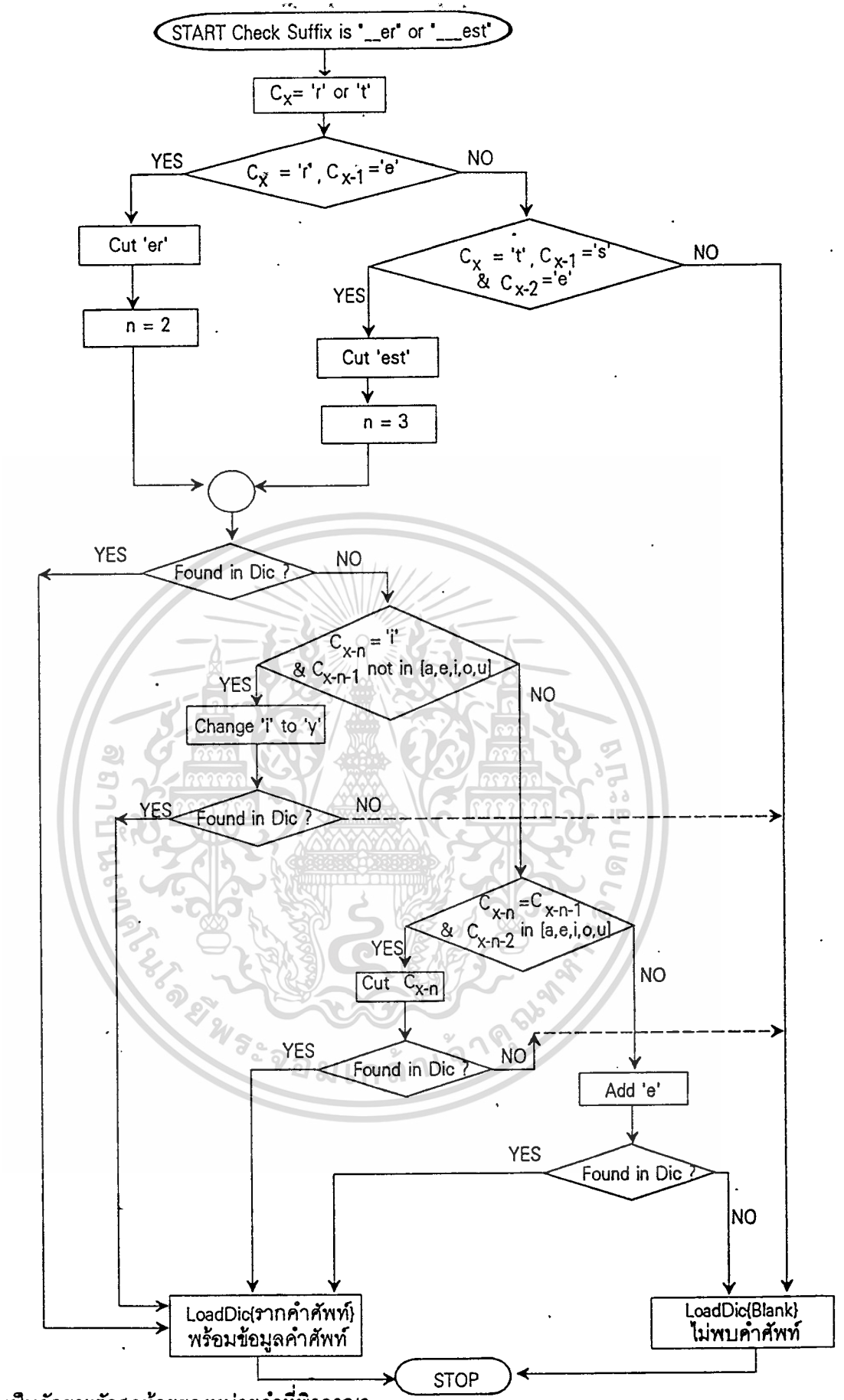
Step 9 : ทำการโหลดข้อมูลของหน่วยคำจากคำศัพท์ที่พบด้วยคำสั่ง LoadDic{รากศัพท์ที่พบ} กลับไปให้ขบวนการวิเคราะห์แล้วหยุดการค้นหา

อัลกอริทึมนี้แสดงในโฟล์วชาร์ตรูปที่ 5.5

---

* หมายเหตุ n=2 เมื่อ Suffix คือ "___er" และ n=3 เมื่อ Suffix คือ "___est"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



หมายเหตุ

C_x : เป็นอักขระตัวสุดท้ายของหน่วยคำที่พิจารณา

x : เป็นตำแหน่งอักขระตัวสุดท้ายของหน่วยคำที่พิจารณา โดยที่ x เป็นเลขจำนวนเต็มบวก

Dic : หมายถึงพจนานุกรมคำศัพท์ส่วนโครงสร้างและความหมาย

n = 2 เมื่อ Suffix คือ "__er" และ n=3 เมื่อ Suffix คือ "___est"

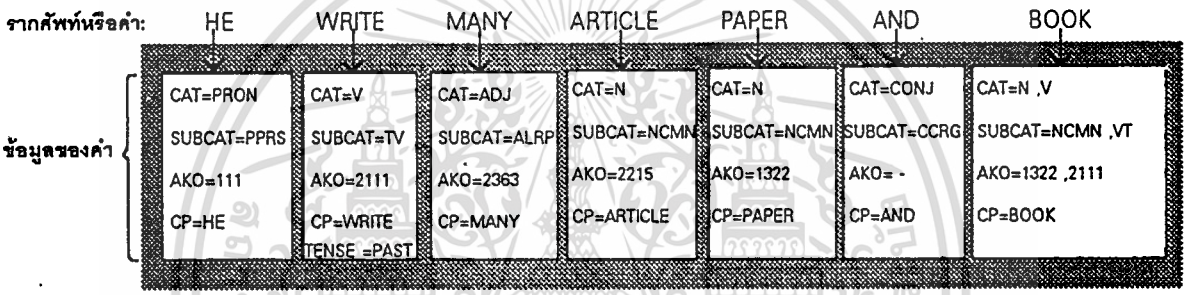
รูปที่ 5.5 แสดงโฟล์วชาร์ตของอัลกอริทึมในกลุ่มที่ 4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.3 ส่วนการวิเคราะห์ในระดับวากยสัมพันธ์ (Syntax Analysis)

เป็นขบวนการวิเคราะห์โครงสร้างของประโยคภาษาอังกฤษจากผลลัพธ์ข้อมูลทางภาษาที่ได้จากขบวนการวิเคราะห์ในระดับหน่วยคำ ซึ่งจะเริ่มวิเคราะห์จากคำรากศัพท์และข้อมูลแบบโครงสร้างและความหมายของคำรากศัพท์นั้น ๆ เพื่อค้นหาความสัมพันธ์ระหว่างรากศัพท์ของหน่วยคำ โดยอาศัยกฎไวยากรณ์ทางภาษาอังกฤษที่กำหนดขึ้นประกอบการประมวลผล และแสดงผลของความสัมพันธ์ในรูปของโครงสร้างต้นไม้ความสัมพันธ์แบบพึ่งพา (Dependency Relation Tree Structure) ต่อไป

ตัวอย่างข้อมูลเริ่มต้นของการประมวลผลในส่วนนี้ เช่น ผลจากการวิเคราะห์ในระดับหน่วยคำของประโยค "He wrote many articles, papers and books." เมื่อผ่านจะได้ผลลัพธ์ดังรูปข้างล่างนี้ ซึ่งจะอยู่ในรูปความสัมพันธ์เชิงเส้น (Linear)



ดังนั้นงานวิจัยนี้จึงได้ออกแบบเครื่องมือของขบวนการวิเคราะห์ (Parser Tool) ทางระบบซอฟท์แวร์ในรูปแบบของการใช้ตัวบ่งชี้ (Pointer หรือ Window) กำหนดตำแหน่งของคำที่จะพิจารณาวิเคราะห์เปรียบเทียบหาความสัมพันธ์กับคำข้างเคียงในลำดับต่าง ๆ กัน จากคำแรกหรือคำตำแหน่งซ้ายสุดของประโยคไปยังคำสุดท้ายหรือคำตำแหน่งขวาสุดของประโยคนั้น

นอกจากนั้นยังได้กำหนดฟังก์ชันและสัญญลักษณ์เพื่อการสร้างกฎไวยากรณ์ในรูปแบบของซอฟท์แวร์ด้วย

5.3.1 การออกแบบเครื่องมือของขบวนการวิเคราะห์ (Parser Tool)

ขบวนการวิเคราะห์หาความสัมพันธ์ระหว่างหน่วยคำในส่วนนี้เป็นการค้นหาหรือประมวลผลจากข้อมูลคำในเชิงเส้นตรงจากซ้ายไปขวา โดยต้องกำหนดคำที่กำลังพิจารณาเพื่อเปรียบเทียบกับคำข้างเคียงภายใต้กฎไวยากรณ์ที่กำหนดโดยจะเริ่มพิจารณาจากคำแรกสุดของประโยค งานวิจัยนี้จึงได้ออกแบบให้ตัวบ่งชี้ (ซึ่งต่อไปในวิทยานิพนธ์นี้จะเรียกว่า window) มี 3 แบบ ได้แก่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

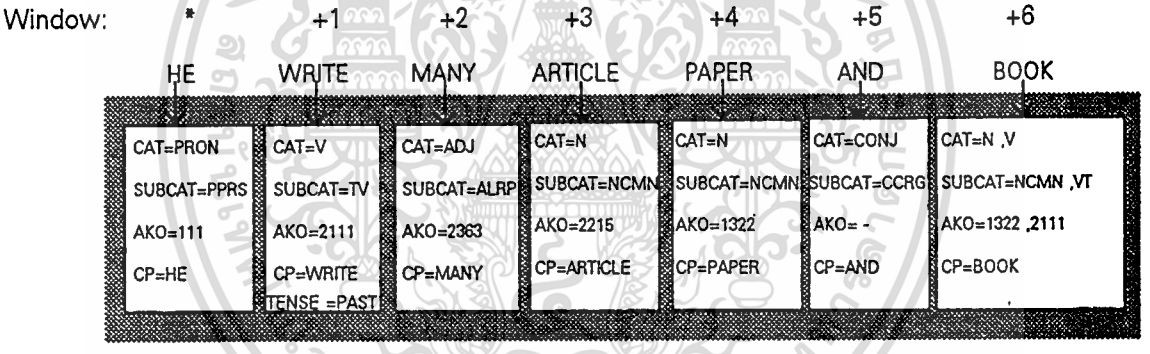
* : เป็น window ที่บ่งชี้คำที่กำลังพิจารณาจะเรียก window นี้ว่า "monitor window"

-n : เป็น window ที่บ่งชี้คำที่อยู่ก่อนคำที่กำลังพิจารณาไปทางซ้ายของประโยค อยู่ n ตำแหน่ง

+n : เป็น window ที่บ่งชี้คำที่อยู่หลังคำที่กำลังพิจารณาไปทางขวาของประโยค อยู่ n ตำแหน่ง

โดยที่ n = 1, 2, 3, ...

ตัวอย่างเช่น จากผลลัพธ์ที่ได้จากการวิเคราะห์ระดับหน่วยคำ ขบวนการวิเคราะห์ในส่วนนี้จะเริ่มจากการประมวลผลคำแรกสุดของประโยค นั่นคือเริ่มจากคำ "HE" ณ ขณะนี้เครื่องมือการวิเคราะห์จะทำการบ่งชี้คำ "HE" ด้วย monitor window (*) เพื่อให้ซอฟท์แวร์ของขบวนการวิเคราะห์ทราบว่าระบบวิเคราะห์กำลังพิจารณาคำที่บ่งชี้นี้เป็นหลักของการค้นหาความสัมพันธ์กับคำข้างเคียง ดังนั้นเครื่องมือวิเคราะห์ของระบบจะสร้าง window อื่น ๆ เพื่อบ่งชี้คำข้างเคียงที่ต้องการเปรียบเทียบกับคำที่พิจารณา ดังแสดงในรูปที่ 5.6

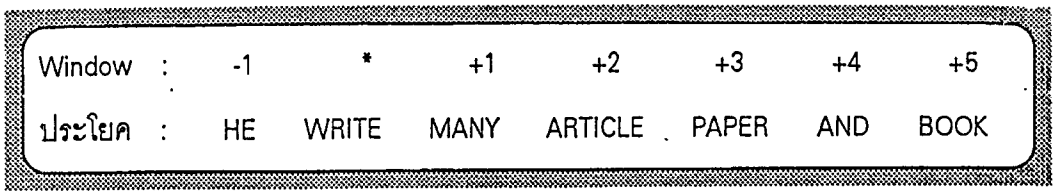


รูปที่ 5.6 แสดงตำแหน่ง window ณ จุดเริ่มต้นของขบวนการวิเคราะห์ในระดับวากยสัมพันธ์ที่คำแรกสุดของประโยค

จากตัวอย่างรูปที่ 5.6 ชี้ให้เห็นว่าขณะนี้ระบบวิเคราะห์กำลังทำการประมวลผลอยู่กับคำที่พิจารณาคือ "HE" โดยมี monitor window (*) บ่งชี้อยู่ ดังนั้น คำที่อยู่ทางขวาทั้งหมดของคำว่า "HE" จะถูกบ่งชี้ด้วย window +1, +2, +3, +4, +5, +6 ตามลำดับดังแสดงในรูป

และเมื่อการประมวลผลที่คำแรกสุดเสร็จสิ้นเรียบร้อยด้วยกฎไวยากรณ์ที่นำมาวิเคราะห์ monitor window(*) จะถูกเลื่อนไปยังคำที่อยู่ถัดไปทางขวาด้วยคำสั่ง ShiftWindow พร้อมกันนั้น window อื่น ๆ ก็จะเลื่อนตามไปด้วยดังแสดงในรูปที่ 5.7

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 5.7 แสดงตำแหน่งของ window ต่าง ๆ เมื่อ monitor window เปลี่ยนตำแหน่งของคำบ่งชี้

เนื่องจากงานวิจัยได้พิจารณาถึงลำดับความสำคัญของกฎไวยากรณ์ในการวิเคราะห์และได้จัดลำดับของการวิเคราะห์ โดยแบ่งกฎไวยากรณ์ออกเป็น 4 กลุ่มใหญ่ (ดังที่จะได้กล่าวในรายละเอียดในหัวข้อ 5.3.2) และกำหนดให้มีการประมวลโดยใช้กฎไวยากรณ์กลุ่มที่ 1 กับคำทุกคำในประโยคเสียก่อน แล้วจึงนำกฎไวยากรณ์ในกลุ่มที่ 2 มาประมวลผลกับคำทุกคำของประโยค โดยเริ่มจากคำแรกจนถึงคำสุดท้ายเช่นเดิม หลังจากนั้นก็นำกฎไวยากรณ์กลุ่มที่ 3 และกลุ่มที่ 4 มาประมวลผลกับคำทุกคำตามลำดับในทำนองเดียวกัน ดังนั้นเมื่อ monitor window(*) เลื่อนไปจนผ่านคำสุดท้ายของประโยคแล้ว (ในทางซอฟท์แวร์กำหนดไว้ว่า เมื่อ monitor window(*) บ่งชี้ความว่างเปล่าคือค่า NIL) ระบบวิเคราะห์จะเคลื่อน monitor window(*) กลับไปบ่งชี้คำแรกสุดของประโยคพร้อมกับเริ่มประมวลผลโดยใช้กฎไวยากรณ์ของกลุ่มถัดไป

<<กฎการเคลื่อนตำแหน่ง monitor window (*)>>

1. เริ่มแรกสุด monitor window จะบ่งชี้คำแรกสุดของประโยค
2. monitor window จะถูกเลื่อนไปบ่งชี้คำถัดไปทางขวา 1 คำ ด้วยคำสั่ง ShiftWindow
3. เมื่อ monitor window ถูกเลื่อนผ่านคำสุดท้ายของประโยค นั่นคือจะไปบ่งชี้ค่าความว่างเปล่า (NIL) ระบบวิเคราะห์จะเลื่อน monitor window กลับไปบ่งชี้คำแรกสุดของประโยคใหม่

5.3.2 การสร้างกฎไวยากรณ์ของขบวนการวิเคราะห์

ดังที่กล่าวมาแล้วว่าในการวิเคราะห์หาความสัมพันธ์ของคำในประโยคในระดับวากยสัมพันธ์นี้จะต้องประมวลผลโดยใช้ฐานข้อมูลทางกฎไวยากรณ์ (Grammar Rules) ของภาษาไปพร้อมกับฐานข้อมูลทางภาษาของคำที่เก็บไว้ในพจนานุกรมมาประกอบการวิเคราะห์ และอาศัยเครื่องมือการวิเคราะห์ (ที่กล่าวมาในข้อ 5.3.1) เป็นส่วนช่วยในการกำหนดคำพิจารณา

งานวิจัยนี้ได้รวบรวมฐานความรู้เกี่ยวกับไวยากรณ์ภาษาอังกฤษที่ใช้ในการวิเคราะห์ในระดับวากยสัมพันธ์นี้ไว้ทั้งหมดในบทที่ 2 ดังที่กล่าวมาแล้ว และจากฐานความรู้เหล่านั้น งานวิจัยนี้ได้สรุปแบ่งกลุ่มของกฎไวยากรณ์ภาษาอังกฤษที่ใช้ในขบวนการวิเคราะห์ออกเป็น 4 กลุ่มใหญ่ ได้แก่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กลุ่มที่ 1 กฎไวยากรณ์ของการวิเคราะห์หานามวลี (Grammar Rules for Noun Phrase Construction)

กลุ่มที่ 2 กฎไวยากรณ์ของการวิเคราะห์หาวลีของกริยาช่วย (Grammar Rules for Auxiliary Phrase Construction)

กลุ่มที่ 3 กฎไวยากรณ์ของการวิเคราะห์ส่วนขยายประโยค (Grammar Rules for Optional Construction)

กลุ่มที่ 4 กฎไวยากรณ์ของการวิเคราะห์หารูปแบบคำกริยา (Grammar Rules for Verb Pattern Construction)

ระบบวิเคราะห์จะเริ่มจากการประมวลผลโดยใช้กฎไวยากรณ์ในกลุ่มที่ 1 พิจารณากับคำทุกคำในประโยคจนสิ้นสุดคำสุดท้าย แล้วจึงนำกฎไวยากรณ์กลุ่มที่ 2 มาพิจารณากับคำทุกคำในประโยคจนสิ้นสุดคำสุดท้ายเช่นกัน แล้วจึงนำกฎไวยากรณ์กลุ่มที่ 3 และกลุ่มที่ 4 มาปฏิบัติในทำนองเดียวกัน

การลำดับความสำคัญของกฎไวยากรณ์ให้มีการประมวลผลก่อน/หลังกันนี้ พิจารณากำหนดขึ้นตามหลักไวยากรณ์ภาษาอังกฤษของหน่วยทางภาษาที่ควรจะมีการวิเคราะห์ให้พบเป็นลำดับไป ก่อนที่จะกล่าวถึงรายละเอียดของกฎไวยากรณ์ในกลุ่มต่าง ๆ ในลำดับต่อไป งานวิจัยได้กำหนดรูปแบบของการสร้างกฎไวยากรณ์ (Rule Formats) และความหมายของฟังก์ชันต่าง ๆ ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

<<รูปแบบของกฎไวยากรณ์>>

Rule No. / Rule Name /

{COND1 [& COND2 & COND3...]}

= {ACTION1 [, ACTION2, ACTION3,...]}

กำหนดให้

Rule No. : หมายถึงลำดับของกฎ เช่น R1 คือกฎไวยากรณ์ที่ 1 เป็นต้น

Rule Name : หมายถึงชื่อของกฎที่แสดงให้เห็นว่าเป็นกฎไวยากรณ์ใน การวิเคราะห์เกี่ยวกับอะไร

COND : หมายถึงเงื่อนไขของการพิจารณา ถ้าเงื่อนไขเป็นจริงทั้งหมดกับคำที่กำลัง พิจารณา แสดงว่าเกิดความสำเร็จของการวิเคราะห์กับคำที่พิจารณา เรียกว่า "SUCCESS" ระบบก็จะทำการสร้างผลลัพธ์ให้กับขบวนการวิเคราะห์ ด้วยการปฏิบัติตามคำสั่ง ACTION

ACTION : หมายถึงการประมวลผลที่จะให้ระบบวิเคราะห์ปฏิบัติในกรณีที่ COND ทั้งหมดเป็นจริง และระบบต้องปฏิบัติตามทุก ACTION ที่อยู่หลัง เครื่องหมาย "="

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

<<นิยามของฟังก์ชันและสัญญลักษณ์>>

- CHKCAT(P,Val1,Val2,Val3) : ตรวจสอบว่าคำ ณ ตำแหน่ง P โดยพิจารณาเงื่อนไขดังนี้

  1) มีประเภทของคำเท่ากับ Val1 และ

  2) มีประเภทของคำย่อยเท่ากับ Val2 และ

  3) มีชนิดของคำเท่ากับ Val3 หรือไม่

  ผลลัพธ์ของฟังก์ชันนี้จะะให้ค่าเป็นจริงก็ต่อเมื่อเงื่อนไขทั้ง 3 ข้อเป็นจริงทั้งหมด

- LINK(P1,P2,Val) : ทำการเชื่อมระหว่างคำ ณ ตำแหน่ง P1 กับ P2 ด้วยค่า Val และฟังก์ชันนี้จะให้ผลลัพธ์ดังนี้

  - ณ คำตำแหน่ง P1 เป็น Head และคำ ณ ตำแหน่ง P2 เป็น depend ในลักษณะ P1—Val—> P2 จะมีการยุบคำรวมระหว่าง P1,P2 เกิดเป็นคำ P1' ใหม่ที่มีข้อมูลมากขึ้น

- DELCAT(P,Val) : ในกรณีที่คำนั้นมีประเภทของคำมากกว่าหนึ่ง เมื่อผ่านฟังก์ชันนี้ จะทำการตัดประเภทของคำอื่นๆ ทิ้งเหลือประเภทของคำเท่ากับ Val ประเภทเดียว ณ คำที่ ตำแหน่ง P

- ASS((P,Var.{Val}) : เป็นฟังก์ชันให้ค่าฟิลด์ Var มีค่าเท่ากับ Val ณ คำตำแหน่งที่ P

- DELNODE(P) : เป็นฟังก์ชันที่ตัดคำ ณ ตำแหน่ง P ทิ้ง

- CHKENTRY(P,Val) : เป็นฟังก์ชันที่ตรวจสอบว่าคำ ณ ตำแหน่ง P มีค่าเท่ากับ Val หรือไม่ซึ่งผลลัพธ์ที่ได้จะให้ค่าฟังก์ชันเป็นค่าจริงหรือเท็จ

- FINDPrep(HEAD,P) : เป็นฟังก์ชันตรวจสอบว่า คำ ณ ตำแหน่งใดมีค่าประเภทของคำเป็นคำบุพบทได้ ซึ่งจะส่งตำแหน่งที่ได้กลับมาเท่ากับตำแหน่ง P

- FINDVerb(HEAD,P) : ตรวจสอบว่าที่ตำแหน่งใดมีค่าประเภทของคำเป็นคำกริยาได้ ซึ่งจะส่งตำแหน่งที่ได้กลับมาเท่ากับตำแหน่ง P

- เครื่องหมาย "&" : เป็นเครื่องหมายที่เชื่อม (and) ระหว่างเงื่อนไขที่พิจารณาว่าเงื่อนไขที่มาเชื่อมเป็นจริงทั้งคู่หรือไม่เช่น ถ้ากำหนดให้จริงมีสัญญลักษณ์ "T" และเท็จมีสัญญลักษณ์ "F" ดังนั้นถ้า X & Y สามารถอธิบายตามหลักตรรกศาสตร์ได้ดังนี้

  -ได้ผลลัพธ์เป็นจริงกรณีเดียว คือ ถ้า X และ Y มีค่าเป็นจริงทั้งคู่ นอกนั้นไม่ว่าค่าของ X หรือ Y ตัวใดตัวหนึ่งมีค่าเป็นเท็จจะได้ผลลัพธ์เป็นเท็จ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

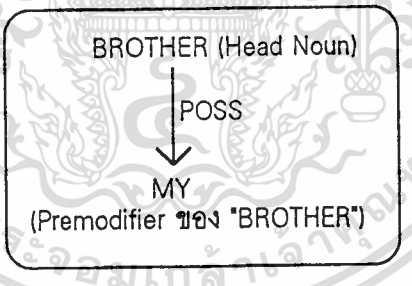
5.3.2.1 กฎไวยากรณ์ของการวิเคราะห์หานามวลี (กลุ่มที่ 1)

กฎไวยากรณ์ในกลุ่มนี้จะทำการเชื่อม (link) คำต่าง ๆ ที่ประกอบเป็นนามวลี (Noun Phrase) ด้วยส่วนแสดงความสัมพันธ์ของหน่วยความหมาย (กล่าวมาแล้วในบทที่ 4 ข้อ4.3) โดยการค้นหาคำที่ทำหน้าที่เป็นคำนามหลักของนามวลีนั้น ๆ เรียกว่า "Head Noun" เช่นถ้ามีนามวลีว่า "My brother" จะเห็นว่าคำนามหลัก(Head Noun) ของนามวลีนี้ก็คือคำว่า "brother" (จะพิจารณาคำที่เมื่อถูกตัดทิ้งแล้วจะทำให้นามวลีนั้นไม่เป็นนามวลีอีกต่อไป) ถ้านามวลีข้างต้นไม่มีคำว่า "My" เหลือแต่คำว่า "brother" ก็ทำให้นามวลีดังกล่าวเป็นนามวลีได้ ในทางตรงกันข้ามถ้าเราตัดคำว่า "brother" ทิ้งเหลือแต่คำว่า "My" คำเดียวในทางไวยากรณ์ภาษาอังกฤษถือว่าผิดหลักไวยากรณ์เพราะคำนี้เป็นคำที่แสดงความเป็นเจ้าของอย่างน้อยต้องมีคำนามตามมา 1 คำ (ซึ่งก็คือคำนามหลักนั่นเอง)

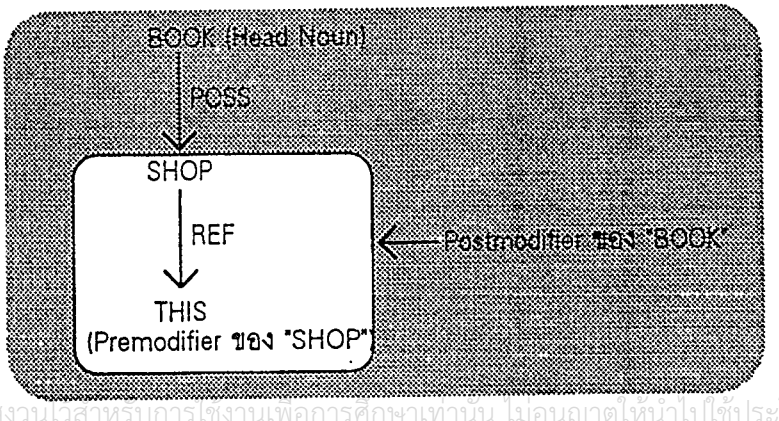
โดยคำนามหลักจะมีส่วนขยายหน้าและส่วนขยายหลัง ประกอบเป็นนามวลี ซึ่งเรียกส่วนขยายหน้าคำนามหลัก(Head Noun) ว่า "Premodifier" และเรียกส่วนที่ขยายหลังคำนามหลัก ว่า "Postmodifier" และเมื่อนำนามวลีมาผ่านกฎไวยากรณ์ในกลุ่มนี้จะทำให้ได้ผลลัพธ์เป็นการเชื่อมกันระหว่างคำที่ประกอบเป็นนามวลี เช่น

ตัวอย่างนามวลี

- "My brother" เมื่อผ่านกฎไวยากรณ์ในกลุ่มนี้จะได้ผลลัพธ์ดังรูปข้างล่างนี้



- "Book of this shop" เมื่อผ่านกฎไวยากรณ์ในกลุ่มนี้จะได้ผลลัพธ์ดังรูปข้างล่างนี้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ดังนั้นกฎไวยากรณ์ในกลุ่มนี้จึงแบ่งเป็น 2 กลุ่มคือ

กลุ่มที่ 1.1 กฎการหาส่วนขยายหน้าคำนามหลัก (Premodifier) หมายถึงคำหรือกลุ่มคำวลีที่ทำหน้าที่เป็นส่วนขยายหน้าคำนามหลัก เช่น คำชี้เฉพาะ (Determiner) คำคุณศัพท์ (Adjective) คำนาม (Noun) และคำกริยาวิเศษณ์ (Adverb) เป็นต้น

รูปแบบคือ : [DET] + [ADV] + [ADJ] + [N_2] + N_1/Pron

กลุ่มที่ 1.2 กฎการหาส่วนขยายหลังคำนามหลัก (Postmodifier) หมายถึงคำหรือกลุ่มคำวลีที่ทำหน้าที่เป็นส่วนขยายหลังคำนามหลัก เช่น คำบุพบท (Preposition), คำนาม (Noun) เป็นต้น ซึ่งมีการหาความสัมพันธ์ระหว่างคำนาม (Head Noun : N_1) คำบุพบท(PREP) และคำนาม (N_2) ดังรูปแบบข้างล่างนี้

รูปแบบคือ : N_1(ซึ่งเป็น head มาจาก Premodifier) + PREP + N_2

ในการสร้างกฎไวยากรณ์จะอาศัยฐานความรู้ทางไวยากรณ์อังกฤษที่กล่าวในบทที่ 2 เกี่ยวกับนามวลีได้แก่

1) NP ⟶ [DET] + N

2) NP ⟶ [DET] + [ADJ] + N

3) NP ⟶ [DET] + N + [PP]

4) NP ⟶ [DET] + [ADJ] + N + [PP]

5) NP ⟶ NP + NP

โดยที่คำนามที่ทำหน้าขยายคำนามด้วยกันนั้น จะให้สัญญลักษณ์เป็น N_1 และ N_2 ตามลำดับ ซึ่งจะปรากฏอยู่ในกฎไวยากรณ์ต่าง ๆข้างล่างนี้

จากฐานความรู้ไวยากรณ์ภาษาอังกฤษข้างต้นนำมาสร้างเป็นกฎไวยากรณ์ดังมีรายละเอียดต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

