

การศึกษาข้อความในคลังข้อมูลการเกษตรสำหรับประยุกต์ใช้กับการประมวลผลคำตอบ
ด้านการเกษตรของระบบถาม-ตอบอัตโนมัติ*

A Study of Discourse in an Agricultural Corpus towards an
Automated Machine for Question-Answering Processing

กรกฤษ มีมงคล**

พุทธชาติ โพธิบาล***

อุมารัตน์ ลังขมาน****

บทคัดย่อ

งานวิจัยนี้มุ่งศึกษารูปแบบโครงสร้างของบทความด้านการเกษตรและการเชื่อมโยงความ รวมทั้งชุดคำบ่งชี้เพื่อใช้ในการตอบคำถามสำหรับระบบถาม-