<<กลุ่มที่ 1.1 กฎการหาส่วนขยายหน้าคำนามหลัก>>

แบ่งเป็นกฎการหาส่วนขยายหน้าคำนามหลักย่อย 6 กฎได้แก่

<กฎย่อย R1 การหาความสัมพันธ์ระหว่างคำกริยาวิเศษณ์ (ADV) กับคำคุณศัพท์ (ADJ)>

ที่คำปัจจุบัน(*) เป็นคำคุณศัพท์และคำทางซ้าย(-1) เป็นคำกริยาวิเศษณ์ได้แล้วเราจะได้ความสัมพันธ์เชิงการกของคำทั้งสองคือ MOD (assign case(*,-1) = MOD) รวมทั้งเชื่อมคำระหว่างคำกริยาวิเศษณ์และคำคุณศัพท์ โดยที่คำคุณศัพท์เป็นคำหลัก (head,ADJ –> ADV) และมีการตัดความกำกวมประเภทของคำ (Disambiguous Category) คือที่คำปัจจุบัน(*) จะให้มีประเภทของคำเป็นคำคุณศัพท์และที่คำทางซ้าย(-1) จะมีประเภทของคำเป็นกริยาวิเศษณ์ เขียนเป็นรูปแบบกฎได้ดังนี้

```
R1. / link ADV, ADJ /
        CHKCAT(*,ADJ,'','') & CHKCAT(-1,ADV,'','')
= LINK(*,-1,MOD), DELCAT(*,ADJ),  DELCAT(-1,ADV)
```

<กฎย่อย R2 การหาความสัมพันธ์ระหว่างคำนาม (N_1) และคำนาม (N_2) >

โดยที่คำปัจจุบัน(*) เป็นคำนามได้ (N_2) และคำทางซ้าย(-1) เป็นคำนามได้ (N_1) ซึ่งจะมีคุณสมบัติดังนี้

<u>กรณี</u>

-N_1 เป็นคำนามประเภท NPSS แล้ว N_2 เป็นคำนามประเภท NCMN และ assign case(*,-1) = POSS

-N_1 เป็นคำนามประเภท NPRP แล้ว N_2 เป็นคำนามประเภท NCMN และ assign case(*,-1) = NAM

-N_1 เป็นคำนามประเภท NALN แล้ว เป็น N_2 เป็นคำนามประเภท NCMN และ assign case(*,-1) =ATT

-N_1 เป็นคำนามประเภท NCNM แล้ว เป็น N_2 เป็นคำนามประเภท NCMN และ assign case(*,-1) =QUAN

-N_1 เป็นคำนามประเภท NCMN มี AKO = 112 แล้ว และ N_2เป็นคำนามประเภท NCMN และ assign case(*,-1) = MOD

-N_1 เป็นคำนามประเภทนอกเหนือจากกฎข้างบนแล้ว N_2 เป็นคำนามประเภท NCMN และ assign case(*,-1) = NMOD

แล้วจึง link ระหว่าง N_1,N_2 โดยที่ N_2 เป็น head (N_2 –> N_1) และ มีการตัดความกำกวมประเภทของคำ (Disambiguous Category) คือ ที่คำปัจจุบัน(*) จะมีประเภทของคำเป็นคำนามและที่คำทางซ้าย(-1) จะมีประเภทของคำเป็นคำนามเขียนเป็นรูปแบบกฎได้ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
R2 / link N1 , N2 /
      R2.1 CHKCAT(*,N2 ,NCMN,'') & CHKCAT(-1,N1 ,NPSS,'')
=LINK(*,-1,POSS),DELCAT(*,N),DELCAT(-1,N)
      R2.2 CHKCAT(*,N2 ,NCMN,'') & CHKCAT(-1,N1 ,NPRP,'')
= LINK(*,-1,NAM), DELCAT(*,N),  DELCAT(-1,N)
      R2.3 CHKCAT(*,N2 ,NCMN,'') & CHKCAT(-1,N1 ,NALN,'')
= LINK(*,-1,ATT), DELCAT(*,N),  DELCAT(-1,N)
      R2.4 CHKCAT(*,N2 ,NCMN,'') & CHKCAT(-1,N1 ,NCNM,'')
= LINK(*,-1,QUAN), DELCAT(*,N),  DELCAT(-1,N)
      R2.5CHKCAT(*,N2,NCMN,'')& CHKCAT(-1,N1 ,NCMN,112)
= LINK(*,-1,MOD), DELCAT(*,N),  DELCAT(-1,N)
      R2.6 CHKCAT(*,N2 ,NCMN,'') & CHKCAT(-1,N1 ,'','')
= LINK(*,-1,NMOD), DELCAT(*,N),  DELCAT(-1,N))
```

<กฎย่อย R3 การหาความสัมพันธ์ระหว่างคำคุณศัพท์ (ADJ) และคำนาม (N)>

ที่คำปัจจุบัน(*)เป็นคำนามและคำทางซ้าย(-1) เป็นคำคุณศัพท์โดยคำทางซ้ายมีคุณสมบัติดังนี้

<u>กรณี</u>

-มี AKO = 2363 หรือ 2361 และ assign case(*,-1) = ATT

-มี AKO = 232 และ assign case(*,-1) = MOD

และจะรวมทั้ง link ระหว่าง ADJ,N โดยที่ N เป็น head (N —>ADJ) และมีการตัดความกำกวมประเภทของคำ (Disambiguous Category) คือ ที่คำปัจจุบัน(*) จะมีประเภทของคำเป็นคำนาม และที่คำทางซ้าย(-1) จะมีประเภทของคำเป็นคำคุณศัพท์ เขียนเป็นรูปแบบกฎได้ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
R3 / link ADJ, N /
        R3.1 CHKCAT(*,N,'','') & CHKCAT(-1,ADJ,'',2363)
= LINK(*,-1,ATT), DELCAT(*,N),  DELCAT(-1,ADJ)
        R3.2 CHKCAT(*,N,'','') & CHKCAT(-1,ADJ,'',2361)
= LINK(*,-1,ATT), DELCAT(*,N),  DELCAT(-1,ADJ)
        R3.3 CHKCAT(*,N,'','') & CHKCAT(-1,ADJ,'',232)
= LINK(*,-1,MOD), DELCAT(*,N),  DELCAT(-1,ADJ)
```

<กฎย่อย R4 การหาความสัมพันธ์ระหว่างคำนาม(N_1) กับคำนาม (N_2)>

คำนามที่มีประเภทของคำนามย่อยเป็นคำที่แสดงความเป็นเจ้าของ (NPSS) และคำนาม (N_2) ที่คำปัจจุบัน(*)เป็นคำนามได้ (N_2)แล้ว และคำทางซ้าย(-1) เป็นคำนามประเภท NPSS ได้แล้ว และคำที่อยู่ก่อนคำทางซ้าย(-2) เป็นคำนาม (N_1) ได้ยกตัวอย่างเช่น

```
Somsak 's  book
-2     -1   *
```

และ assign case(*,-2)= POSS โดยตัดคำ -1 ทิ้ง แล้ว link ระหว่าง N_1,N_2 โดยที่ N_2 เป็น head (N_2 --> N_1) และมีการตัดความกำกวมประเภทของคำ (Disambiguous Category) คือ ที่คำ(*) จะมีประเภทของคำเป็นคำนาม (N) และที่คำก่อนคำทางซ้าย(-2) จะมีประเภทของคำเป็นคำนาม (N) เขียนเป็นรูปแบบกฎได้ดังนี้

```
R4 / link N2, NPSS, N1 /
        CHKCAT(*,N,'','') & CHKCAT(-1,N,NPSS,'') & CHKCAT(-2,N,'','')
= LINK(*,-2,POSS), DELCAT(*,N),  DELCAT(-2,N), DELNODE(-1)
```

<กฎย่อย R5 การหาความสัมพันธ์ระหว่างคำชี้เฉพาะ (DET)กับคำนาม (N) >

ที่คำปัจจุบัน(*) เป็นคำนามได้แล้ว และคำทางซ้าย(-1) เป็นคำชี้เฉพาะได้แล้ว โดยคำชี้เฉพาะมีคุณสมบัติดังนี้

<u>กรณี</u>

- เป็นคำคุณศัพท์ประเภท DPRS และ assign case(*,-1) = POSS หรือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- เป็นคำคุณศัพท์ประเภท DCNM และ assign case(*,-1) = QUAN หรือ
- เป็นคำคุณศัพท์ประเภท DDMN และ assign case(*,-1) = REF หรือ
- เป็นคำคุณศัพท์ประเภท DINA และ assign case(*,-1) = VAL

และจึง link ระหว่าง DET,N โดยที่ N เป็น head (N —> DET) และมีการตัดความกำกวมประเภทของคำ (Disambiguous Category) คือ ที่คำปัจจุบัน(*)จะมีประเภทของคำเป็นคำนาม(N) และที่คำทางซ้าย(-1) มีประเภทของคำเป็นคำชี้เฉพาะ (DET) เป็นคำชี้เฉพาะประเภทย่อย DTHE หรือ ประเภท DAAN และ assign field DET.{THE} หรือ assign field DET.{A} หรือ DET.{AN} ไว้ที่คำปัจจุบัน(*) และตัดคำทางซ้าย(-1) ทิ้ง

เขียนเป็นรูปแบบกฎได้ดังนี้

```
R5 / link DET , N /
     R5.1 CHKCAT(*,N ,'','') & CHKCAT(-1,DET ,DPRS,'')
= LINK(*,-1,POSS), DELCAT(*,N),  DELCAT(-1,DET)
     R5.2 CHKCAT(*,N ,'','') & CHKCAT(-1,DET,DCNM ,'')
= LINK(*,-1,QUAN), DELCAT(*,N),  DELCAT(-1,DET)
     R5.3 CHKCAT(*,N ,'','') & CHKCAT(-1,DET ,DDMN,'')
= LINK(*,-1,REF), DELCAT(*,N),  DELCAT(-1,DET)
     R5.4 CHKCAT(*,N ,'','') & CHKCAT(-1,DET ,DINA,'')
= LINK(*,-1,VAL), DELCAT(*,N),  DELCAT(-1,DET)
     R5.5 CHKCAT(*,N ,'','') & CHKCAT(-1,DET ,DTHE,'')
= ASS(*,DET.{-1}), DELCAT(*,N),  DELNODE(-1)
     R5.6 CHKCAT(*,N ,'','') & CHKCAT(-1,DET ,DAAN,'')
= ASS(*,DET.{-1}), DELCAT(*,N),  DELNODE(-1)
```

<กฎย่อย R6 ที่คำปัจจุบัน(*)เป็นคำสรรพนาม(PRON) ประเภท PPRS ได้ และคำทางซ้าย (-1) ไม่มีคำใดๆ อยู่ >

ในกรณีนี้แสดงว่าเป็น Noun Pharse ได้เลยไม่ต้องรวมกับตัวใดแล้ว และมีการตัดความกำกวมประเภทของคำ (Disambiguous Category) คือที่คำปัจจุบัน (*) จะเป็นคำสรรพนาม (PRON) เขียนเป็นรูปแบบกฎดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
R6. / PRON /
    CHKCAT(*.PRON,PPRS,'') & CHKNODE(-1=NIL)
= DELCAT(*,PRON)
```

<<กลุ่มที่ 1.2 กฎการหาส่วนขยายหลังคำนามหลัก>>

กฎการหาส่วนขยายหลังคำนามหลักได้แก่ กฎการหาความสัมพันธ์ระหว่างคำนาม (N_1) คำบุพบท และคำนาม (N_2)

<กฎย่อย R7 การหาความสัมพันธ์ระหว่างคำนาม (N_1) คำบุพบท และคำนาม (N_2)>

พิจารณาที่คำปัจจุบัน(*) เป็นคำนาม(N_1) ได้แล้ว และคำทางขวา(+1) เป็นคำบุพบท(PREP) ได้ และคำที่ตามหลังคำทางขวา(+2) เป็นคำนาม(N_2) ได้แล้วให้ตัดคำทางขวา(+1) และassign field Subfea มีค่าเท่ากับ SUBFEA.{prep} ให้กับ case ที่จะมา link ระหว่าง (*,+2)

ถ้า PREP เป็น Of และ assign case(*,+2) = POSS โดย รวมทั้ง link ระหว่าง N_1,N_2 โดยที่ N_1 เป็น head (N_1 --> N_2)

และมีการตัดความกำกวมประเภทของคำ (Disambiguous Category) คือ ที่คำปัจจุบัน(*) จะมีประเภทของคำเป็นคำนาม และที่ คำทางขวา (+1) มีประเภทของคำเป็นคำบุพบท

เขียนเป็นรูปแบบกฎได้ดังนี้

```
R7 / link N PREP N /
    CHKCAT(*, N1 ,'','') & CHKCAT(+1, PREP,'','') &
    CHKCAT(+2,N2 ,'','') & CHKENTRY(+1,OF)
 =  LINK(*,+2,POSS)  ,  ASS(POSS,SUBFEA.{of}),  DELCAT(*,N),
DELCAT(+1,PREP),DELCAT(+2,N)
```

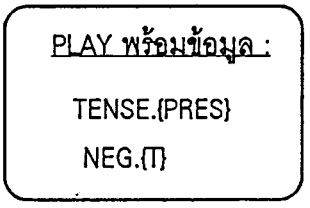
5.3.2.2 กฎไวยากรณ์ของการวิเคราะห์หาวลีของกริยาช่วย (กลุ่มที่ 2)

กฎไวยากรณ์ในกลุ่มจะทำการเพิ่มข้อมูลให้กับคำกริยาหลักของประโยค เช่น คำว่า "not" ที่ทำหน้าที่เป็นคำปฏิเสธ หรือ do, did, have, has เป็นต้น ที่ทำหน้าที่เป็นคำกริยาช่วยของคำกริยาหลักของประโยค

รูปแบบคือ : [AUX] + [NEG] + VP

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวอย่างเช่น จากประโยค "My brother does not play." พิจารณาวลี "does not play" ผ่านกฎไวยากรณ์ในกลุ่มนี้จะให้ข้อมูลแก่คำกริยาคือทำให้ทราบว่าประโยคอยู่ในรูปของกาลเป็นปัจจุบัน (TENSE.{PRES}) และกริยาเป็นกริยาที่เป็นปฏิเสธ(คำว่า "ไม่" ตามด้วยคำกริยา) ดังรูปข้างล่างนี้



โดยในกฎไวยากรณ์ในกลุ่มนี้แบ่งการทำงานเป็น 2 กฎย่อยคือ

กฎย่อย R8 กฎการให้ความหมายปฏิเสธกับคำกริยาหลักของประโยค

กฎย่อย R9 กฎการให้ความหมายของคำกริยาช่วย(AUX) แก่คำกริยาหลักของประโยค

โดยในแต่ละกฎไวยากรณ์มีรายละเอียดต่อไปนี้

<กฎย่อย R8 การให้ความหมายปฏิเสธกับคำกริยาหลักของประโยค

ถ้าคำปัจจุบัน(*) เป็นคำปฏิเสธและ assign field NEG.{T} ไว้ที่คำที่เป็นคำกริยาหลักของประโยค(V) ได้ ตัวสุดท้าย และตัดคำปัจจุบันทิ้ง เขียนเป็นรูปแบบกฎได้ดังนี้

```
R8 / NEG,V /
   FINDVerb(HEAD,P); CHKCAT(*,NEG,'','')
   = ASS(P,NEG.{T}), DELNODE(*) ,DELCAT(P,V)
```

<กฎย่อย R9 การให้ความหมายของกริยาช่วย (AUX) แก่คำกริยาหลักของประโยค>

ถ้าคำปัจจุบัน(*) เป็นกริยาช่วยได้ แล้ว และ มองไปทางขวา จนเจอคำที่เป็นคำกริยาหลักของประโยค(V) ได้

<u>กรณี</u>

-ถ้า * เป็นคำกริยาช่วยประเภท XMDL แล้ว assign XMDL.{ability} และ TENSE.{PRES} ที่คำขวาสุดที่เป็นคำกริยาได้

-ถ้า * เป็นคำกริยาช่วยประเภท XHAV แล้ว assign TENSE.{PPER} ที่คำขวาสุดที่เป็นคำกริยาได้

-ถ้า * เป็นคำกริยาช่วยประเภท XXBE และ มองเจอคำขวาสุดที่เป็นคำกริยาได้ กรณีเป็น

: คำกริยาเติม 'ing' แล้วให้ assign field XMDL.{progress} และ TENSE.{PRES} ที่คำขวาสุดที่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เป็นคำกริยาได้

: คำกริยาในรูปของ Past Tense แล้วให้ assign field XMDL.{pass} และ TENSE.{PAST} ที่คำขวาสุดที่เป็นคำกริยาได้

-ถ้า * เป็นคำกริยาช่วยประเภท XXDO แล้วให้ assign TENSE.{PRES} ที่คำขวาสุดที่เป็นคำกริยาได้ และให้ทำการตัดคำปัจจุบัน(*) ทิ้ง และมีการตัดความกำกวมประเภทของคำ (Disambiguous Category) คือที่คำขวาสุดจะมีประเภทของคำเป็นคำกริยา

เขียนเป็นรูปแบบกฎได้ดังนี้

```
R9 / AUX,V /
    FINDVerb(HEAD,P);
     R9.1 CHKCAT(*,AUX,XMDL,'')
=ASS(P,XMDL.{ability}), ASS(P,TENSE.{PRES}), DELNODE(*) ,DELCAT(P,V)
      R9.2 CHKCAT(*,AUX,XMDL,'')
=ASS(P,XMDL.{ability}), ASS(P,TENSE.{PRES}), DELNODE(*) ,DELCAT(P,V)
      R9.3 CHKCAT(*,AUX,XXBE,'') & ENDING(*,ING)
=ASS(P,XMDL.{progress}),ASS(P,TENSE.{PRES}),DELNODE(*) ,DELCAT(P,V)
      R9.4 CHKCAT(*,AUX,XXBE,'')  & ENDING(*,ED)
 = ASS(P,XMDL.{pass}), ASS(P,TENSE.{PAST}), DELNODE(*) ,DELCAT(P,V)
      R9.5 CHKCAT(*,AUX,XXDO,'')
=  ASS(P,TENSE.{PRES}), DELNODE(*) ,DELCAT(P,V)
```

5.3.2.3 กฎไวยากรณ์ของการวิเคราะห์ส่วนขยายประโยค (กลุ่มที่ 3)

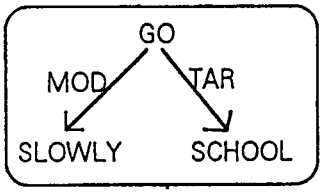
กฎไวยากรณ์ในกลุ่มนี้จะทำการจัดการกับส่วนที่เป็นส่วนขยายของประโยค ซึ่งส่วนขยายเป็นส่วนที่ให้ประโยคมีความชัดเจนมากยิ่งขึ้น ถึงแม้จะไม่มีส่วนนี้ก็ตามประโยคก็ยังได้ใจความเหมือนเดิม เช่น ประโยค "He go slowly to school." ถ้าตัดส่วนขยายออกจะเหลือประโยค "He go." ประโยคก็ยังถูกตามหลักไวยากรณ์ภาษาอังกฤษ

กฎไวยากรณ์ในกลุ่มนี้จะจัดการกับคำบุพบท โดยทำการเชื่อมระหว่างคำกริยาหลักของประโยคกับคำนามหลักที่อยู่หลังคำบุพบทที่พิจารณาด้วยส่วนแสดงความสัมพันธ์ของหน่วยความหมาย (Case Grammar ในบทที่ 4 ข้อ 4.3) และจัดการกับคำกริยาวิเศษณ์ให้กับคำกริยาหลักของประโยค

รูปแบบคือ : V + PREP + N และ [ADV] + V + [ADV] ตามลำดับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

<u>ตัวอย่างเช่น</u> ประโยค "He go slowly to school" เมื่อผ่านกฎไวยากรณ์ในกลุ่มนี้จะได้ผลลัพธ์ดังรูปข้างล่างนี้



ดังนั้นกฎไวยากรณ์ในกลุ่มนี้จึงแบ่งเป็น 2 กลุ่มคือ

กลุ่มที่ 3.1 กฎการจัดการกับคำบุพบท (Prepostion) ให้กับคำกริยาหลักของประโยค

กลุ่มที่ 3.2 กฎการจัดการกับคำกริยาวิเศษณ์ (Adverb)ให้กับคำกริยาหลักของประโยค

ในการสร้างกฎไวยากรณ์จะอาศัยฐานความรู้ทางไวยากรณ์อังกฤษที่กล่าวในบทที่ 2 เกี่ยวกับบุพบทวลีได้แก่

1) PP ——> PREP + [DET] + N

2) PP ——> PREP + [DET] + [ADJ] + N

และกริยาวิเศษณ์วลีได้แก่

ADVP ——> [ADV] + ADV

จากฐานความรู้ไวยากรณ์ภาษาอังกฤษข้างต้นนำมาสร้างเป็นกฎไวยากรณ์ดังมีรายละเอียดต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

<<กลุ่มที่ 3.1 กฎการจัดการกับคำบุพบทให้กับคำกริยาหลักของประโยค>>

กฎไวยากรณ์ในกลุ่มนี้เป็นกฎย่อยที่ R 10 ดังมีรายละเอียดดังนี้

<กฎย่อย R10 การหาความสัมพันธ์ระหว่างคำกริยาหลักของประโยคกับ บุพบท และคำนาม >

ถ้าคำปัจจุบัน(*) เป็นคำกริยามองจนเจอคำที่เป็นคำบุพบทได้(P) และที่คำทางขวาของคำบุพบทเป็นคำนาม(P+1) ได้แล้ว assign คำปัจจุบัน(*) เป็น field SUBFEA.{PREP} ไว้ที่ CASE ที่ link มายังคำ ทางขวาที่เป็น N ได้ คือ assign CASE ระหว่างคำกริยาหลักของประโยค และคำทางขวาของคำบุพบทที่เป็นในกรณี ที่ Cp of Marker = 'TO' แล้ว คำทางขวาของคำบุพบทที่เป็นคำนามได้ จะต้องเป็นคำกริยาไม่ได้ และถ้ากรณีที่ Cp of Marker มีค่าเท่ากับค่าในตารางข้างล่าง(คอลัมน์ที่ 1) นี้ ก็จะ assign CASE ระหว่างคำกริยาหลักของประโยค และคำทางขวาของคำบุพบทที่เป็นคำนาม(N) ได้ ด้วยค่า CASE ในคอลัมน์ที่ 3 ตามเงื่อนไขของ AKO (คอลัมน์ที่ 2 กรณีที่ค่า AKO = '-' หมายถึงไม่มีค่า AKO มาเกี่ยวข้องในการพิจารณาการ assign CASE)

Cp of Marker	คำทางขวาที่เป็นคำนาม ได้มีค่า AKO เท่ากับ	Assign CASE
OF	-	POSS
FROM,IN	-	SOR
WITH,BY,UPON	11	OBJ
WITH,BY	-	INS
TO,FOR	-	TAR
AT	232	TIM
AT	มากกว่า 232	TAR

เขียนเป็นรูปแบบกฎได้ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
R10 / link PREP,N /
    FINDPrep(HEAD,P), CHKCAT(*,V), DELCAT(*,V)
    R10.1 CHKENTRY(P,'OF') & CHKCAT(P+1,N,'','')
= ASS(POSS,SUBFEA.{of}), LINK(*,P+1,POSS), DELNODE(P) ,DELCAT(P+1,N)
    R10.2 CHKENTRY(P,'FROM') & CHKCAT(P+1,N,'','')
= ASS(SOR,SUBFEA.{from}), LINK(*,P+1,SOR), DELNODE(P) ,DELCAT(P+1,N)
    R10.3 CHKENTRY(P,'IN') & CHKCAT(P+1,N,'','')
= ASS(SOR,SUBFEA.{in}), LINK(*,P+1,SOR), DELNODE(P) ,DELCAT(P+1,N)
    R10.4 CHKENTRY(P,'WITH') & CHKCAT(P+1,N,'',11)
= ASS(OBJ,SUBFEA.{with}), LINK(*,P+1,OBJ), DELNODE(P) ,DELCAT(P+1,N)
    R10.5 CHKENTRY(P,'BY') & CHKCAT(P+1,N,'',11)
= ASS(OBJ,SUBFEA.{by}), LINK(*,P+1,OBJ),DELNODE(P) ,DELCAT(P+1,N)
    R10.6 CHKENTRY(P,'UPON') & CHKCAT(P+1,N,'',11)
= ASS(OBJ,SUBFEA.{upon}), LINK(*,P+1,OBJ), DELNODE(P) ,DELCAT(P+1,N)
    R10.7 CHKENTRY(P,'WITH') & CHKCAT(P+1,N,'',)
= ASS(INS,SUBFEA.{with}), LINK(*,P+1,INS), DELNODE(P) ,DELCAT(P+1,N)
    R10.8 CHKENTRY(P,'BY') & CHKCAT(P+1,N,'','')
= ASS(INS,SUBFEA.{by}), LINK(*,P+1,INS), DELNODE(P) DELCAT(P+1,N)
    R10.9 CHKENTRY(P,'TO') & CHKCAT(P+1,N,'','') & Not(CHKCAT(P+1,V,'',''))
= ASS(TAR,SUBFEA.{to}), LINK(*,P+1,TAR), DELNODE(P) ,DELCAT(P+1,N)
    R10.10 CHKENTRY(P,'FOR') & CHKCAT(P+1,N,'','')
= ASS(TAR,SUBFEA.{for}), LINK(*,P+1,TAR), DELNODE(P) ,DELCAT(P+1,N)
    R10.11 CHKENTRY(P,'AT') & CHKCAT(P+1,N,'','232')
= ASS(TIM,SUBFEA.{at}), LINK(*,P+1,TIM), DELNODE(P) ,DELCAT(P+1,N)
    R10.12 CHKENTRY(P,'AT') & CHKCAT(P+1,N,'',)
= ASS(TAR,SUBFEA.{at}), LINK(*,P+1,TAR), DELNODE(P) ,DELCAT(P+1,N)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

<<กลุ่มที่ 3.2 กฎการจัดการกับคำกริยาวิเศษณ์ให้กับคำกริยาหลักของประโยค>>

แบ่งเป็น 2 กฎย่อยได้แก่ กฎย่อย R11 และR12 ดังรายละเอียดต่อไปนี้

<กฎย่อย R11 การหาความสัมพันธ์ระหว่างคำกริยาวิเศษณ์กับคำกริยาหลักของประโยค>

ถ้าคำปัจจุบัน(*) เป็นคำกริยาแล้วคำทางซ้าย(-1) เป็นคำกริยาวิเศษณ์ แล้ว assign case(*,-1) = MOD แล้ว link ระหว่าง ADV,V โดยที่ V เป็น head (V --> ADV) เขียนเป็นรูปแบบกฎได้ดังนี้

```
R11 / link ADV,V /
      CHKCAT(*,V,'','') & CHKCAT(-1,ADV,'','')
= LINK(*,-1,MOD), DELCAT(*,V), DELCAT(-1,ADV)
```

<กฎย่อย R12 การหาความสัมพันธ์ระหว่างคำกริยาหลักของประโยคกับคำกริยาวิเศษณ์>

ถ้าคำปัจจุบัน(*) เป็นคำกริยาแล้วคำทางขวา(+1) เป็นคำกริยาวิเศษณ์แล้ว assign case(*,+1) = MOD แล้ว link ระหว่าง V,ADV โดยที่ V เป็น head (V --> ADV) เขียนเป็นรูปแบบกฎได้ดังนี้

```
R12 / link V,ADV /
      CHKCAT(*,V,'','') & CHKCAT(+1,ADV,'','')
= LINK(*,+1,MOD), DELCAT(*,V), DELCAT(+1,ADV)
```

5.3.2.4 กฎไวยากรณ์ของการวิเคราะห์หารูปแบบคำกริยา (กลุ่มที่ 4)

กฎไวยากรณ์ในกลุ่มนี้จะทำการเชื่อมระหว่างคำกริยาหลักกับคำนามหลักในแต่ละตำแหน่งของประโยค หรือ ทำการเชื่อมระหว่างคำกริยาหลักกับคำที่อยู่หลังคำว่า "to" ด้วยแสดงความสัมพันธ์ของหน่วยความหมาย (เกี่ยวกับไวยากรณ์การกที่กล่าวมาแล้วในบทที่ 4 ข้อ 4.3) หรือส่วนแสดงความสัมพันธ์ทางโครงสร้าง(Surface Meaning หรือ Part of speech ที่กล่าวมาแล้วในบทที่ 3 ข้อ 3.2.3.1) ซึ่งจะขึ้นอยู่กับรูปแบบคำกริยาของประโยค(verb pattern) ดังนี้

vp1 SUB+IV+COMP1

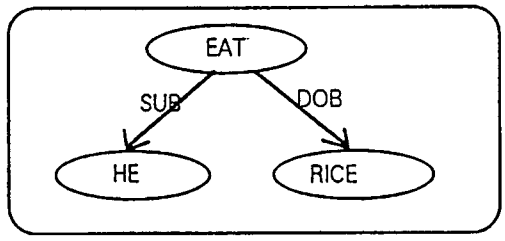
vp2 SUB+IV+TO+INF+COMP2

vp3 SUB+TV+DOB+COMP3

vp4 SUB+TV+DOB+TO+INF+COMP4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวอย่างเช่น ประโยค "He eats rice. " เมื่อผ่านกฎในข้อนี้จะทำให้ได้ผลลัพธ์ดังรูปข้างล่างนี้



จากรูปแบบคำกริยานำมาสร้างเป็นกฎไวยากรณ์เป็นกฎย่อย R13 ดังมีรายละเอียดต่อไปนี้

<กฎย่อย R13 กฎการหารูปแบบคำกริยา>

ถ้า * เป็น V แล้ว ดูว่า V ตัวนี้ใช้ Vp(Verb pattern) ไหนบ้าง และก็เช็คว่ามันต้องใช้ Vp อันไหน โดยสร้างกฎขึ้นมาโดยมี Verb Pattern ดังนี้

vp1 SUB+IV+COMP1

vp2 SUB+IV+TO+INF+COMP2

vp3 SUB+TV+DOB+COMP3

vp4 SUB+TV+DOB+TO+INF+COMP4

<u>กฎที่ได้คือ</u>

1. เริ่มที่ * เป็น V ทางซ้าย(-1) เป็น N จะ link ระหว่าง (*,-1) = SUB มองทางขวา(+1)ไม่มี N และมีการตัดความกำกวม (Disambiguous) ให้ที่ * มี Vp = V1 อย่างเดียว โดยที่ V เป็น head (V --> N)

2. เริ่มที่ * เป็น V_1 ทางซ้าย(-1) เป็น N จะ link ระหว่าง (*,-1) = SUB โดยที่ V_1 เป็น head (V_1 -> N) และ มองทางขวา(+1) เป็น "TO" และ ที่คำ (+2) เป็น V_2 แล้ว จะ link ระหว่าง (*,+2)= PAR และมีการตัดความกำกวม (Disambiguous) ให้ที่ * มี Vp = V2 อย่างเดียว โดยที่ V_1 เป็น head (V_1 --> V_2) และตัดคำ (+1) ทิ้ง

3. เริ่มที่ * เป็น V ทางซ้าย(-1) เป็น N จะ link ระหว่าง (*,-1) = SUB โดยที่ V เป็น head (V -> N) และ มองทางขวา(+1) เป็น N จะ link ระหว่าง (*,+1)= DOB และมีการตัดความกำกวม (Disambiguous) ให้ที่ * มี Vp = V3 อย่างเดียว โดยที่ V เป็น head (V --> N)

4. เริ่มที่ * เป็น V_1 ทางซ้าย(-1) เป็น N จะ link ระหว่าง (*,-1) = SUB โดยที่ V_1 เป็น head (V_1 --> N) และ มองทางขวา(+1) เป็น N จะ link ระหว่าง (*,+1)= DOBเป็น โดยที่ V_1 เป็น head (V_1 -> N) และคำ (+2) เป็น "TO" และ ที่คำ (+3) เป็น V_2 แล้ว จะ link ระหว่าง (*,+3)= PAR และมีการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัดความกำกวม (Disambiguous) ให้ที่ * มี Vp = V4 อย่างเดียว โดยที่ V_1 เป็น head (V_1 --> V_2) และ ตัดคำ (+2) ทิ้ง

ทำส่วนนี้สำเร็จจึงทำการ Map Case Semantic โดยดูจาก EMAP เช่น SUB=AGT เป็นต้น

<u>หมายเหตุ</u>

[ ] หมายถึง จะมีหรือไม่มีสิ่งที่อยู่ในเครื่องหมายก้ามปูนี้ก็ได้ เรียกว่า Optional เขียนเป็นรูปแบบกฎได้ดังนี้

```
R13 / Vp1 /
     CHKCAT(*,V,'','') & CHKCAT(-1,N,'','') = LINK(*,-1,SUB),
    R13.1 NOT(CHKCAT(+1,N)) & NOT(CHKENTRY(+1,'TO')) = VP(*,1)
    R13.2 CHKCAT(+1,N) & NOT(CHKENTRY(+2,'TO'))
= LINK(*,+1,DOB), VP(*,2), DELNODE(+2)
    R13.3 CHKENTRY(+1,'TO) & CHKCAT(+2,V) & CHKCAT(+3,N)
= LINK(*,1,PAR), LINK(+2,+3,DOB), VP(*,3), DELCAT(+2,V), DELCAT(+3,N),
DELNODE(+1), DELNODE(+2), DELNODE(3)
    R13.4 CHKENTRY(+1,'TO') & CHKCAT(+2,V)
= LINK(*,+1,PAR), VP(*,3),DELCAT(+2,V), DELNODE(+1), DELNODE(+2)
    R13.5 CHKCAT(+1,N) & CHKENTRY(+2,'TO') & CHKCAT(+3,V) &
CHKCAT(+4,N)
= LINK(*,+1,DOB), LINK(*,+3,PAR), LINK(+3,+4,DOB), VP(*,4), DELCAT(+1,N),
DELCAT(+3,V), DELCAT(+4,N),  DELNODE(+1), DELNODE(+2), DELNODE(+3),
DELNODE(+4)
     R13.6 CHKCAT(+1,N) & CHKENTRY(+2,'TO') & CHKCAT(+3,V)
= LINK(*,+1,DOB), LINK(*,+3,PAR), VP(*,4),  DELCAT(+1,N), DELCAT(+3,V),
DELNODE(+1), DELNODE(+2), DELNODE(+3)
```

เพิ่มกฎการเคลื่อนที่ของ monitor window(*)

R14 / Shift Window / RS ซึ่งกฎนี้จะต้องอยู่ในกลุ่มที่ 1 ถึง กลุ่มที่ 4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.3.2.5 สรุปกฎไวยากรณ์ทั้งหมด

กลุ่มที่ 1 Grammar Rules for Noun Phrase Construction

กลุ่มที่ 1.1 Premodifier

R1. / link ADV, ADJ /

CHKCAT(*,ADJ,'','') & CHKCAT(-1,ADV,'','')

= LINK(*,-1,MOD), DELCAT(*,ADJ), DELCAT(-1,ADV)

R2 / link N_1 , N_2 /

R2.1 CHKCAT(*,N_2 ,NCMN,'') & CHKCAT(-1,N_1 ,NPSS,'')

= LINK(*,-1,POSS), DELCAT(*,N), DELCAT(-1,N)

R2.2 CHKCAT(*,N_2 ,NCMN,'') & CHKCAT(-1,N_1 ,NPRP,'')

= LINK(*,-1,NAM), DELCAT(*,N), DELCAT(-1,N)

R2.3 CHKCAT(*,N_2 ,NCMN,'') & CHKCAT(-1,N_1 ,NALN,'')

= LINK(*,-1,ATT), DELCAT(*,N), DELCAT(-1,N)

R2.4 CHKCAT(*,N_2 ,NCMN,'') & CHKCAT(-1,N_1 ,NCNM,'')

= LINK(*,-1,QUAN), DELCAT(*,N), DELCAT(-1,N)

R2.5 CHKCAT(*,N_2 ,NCMN,'') & CHKCAT(-1,N_1 ,NCMN,112)

= LINK(*,-1,MOD), DELCAT(*,N), DELCAT(-1,N)

R2.6 CHKCAT(*,N_2 ,NCMN,'') & CHKCAT(-1,N_1 ,'','')

= LINK(*,-1,NMOD), DELCAT(*,N), DELCAT(-1,N)

R3 /link ADJ, N/

R3.1 CHKCAT(*,N,'','') & CHKCAT(-1,ADJ,'',2363)

= LINK(*,-1,ATT), DELCAT(*,N), DELCAT(-1,ADJ)

R3.2 CHKCAT(*,N,'','') & CHKCAT(-1,ADJ,'',2361)

= LINK(*,-1,ATT), DELCAT(*,N), DELCAT(-1,ADJ)

R3.3 CHKCAT(*,N,'','') & CHKCAT(-1,ADJ,'',232)

= LINK(*,-1,MOD), DELCAT(*,N), DELCAT(-1,ADJ)

R4 / link N_2, NPSS, N_1 /

CHKCAT(*,N,'','') & CHKCAT(-1,N,NPSS,'') & CHKCAT(-2,N,'','')

= LINK(*,-2,POSS), DELCAT(*,N), DELCAT(-2,N), DELNODE(-1)

R5 / link DET , N /

R5.1 CHKCAT(*,N ,'','') & CHKCAT(-1,DET ,DPRS,'')

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

= LINK(*,-1,POSS), DELCAT(*,N), DELCAT(-1,DET)

R5.2 CHKCAT(*,N ,'','') & CHKCAT(-1,DET,DCNM ,'')

= LINK(*,-1,QUAN), DELCAT(*,N), DELCAT(-1,DET)

R5.3 CHKCAT(*,N ,'','') & CHKCAT(-1,DET ,DDMN,'')

= LINK(*,-1,REF), DELCAT(*,N), DELCAT(-1,DET)

R5.4 CHKCAT(*,N ,'','') & CHKCAT(-1,DET ,DINA,'')

= LINK(*,-1,VAL), DELCAT(*,N), DELCAT(-1,DET)

R5.5 CHKCAT(*,N ,'','') & CHKCAT(-1,DET ,DTHE,'')

= ASS(*,DET.{-1}), DELCAT(*,N), DELNODE(-1)

R5.6 CHKCAT(*,N,'','') & CHKCAT(-1,DET ,DAAN,'')

= ASS(*,DET.{-1}), DELCAT(*,N), DELNODE(-1)

R6. / PRON /

CHKCAT(*.PRON,PPRS,'') & CHKNODE(-1=NIL) = DELCAT(*,PRON)

R14 / Shift Window / RS

กลุ่มที่ 1.2 Postmodifier

R7 / link N PREP N /

CHKCAT(*, N_1 ,'','') & CHKCAT(+1, PREP,'','') &

CHKCAT(+2,N_2 ,'','') & CHKENTRY(+1,OF)

= LINK(*,+2,POSS) , ASS(POSS,SUBFEA.{of}), DELCAT(*,N),

DELCAT(+1,PREP),DELCAT(+2,N)

R14/ Shift Window / RS

กลุ่มที่ 2 Grammar Rules for Auxiliary Phrase Construction

R8 / NEG,V /

FINDVerb(HEAD,P);

CHKCAT(*,NEG,'','') = ASS(P,NEG.{T}), DELNODE(*) ,DELCAT(P,V)

R9 / AUX,V /

FINDVerb(HEAD,P);

R9.1 CHKCAT(*,AUX,XMDL,'')

= ASS(P,XMDL.{ability}), ASS(P,TENSE.{PRES}), DELNODE(*) ,DELCAT(P,V)

R9.2 CHKCAT(*,AUX,XMDL,'')

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

= ASS(P,XMDL.{ability}), ASS(P,TENSE.{PRES}), DELNODE(*) ,DELCAT(P,V)

R9.3 CHKCAT(*,AUX,XXBE,'') & ENDING(*,ING)

= ASS(P,XMDL.{progress}), ASS(P,TENSE.{PRES}), DELNODE(*) ,DELCAT(P,V)

R9.4 CHKCAT(*,AUX,XXBE,'') & ENDING(*,ED)

= ASS(P,XMDL.{pass}), ASS(P,TENSE.{PAST}), DELNODE(*) ,DELCAT(P,V)

R9.5 CHKCAT(*,AUX,XXDO,'')

= ASS(P,TENSE.{PRES}) DELNODE(*) ,DELCAT(P,V)

R14 / Shift Window / RS

กลุ่มที่ 3 Grammar Rules for Optional Construction

R10 / link PREP,N /

FINDPrep(HEAD,P), CHKCAT(*,V), DELCAT(*,V)

R10.1 CHKENTRY(P,'OF')&CHKCAT(P+1,N,'','')

= ASS(POSS,SUBFEA.{of}), LINK(*,P+1,POSS), DELNODE(P) ,DELCAT(P+1,N)

R10.2 CHKENTRY(P,'FROM') & CHKCAT(P+1,N,'','')

= ASS(SOR,SUBFEA.{from}), LINK(*,P+1,SOR), DELNODE(P) ,DELCAT(P+1,N)

R10.3 CHKENTRY(P,'IN') & CHKCAT(P+1,N,'','')

= ASS(SOR,SUBFEA.{in}), LINK(*,P+1,SOR), DELNODE(P) ,DELCAT(P+1,N)

R10.4 CHKENTRY(P,'WITH') & CHKCAT(P+1,N,'',11)

= ASS(OBJ,SUBFEA.{with}), LINK(*,P+1,OBJ), DELNODE(P) ,DELCAT(P+1,N)

R10.5 CHKENTRY(P,'BY') & CHKCAT(P+1,N,'',11)

= ASS(OBJ,SUBFEA.{by}), LINK(*,P+1,OBJ),DELNODE(P) ,DELCAT(P+1,N)

R10.6 CHKENTRY(P,'UPON') & CHKCAT(P+1,N,'',11)

= ASS(OBJ,SUBFEA.{upon}), LINK(*,P+1,OBJ), DELNODE(P) ,DELCAT(P+1,N)

R10.7 CHKENTRY(P,'WITH') & CHKCAT(P+1,N,'',)

= ASS(INS,SUBFEA.{with}), LINK(*,P+1,INS), DELNODE(P) ,DELCAT(P+1,N)

R10.8 CHKENTRY(P,'BY') & CHKCAT(P+1,N,'','')

= ASS(INS,SUBFEA.{by}), LINK(*,P+1,INS), DELNODE(P) ,DELCAT(P+1,N)

R10.9 CHKENTRY(P,'TO') & CHKCAT(P+1,N,'','') & CHKCAT(P+1,V,'',''=FALSE)

= ASS(TAR,SUBFEA.{to}, LINK(*,P+1,TAR), DELNODE(P), DELCAT(P+1,N)

R10.10 CHKENTRY(P,'FOR') & CHKCAT(P+1,N,'','')

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

= ASS(TAR,SUBFEA.{for}), LINK(*,P+1,TAR), DELNODE(P) ,DELCAT(P+1,N)

R10.11 CHKENTRY(P,'AT') & CHKCAT(P+1,N,'','232')

= ASS(TIM,SUBFEA.{at}), LINK(*,P+1,TIM), DELNODE(P) ,DELCAT(P+1,N)

R10.12 CHKENTRY(P,'AT') & CHKCAT(P+1,N,'',)

= ASS(TAR,SUBFEA.{at}), LINK(*,P+1,TAR), DELNODE(P) ,DELCAT(P+1,N)

R11 / link ADV,V /

CHKCAT(*,V,'','') & CHKCAT(-1,ADV,'','')

= LINK(*,-1,MOD), DELCAT(*,V), DELCAT(-1,ADV)

R12 / link ADV,V /

CHKCAT(*,V,'','') & CHKCAT(+1,ADV,'','')

= LINK(*,+1,MOD), DELCAT(*,V), DELCAT(+1,ADV)

R14 / Shift Window / RS

กลุ่มที่ 4 Grammar Rules for Verb Pattern Construction

R13 / Vp1 /

CHKCAT(*,V,'','') & CHKCAT(-1,N,'','') = LINK(*,-1,SUB),

R13.1 NOT(CHKCAT(+1,N)) & NOT(CHKENTRY(+1,'TO'))

= VP(*,1)

R13.2 CHKCAT(+1,N) & NOT(CHKENTRY(+2,'TO'))

= LINK(*,+1,DOB), VP(*,2), DELNODE(+2)

R13.3 CHKENTRY(+1,'TO') & CHKCAT(+2,V) & CHKCAT(+3,N)

= LINK(*,1,PAR), LINK(+2,+3,DOB), VP(*,3),

DELCAT(+2,V), DELCAT(+3,N),

DELNODE(+1), DELNODE(+2), DELNODE(3)

R13.4 CHKENTRY(+1,'TO') & CHKCAT(+2,V)

= LINK(*,+1,PAR), VP(*,3),DELCAT(+2,V),

DELNODE(+1), DELNODE(+2)

R13.5 CHKCAT(+1,N) & CHKENTRY(+2,'TO') & CHKCAT(+3,V) & CHKCAT(+4,N)

= LINK(*,+1,DOB), LINK(*,+3,PAR), LINK(+3,+4,DOB), VP(*,4),

DELCAT(+1,N), DELCAT(+3,V), DELCAT(+4,N),

DELNODE(+1), DELNODE(+2), DELNODE(+3), DELNODE(+4)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

R13.6 CHKCAT(+1,N) & CHKENTRY(+2,'TO') & CHKCAT(+3,V)

= LINK(*,+1,DOB), LINK(*,+3,PAR), VP(*,4),

DELCAT(+1,N), DELCAT(+3,V),

DELNODE(+1), DELNODE(+2), DELNODE(+3)

R14 / Shift Window / RS

5.3.2.6 ตัวอย่างของการใช้กฎการวิเคราะห์

ตัวอย่างประโยค : My brother wants to play football in the filed.

การวิเคราะห์นี้มีทั้งหมด 4 กลุ่ม โดยในแต่ละขั้นตอนจะมีจำนวนกฎไวยากรณ์ที่แตกต่างกันไป แต่จะผ่านกฎไวยากรณ์ของขั้นกลุ่มที่ 1 ก่อนเรียงตามลำดับไปจนถึง กฎไวยากรณ์ของกลุ่มที่ 4 โดยในแต่ละกลุ่ม เริ่มแรก monitor window(*) จะอยู่ที่คำแรก เมื่อผ่านกฎไวยากรณ์ข้อใดได้สำเร็จ (Success) จะทำการย้อนกลับ (Backtrack) กลับไปกฎไวยากรณ์เริ่มต้นของแต่ละกลุ่มทุกครั้ง หากผลลัพธ์ที่ได้ไม่ประสบผลสำเร็จ (Fail) จะผ่านไปยังกฎไวยากรณ์ถัดไป และจะหยุดการประมวลผลในแต่ละกลุ่มเมื่อ monitor window(*) เลื่อนไปถึงคำที่มีค่าเท่ากับ NIL หมายถึง วิเคราะห์หมดทุกคำแล้ว

1	2	3	4	5	6	7	8	9	10
My	brother	wants	to	play	football	in	the	filed.	NIL
PRON	N	N,V	PREP	V,N	N	PREP	DET	N	
PPRS							DTHE		
*	+1								

โดยมีขั้นตอนการประมวลผลดังนี้

1) ผ่านกฎไวยากรณ์กลุ่มที่ 1 (Grammar Rules for Noun Phrase Construction) ที่มีกฎตั้งแต่กฎไวยากรณ์ที่ 1 ถึงกฎไวยากรณ์ที่ 7 และกฎไวยากรณ์ที่ 14 โดยจะเริ่มผ่านกฎไวยากรณ์ของกลุ่มที่ 1.1 PreModifier ก่อน โดยเริ่มในกฎไวยากรณ์ที่ 1 ถึงกฎไวยากรณ์ที่ 6 และ กฎไวยากรณ์ที่ 14 ตามลำดับดังนี้

R1. / link ADV, ADJ /

CHKCAT(*,ADJ,'','') & CHKCAT(-1,ADV,'','')

= LINK(*,-1,MOD), DELCAT(*,ADJ), DELCAT(-1,ADV)

(FAIL)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

R2 / link N_1 , N_2 /

R2.1 CHKCAT(*,N_2 ,NCMN,'') & CHKCAT(-1,N_1 ,NPSS,'')

= LINK(*,-1,POSS), DELCAT(*,N), DELCAT(-1,N)

R2.2 CHKCAT(*,N_2 ,NCMN,'') & CHKCAT(-1,N_1 ,NPRP,'')

= LINK(*,-1,NAM), DELCAT(*,N), DELCAT(-1,N)

R2.3 CHKCAT(*,N_2 ,NCMN,'') & CHKCAT(-1,N_1 ,NALN,'')

= LINK(*,-1,ATT), DELCAT(*,N), DELCAT(-1,N)

R2.4 CHKCAT(*,N_2 ,NCMN,'') & CHKCAT(-1,N_1 ,NCNM,'')

= LINK(*,-1,QUAN), DELCAT(*,N), DELCAT(-1,N)

R2.5 CHKCAT(*,N_2 ,NCMN,'') & CHKCAT(-1,N_1 ,NCMN,112)

= LINK(*,-1,MOD), DELCAT(*,N), DELCAT(-1,N)

R2.6 CHKCAT(*,N_2 ,NCMN,'') & CHKCAT(-1,N_1 ,'','')

= LINK(*,-1,NMOD), DELCAT(*,N), DELCAT(-1,N)

R2.1- R2.6 (FAIL)

R3 /link ADJ, N/

R3.1 CHKCAT(*,N,'','') & CHKCAT(-1,ADJ,'',2363)

= LINK(*,-1,ATT), DELCAT(*,N), DELCAT(-1,ADJ)

R3.2 CHKCAT(*,N,'','') & CHKCAT(-1,ADJ,'',2361)

= LINK(*,-1,ATT), DELCAT(*,N), DELCAT(-1,ADJ)

R3.3 CHKCAT(*,N,'','') & CHKCAT(-1,ADJ,'',232)

= LINK(*,-1,MOD), DELCAT(*,N), DELCAT(-1,ADJ)

R3.1 - R3.3 (FAIL)

R4 / link N_2, NPSS, N_1 /

CHKCAT(*,N,'','') & CHKCAT(-1,N,NPSS,'') & CHKCAT(-2,N,'','')

= LINK(*,-2,POSS), DELCAT(*,N), DELCAT(-2,N), DELNODE(-1)

(FAIL)

R5 / link DET , N /

R5.1 CHKCAT(*,N ,'','') & CHKCAT(-1,DET ,DPRS,'')

= LINK(*,-1,POSS), DELCAT(*,N), DELCAT(-1,DET)

R5.2 CHKCAT(*,N ,'','') & CHKCAT(-1,DET,DCNM ,'')

= LINK(*,-1,QUAN), DELCAT(*,N), DELCAT(-1,DET)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

R5.3 CHKCAT(*,N ,'','') &.CHKCAT(-1,DET ,DDMN,'')

= LINK(*,-1,REF), DELCAT(*,N), DELCAT(-1,DET)

R5.4 CHKCAT(*,N ,'','') & CHKCAT(-1,DET ,DINA,'')

= LINK(*,-1,VAL), DELCAT(*,N), DELCAT(-1,DET)

R5.5 CHKCAT(*,N ,'','') & CHKCAT(-1,DET ,DTHE,'')

= ASS(*,DET.(-1)), DELCAT(*,N), DELNODE(-1)

R5.6 CHKCAT(*,N ,'','') & CHKCAT(-1,DET ,DAAN,'')

= ASS(*,DET.(-1)), DELGAT(*,N), DELNODE(-1)

R5.1 - R5.6 (FAIL)

R6. / PRON /

CHKCAT(*.PRON,PPRS,'') & CHKNODE(-1=NIL) = DELCAT(*,PRON)

(FAIL)

R14 / Shift Window / RS

(SUCCESS) จะทำการเลื่อน monitor window (*) ไปทางซ้ายแล้วทำการย้อนกลับไปกฎไวยากรณ์ที่ 1 ใหม่อีกได้ผลลัพธ์ดังข้างล่างนี้

1	2	3	4	5	6	7	8	9	10
My	brother	wants	to	play	football	in	the	filed.	NIL
DET	N	N,V	PREP	V,N	N	PREP	DET	N	
DPRS							DTHE		
-1	*	+1							

R1. / link ADV, ADJ /

CHKCAT(*,ADJ,'','') & CHKCAT(-1,ADV,'','')

= LINK(*,-1,MOD), DELCAT(*,ADJ), DELCAT(-1,ADV)

(FAIL)

R2 / link N_1 , N_2 /

R2.1 CHKCAT(*,N_2 ,NCMN,'') & CHKCAT(-1,N_1 ,NPSS,'')

= LINK(*,-1,POSS), DELCAT(*,N), DELCAT(-1,N)

R2.2 CHKCAT(*,N_2 ,NCMN,'') & CHKCAT(-1,N_1 ,NPRP,'')

= LINK(*,-1,NAM), DELCAT(*,N), DELCAT(-1,N)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

R2.3 CHKCAT(*,N_2 ,NCMN,'') & CHKCAT(-1,N_1 ,NALN,'')

= LINK(*,-1,ATT), DELCAT(*,N), DELCAT(-1,N)

R2.4 CHKCAT(*,N_2 ,NCMN,'') & CHKCAT(-1,N_1 ,NCNM,'')

= LINK(*,-1,QUAN), DELCAT(*,N), DELCAT(-1,N)

R2.5 CHKCAT(*,N_2 ,NCMN,'') & CHKCAT(-1,N_1 ,NCMN,112)

= LINK(*,-1,MOD), DELCAT(*,N), DELCAT(-1,N)

R2.6 CHKCAT(*,N_2 ,NCMN,'') & CHKCAT(-1,N_1 ,'','')

= LINK(*,-1,NMOD), DELCAT(*,N), DELCAT(-1,N)

R2.1- R2.6 (FAIL)

R3 /link ADJ, N/

R3.1 CHKCAT(*,N,'','') & CHKCAT(-1,ADJ,'',2363)

= LINK(*,-1,ATT), DELCAT(*,N), DELCAT(-1,ADJ)

R3.2 CHKCAT(*,N,'','') & CHKCAT(-1,ADJ,'',2361)

= LINK(*,-1,ATT), DELCAT(*,N), DELCAT(-1,ADJ)

R3.3 CHKCAT(*,N,'','') & CHKCAT(-1,ADJ,'',232)

= LINK(*,-1,MOD), DELCAT(*,N), DELCAT(-1,ADJ)

R3.1 - R3.3 (FAIL)

R4 / link N_2, NPSS, N_1 /

CHKCAT(*,N,'','') & CHKCAT(-1,N,NPSS,'') & CHKCAT(-2,N,'','')

= LINK(*,-2,POSS), DELCAT(*,N), DELCAT(-2,N), DELNODE(-1)

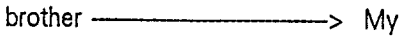
(FAIL)

R5 / link DET , N /

R5.1 CHKCAT(*,N ,'','') & CHKCAT(-1,DET ,DPRS,'')

= LINK(*,-1,POSS), DELCAT(*,N), DELCAT(-1,DET)

(SUCCESS) ทำการเชื่อมระหว่างคำว่า "brother" กับ "My" ด้วยการก "POSS" แล้วจะทำการตัดคำ(-1)ทิ้ง (ข้อมูลของคำ "My" นี้จะถูกยุบรวมเก็บไว้ในคำ "brother") แล้วทำการย้อน กลับไปที่กฎไวยากรณ์ที่ 1 อีก โดยผลลัพธ์ที่ได้คือ



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2	3	4	5	6	7	8	9	10
brother	wants	to	play	football	in	the	filed.	NIL
N	N,V	PREP	V,N	N	PREP	DET	N	
						DTHE		
*	+1							

R1. / link ADV, ADJ /

CHKCAT(*,ADJ,'','') & CHKCAT(-1,ADV,'','')

= LINK(*,-1,MOD), DELCAT(*,ADJ), DELCAT(-1,ADV)

(FAIL)

R2 / link N_1 , N_2 /

R2.1 CHKCAT(*,N_2 ,NCMN,'') & CHKCAT(-1,N_1 ,NPSS,'')

= LINK(*,-1,POSS), DELCAT(*,N), DELCAT(-1,N)

R2.2 CHKCAT(*,N_2 ,NCMN,'') & CHKCAT(-1,N_1 ,NPRP,'')

= LINK(*,-1,NAM), DELCAT(*,N), DELCAT(-1,N)

R2.3 CHKCAT(*,N_2 ,NCMN,'') & CHKCAT(-1,N_1 ,NALN,'')

= LINK(*,-1,ATT), DELCAT(*,N), DELCAT(-1,N)

R2.4 CHKCAT(*,N_2 ,NCMN,'') & CHKCAT(-1,N_1 ,NCNM,'')

= LINK(*,-1,QUAN), DELCAT(*,N), DELCAT(-1,N)

R2.5 CHKCAT(*,N_2 ,NCMN,'') & CHKCAT(-1,N_1 ,NCMN,112)

= LINK(*,-1,MOD), DELCAT(*,N), DELCAT(-1,N)

R2.6 CHKCAT(*,N_2 ,NCMN,'') & CHKCAT(-1,N_1 ,'','')

= LINK(*,-1,NMOD), DELCAT(*,N), DELCAT(-1,N)

R2.1- R2.6 (FAIL)

R3 /link ADJ, N/

R3.1 CHKCAT(*,N,'','') & CHKCAT(-1,ADJ,'',2363)

= LINK(*,-1,ATT), DELCAT(*,N), DELCAT(-1,ADJ)

R3.2 CHKCAT(*,N,'','') & CHKCAT(-1,ADJ,'',2361)

= LINK(*,-1,ATT), DELCAT(*,N), DELCAT(-1,ADJ)

R3.3 CHKCAT(*,N,'','') & CHKCAT(-1,ADJ,'',232)

= LINK(*,-1,MOD), DELCAT(*,N), DELCAT(-1,ADJ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

R3.1 - R3.3 (FAIL)

R4 / link N_2, NPSS, N_1 /

CHKCAT(*,N,'','') & CHKCAT(-1,N,NPSS,'') & CHKCAT(-2,N,'','')

= LINK(*,-2,POSS), DELCAT(*,N), DELCAT(-2,N), DELNODE(-1)

(FAIL)

R5 / link DET , N /

R5.1 CHKCAT(*,N ,'','') & CHKCAT(-1,DET ,DPRS,'')

= LINK(*,-1,POSS), DELCAT(*,N), DELCAT(-1,DET)

R5.2 CHKCAT(*,N ,'','') & CHKCAT(-1,DET,DCNM ,'')

= LINK(*,-1,QUAN), DELCAT(*,N), DELCAT(-1,DET)

R5.3 CHKCAT(*,N ,'','') & CHKCAT(-1,DET ,DDMN,'')

= LINK(*,-1,REF), DELCAT(*,N), DELCAT(-1,DET)

R5.4 CHKCAT(*,N ,'','') & CHKCAT(-1,DET ,DINA,'')

= LINK(*,-1,VAL), DELCAT(*,N), DELCAT(-1,DET)

R5.5 CHKCAT(*,N ,'','') & CHKCAT(-1,DET ,DTHE,'')

= ASS(*,DET,{-1}), DELCAT(*,N), DELNODE(-1)

R5.6 CHKCAT(*,N ,'','') & CHKCAT(-1,DET ,DAAN,'')

= ASS(*,DET,{-1}), DELCAT(*,N),DELNODE(-1)

R5.1 - R5.6 (FAI

R6. / PRON /

CHKCAT(*,PRON,PPRS,'') & CHKNODE(-1=NIL) = DELCAT(*,PRON)

(FAIL)

R14 / Shift Window / RS

(SUCCESS) ทำการเลื่อน monitor window (*) ไปทางซ้าย แล้วทำการย้อนกลับไปกฎไวยากรณ์ที่ 1 อีก

2	3	4	5	6	7	8	9	10
brother	wants	to	play	football	in	the	filed.	NIL
N	N,V	PREP	V,N	N	PREP	DET	N	
							DTHE	
-1	*	+1						

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ก็จะผ่านกฎไวยากรณ์ตั้งแต่ข้อ R1 - R6 จะ FAIL หมด และผ่านกฎไวยากรณ์ R14 (SUCCESS) ก็จะทำการเลื่อน monitor window (*) ไปทางซ้ายจนกระทั่งได้ผลลัพธ์ดังนี้

2	3	4	5	6	7	8	9	10
brother	wants	to	play	football	in	the	filed.	NIL
N	N,V	PREP	V,N	N	PREP	DET	N	
						DTHE		
							-1	*

จะเริ่มผ่านกฎไวยากรณ์ที่ 1 ใหม่อีกดังนี้

R1. / link ADV, ADJ /

CHKCAT(*,ADJ,'','') & CHKCAT(-1,ADV,'','')

= LINK(*,-1,MOD), DELCAT(*,ADJ), DELCAT(-1,ADV)

(FAIL)

R2 / link N_1 , N_2 /

R2.1 CHKCAT(*,N_2 ,NCMN,'') & CHKCAT(-1,N_1 ,NPSS,'')

= LINK(*,-1,POSS), DELCAT(*,N), DELCAT(-1,N)

R2.2 CHKCAT(*,N_2 ,NCMN,'') & CHKCAT(-1,N_1 ,NPRP,'')

= LINK(*,-1,NAM), DELCAT(*,N), DELCAT(-1,N)

R2.3 CHKCAT(*,N_2 ,NCMN,'') & CHKCAT(-1,N_1 ,NALN,'')

= LINK(*,-1,ATT), DELCAT(*,N), DELCAT(-1,N)

R2.4 CHKCAT(*,N_2 ,NCMN,'') & CHKCAT(-1,N_1 ,NCNM,'')

= LINK(*,-1,QUAN), DELCAT(*,N), DELCAT(-1,N)

R2.5 CHKCAT(*,N_2 ,NCMN,'') & CHKCAT(-1,N_1 ,NCMN,112)

= LINK(*,-1,MOD), DELCAT(*,N), DELCAT(-1,N)

R2.6 CHKCAT(*,N_2 ,NCMN,'') & CHKCAT(-1,N_1 ,'','')

= LINK(*,-1,NMOD), DELCAT(*,N), DELCAT(-1,N)

R2.1- R2.6 (FAIL)

R3 /link ADJ, N/

R3.1 CHKCAT(*,N,'','') & CHKCAT(-1,ADJ,'',2363)

= LINK(*,-1,ATT), DELCAT(*,N), DELCAT(-1,ADJ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

R3.2 CHKCAT(*,N,'','') & CHKCAT(-1,ADJ,'',2361)

= LINK(*,-1,ATT), DELCAT(*,N), DELCAT(-1,ADJ)

R3.3 CHKCAT(*,N,'','') & CHKCAT(-1,ADJ,'',232)

= LINK(*,-1,MOD), DELCAT(*,N), DELCAT(-1,ADJ)

R3.1 - R3.3 (FAIL)

R4 / link N_2, NPSS, N_1 /

CHKCAT(*,N,'','') & CHKCAT(-1,N,NPSS,'') & CHKCAT(-2,N,'','')

= LINK(*,-2,POSS), DELCAT(*,N), DELCAT(-2,N), DELNODE(-1)

(FAIL)

R5 / link DET , N /

R5.1 CHKCAT(*,N ,'','') & CHKCAT(-1,DET ,DPRS,'')

= LINK(*,-1,POSS), DELCAT(*,N), DELCAT(-1,DET)

R5.2 CHKCAT(*,N ,'','') & CHKCAT(-1,DET,DCNM ,'')

= LINK(*,-1,QUAN), DELCAT(*,N), DELCAT(-1,DET)

R5.3 CHKCAT(*,N ,'','') & CHKCAT(-1,DET ,DDMN,'')

= LINK(*,-1,REF), DELCAT(*,N), DELCAT(-1,DET)

R5.4 CHKCAT(*,N ,'','') & CHKCAT(-1,DET ,DINA,'')

= LINK(*,-1,VAL), DELCAT(*,N), DELCAT(-1,DET)

R5.5 CHKCAT(*,N ,'','') & CHKCAT(-1,DET ,DTHE,'')

= ASS(*,DET.(-1)), DELCAT(*,N), DELNODE(-1)

(SUCCESS) จะทำให้คำที่ 9 มีฟีลด์ DET.(THE) เกิดขึ้น และตัดคำ(-1) ทิ้งแล้วทำการย้อนกลับไปกฎไวยากรณ์ที่ 1 อีก ผลลัพธ์ที่ได้คือ

-	2	3	4	5	6	8	9	10
	brother	wants	to	play	football	in	filed.	NIL
	N	N,V	PREP	V,N	N	PREP	N	
							DET.(THE)	
						-1	*	

เมื่อผ่านกฎที่ 1 ถึง กฎที่ 6 แล้วจะ FAIL แล้วผ่านกฎที่ 14

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

R14/ Shift Window / RS

(SUCCESS) จะทำการเลื่อน monitor window(*) ไปทางซ้ายอีกเจอคำว่า "NIL" แสดงว่าจบขั้นตอนประมวลผลในส่วนนี้ผลลัพธ์ที่ได้คือ

2	3	4	5	6	8	9	10
brother	wants	to	play	football	in	filed.	NIL
N	N,V	PREP	V,N	N	PREP	N	
						DET.{THE}	
						-1	*

เมื่อจบในส่วน Premodifier จะเข้าสู่กฎไวยากรณ์ที่ 7 และกฎไวยากรณ์ที่ 14 ของส่วน Postmodifier ซึ่ง monitor window(*) จะกลับไปอยู่ที่ต้นประโยคดังรูปข้างล่างนี้

2	3	4	5	6	8	9	10
brother	wants	to	play	football	in	filed.	NIL
N	N,V	PREP	V,N	N	PREP	N	
						DET.{THE}	
*	+1						

R7 / link N PREP N /

CHKCAT(*, N_1 ,'','') & CHKCAT(+1, PREP,'','') &

CHKCAT(+2,N_2 ,'','') & CHKENTRY(+1,OF)

= LINK(*,+2,POSS) , ASS(POSS,SUBFEA.{of}), DELCAT(*,N),

DELCAT(+1,PREP),DELCAT(+2,N)

(FAIL)

R14/ Shift Window / RS

(SUCCESS) จะทำการเลื่อน monitor window(*) ไปทางซ้ายแล้วย้อนกลับไปกฎไวยากรณ์ที่ 7 และกฎไวยากรณ์ที่ 14 อีกก็จะ FAIL จนกระทั่งเลื่อน monitor window(*) จนเจอคำว่า "NIL" แสดงว่าจบขั้นตอนประมวลผลในส่วนนี้ ดังรูปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2	3	4	5	6	8	9	10
brother	wants	to	play	football	in	filed.	NIL
N	N,V	PREP	V,N	N	PREP	N	
						DET.(THE)	
						-1	*

2) ผ่านกฎไวยากรณ์กลุ่มที่ 2 (Grammar Rules for Auxiliary Phrase Construction) ที่มีกฎไวยากรณ์ที่ 8 กฎไวยากรณ์ที่ 9 และกฎไวยากรณ์ที่ 14 ซึ่ง monitor window(*) จะอยู่ที่ต้นประโยคดังรูปข้างล่างนี้ โดยผ่านกฎไวยากรณ์ที่ 8 กฎไวยากรณ์ที่ 9 และกฎไวยากรณ์ที่ 14 ตามลำดับ

2	3	4	5	6	8	9	10
brother	wants	to	play	football	in	filed.	NIL
N	N,V	PREP	V,N	N	PREP	N	
						DET.(THE)	
*	+1						

R8 / NEG,V /

FINDVerb(HEAD,P);

CHKCAT(*,NEG,'',') = ASS(P,NEG.(T)), DELNODE(*) ,DELCAT(P,V)

(FAIL)

R9 / AUX,V /

FINDVerb(HEAD,P);

R9.1 CHKCAT(*,AUX,XMDL,'')

= ASS(P,XMDL.{ability}), ASS(P,TENSE.{PRES}), DELNODE(*) ,DELCAT(P,V)

R9.2 CHKCAT(*,AUX,XMDL,'')

= ASS(P,XMDL.{ability}), ASS(P,TENSE.{PRES}), DELNODE(*) ,DELCAT(P,V)

R9.3 CHKCAT(*,AUX,XXBE,'') & ENDING(*,ING)

= ASS(P,XMDL.{progress}), ASS(P,TENSE.{PRES}), DELNODE(*) ,DELCAT(P,V)

R9.4 CHKCAT(*,AUX,XXBE,'') & ENDING(*,ED)

= ASS(P,XMDL.{pass}), ASS(P,TENSE.{PAST}), DELNODE(*) ,DELCAT(P,V)

R9.5 CHKCAT(*,AUX,XXDO,'')

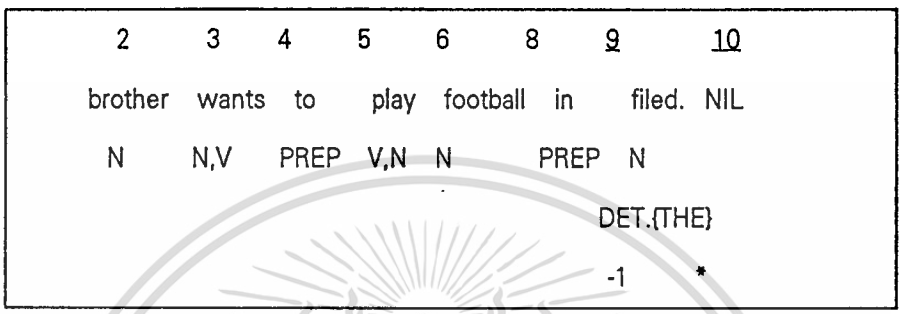
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

= ASS(P,TENSE.{PRES}) DELNODE(*) ,DELCAT(P,V)

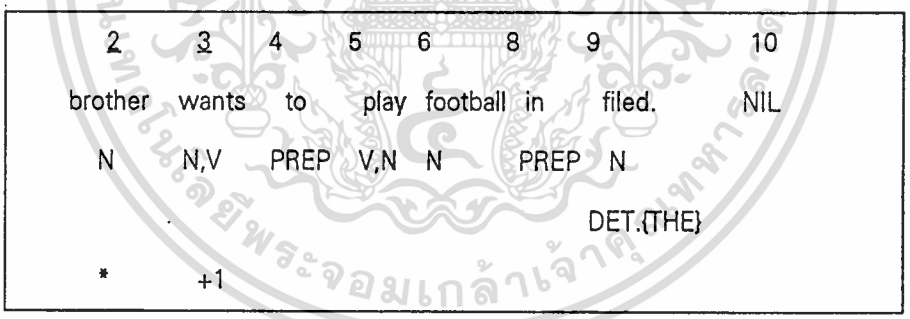
R9.1 - R9.5 (FAIL)

R14/ Shift Window / RS

(SUCCESS) จะทำการเลื่อน monitor window(*) ไปทางซ้ายแล้วย้อนกลับไปกฎไวยากรณ์ที่ 8 กฎไวยากรณ์ที่ 9 และกฎไวยากรณ์ที่ 14 อีกก็จะ FAIL จนกระทั่งเลื่อน monitor window(*) จนเจอคำว่า "NIL" แสดงว่าจบขั้นตอนประมวลผลในส่วนนี้ดังรูปข้างล่างนี้



3) ผ่านกฎไวยากรณ์กลุ่มที่ 3 (Grammar Rules for Optional Construction) ที่มีกฎไวยากรณ์ที่ 10, 11 และกฎไวยากรณ์ที่ 12 และกฎไวยากรณ์ที่ 14 ซึ่ง monitor window(*) จะอยู่ที่ต้นประโยคดังรูปข้างล่างนี้ โดยผ่านกฎไวยากรณ์ที่ 10, 11, 12 และกฎที่ 14 ตามลำดับ



R10 / link Prep,N /

FINDPrep(HEAD,P), CHKCAT(*,V), DELCAT(*,V)

R10.1 CHKENTRY(P,'OF') & CHKCAT(P+1,N,'','')

= ASS(POSS,SUBFEA.{of}), LINK(*,P+1,POSS), DELNODE(P) ,DELCAT(P+1,N)

R10.2 CHKENTRY(P,'FROM') & CHKCAT(P+1,N,'','')

= ASS(SOR,SUBFEA.{from}), LINK(*,P+1,SOR), DELNODE(P) ,DELCAT(P+1,N)

R10.3 CHKENTRY(P,'IN') & CHKCAT(P+1,N,'','')

= ASS(SOR,SUBFEA.{in}), LINK(*,P+1,SOR), DELNODE(P) ,DELCAT(P+1,N)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

R10.4 CHKENTRY(P,'WITH') & CHKCAT(P+1,N,'',11)

= ASS(OBJ,SUBFEA.{with}), LINK(*,P+1,OBJ), DELNODE(P) ,DELCAT(P+1,N)

R10.5 CHKENTRY(P,'BY') & CHKCAT(P+1,N,'',11)

= ASS(OBJ,SUBFEA.{by}), LINK(*,P+1,OBJ),DELNODE(P) ,DELCAT(P+1,N)

R10.6 CHKENTRY(P,'UPON') & CHKCAT(P+1,N,'',11)

= ASS(OBJ,SUBFEA.{upon}), LINK(*,P+1,OBJ), DELNODE(P) ,DELCAT(P+1,N)

R10.7 CHKENTRY(P,'WITH') & CHKCAT(P+1,N,'',)

= ASS(INS,SUBFEA.{with}), LINK(*,P+1,INS), DELNODE(P) ,DELCAT(P+1,N)

R10.8 CHKENTRY(P,'BY') & CHKCAT(P+1,N,'','')

= ASS(INS,SUBFEA.{by}), LINK(*,P+1,INS), DELNODE(P) ,DELCAT(P+1,N)

R10.9 CHKENTRY(P,'TO') & CHKCAT(P+1,N,'','') & CHKCAT(P+1,V,'',''=FALSE)

= ASS(TAR,SUBFEA.{to}), LINK(*,P+1,TAR), DELNODE(P) ,DELCAT(P+1,N)

R10.10 CHKENTRY(P,'FOR') & CHKCAT(P+1,N,'','')

= ASS(TAR,SUBFEA.{for}), LINK(*,P+1,TAR), DELNODE(P) ,DELCAT(P+1,N)

R10.11 CHKENTRY(P,'AT') & CHKCAT(P+1,N,'','232')

= ASS(TIM,SUBFEA.{at}), LINK(*,P+1,TIM), DELNODE(P) ,DELCAT(P+1,N)

R10.12 CHKENTRY(P,'AT') & CHKCAT(P+1,N,'',)

= ASS(TAR,SUBFEA.{at}), LINK(*,P+1,TAR), DELNODE(P) ,DELCAT(P+1,N)

R10.1 - R10.12 (FAIL)

R11 / link ADV,V /

CHKCAT(*,V,'','') & CHKCAT(-1,ADV,'','') = LINK(*,-1,MOD), DELCAT(*,V),

DELCAT(-1,ADV)

R12 / link V,ADV /

CHKCAT(*,V,'','') & CHKCAT(+1,ADV,'','') = LINK(*,+1,MOD), DELCAT(*,V),

DELCAT(+1,ADV)

(FAIL)

R14/ Shift Window / RS

(SUCCESS) จะทำการเลื่อน monitor window(*) ไปทางซ้ายแล้วย้อนกลับไปกฎไวยากรณ์ที่ 10, 11, 12 และกฎที่ 14 อีกก็จะ FAIL จนกระทั่งเลื่อน monitor window(*) ดังรูปข้างล่างนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2	3	4	5	6	8	9	10
brother	wants	to	play	football	in	filed.	NIL
N	N,V	PREP	V,N	N	PREP	N	
						DET.{THE}	
		-1	*		P	P+1	

ก็จะย้อนกลับไปกฎไวยากรณ์ที่ 10 ใหม่อีกครั้งหนึ่ง

R10 / link PREP,N /

FINDPrep(HEAD,P), CHKCAT(*,V), DELCAT(*,V)

R10.1 CHKENTRY(P,'OF') & CHKCAT(P+1,N,'','')

= ASS(POSS,SUBFEA.{of}), LINK(*,P+1,POSS), DELNODE(P) ,DELCAT(P+1,N)

R10.2 CHKENTRY(P,'FROM') & CHKCAT(P+1,N,'','')

= ASS(SOR,SUBFEA.{from}), LINK(*,P+1,SOR), DELNODE(P) ,DELCAT(P+1,N)

R10.3 CHKENTRY(P,'IN') & CHKCAT(P+1,N,'','')

= ASS(SOR,SUBFEA.{in}), LINK(*,P+1,SOR), DELNODE(P) ,DELCAT(P+1,N)

(SUCCESS) จะได้ผลลัพธ์ดังรูปข้างล่างนี้

SOR (SUBFEA.{in})

play ----------> field

2	3	4	5	6	10
brother	wants	to	play	football	NIL
N	V	PREP	V,N	N	
-1	*				

แล้วจะย้อนกลับไปกฎไวยากรณ์ที่10, 11, 12 และ ที่14 อีกครั้ง

R10 / link PREP,N /

FINDPrep(HEAD,P), CHKCAT(*,V), DELCAT(*,V)

R10.1 CHKENTRY(P,'OF') & CHKCAT(P+1,N,'','')

= ASS(POSS,SUBFEA.{of}) ,LINK(*,P+1,POSS), DELNODE(P), DELCAT(P+1,N)

R10.2 CHKENTRY(P,'FROM') & CHKCAT(P+1,N,'','')

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

= ASS(SOR,SUBFEA.{from}), LINK(*,P+1,SOR), DELNODE(P) ,DELCAT(P+1,N)

R10.3 CHKENTRY(P,'IN') & CHKCAT(P+1,N,'','')

= ASS(SOR,SUBFEA.{in}), LINK(*,P+1,SOR), DELNODE(P) ,DELCAT(P+1,N)

R10.4 CHKENTRY(P,'WITH') & CHKCAT(P+1,N,'',11)

= ASS(OBJ,SUBFEA.{with}), LINK(*,P+1,OBJ), DELNODE(P) ,DELCAT(P+1,N)

R10.5 CHKENTRY(P,'BY') & CHKCAT(P+1,N,'',11)

= ASS(OBJ,SUBFEA.{by}), LINK(*,P+1,OBJ),DELNODE(P) ,DELCAT(P+1,N)

R10.6 CHKENTRY(P,'UPON') & CHKCAT(P+1,N,'',11)

= ASS(OBJ,SUBFEA.{upon}), LINK(*,P+1,OBJ), DELNODE(P) ,DELCAT(P+1,N).

R10.7 CHKENTRY(P,'WITH') & CHKCAT(P+1,N,'',)

= ASS(INS,SUBFEA.{with}), LINK(*,P+1,INS), DELNODE(P) ,DELCAT(P+1,N)

R10.8 CHKENTRY(P,'BY') & CHKCAT(P+1,N,'','')

= ASS(INS,SUBFEA.{by}), LINK(*,P+1,INS), DELNODE(P) ,DELCAT(P+1,N)

R10.9 CHKENTRY(P,'TO') & CHKCAT(P+1,N,'','') & CHKCAT(P+1,V,'',''=FALSE)

= ASS(TAR,SUBFEA.{to}), LINK(*,P+1,TAR), DELNODE(P) ,DELCAT(P+1,N)

R10.10 CHKENTRY(P,'FOR') & CHKCAT(P+1,N,'','')

= ASS(TAR,SUBFEA.{for}), LINK(*,P+1,TAR), DELNODE(P) ,DELCAT(P+1,N)

R10.11 CHKENTRY(P,'AT') & CHKCAT(P+1,N,'','232')

= ASS(TIM,SUBFEA.{at}), LINK(*,P+1,TIM), DELNODE(P) ,DELCAT(P+1,N)

R10.12 CHKENTRY(P,'AT') & CHKCAT(P+1,N,'',)

= ASS(TAR,SUBFEA.{at}), LINK(*,P+1,TAR), DELNODE(P) ,DELCAT(P+1,N)

R10.1 - R10.12 (FAIL)

R11 / link ADV,V /

CHKCAT(*,V,'','') & CHKCAT(-1,ADV,'','') = LINK(*,-1,MOD), DELCAT(*,V),

DELCAT(-1,ADV)

R12 / link V,ADV /

CHKCAT(*,V,'','') & CHKCAT(+1,ADV,'','') = LINK(*,+1,MOD), DELCAT(*,V),

DELCAT(+1,ADV)

(FAIL)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

R14 / Shift Window / RS

(SUCCESS) จะทำการเลื่อน monitor window(*) ไปทางซ้ายแล้วย้อนกลับไปกฎไวยากรณ์ที่ 10, 11, 12 และกฎที่ 14 อีกก็จะ FAIL จนกระทั่งเลื่อน monitor window(*) จนเจอคำว่า "NIL" แสดงว่าจบขั้นตอนประมวลผลในส่วนนี้ดังรูปข้างล่างนี้

2	3	4	5	6	10
brother	wants	to	play	football	NIL
N	V	PREP	V,N	N	
				-1	*

4) ผ่านกฎไวยากรณ์ในกลุ่มที่ 4 (Grammar Rules for Verb Pattern Construction) ที่มีกฎไวยากรณ์ที่ 13 และกฎที่ 14 ซึ่ง monitor window(*) จะอยู่ที่ต้นประโยคดังรูปข้างล่างนี้ โดยผ่านกฎไวยากรณ์ที่ 13 และกฎที่ 14 ตามลำดับ

2	3	4	5	6	10
brother	wants	to	play	football	NIL
N	V	PREP	V,N	N	
*	+1				

R13 / Vp1 /

CHKCAT(*,V,'','') & CHKCAT(-1,N,'','') = LINK(*,-1,SUB),

R13.1 NOT(CHKCAT(+1,N)) & NOT(CHKENTRY(+1,'TO'))

= VP(*,1)

R13.2 CHKCAT(+1,N) & NOT(CHKENTRY(+2,'TO'))

= LINK(*,+1,DOB), VP(*,2), DELNODE(+2)

R13.3 CHKENTRY(+1,'TO) & CHKCAT(+2,V) & CHKCAT(+3,N)

= LINK(*,1,PAR), LINK(+2,+3,DOB), VP(*,3),

DELCAT(+2,V), DELCAT(+3,N),

DELNODE(+1), DELNODE(+2), DELNODE(3)

R13.4 CHKENTRY(+1,'TO') & CHKCAT(+2,V)

= LINK(*,+1,PAR), VP(*,3),DELCAT(+2,V),

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

DELNODE(+1), DELNODE(+2)

R13.5 CHKCAT(+1,N) & CHKENTRY(+2,'TO') & CHKCAT(+3,V) & CHKCAT(+4,N)

= LINK(*,+1,DOB), LINK(*,+3,PAR), LINK(+3,+4,DOB), VP(*,4),

DELCAT(+1,N), DELCAT(+3,V), DELCAT(+4,N),

DELNODE(+1), DELNODE(+2), DELNODE(+3), DELNODE(+4)

R13.6 CHKCAT(+1,N) & CHKENTRY(+2,'TO') & CHKCAT(+3,V)

= LINK(*,+1,DOB), LINK(*,+3,PAR), VP(*,4),

DELCAT(+1,N), DELCAT(+3,V),

DELNODE(+1), DELNODE(+2), DELNODE(+3)

R13.1 - R13.4 (FAIL)

R14 / Shift Window / RS

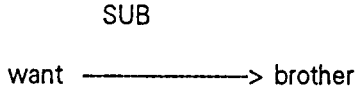
(SUCCESS) จะทำการเลื่อน monitor window(*) ไปทางซ้ายแล้วย้อนกลับไปกฎที่ 12 และกฎที่ 13 อีกก็จะ FAIL จนกระทั่งเลื่อน monitor window(*) ดังรูปข้างล่างนี้ และผ่านกฎไวยากรณ์ที่ 12 และ13 อีกครั้ง

2	3	4	5	6	10
brother	wants	to	play	football	NIL
N	V	PREP	V,N	N	
-1	*	+1	+2		

R13 / Vp1 /

CHKCAT(*,V,'','') & CHKCAT(-1,N,'','') = LINK(*,-1,SUB), DELNODE(-1)

(SUCCESS) จะได้ผลลัพธ์ดังนี้



monitor window(*) จะอยู่ในตำแหน่งดังรูปข้างล่างนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3	4	5	6	10
wants	to	play	football	NIL
V	PREP	V,N	N	
*	+1	+2	+3	

ก็จะย้อนไปกฎไวยากรณ์ที่ 12 และที่ 13ใหม่อีก

R13 / Vp1 /

CHKCAT(*,V,'','') & CHKCAT(-1,N,'','') = LINK(*,-1,SUB), DELNODE(-1)

R13.1 NOT(CHKCAT(+1,N)) & NOT(CHKENTRY(+1,'TO'))

= VP(*,1)

R13.2 CHKCAT(+1,N) & NOT(CHKENTRY(+2,'TO'))

= LINK(*,+1,DOB), VP(*,2), DELNODE(+2)

R13.3 CHKENTRY(+1,'TO) & CHKCAT(+2,V) & CHKCAT(+3,N)

= LINK(*,1,PAR), LINK(+2,+3,DOB), VP(*,3),

DELCAT(+2,V), DELCAT(+3,N),

DELNODE(+1), DELNODE(+2), DELNODE(3)

(SUCCESS) จะได้ผลดังนี้

PAR

want ——————> play และ

DOB

play ——————> football

ซึ่งจะแสดง monitor window(*) ดังรูปข้างล่างนี้

3	10
wants	NIL
V	
*	+1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ก็จะย้อนไปกฎไวยากรณ์ที่ 12 และ 13 อีกครั้ง

R13 / Vp1 /

CHKCAT(*,V,'','') & CHKCAT(-1,N,'','') = LINK(*,-1,SUB), DELNODE(-1)

R13.1 NOT(CHKCAT(+1,N)) & NOT(CHKENTRY(+1,'TO'))

= VP(*,1)

R13.2 CHKCAT(+1,N) & NOT(CHKENTRY(+2,'TO'))

= LINK(*,+1,DOB), VP(*,2), DELNODE(+2)

R13.3 CHKENTRY(+1,'TO) & CHKCAT(+2,V) & CHKCAT(+3,N)

= LINK(*,1,PAR), LINK(+2,+3,DOB), VP(*,3),

DELCAT(+2,V), DELCAT(+3,N),

DELNODE(+1), DELNODE(+2), DELNODE(3)

R13.4 CHKENTRY(+1,'TO') & CHKCAT(+2,V)

= LINK(*,+1,PAR), VP(*,3),DELCAT(+2,V),

DELNODE(+1), DELNODE(+2)

R13.5 CHKCAT(+1,N) & CHKENTRY(+2,'TO') & CHKCAT(+3,V) & CHKCAT(+4,N)

= LINK(*,+1,DOB), LINK(*,+3,PAR), LINK(+3,+4,DOB), VP(*,4),

DELCAT(+1,N), DELCAT(+3,V), DELCAT(+4,N),

DELNODE(+1), DELNODE(+2), DELNODE(+3), DELNODE(+4)

R13.6 CHKCAT(+1,N) & CHKENTRY(+2,'TO') & CHKCAT(+3,V)

= LINK(*,+1,DOB), LINK(*,+3,PAR), VP(*,4),

DELCAT(+1,N), DELCAT(+3,V),

DELNODE(+1), DELNODE(+2), DELNODE(+3)

R13.1 - R13.4 (FAIL)

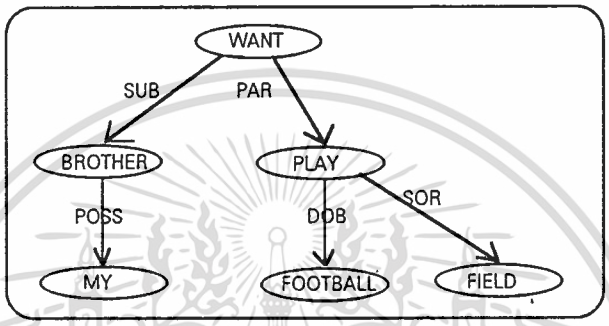
R14 / Shift Window / RS

(SUCCESS) จะทำการเลื่อน monitor window(*) ไปทางซ้ายแล้วย้อนกลับไปกฎไวยากรณ์ที่ 12 และ 13 อีกก็จะ FAIL จนกระทั่งเลื่อน monitor window(*) จนเจอคำว่า "NIL" แสดงว่าจบขั้นตอนประมวลผลในส่วนนี้ดังรูปข้างล่างนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3	10
wants	NIL
V	
-1	*

ซึ่งเอาต์พุตที่ได้ของประโยค ตัวอย่าง "My brother wants to play football in the filed." แสดงดังรูปที่ 5.8 ข้างล่างนี้



รูปที่ 5.8 แสดงรูปแบบภาษากลางจากตัวอย่างประโยค
"My brother wants to play football in the filed."

5.3.3 โครงสร้างเอาต์พุต

โครงสร้างของเอาต์พุตจะอยู่ในรูปแบบความสัมพันธ์ทางโครงสร้างที่แสดงหน้าที่ของคำทางความหมายระดับผิว (Surface Meaning หรือ Part of speech) ซึ่งมีทิศทางการเชื่อมความสัมพันธ์ที่แสดงถึง syntax head และ dependent ของมันดังแสดงในรูปที่ 5.8

5.4 ส่วนการวิเคราะห์ในระดับอรรถศาสตร์ (Semantic Analysis)

ผลที่ได้จากขบวนการวิเคราะห์ระดับวากยสัมพันธ์ที่ออกมาเป็นความสัมพันธ์ทางโครงสร้างระดับผิว ซึ่งเป็นโครงสร้างที่ขึ้นอยู่กับภาษาต้นฉบับ ดังนั้นขบวนการวิเคราะห์ในระดับอรรถศาสตร์ จะต้องถอดส่วนที่ขึ้นกับภาษาต้นฉบับออกให้กลายเป็นโครงสร้างทางความหมายที่ไม่ขึ้นต่อภาษาต้นฉบับ (Language-Independent Conceptual Structure หรือ ภาษากลาง)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.4.1 โครงสร้างอินพุต

โครงสร้างของอินพุตของการวิเคราะห์ทางความหมายเป็นความสัมพันธ์ทางโครงสร้างระดับผิวที่ยังขึ้นกับภาษาต้นฉบับดังแสดงในรูปที่ 5.8

5.4.2 การประมวลผล

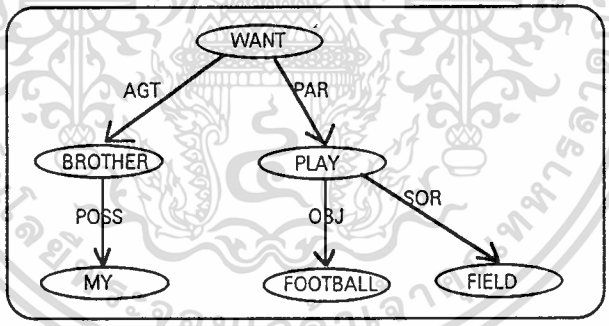
จะทำการ Mapping CASE Relation ดังนี้

SURFACE RELATION		CASE RELATION
SUB	---------- Mapping ---------->	AGT
DOB	---------- Mapping ---------->	OBJ

จะทำให้ได้เอาต์พุตเป็นรูปแบบภาษากลางที่ไม่ขึ้นกับภาษาต้นฉบับ

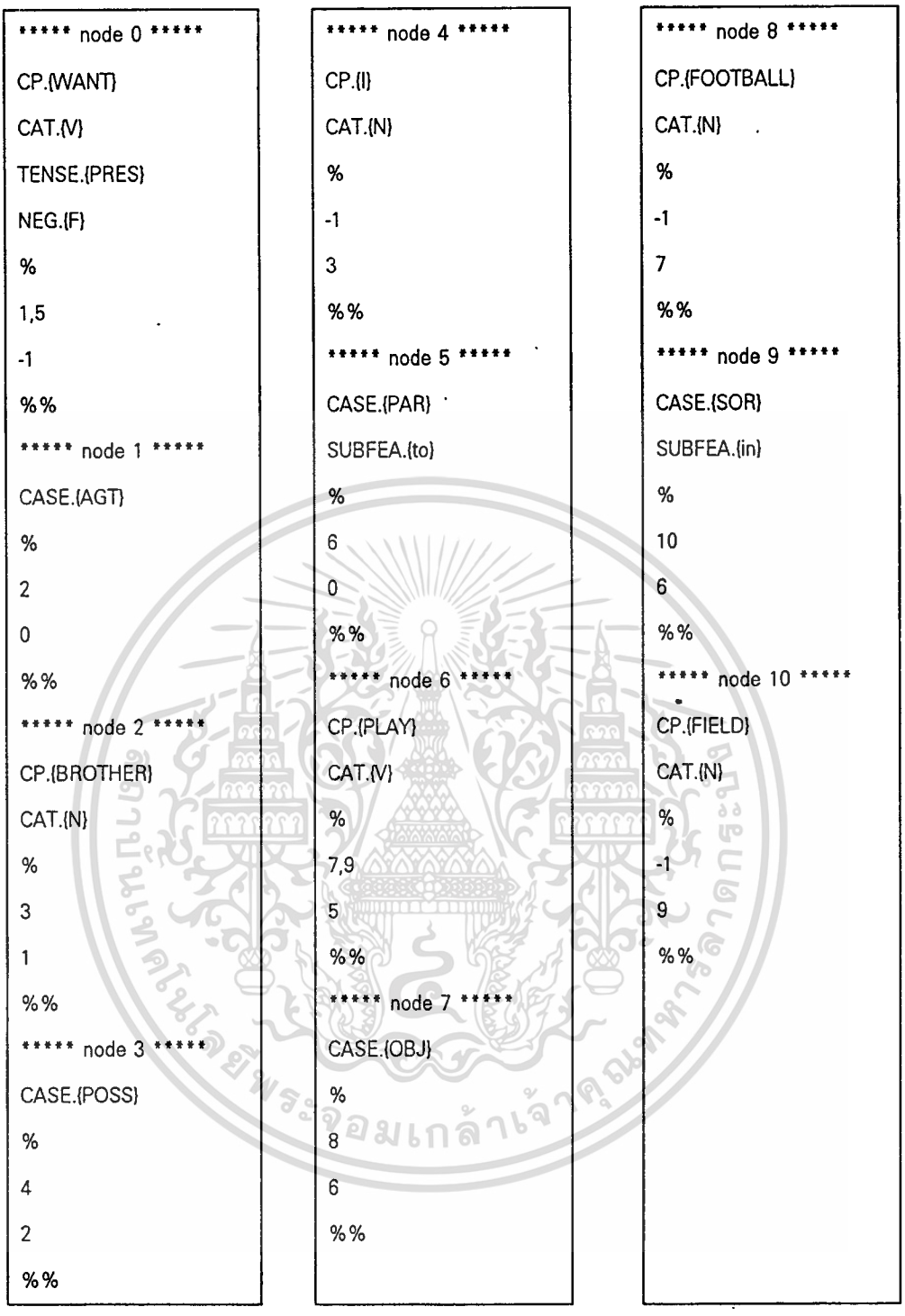
5.4.3 โครงสร้างเอาต์พุต

จากตัวอย่างประโยค "My brother wants to play football in the filed." เมื่อผ่านขบวนการวิเคราะห์ในระดับอรรถศาสตร์จะได้เอาต์พุตดังแสดงในรูปแบบของต้นไม้หรือรูปแบบทางโครงสร้างดังในรูปที่ 5.9 และแสดงรูปแบบข้อมูลเป็นบล็อคดังรูปที่ 5.10



รูปที่ 5.9 แสดงโครงสร้างภาษากลาง
ของตัวอย่างประโยค "My brother wants to play football in the filed."

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 5.10 แสดงรูปแบบข้อมูลเป็นบล็อคของ

ตัวอย่างประโยค"My brother wants to play football in the filed."

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 6
ซอฟท์แวร์ของระบบวิเคราะห์ประโยคภาษาอังกฤษไปสู่ภาษากลาง

6.1 บทนำ

ในการเขียนซอฟท์แวร์เพื่อรองรับการทำงานใดงานหนึ่งในแต่ละครั้งมีสิ่งที่ต้องคำนึงถึงก็คือ อัลกอริทึมของซอฟท์แวร์ ว่าจะมีขั้นตอนการทำงานอย่างไร มีลักษณะโครงสร้างข้อมูล (Data Structure) ของข้อมูลเข้า (Input Data) และข้อมูลออก (Output Data) และมีการประมวลผลอย่างไรบ้าง เป็นต้น ซึ่งสิ่งเหล่านี้เป็นส่วนสำคัญในการสร้างซอฟท์แวร์มาก จากฐานความรู้เกี่ยวกับไวยากรณ์ภาษาอังกฤษที่กล่าวโดยละเอียดในบทที่ 2 รายละเอียดเกี่ยวกับพจนานุกรมคำศัพท์ของระบบวิเคราะห์ในบทที่ 3 โครงสร้างภาษากลางในบทที่ 4 และอัลกอริทึมการทำงานของการประมวลผลการวิเคราะห์โครงสร้างประโยคในบทที่ 5 เราได้นำอัลกอริทึม ฐานความรู้ต่างๆ และฐานข้อมูลเหล่านี้มาออกแบบและสร้างเป็นซอฟท์แวร์คอมพิวเตอร์ ซึ่งจะมีลำดับการทำงานเริ่มจากการรับข้อมูลประโยคภาษาอังกฤษทางแป้นพิมพ์ แล้วผ่านขบวนการตัดคำเพื่อหารากศัพท์ของคำนั้นๆ ต่อจากนั้นก็นำคำที่ได้ไปค้นหาในพจนานุกรมคำศัพท์เพื่อได้ข้อมูลของคำๆ นั้น เช่น ประเภทของคำ เป็นต้น แล้วจึงนำไปวิเคราะห์ทางด้านวากยสัมพันธ์ (ไวยากรณ์) และอรรถศาสตร์ (ความหมาย) ต่อไป ในบทนี้จะกล่าวถึงการออกแบบซอฟท์แวร์ของส่วนต่างๆ ของระบบทั้งหมด รวมทั้งการสร้างโปรแกรมและการใช้งานจริงของโปรแกรมทั้งหมด

6.2 การออกแบบและการเขียนซอฟท์แวร์ของระบบวิเคราะห์ประโยคภาษาอังกฤษ

ดังที่ได้กล่าวมาทั้งหมดในบทที่ 5 แล้วว่าการวิเคราะห์ประโยคภาษาอังกฤษของงานวิจัยนี้แบ่งออกเป็น 3 ส่วนใหญ่ๆ ซึ่งมีการทำงานร่วมกันตามลำดับก่อน/หลังดังนี้

ลำดับที่ 1 : ส่วนการวิเคราะห์ในระดับคำ
ลำดับที่ 2 : ส่วนการวิเคราะห์ในระดับวากยสัมพันธ์
ลำดับที่ 3 : ส่วนการวิเคราะห์ในระดับอรรถศาสตร์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แต่ในด้านการออกแบบและเขียนซอฟท์แวร์เพื่อให้การทำงานเป็นระบบขั้นตอน และสะดวกต่อการทำงานเป็นระบบงานวิจัยนี้จึงได้เพิ่มส่วนการแสดงผลภาษากลางเป็นลำดับที่ 4 แยกออกมาจากส่วนการวิเคราะห์ในระดับความหมายซึ่งผลลัพธ์ที่ได้จากส่วนการวิเคราะห์ในระดับความหมายมาแสดงในรูปแบบของภาษากลางที่เหมาะสมและสะดวกต่อการนำไปใช้กับการแปลภาษาด้วยเครื่องคอมพิวเตอร์ต่อไป โดยส่วนที่เพิ่มนี้จะเป็นซอฟท์แวร์ที่แสดงรูปแบบภาษากลางในรูปแบบต้นไม้(เป็นรูปแบบที่เข้าใจง่าย) และรูปแบบข้อมูลภาษากลางที่กล่าวมาแล้วในบทที่ 4 ดังนั้นในบทที่ 6 นี้จึงแบ่งการออกแบบและเขียนซอฟท์แวร์ของระบบวิเคราะห์ประโยคภาษาอังกฤษเป็น 4 ลำดับดังนี้

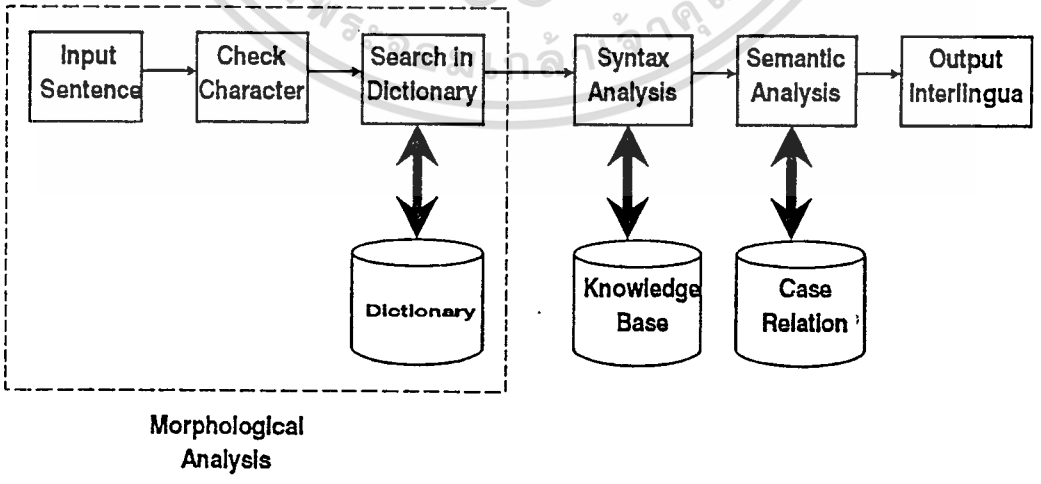
ลำดับที่ 1 : ซอฟท์แวร์ของระบบส่วนการวิเคราะห์ในระดับคำ

ลำดับที่ 2 : ซอฟท์แวร์ของระบบส่วนการวิเคราะห์ในระดับวากยสัมพันธ์

ลำดับที่ 3 : ซอฟท์แวร์ของระบบส่วนการวิเคราะห์ในระดับอรรถศาสตร์

ลำดับที่ 4 : ซอฟท์แวร์ของระบบการแสดงผลภาษากลาง

โดยจะเริ่มจากการหาหน่วยคำโดยอ้างอิงจากพจนานุกรมคำศัพท์ เมื่อได้หน่วยคำแล้วก็จะหาประเภทของคำที่จะทำการประมวลผล ต่อจากนั้นส่งผลลัพธ์ไปยังส่วนการวิเคราะห์ระดับไวยากรณ์ว่าถูกต้องตามหลักไวยากรณ์ภาษาอังกฤษหรือไม่ แล้วส่งต่อมายังขั้นตอนการวิเคราะห์ระดับความหมายเพื่อให้ได้ความสัมพันธ์ระหว่างหน่วยคำแล้วจะผ่านกระบวนการประมวลผลออกมาเป็นรูปแบบภาษากลาง เรานำมาสร้างเป็นระบบซอฟท์แวร์คอมพิวเตอร์ โดยมีขั้นตอนการทำงานดังรูปที่ 6.1



รูปที่ 6.1 แสดงขั้นตอนการทำงานของซอฟท์แวร์ระบบวิเคราะห์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

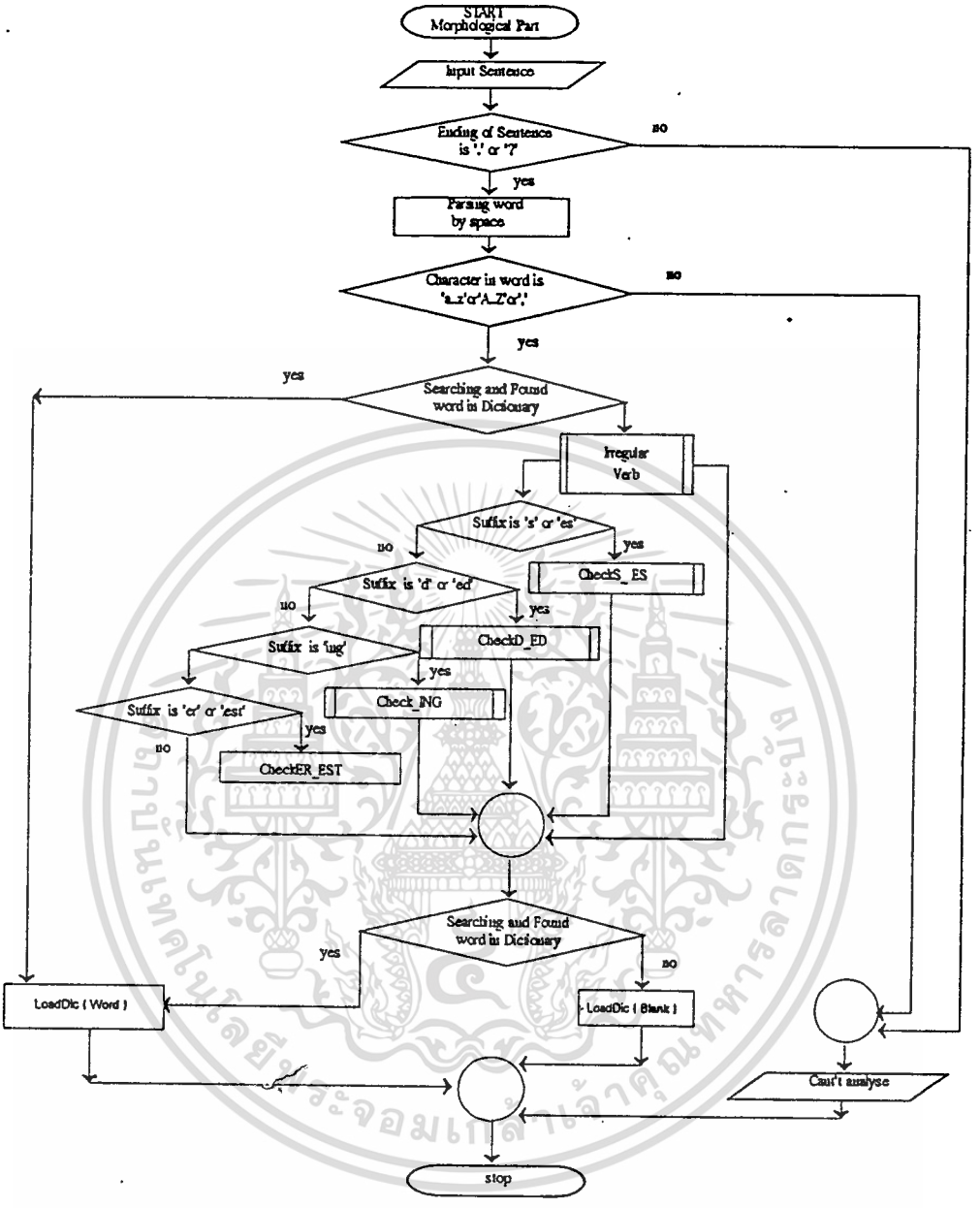
ในรูปที่ 6.1 มีการทำงานดังนี้คือ เมื่อป้อนข้อมูลเป็นประโยคภาษาอังกฤษ (English Sentence) ซึ่งเป็นอักขระที่ได้จากทางแป้นพิมพ์แล้วนำประโยคที่ได้มาตัดคำเพื่อไปตรวจสอบอักขระในแต่ละคำถ้าถูกต้องก็จะนำคำนั้นๆ ไปตรวจสอบกับพจนานุกรมคำศัพท์ว่ามีอยู่ในฐานข้อมูลหรือไม่ หลังจากนั้น ก็นำคำทุกๆ คำในประโยคผ่านกระบวนการฐานความรู้ทางไวยากรณ์ภาษาอังกฤษ และทางด้านความหมาย เพื่อหาความสัมพันธ์ออกมาในรูปแบบภาษากลางโดยเก็บอยู่ในไฟล์ข้อมูล (Text File) ซึ่งจะเป็นผลลัพธ์ของระบบวิเคราะห์โครงสร้างประโยคภาษาอังกฤษ

จากที่กล่าวมาข้างต้น งานวิจัยนี้ได้ทำการออกแบบและสร้างซอฟท์แวร์ในแต่ละส่วนดังมีรายละเอียดต่อไปนี้

6.2.1 ซอฟท์แวร์ของระบบส่วนการวิเคราะห์ในระดับคำ

ในส่วนนี้ข้อมูลอินพุตคือ ประโยคภาษาอังกฤษ ดังนั้นจุดประสงค์ในส่วนนี้คือการหาหน่วยคำ โดยที่การหาหน่วยคำอันดับแรกก็จะใช้ช่องว่างระหว่างหน่วยคำเป็นตัวตรวจสอบดังที่ได้กล่าวรายละเอียดมาแล้วในบทที่ 5 ข้อ 5.2 ซึ่งได้นำมาออกแบบซอฟท์แวร์ และเขียนเป็นโฟล์วชาร์ตของขบวนการประมวลผลทางซอฟท์แวร์ดังรูปที่ 6.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 6.2 แสดงโฟล์วชาร์ตของซอฟท์แวร์ของระบบส่วนการวิเคราะห์ในระดับคำ

จากโฟล์วชาร์ตในรูปที่ 6.2 นำมาออกแบบและสร้างซอฟท์แวร์ ซึ่งจะเริ่มทำงานได้นั้น จะต้องมีการป้อนข้อมูลที่เป็นประโยคภาษาอังกฤษ โดยมีรูปแบบของข้อมูลจะต้องเป็นอักขระภาษาอังกฤษที่ป้อนเข้ามาทางแป้นพิมพ์ในรูปแบบประโยคภาษาอังกฤษทั่วๆ ไป เมื่อประโยคป้อนเข้าสู่ระบบการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วิเคราะห์ในระดับหน่วยคำ ข้อมูลนั้นจะถูกเก็บไว้ในหน่วยความจำชั่วคราว (คือส่วน memory ของเครื่องคอมพิวเตอร์ เรียกว่า Key Buffer) มีลักษณะเรียงลำดับของอักขระตามลำดับของการป้อนเข้าสู่ระบบจากแป้นพิมพ์ ตัวอย่างเช่นป้อนประโยค "I have a house." จะมีการเรียงลำดับของอักขระและสัญญลักษณ์เช่น ช่องว่าง(ใช้เครื่องหมาย SM: Space Marker), .(full stop), ใน Key Buffer ดังแสดงในรูปที่ 6.3

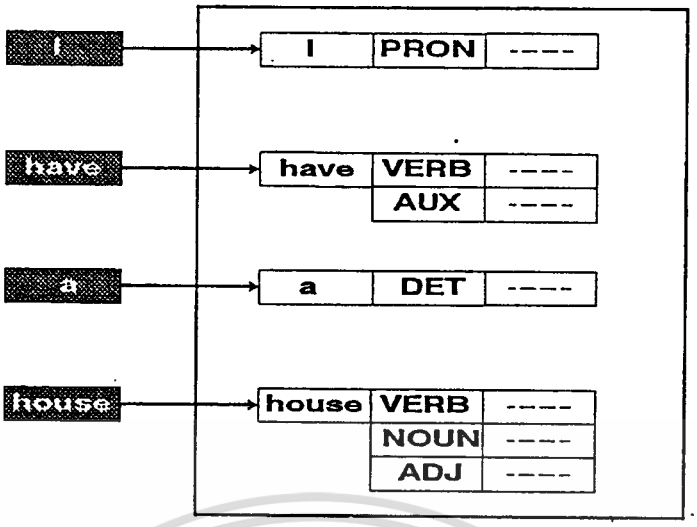
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
I	SM	h	a	v	e	SM	a	SM	h	o	u	s	e	.	enter

รูปที่ 6.3 แสดงลักษณะการเรียงลำดับของอักขระในประโยคตัวอย่าง "I have a house."
(SM : Space Marker เครื่องหมายช่องว่างระหว่างหน่วยคำ)

เมื่อผ่านกระบวนการการตัดคำโดยใช้ช่องว่างระหว่างคำเป็นตัวตรวจสอบหาหน่วยคำและตรวจสอบว่าอักขระอยู่ในเซ็ท [a..z], [A..Z] หรือ คือสัญญลักษณ์ เช่น ; หรือ , หรือ : หรือไม่ ถ้าขั้นตอนดังกล่าวสำเร็จ ก็จะนำหน่วยคำนั้นๆ ที่เก็บอยู่ในหน่วยความจำชั่วคราวเป็นดรรชนีในการเข้าค้นหาคำศัพท์ในพจนานุกรมคำศัพท์ทางโครงสร้างและความหมาย ถ้าหาไม่พบก็จะนำหน่วยคำนั้นๆ ไปผ่านบวนหารากคำศัพท์ โดยจะนำหน่วยคำนั้นเป็นดรรชนีในการค้นหากับไฟล์ที่เป็นฐานข้อมูลของกริยาที่มีรูปไม่ปกติซึ่งจะเปลี่ยนรูปตามกาลของประโยค (Irregular Verb) และผลลัพธ์ที่ได้ก็คือ หน่วยคำที่เป็นดรรชนีในการเข้าค้นหาในพจนานุกรมคำศัพท์ทางโครงสร้างและความหมาย แต่ถ้าไม่มีหน่วยคำนี้ในพจนานุกรมดังกล่าว ขั้นต่อไปก็คือ ถ้าหน่วยคำนั้นลงท้ายด้วย " s หรือ es, d หรือ ed, ing, er หรือ est" จะนำมาผ่านอัลกอริทึมในการตัดส่วนเสริมหน่วยคำ (กล่าวมาแล้วในบทที่ 5 ข้อ 5.2.2.4) ถ้าผ่านขั้นตอนนี้แล้ว เราก็จะได้หน่วยคำเป็นดรรชนีในการเข้าหาคำศัพท์ในพจนานุกรมคำศัพท์ทางโครงสร้างและความหมาย ดังแสดงในรูปที่ 6.4

ซึ่งในส่วนนี้จะทำให้ได้ข้อมูลของหน่วยคำนั้น ๆ ติดมาด้วย เช่น การเป็นคำนาม, คำกริยา เป็นต้น ดังตัวอย่างหน่วยคำที่แสดงในรูปที่ 6.5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 6.4 แสดงตัวอย่างการเข้าหาคำศัพท์ในพจนานุกรมคำศัพท์ของประโยคตัวอย่าง "I have a house."

I	PRON	PPRS	—	—	111	I

have	VERB	VT	2	SUB=AGT, DOB=OBJ	2111	HAVE
	AUX	XHAV	—	—	—	HAVE

a	DET	DDAN	—	—	—	A

house	ADJ	ALRP	—	—	2322	HOUSE
	NOUN	NCMN	—	—	1322	HOUSE
	VERB	VT	2	SUB=AGT, DOB=OBJ	2111	HOUSE

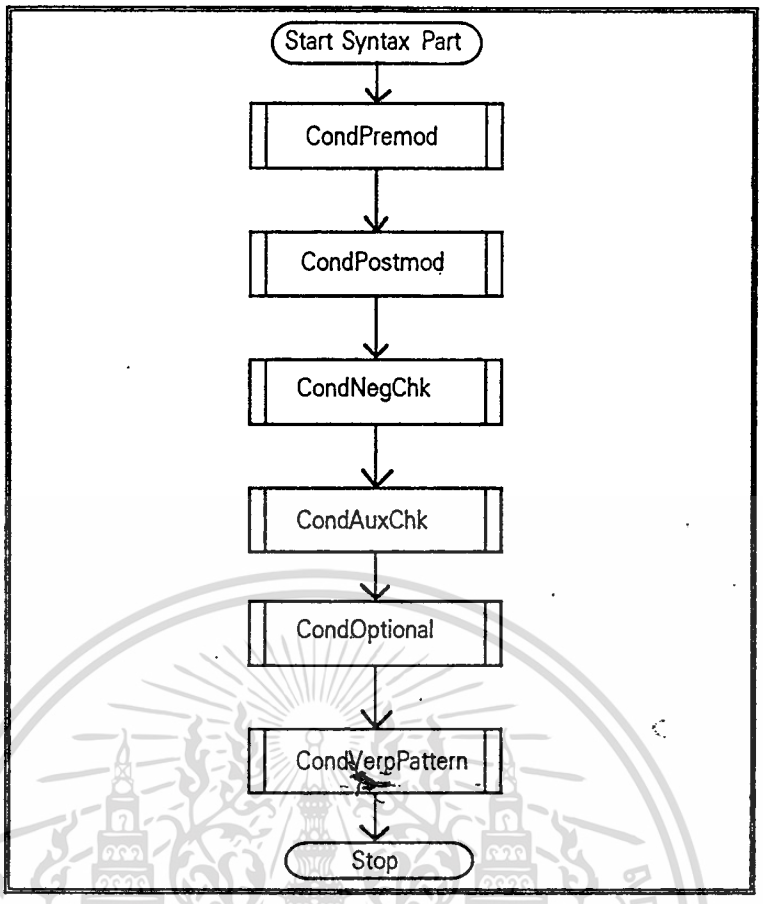
รูปที่ 6.5 แสดงข้อมูลที่ติดมากับหน่วยคำของประโยคตัวอย่าง "I have a house."

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6.2.2 ซอฟท์แวร์ระบบของส่วนการวิเคราะห์ในระดับวากยสัมพันธ์(ไวยากรณ์)

อัลกอริทึมในส่วนนี้จะรับข้อมูลต่อจากระบบของส่วนการวิเคราะห์ในระดับหน่วยคำซึ่งจะมีข้อมูลของหน่วยคำเช่น ประเภทของคำ ซึ่งอาจจะมีประเภทของคำมากกว่า 1 ประเภทก็ได้ เช่นจากผลลัพธ์ของตัวอย่างประโยคข้างต้นในรูปที่ 6.5 จะได้ "have" มีประเภทของคำเป็นทั้งคำกริยา(VERB) และคำกริยาช่วย(AUX) หรือ "house" จะมีประเภทของคำเป็นทั้งคำคุณศัพท์(ADJ) หรือคำนาม (NOUN) หรือคำกริยา(VERB) เป็นต้น จากการที่หน่วยคำหนึ่งๆ สามารถมีประเภทของคำมากกว่าหนึ่ง ดังนั้นจึงได้สร้างกฎไวยากรณ์ขึ้น ในส่วนการวิเคราะห์ระดับวากยสัมพันธ์ดังได้กล่าวรายละเอียดมาแล้วในบทที่ 5 ข้อ 5.3 โดยอาศัยฐานความรู้ทางไวยากรณ์ของภาษาอังกฤษทางด้านกฎโครงสร้างวลี (Phrase Structure Rule) ที่กล่าวรายละเอียดในบทที่ 2 มาใช้ตรวจสอบว่าคำนั้นๆ ควรจะมีประเภทเป็นอะไรตามหน้าที่ของมันในประโยคที่ปรากฏ จากขบวนการและกฎไวยากรณ์ต่างๆ ที่ได้วิจัยและออกแบบในส่วนการวิเคราะห์ระดับวากยสัมพันธ์ที่กล่าวมาในบทที่ 5 ข้อ 5.3 วิทยานิพนธ์นี้ได้นำมาออกแบบซอฟท์แวร์ซึ่งแสดงเป็นโฟล์วชาร์ตของขบวนการประมวลผลทางซอฟท์แวร์ได้ดังรูปที่ 6.6

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



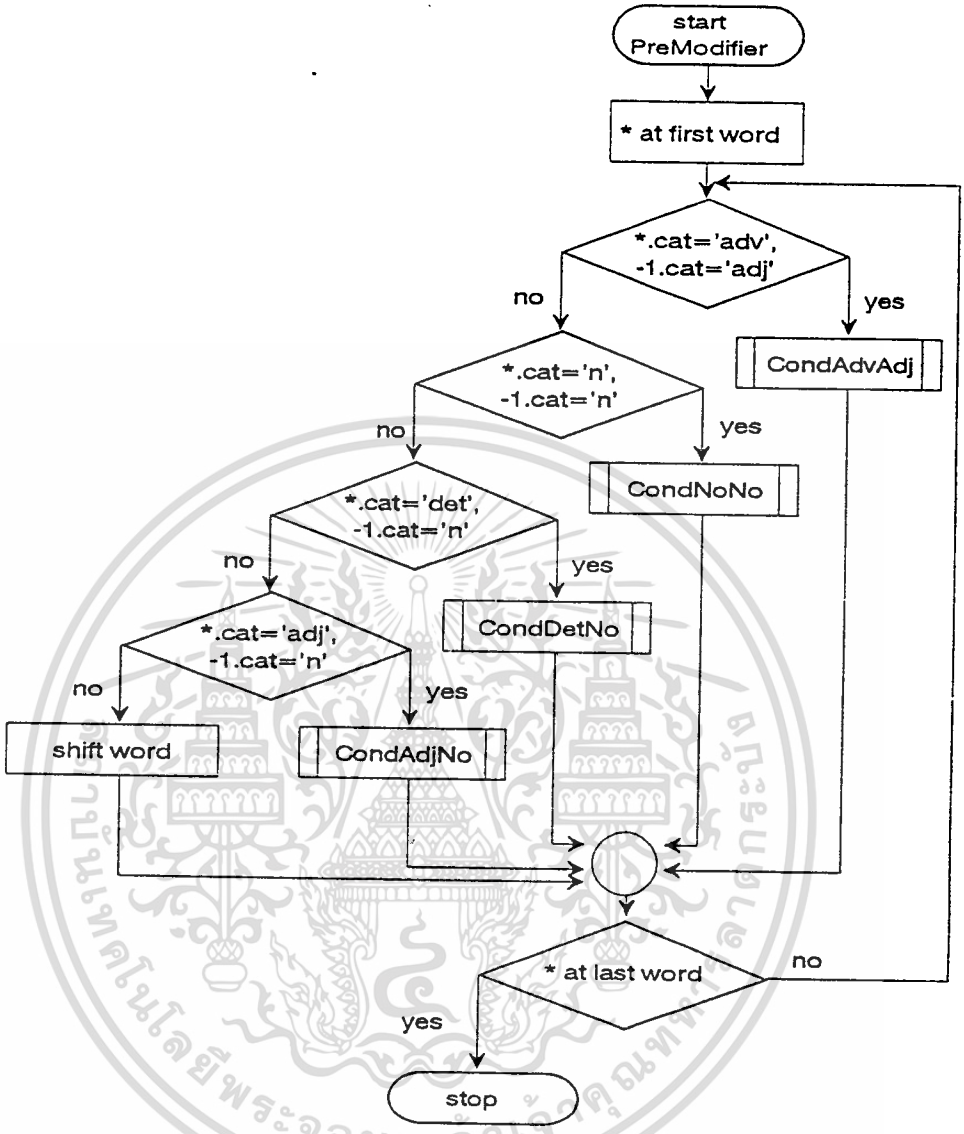
รูปที่ 6.6 แสดงโฟล์วชาร์ตของซอฟท์แวร์ของระบบส่วนการวิเคราะห์ในระดับวากยสัมพันธ์

จากโฟล์วชาร์ตรูปที่ 6.6 แสดงขั้นตอนการทำงานในส่วนการประมวลผลทางวากยสัมพันธ์ (ไวยากรณ์) โดยแบ่งเป็น 6 ขั้นตอน ดังนี้

1) การประมวลผลส่วนขยายหน้าคำนามหลัก (Premodifier) ดังแสดงในโฟล์วชาร์ตในรูปที่ 6.7
2) การประมวลผลส่วนขยายหลังคำนามหลัก (Postmodifier) ดังแสดงในโฟล์วชาร์ตในรูปที่ 6.8
3) การประมวลผลคำปฏิเสธ (Negative) กับคำกริยาหลักของประโยค ดังแสดงขั้นตอนการทำงานในโฟล์วชาร์ตรูปที่ 6.9
4) การประมวลผลคำกริยาช่วย (Auxiliary Verb) กับคำกริยาหลักของประโยค ดังแสดงขั้นตอนการทำงานในโฟล์วชาร์ตรูปที่ 6.10
5) การประมวลผลส่วนขยายประโยค (Optional) กับคำกริยาหลักของประโยค ดังแสดงขั้นตอนการทำงานในโฟล์วชาร์ตรูปที่ 6.11
6) การประมวลผลหารูปแบบคำกริยา (Verp Pattern : vp) กับคำกริยาหลักของประโยค ดังแสดงขั้นตอนการทำงานในโฟล์วชาร์ตรูปที่ 6.12

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในที่นี้เราจะเรียกหน่วยคำว่า คำ (word) ซึ่งจะปรากฏในโฟล์วชาร์ตรูปที่ 6.7 ถึง 6.12

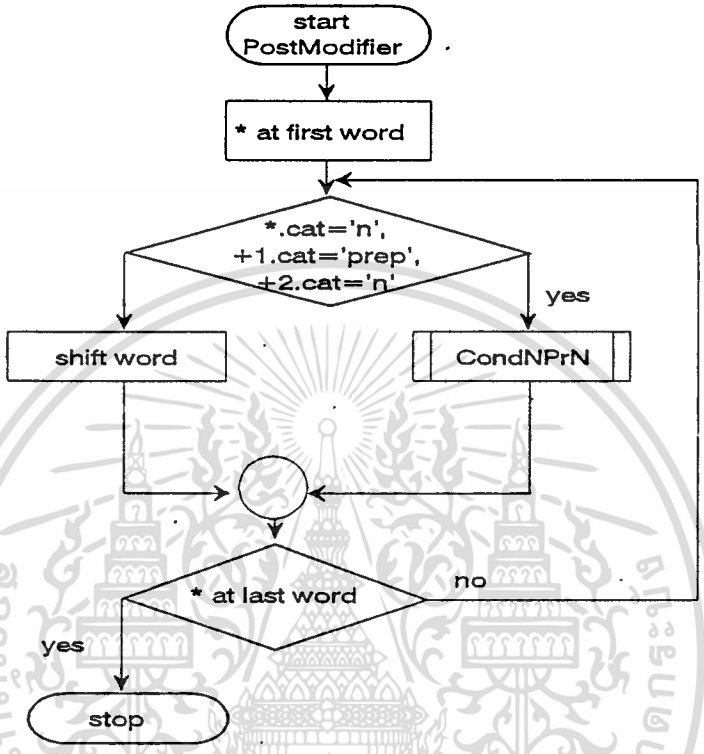


รูปที่ 6.7 แสดงโฟล์วชาร์ตการประมวลผลส่วนขยายหน้าคำนามหลัก
(* คือ monitor window)

จากโฟล์วชาร์ตรูปที่ 6.7 แสดงขั้นตอนการทำงานในส่วนการประมวลผลคำขยายหน้าคำนามหลักเมื่อ

*	หมายถึง คำที่กำลังพิจารณาอยู่
-1	หมายถึง คำที่อยู่ข้างหน้าคำกำลังพิจารณา
*.cat = ' '	หมายถึง ประเภทของคำที่กำลังพิจารณามีประเภทของคำเท่ากับค่าที่อยู่ในเครื่องหมาย ' '

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 6.8 แสดงโฟล์วชาร์ตการประมวลผลส่วนขยายหลังคำนามหลัก
(* คือ monitor window)

จากโฟล์วชาร์ตรูปที่ 6.8 แสดงขั้นตอนการทำงานในส่วนการประมวลผลคำขยายหลังคำนามหลักเมื่อ

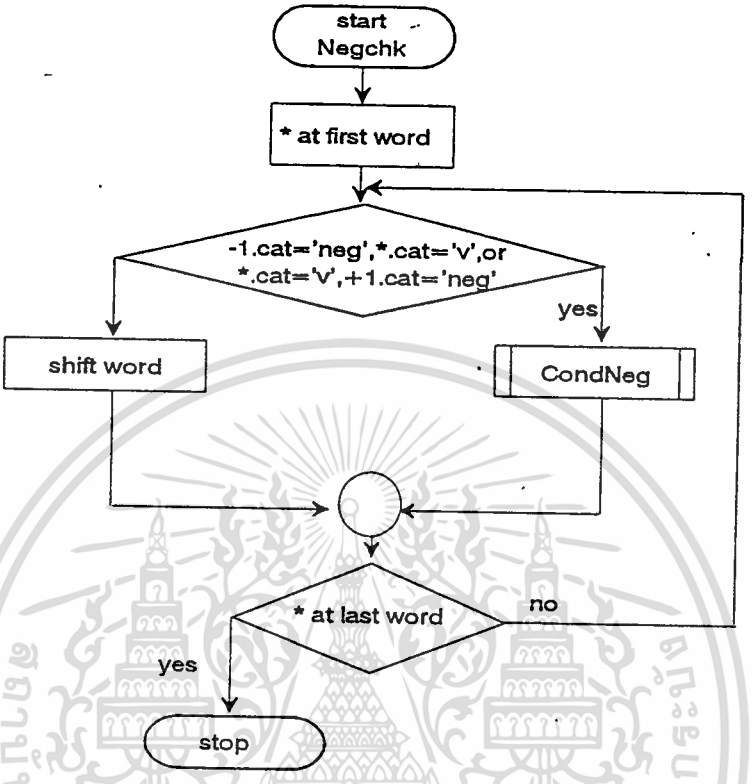
* หมายถึง คำที่กำลังพิจารณาอยู่

+1 หมายถึง คำที่อยู่ข้างหลังคำกำลังพิจารณา

+2 หมายถึง คำที่อยู่ข้างหลังคำ +1

*.cat = ' ' หมายถึง ประเภทของคำที่กำลังพิจารณามีประเภทของคำเท่ากับ ค่าที่อยู่ในเครื่องหมาย ' '

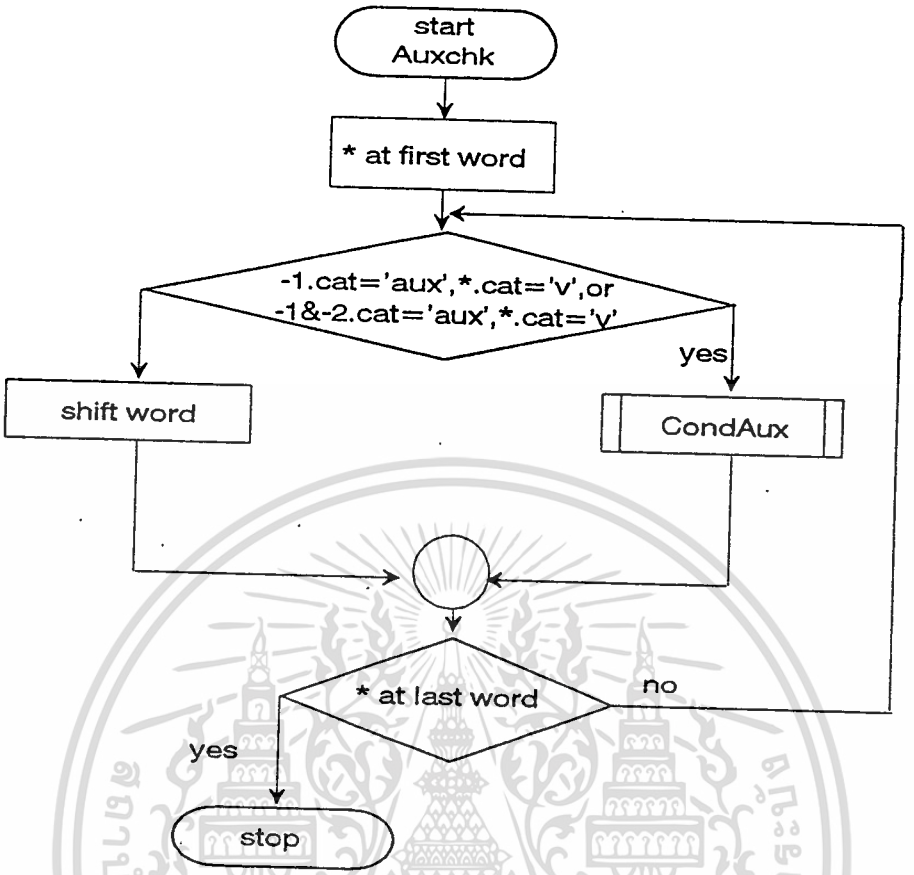
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 6.9 แสดงโฟล์วชาร์ตการประมวลผลคำปฏิเสธกับคำกริยาหลัก
(* คือ monitor window)

จากโฟล์วชาร์ตรูปที่ 6.9 เป็นการประมวลผลระหว่างคำปฏิเสธกับคำกริยาหลักโดยที่ถ้าในประโยคนั้นมีคำปฏิเสธอยู่ก็จะต้องผ่านขั้นตอนนี้ เมื่อผ่านขั้นตอนนี้แล้วคำปฏิเสธจะถูกตัดออกจากประโยคแต่จะมีความหมายทางปฏิเสธให้กับกริยาหลักจะอยู่ในส่วน CondNeg ดังแสดงในรูปที่ 6.9

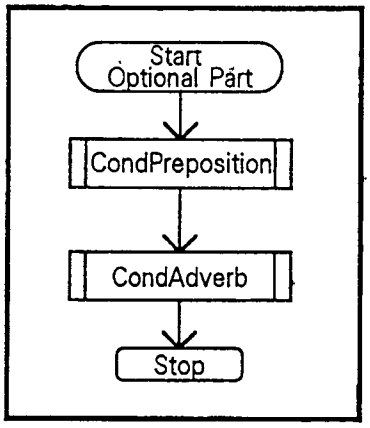
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 6.10 แสดงโฟล์วชาร์ตการประมวลผลคำกริยาช่วยกับคำกริยาหลัก
(* คือ monitor window)

จากโฟล์วชาร์ตรูปที่ 6.10 เป็นการประมวลผลระหว่างคำกริยาช่วยกับคำกริยาหลักโดยที่ถ้าในประโยคนั้นมีคำกริยาช่วยอยู่ก็จะต้องผ่านขั้นตอนนี้ เมื่อผ่านขั้นตอนนี้แล้วคำกริยาช่วยก็จะถูกตัดออกไปจากประโยคโดยจะมีความหมายให้กับคำกริยาหลักในเรื่องกาลของประโยค (Tense) ซึ่งจะอยู่ในส่วน CondAux ในรูปที่ 6.10

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



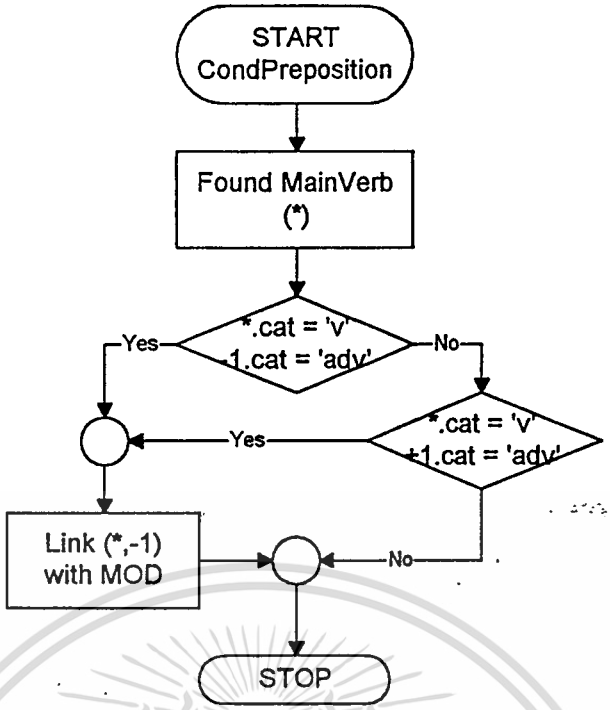
รูปที่ 6.11 แสดงโฟล์วชาร์ตการประมวลผลส่วนขยายประโยคกับคำกริยาหลัก

ในส่วนนี้จะแบ่งเป็น 2 ส่วน

-ส่วนแรกจะประมวลคำกริยาวิเศษณ์กับคำกริยาหลักของประโยคดังโฟล์วชาร์ตในรูปที่ 6.12

-ส่วนที่สองจะประมวลผลคำกริยาหลักของประโยคกับคำบุพบทและคำนามที่ตามหลังคำบุพบทดังแสดงในโฟล์วชาร์ตรูปที่ 6.13

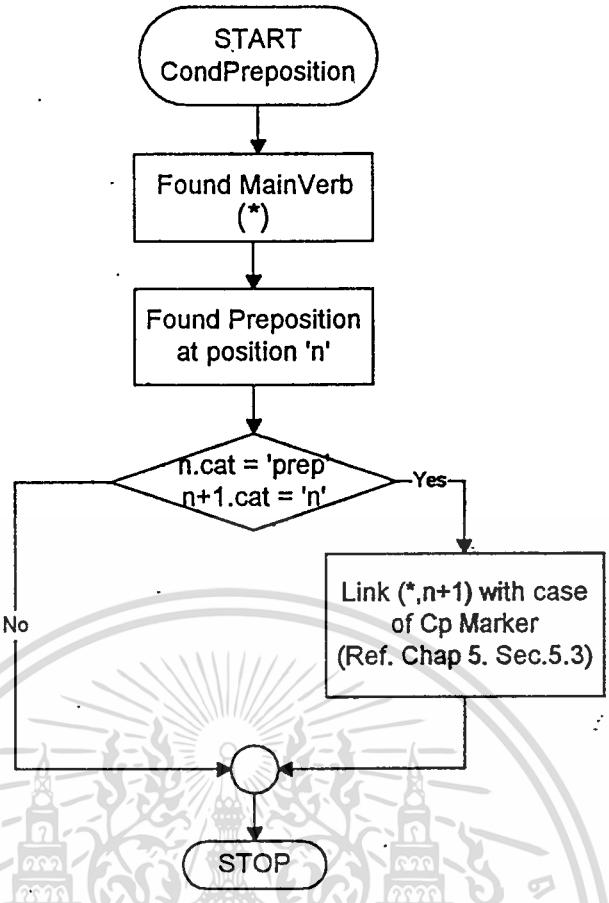
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 6.12 แสดงโฟล์วชาร์ตการประมวลผลคำกริยาวิเศษณ์กับคำกริยาหลัก
(* คือ monitor window)

จากโฟล์วชาร์ตรูปที่ 6.12 เป็นการประมวลผลระหว่างคำกริยาวิเศษณ์กับคำกริยาหลักโดยที่ถ้าในประโยคนั้นมีคำกริยาวิเศษณ์อยู่ก็จะต้องผ่านขั้นตอนนี้ เมื่อผ่านขั้นตอนนี้แล้วคำกริยาวิเศษณ์จะถูกตัดออกจากประโยคแต่จะแสดงความสัมพันธ์ของหน่วยความหมายกับกริยาหลัก(CASE = MOD กล่าวรายละเอียดใน บทที่ 4 ข้อ 4.3)

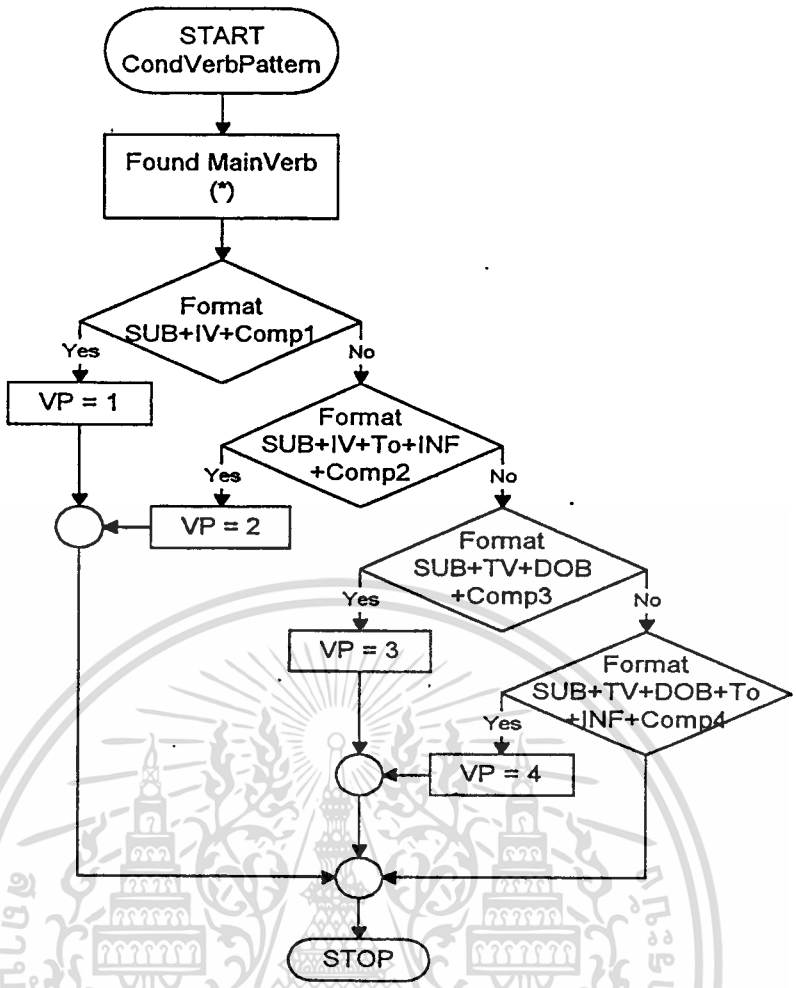
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 6.13 แสดงโฟล์วชาร์ตการประมวลผลคำกริยาหลักกับคำบุพบทและคำนามที่อยู่หลังคำบุพบท
(* คือ monitor window)

จากโฟล์วชาร์ตรูปที่ 6.13 เป็นการประมวลผลระหว่างคำกริยาหลักกับคำบุพบทและคำนามที่อยู่หลังคำบุพบท ในการประมวลนี้จะทำการหาว่ามีคำที่เป็นประเภทของคำเป็นคำบุพบทได้หรือไม่ จากนั้น ก็จะนำคำนามที่อยู่หลังคำบุพบทนั้นมาเป็น Cp of Marker (กล่าวมาแล้วใน บทที่ 5 ข้อ 5.3) เพื่อจะได้ความสัมพันธ์ทางการกระหว่างคำนามนั้นกับคำกริยาหลักของประโยค

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

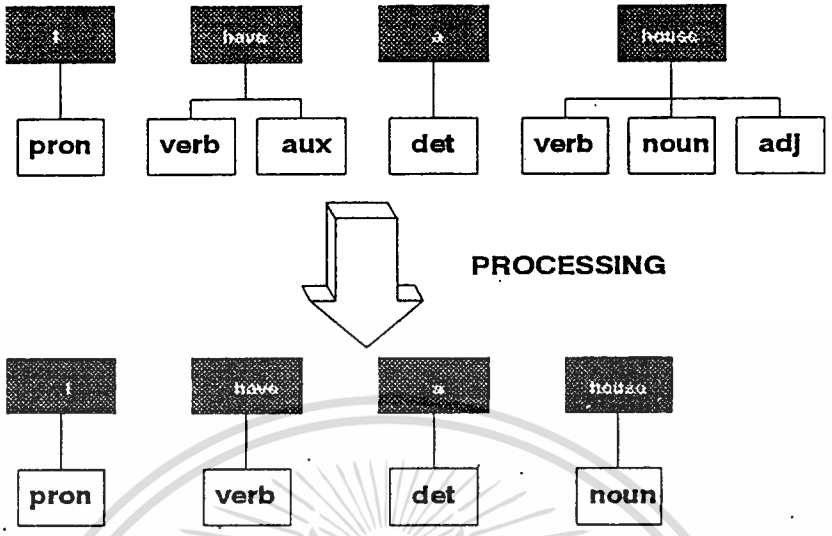


รูปที่ 6.14 แสดงโฟล์วชาร์ตการประมวลผลหารูปแบบของคำกริยา
(* คือ monitor window)

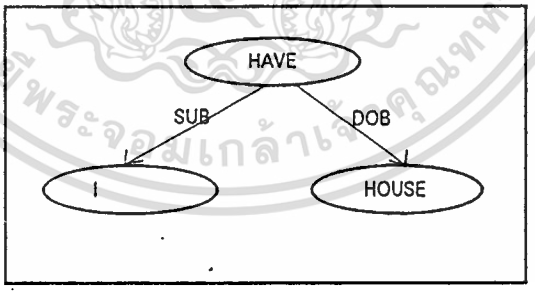
จากโฟล์วชาร์ตรูปที่ 6.14 เป็นการประมวลเพื่อหารูปแบบของคำกริยา(Verp Pattern) ที่ถูกต้องให้กับคำกริยาซึ่งอาจจะมีรูปแบบของคำกริยาได้มากกว่า 1 รูปแบบก็ได้เมื่อผ่านขั้นตอนนี้จะทำให้เหลือรูปแบบของคำกริยาเพียงรูปแบบเดียว และมีการประมวลผลโครงสร้างระดับผิว(Surface Meaning หรือ Part of speech) ให้กับประโยค คือ จะทำให้ทราบว่าคำใดในประโยคทำหน้าที่เป็นประธาน(SUB) หรือถ้าในประโยคนั้นมีกรรมของประโยคเมื่อผ่านขั้นตอนนี้ก็จะทำให้ได้คำที่ทำหน้าที่เป็นกรรมของประโยค(DOB)

จากโฟล์วชาร์ตในรูปที่ 6.7 ถึง 6.14 ได้นำมาสร้างซอฟท์แวร์ซึ่งจะมีกระบวนการประมวลผลทางไวยากรณ์กับทุกหน่วยคำเพื่อหาความสัมพันธ์ทางไวยากรณ์ระหว่างคำซึ่งในขั้นนี้จะมีการจัดการกับหน่วยคำบางหน่วยคำที่มีประเภทของคำมากกว่าหนึ่ง โดยในขั้นนี้สามารถจัดการได้ระดับหนึ่งเพื่อลด ความกำกวมได้ดังแสดงในรูปที่ 6.15 และแสดงโครงสร้างทางต้นไม้ได้ดังรูปที่ 6. 16

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 6.15 แสดงตัวอย่างผลลัพธ์ของหน่วยคำที่ผ่านการประมวลผล
ทางด้านประเภทของหน่วยคำที่มีมากกว่าหนึ่งของประโยคตัวอย่าง "I have a house."



รูปที่ 6.16 แสดงผลลัพธ์ในรูปแผนภาพต้นไม้ของส่วนการวิเคราะห์
ในระดับวากยสัมพันธ์ของประโยคตัวอย่าง "I have a house."

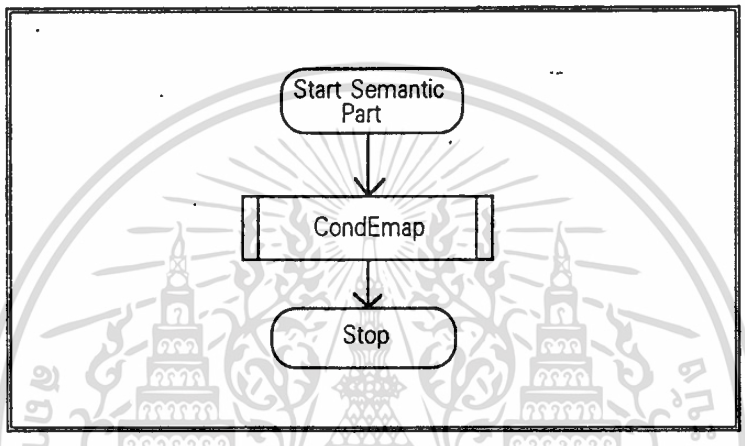
เมื่อ SUB หมายถึง คำที่ถูกลูกศรชี้แสดงหน้าที่เป็นประธานของประโยค
DOB หมายถึง คำที่ถูกลูกศรชี้แสดงหน้าที่เป็นกรรมของประโยค

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6.2.3 ซอฟท์แวร์ระบบของส่วนการวิเคราะห์ในระดับอรรถศาสตร์(ความหมาย)

อัลกอริทึมในส่วนนี้จะรับข้อมูลอินพุตจากส่วนการวิเคราะห์ในระดับวากยสัมพันธ์(ไวยากรณ์) ที่มีรูปแบบทางโครงสร้างต้นไม้มาผ่านการประมวลผลดังโฟล์วชาร์ตในรูปที่ 6.17 โดยทำการ Mapping CASE Relation ดังนี้

SURFACE RELATION		CASE RELATION
SUB	Mapping ——>	AGT
DOB	Mapping ——>	OBJ

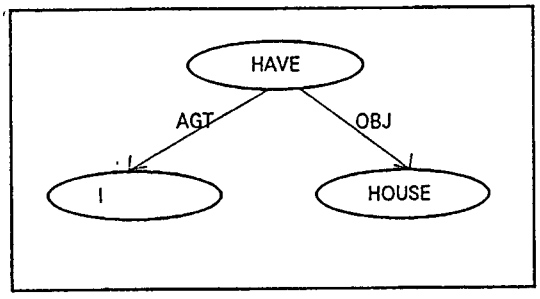


รูปที่ 6.17 แสดงโฟล์วชาร์ตของซอฟท์แวร์ระบบส่วนการวิเคราะห์ในระดับอรรถศาสตร์

จากโฟล์วชาร์ตรูปที่ 6.17 การพิจารณากริยาหลักของประโยคอินพุต ว่าเป็นกริยาที่ต้องมีกรรมมารับหรือไม่ ซึ่งในส่วนนี้จะมีข้อมูลที่ได้จากพจนานุกรมคำศัพท์ เช่น คำว่า "have" มีฟิลด์ EMAP คือ SUB=AGT , DOB=OBJ เป็นต้น เราก็สามารถแปลงความหมายทางด้านวากยสัมพันธ์(ไวยากรณ์) มาเป็นระดับอรรถศาสตร์ (ความหมาย) ซึ่งความหมายต่างๆ ได้กล่าวโดยรายละเอียดในเรื่องการกสัมพันธ์ที่แสดงโดยตัวแปร เช่น AGT, OBJ ว่ามีความหมายอย่างไรบ้าง

ซึ่งจากขั้นตอนดังกล่าวเราจะได้ผลลัพธ์จะแสดงในรูปแผนภาพต้นไม้ดังรูปที่ 6.18

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 6.18 แสดงผลลัพธ์ในรูปแผนภาพต้นไม้ของการวิเคราะห์ระดับอรรถศาสตร์ของประโยคตัวอย่าง "I have a house."

6.2.4. ซอฟท์แวร์ของระบบส่วนการแสดงผลภาษากลาง

อัลกอริทึมในส่วนนี้จะรับข้อมูลจากส่วนการวิเคราะห์ระดับอรรถศาสตร์(ความหมาย) เพื่อแปลงข้อมูลให้อยู่ในรูปแบบภาษากลางที่กำหนดรายละเอียดดังกล่าวแล้วในบทที่ 4 เก็บข้อมูลลงไฟล์เป็นไฟล์ข้อมูลเพื่อนำไปใช้งานต่อไปได้ ดังตัวอย่างรูปแบบภาษากลางของประโยคในรูปที่ 6.19

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
****** node 0 *****
   CP.{HAVE}
   1,3
   -1
***** node 1 *****
   CASE.{AGT}
   2
   0
***** node 2 *****
   CP.{I}
   -1
   1
***** node 3 *****
   CASE.{OBJ}
   4
   0
***** node 4 *****
   CP.{HOUSE}
   -1
   3
```

รูปที่ 6.19 รูปแบบภาษากลางของประโยคตัวอย่าง "I have a house"

6.3 การใช้งานซอฟท์แวร์ของระบบการวิเคราะห์โครงสร้างประโยคภาษาอังกฤษ

ในการใช้งานซอฟท์แวร์นี้จะต้องประกอบไปด้วยไฟล์ข้อมูลดังต่อไปนี้

1) *.DIC เป็นไฟล์ข้อมูลสำหรับเก็บข้อมูลทางพจนานุกรมคำศัพท์

2) VERB.NTX เป็นไฟล์ข้อมูลสำหรับเก็บข้อมูลกริยาที่มีรูปไม่ปกติ (Irregular Verb)

3) *.DAT เป็นไฟล์ข้อมูลสำหรับเก็บเอาท์พุตที่ได้จากการวิเคราะห์

4) MOR4.EXE เป็นซอฟท์แวร์ของการประมวลผลการวิเคราะห์โครงสร้างประโยคภาษาอังกฤษ

5) *.TPU เป็นซอฟท์แวร์ย่อยๆ สร้างเป็นยูนิตเพื่อนำไปใช้งานได้สะดวกขึ้น

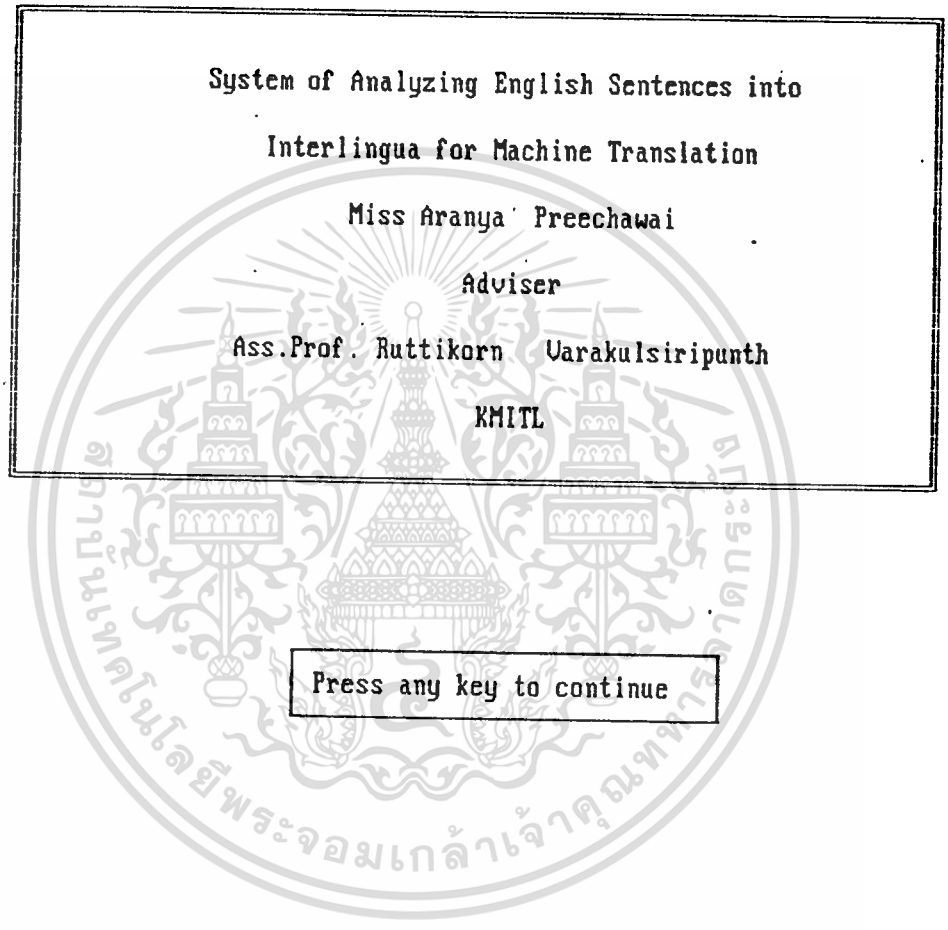
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การใช้งานซอฟท์แวร์นี้จะมีขั้นตอนดังต่อไปนี้

1) พิมพ์คำว่า "MOR4" แล้วกดแป้น Enter หรือ Return ดังนี้

C:> MOR4

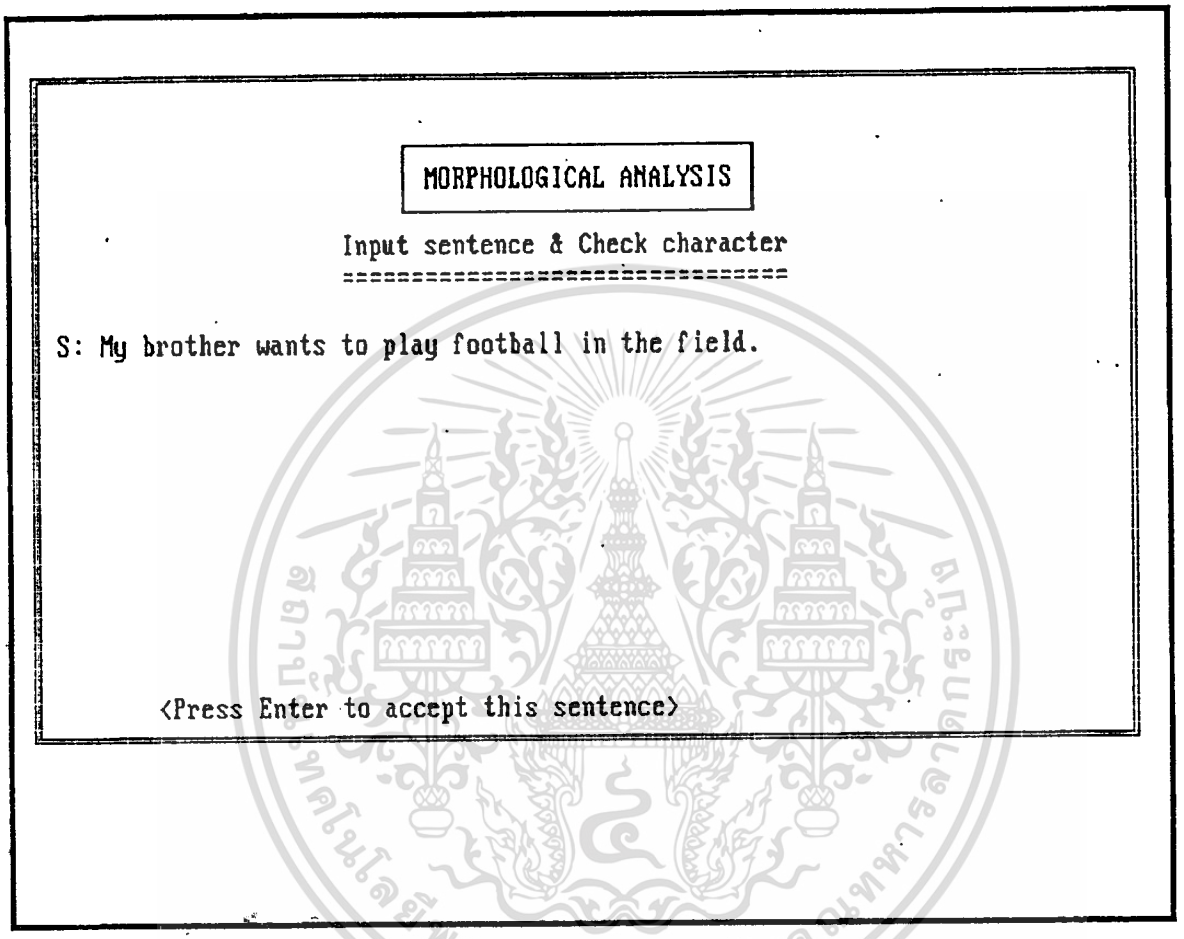
2) จะปรากฏข้อความดังรูปที่ 6.20 ดังนี้



รูปที่ 6.20 แสดงหน้าจอแรกของระบบ

โดยระบบจะรอรับการกดคีย์ใดๆ และเมื่อเรากดคีย์ใดๆ จะปรากฏจอภาพดังรูปที่ 6.21

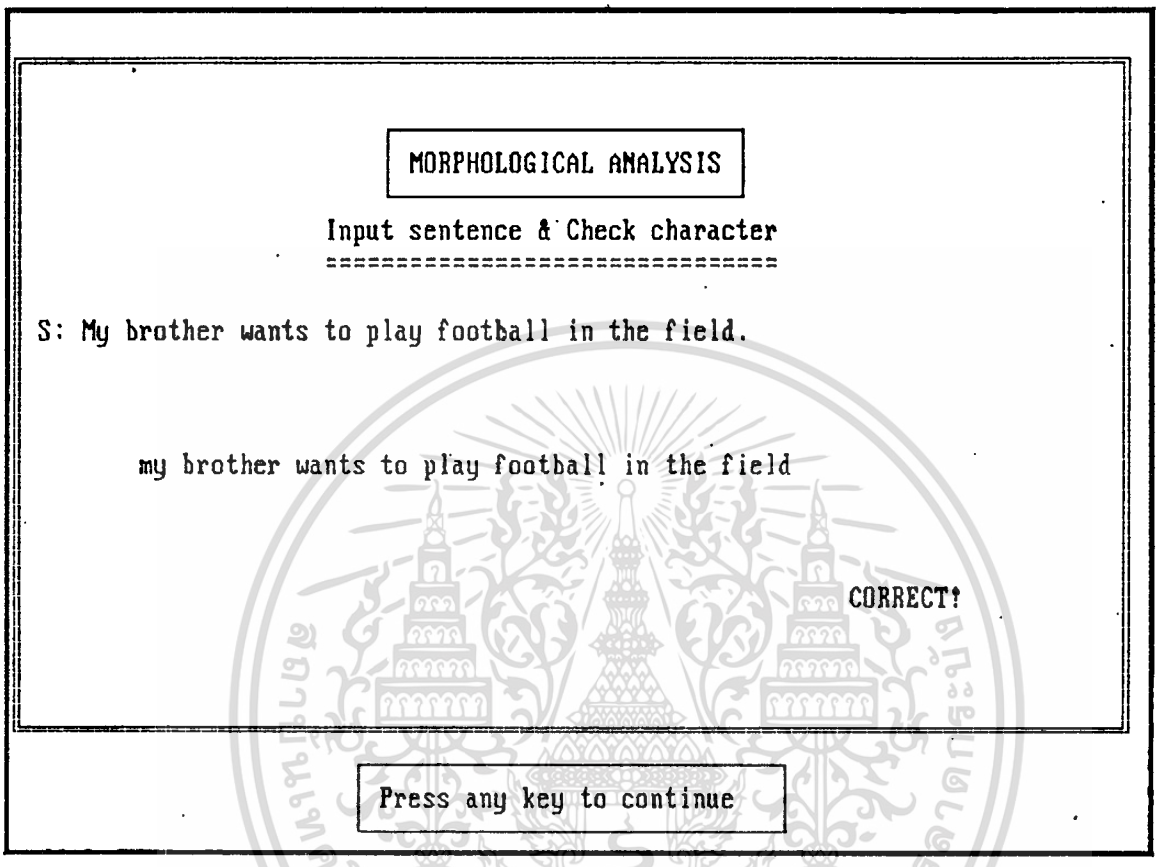
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 6.21 แสดงหน้าจอการรับข้อมูลทางแป้นพิมพ์

ระบบจะรอรับการป้อนประโยคภาษาอังกฤษทางแป้นพิมพ์ที่ช่อง S: เมื่อป้อนประโยคเสร็จแล้วกดคีย์ Enter หรือ Return จะผ่านกระบวนการตรวจสอบอักขระ ซึ่งจะปรากฏจอภาพดังรูปที่ 6.22

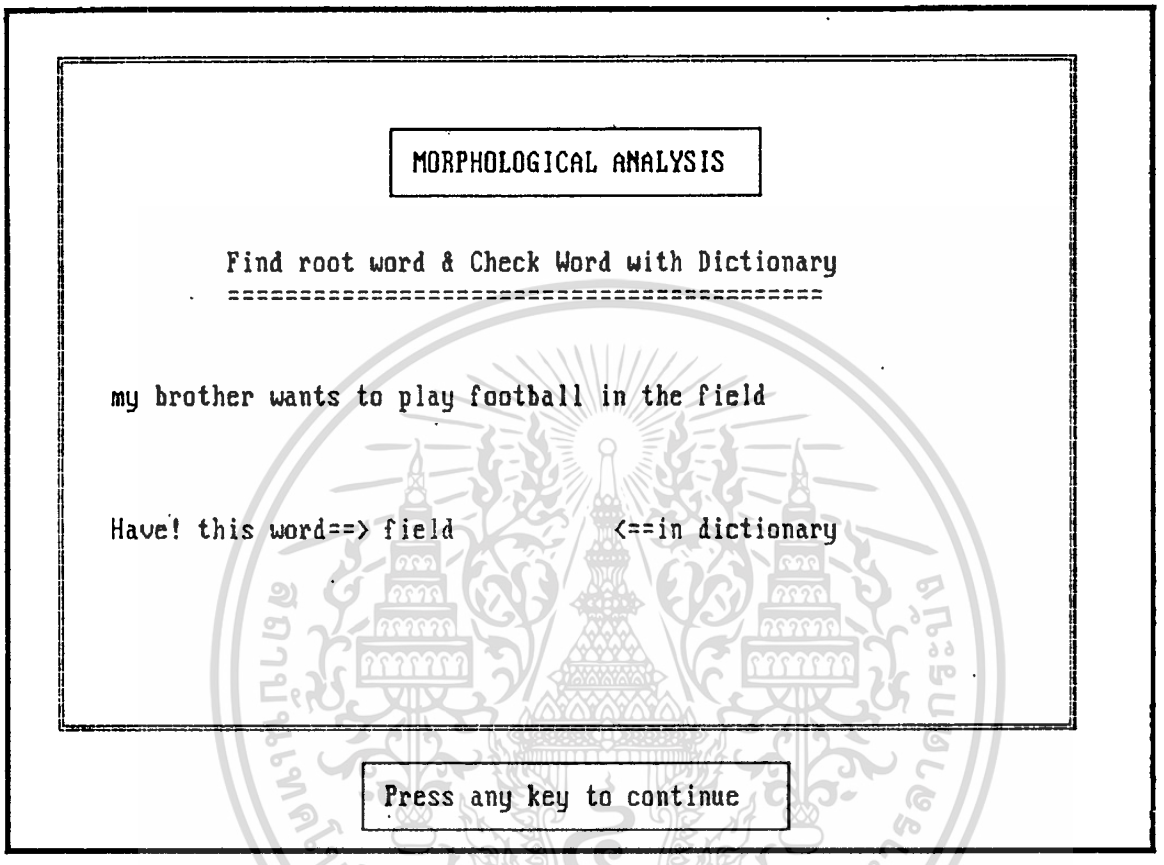
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 6.22 แสดงผลการตรวจสอบอักขระ

จากรูปที่ 6.22 เมื่อผ่านขั้นตอนการตรวจสอบอักขระแล้ว ระบบจะรอรับการกดคีย์ใดๆ เมื่อกดคีย์ใด ๆ แล้ว ระบบจะเข้าสู่กระบวนการค้นหาคำในพจนานุกรมคำศัพท์ ดังแสดงในรูปที่ 6.23

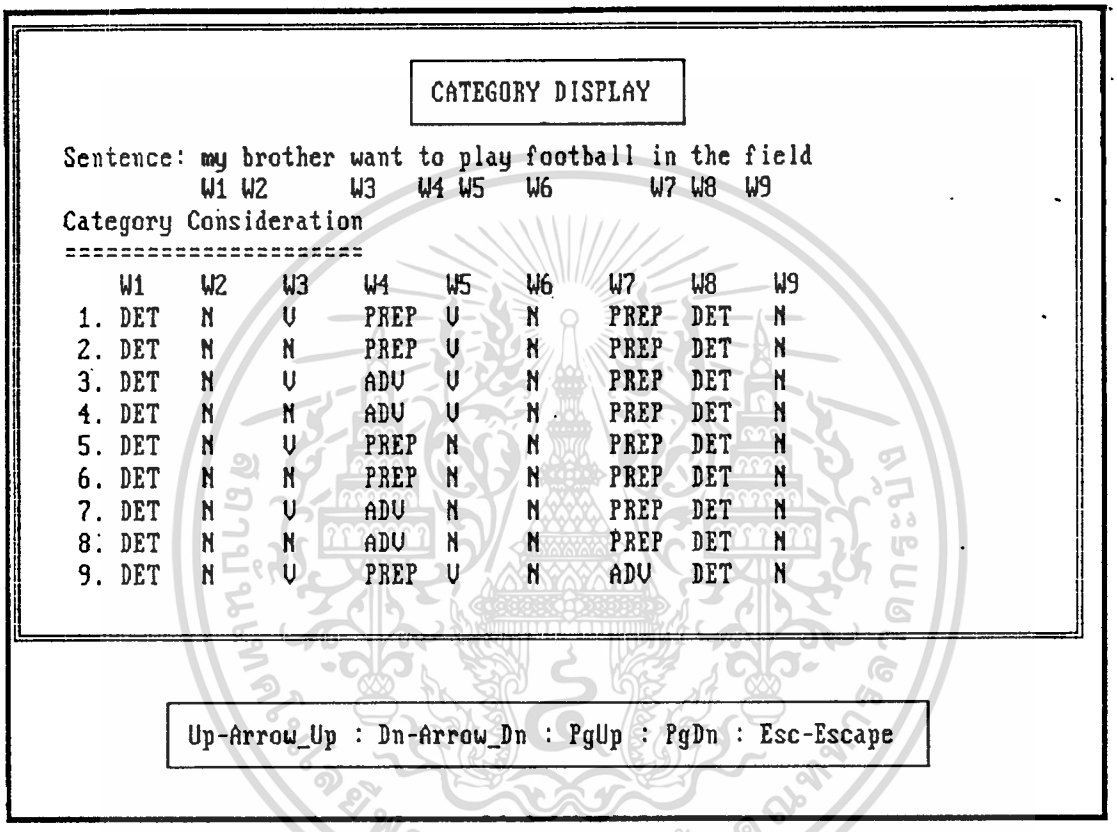
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 6.23 แสดงการค้นหาคำศัพท์ในพจนานุกรมคำศัพท์

เมื่อระบบค้นหาคำศัพท์เสร็จแล้ว ระบบจะรอรับการกดคีย์ใดๆ เมื่อเรากดคีย์ใด ๆ แล้วระบบจะเข้าสู่การประมวลผลทางด้านประเภทของคำทุกคำในประโยค ดังแสดงในรูปที่ 6.24

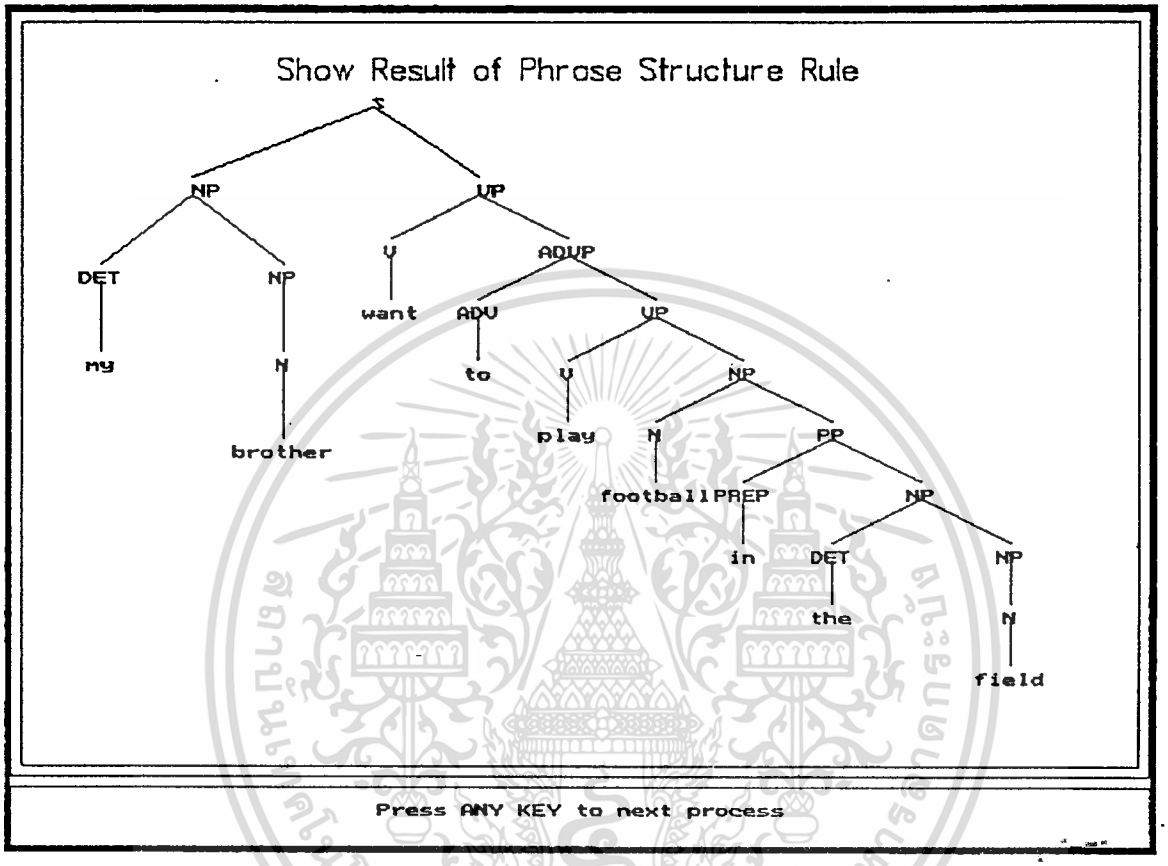
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 6.24 แสดงประเภทของคำในประโยค

จากรูปที่ 6.24 ระบบจะรอการกดคีย์ใด ๆ เมื่อเสร็จกระบวนการดังกล่าวแล้วจะเข้าสู่การวิเคราะห์ทางด้านกฎโครงสร้างวลีดังแสดงในรูปที่ 6.25

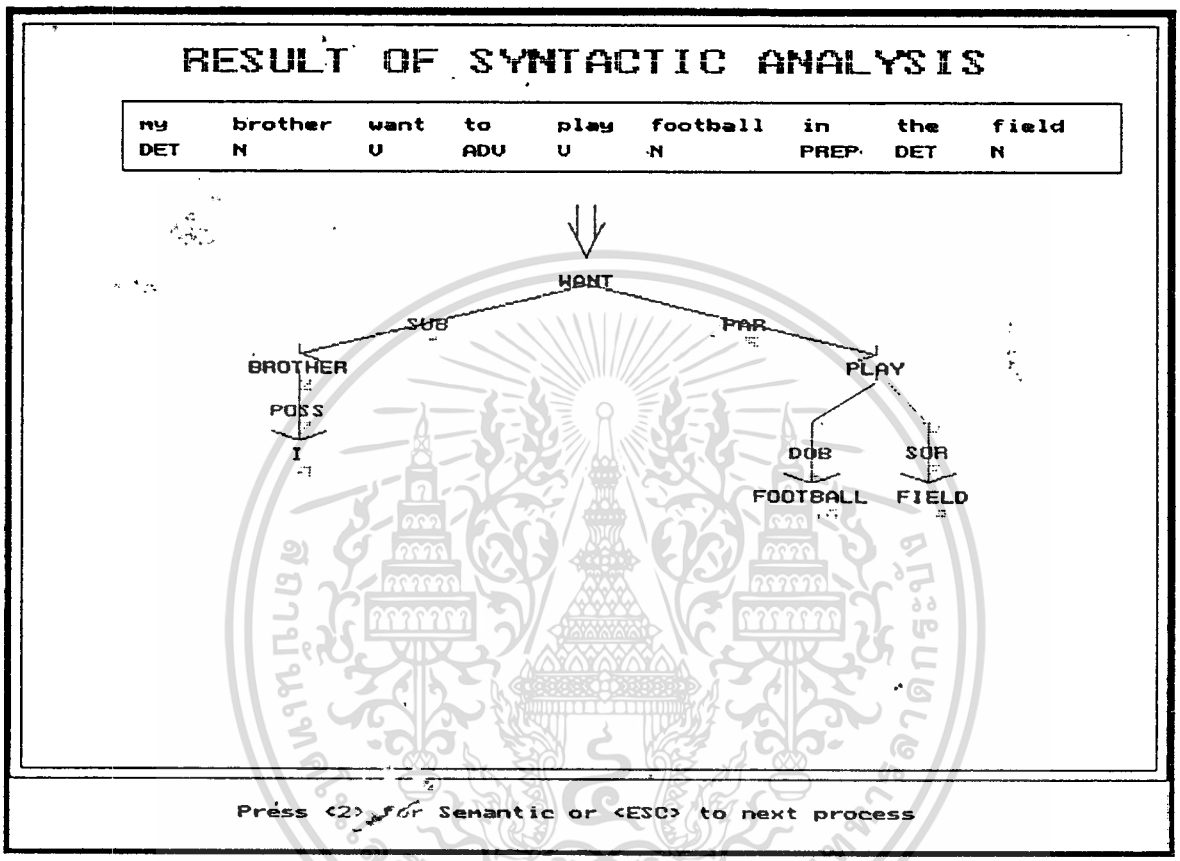
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 6.25 แสดงผลลัพธ์ของกฎโครงสร้างวลี

จากรูปที่ 6.25 ระบบจะรอการกดคีย์ใด และเมื่อเรากดคีย์ใด ๆ แล้วระบบจะผ่านการประมวลผลทางด้านวากยสัมพันธ์ จะได้ผลลัพธ์ดังแสดงในรูปที่ 6.26

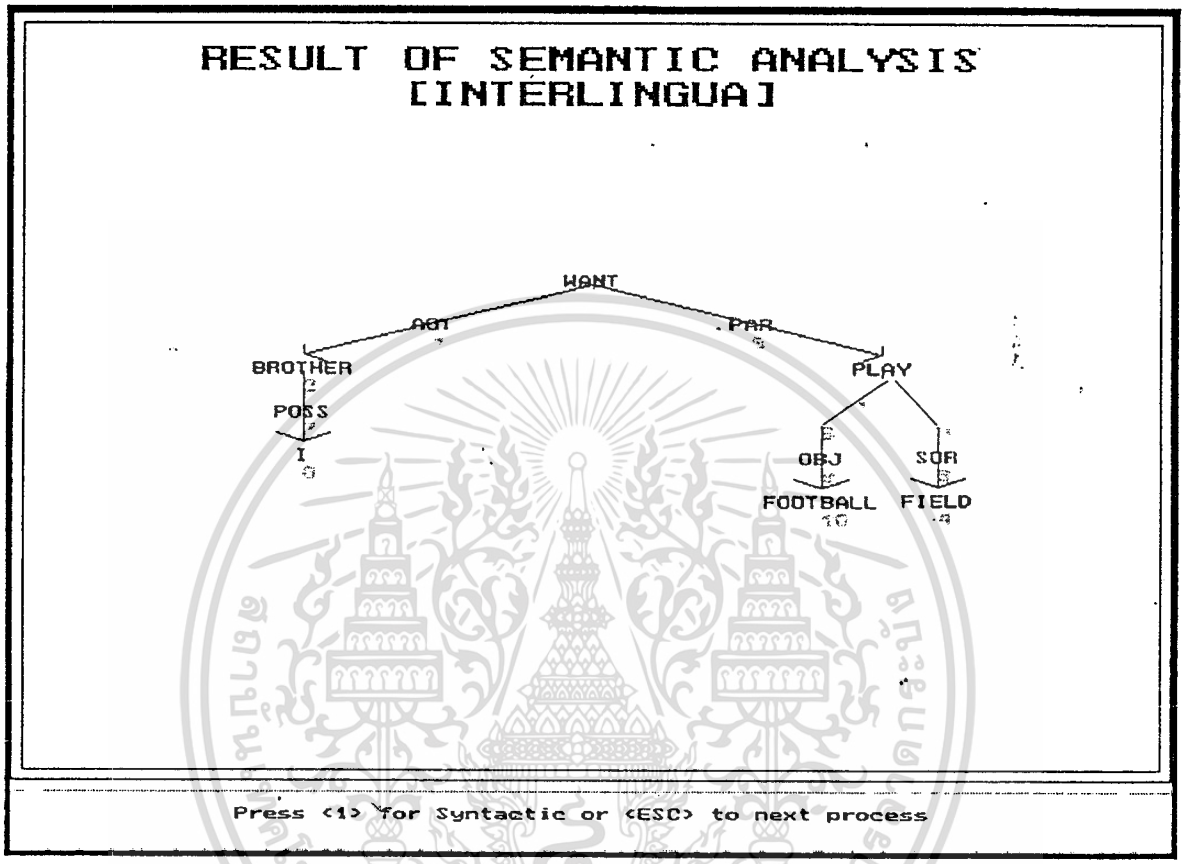
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 6.26 แสดงผลลัพธ์ทางด้านวากยสัมพันธ์

จากรูปที่ 6.26 ระบบจะรอการกดแป้น ESC เพื่อเข้าสู่ระบบการประมวลผลทางด้านอรรถศาสตร์ (ความหมาย) ดังแสดงในรูปที่ 6.27

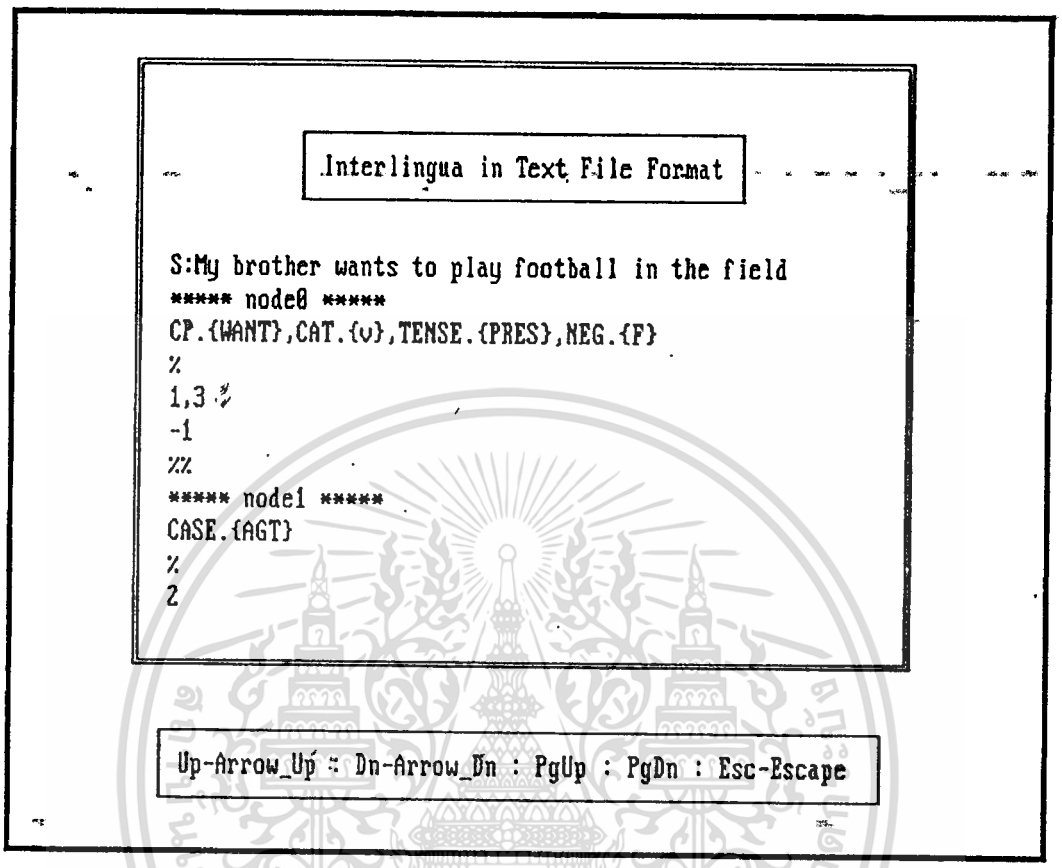
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 6.27 แสดงผลลัพธ์ทางด้านอรรถศาสตร์

โดยระบบจะรอรับการกดแป้น ESC เพื่อเข้าสู่การประมวลผลแปลงข้อมูลทางด้านอรรถศาสตร์เป็นรูปภาษากลาง ดังแสดงในรูปที่ 6.28

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 6.28 แสดงรูปแบบภาษากลาง

จากรูปที่ 6.28 ระบบจะรอการกดแป้น ESC เพื่อที่จะแสดงหน้าจอว่าต้องการทำงานต่อไปหรือไม่ดังแสดงในรูปที่ 6.29

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

7.2.3 โครงสร้างที่ 3 SUB + TV + DOB + COMP3

สำหรับประโยคภายใต้โครงสร้างนี้มีใจความเดียวได้แก่ มีประธานเป็นคนเดียวหรือสิ่งเดียว มีกริยาที่กระทำอาการต่อกรรมของประโยค และมีส่วนขยาย (COMP3) เชื่อมกับคำกริยาหลักของประโยค

ประโยคในรูปลักษณะดังกล่าวนี้จะถูกวิเคราะห์ภายใต้ 46 กฎ

<u>ตัวอย่างประโยค:</u> He gives me some gifts.

เมื่อประโยคผ่านขบวนการวิเคราะห์จะได้ผลลัพธ์ของภาษากลางที่ถูกต้องคือ

```
S:he gives me some gifts
***** node0 *****
CP.(GIVE),CAT.(v),TENSE.(PRES),NEG.(F)
%
1,3,5
-1
%%
***** node1 *****
CASE.(AGT)
%
2
0
%%
***** node2 *****
CP.(HE),CAT.(pron)
%
-1
1
%%
***** node3 *****
```

```
CASE.(TAR)
%
4
0
%%
***** node4 *****
CP.(I),CAT.(pron)
%
-1
3
%%
***** node5 *****
CASE.(OBJ)
%
6
0
%%
***** node6 *****
CP.(GIFT),CAT.(n)
%
```

```
7
5
%%
***** node7 *****
CASE.(VAL)
%
8
6
%%
***** node8 *****
CP.(SOME),CAT.(det)
%
-1
7
%%
```

<u>ข้อมูลทางภาษา</u> : ของประโยคที่ได้จากภาษากลางคือ ทราบว่า He เป็นผู้กระทำกริยา GIVE ต่อกรรมตรงของประโยคคือ ME และกรรมรองของประโยคคือ GIFTS ในรูปของปัจจุบันกาล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

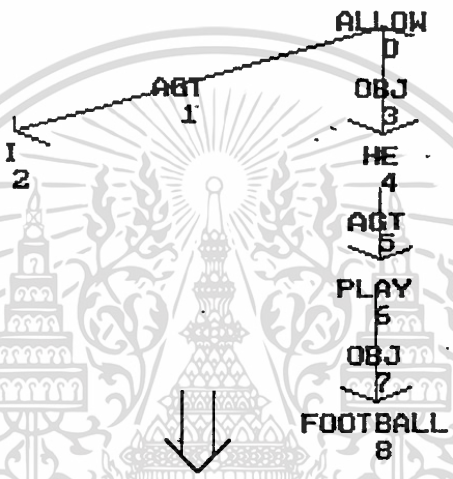
7.2.4 โครงสร้างที่ 4 SUB + TV + DOB + TO INF + COMP4

สำหรับประโยคภายใต้โครงสร้างนี้มีใจความเดียวได้แก่ มีประธานเป็นคนเดียวหรือสิ่งเดียว มีกริยาที่กระทำอาการต่อกรรมของประโยคและกรรมของประโยคจะกระทำอาการด้วยคำกริยาที่อยู่หลังคำว่า "TO" และมีส่วนขยาย (COMP4) เชื่อมกับคำกริยาที่อยู่หลัง "TO"

ประโยคในรูปลักษณะดังกล่าวนี้จะถูกวิเคราะห์ภายใต้ 46 กฎ

ตัวอย่างประโยค: I allowed him to play football.

เมื่อประโยคผ่านขบวนการวิเคราะห์จะได้ผลลัพธ์ของภาษากลางที่ถูกต้องคือ




```
S:I allowed him to play football
***** node0 *****
CP.(ALLOW),CAT.(v),TENSE.(PAST),NEG.(F)
%
1,3
-1
%%
***** node1 *****
CASE.(AGT)
%
2
0
%%
***** node2 *****
CP.(I),CAT.(pron)
%
-1
1
%%
***** node3 *****
CASE.(OBJ)
%
4
0
%%
***** node4 *****
CP.(HE),CAT.(pron)
%
5
3
%%
***** node5 *****
CASE.(AGT)
%
6
4
%%
***** node6 *****
CP.(PLAY),CAT.(v),TENSE.(),NEG.(F)
%
7
5
%%
***** node7 *****
CASE.(OBJ)
%
8
6
%%
***** node8 *****
CP.(FOOTBALL),CAT.(n)
%
-1
7
%%
```

ข้อมูลทางภาษา : ของประโยคที่ได้จากภาษากลางคือ ทราบว่า I เป็นผู้กระทำกริยา ALLOW ต่อผู้ถูกกระทำคือ HIM เพื่อให้ HIM กระทำกริยา PLAY กับ FOOTBALL ในรูปของอดีตกาล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

7.3 ตัวอย่างประโยคอื่น ๆ

สำหรับประโยคที่มีโครงสร้างนอกเหนือจากรูปแบบโครงสร้างทั้ง 4 ที่กล่าวมาในข้อ 7.2 จะยังไม่สามารถวิเคราะห์ด้วยระบบ ณ ขณะนี้ เพราะยังมิได้มีการบรรจุกฎไวยากรณ์ เช่นตัวอย่างประโยคข้างล่างนี้

- John and Mery want to go to school.
- My sister plays and runs in the field.
- She said that she has seen this movie.

ถ้ามีการเพิ่มกฎไวยากรณ์ที่รองรับโครงสร้างอื่นที่กล่าวมาโดยใช้วิธีการของวิทยานิพนธ์นี้เพื่อขยายขอบเขตของการวิเคราะห์ก็จะสามารถประมวลผลประโยคต่าง ๆ ดังกล่าวไปสู่ภาษากลางที่ถูกต้องสมบูรณ์ได้

7.4 การนำผลงานวิจัยไปประยุกต์ใช้กับงานอื่น ๆ

ผลงานวิจัยในวิทยานิพนธ์ที่กล่าวมาทั้งหมดนี้ สามารถที่จะนำไปประยุกต์ใช้กับงานด้านการประมวลผลภาษาธรรมชาติ (Natural Languge Processing) สาขาต่าง ๆ ทั้งในด้านภาษาศาสตร์ และด้านการประมวลผลภาษาทางคอมพิวเตอร์ต่อไป ในวิทยานิพนธ์นี้จะขอแนะนำการประยุกต์ใช้ 3 ตัวอย่างดังนี้

7.4.1 การประยุกต์ใช้กับระบบการแปลหลายภาษาด้วยเครื่องคอมพิวเตอร์ (Multilingual Machine Translation)

ในการแปลภาษาอังกฤษไปเป็นภาษาอื่นๆ นั้น สิ่งที่สำคัญคือ การหาหน่วยคำที่ถูกต้อง แล้วนำมาประมวลผลทางโครงสร้างและประมวลผลทางความหมายเพื่อไปสู่ภาษากลาง ซึ่งระบบนี้สามารถนำไปใช้พัฒนาได้ จะเห็นว่าระบบนี้เป็นส่วนหนึ่งของการแปลภาษาด้วยเครื่องคอมพิวเตอร์

7.4.2 การประยุกต์ใช้กับงานตรวจสอบการสะกดคำในภาษาอังกฤษ (Spelling Check)

ในการตรวจสอบการสะกดคำ หรือตรวจสอบคำในประโยคที่ถูกพิมพ์เข้าสู่คอมพิวเตอร์ที่ใช้ editor ในรูปแบบต่าง ๆ ทาง word processing system สามารถนำขบวนการประมวลผลของการวิเคราะห์ในระดับหน่วยคำที่เสนอในวิทยานิพนธ์นี้ไปประยุกต์ใช้ได้อย่างเหมาะสม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

7.4.3 การประยุกต์ใช้กับงานตรวจสอบโครงสร้างประโยคภาษาอังกฤษ

ขบวนการวิเคราะห์โครงสร้างประโยคภาษาอังกฤษไปสู่ภาษากลาง จะมีประโยชน์ต่อนักภาษาศาสตร์ที่จะใช้ผลลัพธ์ทางภาษากลางในการตรวจสอบและวิเคราะห์โครงสร้างทางไวยากรณ์ และความหมายของภาษาอังกฤษในแง่ภาษาศาสตร์ต่อไป

7.5 แนวทางในการพัฒนาต่อของระบบ

งานวิจัยได้วางแนวทางในการพัฒนาความสามารถของระบบการวิเคราะห์ประโยคภาษาอังกฤษไปสู่ภาษากลางที่เสนอเป็นต้นแบบในวิทยานิพนธ์นี้ต่อไปดังนี้

7.5.1 การวิเคราะห์ประโยคความซ้อน

เนื่องจากระบบนี้สามารถวิเคราะห์ประโยคภาษาอังกฤษที่มีใจความเดียว แต่เราทราบมาว่าประโยคความซ้อนในภาษาอังกฤษจะประกอบไปด้วยประโยคความเดียว 2 ประโยคขึ้นไป ซึ่งเราสามารถเพิ่มกฎการแยกประโยคความซ้อนนี้ให้เป็นประโยคความเดียว 2 ประโยคได้ ดังนั้นจึงทำให้เราสามารถจะวิเคราะห์ประโยคดังกล่าวได้

7.5.2 การวิเคราะห์ให้ได้ครั้งละหลายๆ ประโยค

เนื่องจากระบบนี้จะวิเคราะห์ประโยคภาษาอังกฤษได้ครั้งละประโยค โดยรับข้อมูลทางแป้นพิมพ์ทางเดียวเท่านั้น ระบบนี้สามารถพัฒนาต่อได้โดยให้รับข้อมูลเข้าจาก Floppy Disk , Hard Disk เป็นต้น โดยการอ่านข้อมูลเข้ามาเก็บไว้ในหน่วยความจำชั่วคราว หลังจากนั้นนำข้อมูลทีละประโยคจากหน่วยความจำชั่วคราวมาทำการวิเคราะห์ทีละประโยค

7.5.3 จากข้อจำกัดที่ว่าข้อมูลคำในประโยคนั้นจะต้องมีหน่วยคำในพจนานุกรมคำศัพท์

เราสามารถสร้างซอฟท์แวร์ที่ระบบสามารถแก้ไข, เพิ่ม หรือลบข้อมูลในพจนานุกรมคำศัพท์ได้เข้ามาอีกส่วนหนึ่ง ก็จะทำให้ระบบสามารถทำการวิเคราะห์ต่อไปได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก

ความเป็นมาของการแปลภาษาด้วยระบบคอมพิวเตอร์

การแปลภาษาด้วยคอมพิวเตอร์ เป็นสาขาวิชาหนึ่งในการศึกษาด้านการประมวลผลภาษาธรรมชาติ (Natural Language Processing) หลักฐานชิ้นแรกที่ชี้ให้เห็นถึงความเป็นไปได้ที่จะนำเครื่องคอมพิวเตอร์มาใช้แปลภาษาได้ คือ การยื่นขอจดทะเบียนลิขสิทธิ์ในกรุง มอสโคว์ เมื่อปี 1933 สำหรับวิวัฒนาการของการแปลภาษาด้วยเครื่องคอมพิวเตอร์ (Machine Translation) นั้น สรุปได้ ดังนี้

สมัยที่ 1 ระหว่าง ค.ศ.1946 - ค.ศ.1952 แนวความคิดในการสร้างระบบในขณะนั้น คือ แปลแบบคำต่อคำ อาศัยพจนานุกรมคำศัพท์ (dictionary) ขนาดใหญ่ ต่อมาจึงพัฒนาลักษณะพจนานุกรมคำศัพท์ใหม่ โดยแบ่งเป็น 2 ส่วน ก.พจนานุกรมรากศัพท์ (stem dictionary) ข.พจนานุกรมวิภัติปัจจัย (ending dictionary) และมีการพัฒนาอีกในเรื่อง อินพุต (input) และเอาท์พุต (output) เพื่อการแปลที่สละสลวยและมีใจความที่สมบูรณ์ขึ้น โดยจัดให้มี pre-editor และ post-editor ในช่วงปลายได้เกิดการแตกแยกทางความคิดเป็น 2 ทาง คือ 1.เห็นว่าควรแปลภาษาเป็นคู่ ๆ 2.เห็นว่าควรแปลเป็นภาษากลางก่อน (pivot information, logical หรือ interlanguage)

สมัยที่ 2 ระหว่าง ค.ศ.1952 - ค.ศ.1966 ในช่วงนี้แนวความคิดได้ขยายกว้างขึ้น แต่ยังแตกออกเป็น 2 แนวทาง 1. เห็นว่าต้องใช้อรรถศาสตร์ (Semantic) มาแก้เรื่องความกำกวม 2. เห็นว่าต้องมีการร่วมมือกันระหว่างสาขาต่างๆ ของนักวิจัย เช่น นักปรัชญา นักมานุษยวิทยา นักจิตวิทยา นักภาษาศาสตร์ ผู้เชี่ยวชาญการสื่อสาร นักคณิตศาสตร์ นักสรีรศาสตร์ วิศวกรจัดระบบ

สมัยที่ 3 ระหว่าง ค.ศ.1966 ถึงปัจจุบัน การศึกษาได้แผ่ขยายจนมีศาสตร์แขนงใหม่เกิดขึ้นเรียกว่า "ภาษาศาสตร์คอมพิวเตอร์ " แบ่งย่อยเป็น 8 สาขา มีการแปลเป็น 1 ใน 8 สาขานี้

ซึ่งในขณะนี้สถานภาพทางเทคโนโลยีในต่างประเทศ นั้นในประเทศญี่ปุ่น บริษัทเอกชนหลายแห่งที่ทำงานเกี่ยวข้องกับคอมพิวเตอร์ได้ผลิต Software ระบบการแปลภาษาออกจำหน่ายหลายระบบ เช่น ATLAS ของบริษัท Fujitsu HICATS ของบริษัท Hitachi PANSEE ของบริษัท OKI หรือ PIVOT ของบริษัท NEC ซึ่งแต่ละระบบมีการทำงานแตกต่างกัน นอกจากนั้นรัฐบาลญี่ปุ่นโดยผ่านทาง CICC ของกระทรวง MITI ยังสนับสนุนโครงการร่วมมือระหว่างชาติในการแปลภาษาด้วยคอมพิวเตอร์นี้มาก หน่วยงานต่าง ๆ ที่เกี่ยวข้องกับงานแปลภาษา เช่น หน่วยงานทำงานเกี่ยวกับพจนานุกรมอิเล็กทรอนิกส์ได้รับการจัดตั้งขึ้นเป็นหน่วยงานถาวร เช่น The Japan Electronic Dictionary Research Institute Ltd. เพื่อ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ทำการ Compile พจนานุกรมขนาดใหญ่สำหรับการแปลภาษา และการประมวลผลภาษาธรรมชาติด้วยคอมพิวเตอร์ หรือ หน่วยงานทำการวิจัยเกี่ยวกับโครงสร้างของ Concept โดยใช้ ปริบท (Context) ของ Concept ดังเช่นที่ Electrotechnical Laboratory (ETL) ซึ่งได้พยายามสร้างระบบ CONTRAST (Context Translator) เพื่อให้เครื่องคอมพิวเตอร์เรียนรู้ Contextual Structure ของ Concept เพื่อความเข้าใจภาษาธรรมชาติมากขึ้น เป็นต้น

สำหรับในยุโรปนั้น มีระบบการแปลหลายระบบ เช่น SYSTRAN LOGOS หรือ EUROTA มีการใช้กันอย่างแพร่หลาย รวมทั้งได้มีการนำบางระบบ เช่น SYSTRAN และ LOGOS ไปใช้ในสหรัฐอเมริกาด้วย ในประเทศแคนาดา นอกจากความสำเร็จของระบบ TAUM METEO ซึ่งแปลข่าวพยากรณ์อากาศ อย่างได้ผลแล้ว ยังมีการนำเอาระบบ LOGOS เข้าไปใช้ในหน่วยงานหลายแห่งทั้งภาครัฐบาลและเอกชน เมื่อการพัฒนาการแปลภาษาดำเนินก้าวหน้าไปอย่างไม่หยุดยั้งในส่วนต่าง ๆ ของโลกเช่นนี้จึงได้เกิดมีความคิดว่า ผู้ทำงานทางด้านการแปลภาษาด้วยคอมพิวเตอร์ และผู้ใช้ควรจะได้มีการแลกเปลี่ยนความคิดและข้อมูลซึ่งกันและกัน เพื่อประโยชน์ทุกฝ่าย ดังนั้น จึงมีการตั้งองค์การระหว่างชาติชื่อ An International Association of Translation Technology หรือ An International Association of Machine Translation ขึ้นในเร็วๆ นี้

สำหรับในประเทศไทยนั้นมี งานวิจัยด้าน Machine Translation มาตั้งแต่ปี พ.ศ.2516 โดย ดร. นิตยา กาญจนะวรรณ และโครงการความร่วมมือระหว่างรัฐบาลญี่ปุ่นผ่านกระทรวงการค้าระหว่างประเทศ และอุตสาหกรรม (MITI) โดยมีหน่วยงานชื่อ Center for International Cooperation for Computerization (CICC) เป็นหน่วยงานประสานงานกับรัฐบาลไทยผ่านกระทรวงวิทยาศาสตร์เทคโนโลยีและการพลังงาน โดยมีศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ (NECTEC) เป็นหน่วยประสานงาน ซึ่งมีผู้ทำงานเป็นอาจารย์และนักวิจัยจากมหาวิทยาลัยต่างๆ นอกจากนี้ ยังมีการ วิจัยเกี่ยวกับการแปลภาษาด้วยคอมพิวเตอร์ และการพัฒนาพจนานุกรมอิเล็กทรอนิกส์อีกหลายหน่วยงาน เช่น ที่สถาบันเทคโนโลยีแห่งเอเชีย มีโครงการ "การวิเคราะห์โครงสร้างประโยคภาษาไทยด้วยคอมพิวเตอร์" ของ น.ส. อำไพ พรประเสริฐสกุล หรืองานพัฒนาและสร้างระบบพจนานุกรมที่จะช่วยในการแปลภาษา โดยหาคำที่มีความหมายร่วมกัน (Thesaurus) ของ ร.ศ. ยืน ภู่วรวรรณ และคณะจากมหาวิทยาลัยเกษตรศาสตร์ งานพัฒนาระบบการแปลภาษาอังกฤษ-ไทย โดยใช้ระบบการแปลหลากภาษาที่สถาบันเทคโนโลยีพระจอมเกล้าธนบุรี จากโครงการแปลภาษาที่พัฒนาขึ้นหลายโครงการนี้ ทำให้ศูนย์เทคโนโลยีอิเล็กทรอนิกส์ฯ มีโครงการจะจัดตั้งศูนย์การแปลภาษา หรือศูนย์แลกเปลี่ยนข้อมูลขึ้นเพื่อร่วมงานกับประเทศญี่ปุ่น จีน อินโดนีเซีย และมาเลเซีย ในการให้บริการการแปลภาษา และแลกเปลี่ยนข้อมูลกับเอกชน หรือหน่วยราชการอื่นๆ ซึ่งโครงการจัดตั้งศูนย์การแปลภาษา หรือศูนย์แลกเปลี่ยนข้อมูลนี้ อาจจะเริ่มดำเนินการในปีงบประมาณ 2536 นี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ข
ข้อมูลของคำกริยาที่ผันรูปไปตามกาล

Word Order	Irregular Verb	Entry	Tense Marker
1	abode	abide	PAST,PPER
2	arisen	arise	PPER
3	arose	arise	PAST
4	ate	eat	PAST
5	awoke	awake	PAST
6	awaken	awake	PPER
7	bade	bid	PAST
8	beaten	beat	PPER
9	became	become	PAST
10	been	be	PPER
11	befallen	befall	PPER
12	befell	befall	PAST
13	began	begin	PAST
14	begot	beget	PAST
15	begotten	beget	PPER
16	begun	begin	PPER
17	beheld	behold	PAST,PPER
18	bent	bend	PAST,PPER
19	bereft	bereave	PAST,PPER
20	besought	beseech	PAST,PPER
21	bestridden	bestride	PPER
22	bestrode	bestride	PAST
23	betaken	betake	PPER
24	bethought	bethink	PAST
25	betook	betake	PAST
26	bidden	bid	PPER
27	bit	bite	PAST
28	bitten	bite	PPER

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Word Order	Irregular Verb	Entry	Tense Marker
29	bled	bleed	PAST,PPER
30	blest	bless	PAST,PPER
31	blew	blow	PAST
32	blown	blow	PPER
33	bore	bear	PAST
34	borne	bear	PPER
35	bought	buy	PAST,PPER
36	boun	bind	PAST,PPER
37	bred	breed	PAST,PPER
38	brought	bring	PAST,PPER
39	built	build	PAST,PPER
40	burnt	burn	PAST,PPER
41	came	come	PAST
42	caught	catch	PAST,PPER
43	chid	chide	PAST
44	chidden	chid	PPER
45	chose	choose	PAST
46	chosen	choose	PPER
47	cleft	cleave	PAST
48	clove	cleave	PAST
49	cloven	cleave	PPER
50	clung	cling	PAST,PPER
51	could	can	PAST
52	crept	creep	PAST,PPER
53	dealt	deal	PAST,PPER
54	did	do	PAST
55	done	do	PPER
56	dove	dive	PAST
57	drank	drink	PAST
58	drawn	draw	PPER

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Word Order	Irregular Verb	Entry	Tense Marker
59	dreamt	dream	PAST,PPER
60	drew	draw	PAST
61	driven	drive	PPER
62	drove	drive	PAST
63	drunk	drink	PPER
64	dug	dig	PAST,PPER
65	dwelt	dwell	PAST,PPER
66	eaten	eat	PPER
67	fallen	fall	PPER
68	fed	feed	PAST,PPER
69	fell	fall	PAST
70	felt	feel	PAST,PPER
71	fled	flew	PAST,PPER
72	flew	fly	PAST
73	flown	fly	PPER
74	flung	fling	PAST,PPER
75	forbade	forbid	PAST
76	forbidden	forbid	PPER
77	forbore	forbear	PAST
78	forborne	forbear	PPER
79	foresaw	foresee	PAST
80	foreseen	foresee	PPER
81	foretold	foretell	PAST,PPER
82	forgave	forgive	PAST
83	forgiven	forgive	PPER
84	forgone	forgo	PPER
85	forgot	forget	PAST
86	forgotten	forget	PPER
87	forsaken	forsake	PPER
88	forsook	forsake	PAST

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Word Order	Irregular Verb	Entry	Tense Marker
89	forswore	forswear	PAST
90	forsworn	forswear	PPER
91	forwent	forgo	PAST
92	fought	fight	PAST,PPER
93	found	find	PAST,PPER
94	froze	freeze	PAST
95	frozen	freeze	PPER
96	gainsaid	gainsay	PAST,PPER
97	gave	give	PAST
98	given	give	PPER
99	gone	go	PPER
100	got	get	PAST
101	gotten	get	PPER
102	grew	grow	PAST
103	ground	grind	PAST,PPER
104	grown	grow	PPER
105	had	have	PAST,PPER
106	hamstrung	hamstring	PAST,PPER
107	heard	hear	PAST,PPER
108	held	hold	PAST,PPER
109	hewn	hew	PPER
110	hid	hide	PAST
111	hidden	hide	PPER
112	hove	heave	PAST,PPER
113	hung	hang	PAST,PPER
114	kept	keep	PAST,PPER
115	knelt	knell	PAST,PPER
116	knew	know	PAST
117	known	know	PPER
118	laid	lay	PAST,PPER

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Word Order	Irregular Verb	Entry	Tense Marker
119	lain	lie	PPER
120	lay	lie	PAST
121	leant	lean	PAST,PPER
122	leapt	leap	PAST,PPER
123	learnt	learn	PAST,PPER
124	led	lead	PAST,PPER
125	left	leave	PAST,PPER
126	lent	lend	PAST,PPER
127	lit	light	PAST,PPER
128	lost	lose	PAST,PPER
129	made	make	PAST,PPER
130	meant	mean	PAST,PPER
131	met	meet	PAST,PPER
132	mislaid	mislay	PAST,PPER
133	misled	mislead	PAST,PPER
134	misspelt	misspell	PAST,PPER
135	misspent	misspend	PAST,PPER
136	mistaken	mistake	PPER
137	mistooke	mistake	PAST
138	misunderstood	misunderstand	PAST,PPER
139	mown	mow	PPER
140	outdid	outdo	PAST
141	outdone	outdo	PPER
142	outgrew	outgrow	PAST
143	outgrown	outgrow	PPER
144	outran	outrun	PAST
145	outshone	outshine	PAST,PPER
146	overbore	overbear	PAST
147	oberborne	overbear	PPER

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Word Order	Irregular Verb	Entry	Tense Marker
148	overcame	overcome	PAST
149	overdid	overdo	PAST
150	overdone	overdo	PPER
151	overhung	overhang	PAST,PPER
152	overran	overrun	PAST
153	oversaw	oversee	PAST
154	overseen	oversee	PPER
155	overslept	oversleep	PAST,PPER
156	overtaken	overtake	PPER
157	overtook	overtake	PAST
158	paid	pay	PAST,PPER
159	ran	run	PAST
160	rang	ring	PAST
161	rebound	rebind	PAST,PPER
162	rebuilt	rebuild	PAST,PPER
163	redid	redo	PAST
164	redone	redo	PPER
165	relaid	relay	PAST,PPER
166	remade	remake	PAST,PPER
167	rent	rend	PAST,PPER
168	repaid	repay	PAST,PPER
169	reran	rerun	PAST
170	retold	retell	PAST,PPER
171	rewound	rewind	PAST,PPER
172	rewritten	rewrite	PPER
173	rewrote	rewrite	PAST
174	ridden	ride	PPER
175	risen	rise	PPER
176	rode	ride	PAST
177	rose	rise	PAST
178	rung	ring	PPER

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Word Order	Irregular Verb	Entry	Tense Marker
179	said	say	PAST,PPER
180	sang	sing	PAST
181	sank	sink	PAST
182	sat	sit	PAST,PPER
183	saw	see	PAST
184	sawn	saw	PPER
185	seen	see	PPER
186	sent	send	PAST,PPER
187	sewn	sew	PPER
188	shaken	shake	PPER
189	shod	shoe	PAST,PPER
190	shone	shine	PAST,PPER
191	shook	shake	PAST
192	shorn	shear	PPER
193	shot	shoot	PAST,PPER
194	should	shall	PAST
195	shown	show	PPER
196	shrank	shrink	PAST
197	shrunk	shrink	PPER
198	slain	slay	PPER
199	slept	sleep	PAST,PPER
200	slew	slay	PAST
201	slid	slide	PAST,PPER
202	slung	sling	PAST,PPER
203	slunk	slink	PAST,PPER
204	smelt	smell	PAST,PPER
205	smitten	smite	PAST
206	smote	smite	PPER
207	sold	sell	PAST,PPER
208	sought	seek	PAST,PPER

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Word Order	Irregular Verb	Entry	Tense Marker
210	sown	sow	PPER
211	span	spin	PAST
212	spat	spit	PAST,PPER
213	sped	speed	PAST,PPER
214	spelt	spell	PAST,PPER
215	spent	spend	PAST,PPER
216	spilt	spill	PAST,PPER
217	spoilt	spoil	PAST,PPER
218	spoke	speak	PAST
219	spoken	speak	PPER
220	sprang	spring	PAST
221	sprung	spring	PPER
222	spun	spin	PPER
223	stank	stink	PAST
224	stole	steal	PAST
225	stolen	steal	PPER
226	stood	stand	PAST,PPER
227	strewn	strew	PPER
228	striddn	stride	PPER
229	striven	strive	PPER
230	strode	stride	PAST
231	strove	strive	PAST
232	struck	strike	PAST,PPER
233	strung	string	PAST,PPER
234	stuck	stick	PAST,PPER
235	stung	sting	PAST,PPER
236	stunk	stink	PPER
237	sung	sing	PPER
238	sunk	sink	PPER
239	swam	swim	PAST
240	swept	sweep	PAST,PPER
241	swollen	swell	PPER

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Word Order	Irregular Verb	Entry	Tense Marker
242	swore	swear	PAST
243	sworn	swear	PPER
244	swum	swim	PPER
245	swung	swing	PAST,PPER
246	taken	take	PPER
247	taught	teach	PAST,PPER
248	thought	think	PAST,PPER
249	threw	throw	PAST
250	throve	thrive	PAST
251	thrown	throw	PPER
252	told	tell	PAST,PPER
253	took	take	PAST
254	tore	tear	PAST
255	torn	tear	PPER
256	trod	tread	PAST
257	trodden	tread	PPER
258	unbent	unbend	PAST,PPER
259	unbound	unbind	PAST,PPER
260	undergone	undergo	PPER
261	understood	understand	PAST,PPER
262	undertaken	undertake	PPER
263	undertook	undertake	PAST
264	underwent	undergo	PAST
265	undid	undo	PAST
266	undone	undo	PPER
267	unwound	unwind	PAST,PPER
268	upheld	uphold	PAST,PPER
269	was	be	PAST
270	waylaid	waylay	PAST,PPER

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Word Order	Irregular Verb	Entry	Tense Marker
271	went	go	PAST
272	wept	weep	PAST,PPER
273	were	be	PAST
274	withdrawn	withdraw	PPER
275	withdrew	withdraw	PAST
276	withheld	withhold	PAST,PPER
277	withstood	withstand	PAST,PPER
278	woke	wake	PAST
279	woken	wake	PPER
280	won	win	PAST,PPER
281	wore	wear	PAST
282	worn	wear	PPER
283	would	will	PAST
284	wound	wind	PAST,PPER
285	wove	weave	PAST
286	woven	weave	PPER
287	wrote	write	PAST
288	wrung	wring	PAST,PPER

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง

[1] A.B. Tucker, "Current Strategies in Machine Translation Research and Development ", in Machine Translation Theoritical and Methodological Issues, Cambridge University Press, London, 1987,pp.22-44.

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

[3] ห้องปฏิบัติการปัญญาประดิษฐ์ สถาบันเทคโนโลยีพระจอมเกล้า ธนบุรี, "รายงานการวิจัยและพัฒนาปัญญาประดิษฐ์" 2532 ,หน้า 42-43.

[4] สะรัช บุณยรัตพันธุ์, "An Easy Guide to VERB PATTERNS" บริษัทโรงพิมพ์ไทยวัฒนาพานิช พิมพ์ครั้งที่ 2 พ.ศ. 2527, 317 หน้า.

[5] Harold E. Palmer, D.Litt., "A Grammar of English Words" ,Lowe and Brydone(Printers) limited, London, 300 pages.

[6] อรัญญา ปรีชาไว, ผศ. ดร. รัตติกร วรากูลศิริพันธุ์ และนาย สมศักดิ์ จันวัน, "การสังเคราะห์ภาษาไทยจากภาษากลางในระบบการแปลหลายหลายภาษา" วารสารวิจัยและพัฒนา สถาบันเทคโนโลยีพระจอมเกล้า ธนบุรี ปีที่ 12 ฉบับที่ 2, ธันวาคม 2532, หน้า 1-18.

[7] ยืน ภู่วรวรรณ, ชัยยงค์ วงศ์ชัยสุวัฒน์, "การประมวลผลภาษาธรรมชาติ(Natural Language Processing)" สถาบันเทคโนโลยีพระจอมเกล้า ธนบุรี พิมพ์ครั้งที่ 1, 2535, 241 หน้า.

[8] มนทิรา ประชาฤทธิ์ภักดี, "การพัฒนารูปแบบภาษากลางเพื่อการแปลภาษาอังกฤษเป็นภาษาไทย"(วิทยานิพนธ์ปริญญามหาบัณฑิต, คณะศิลปศาสตร์ มหาวิทยาลัยธรรมศาสตร์), 2533.

[9] อรัญญา ปรีชาไว และ ผศ. ดร. รัตติกร วรากูลศิริพันธุ์, "การวิเคราะห์โครงสร้างประโยคภาษาอังกฤษเพื่อการแปลภาษาด้วยคอมพิวเตอร์" วารสารคอมพิวเตอร์ สมาคมคอมพิวเตอร์แห่งประเทศไทย ปีที่ 19 ฉบับที่ 97, พฤษภาคม - มิถุนายน 2535, หน้า 14-21.

[10] Nancy E. Miller, "File Structures Using Pascal", The Benjamin/Cummings Publishing Company,Inc. 1987 ,487 pages.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

ชื่อผู้เขียน	นางสาวอรัญญา ปรีชาไว
วันเดือนปีเกิด	วันที่ 29 ธันวาคม พ.ศ. 2509
สถานที่เกิด	จังหวัดสุพรรณบุรี
วุฒิการศึกษาระดับปริญญาตรี	เกียรตินิยม อันดับ 2 วิทยาศาสตรบัณฑิต สาขาฟิสิกส์
สถานที่สำเร็จการศึกษา	มหาวิทยาลัยศรีนครินทรวิโรฒ บางแสน
ปีที่สำเร็จการศึกษา	ปีการศึกษา 2531
ผลงานทางวิชาการที่ได้รับการตีพิมพ์	เรื่องการสังเคราะห์ภาษาไทยจากภาษากลาง ในระบบการแปลหลายภาษา
อาชีพปัจจุบัน	รับราชการตำแหน่งเจ้าหน้าที่ระบบงานคอมพิวเตอร์ 3 สำนักวิจัยและบริการคอมพิวเตอร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้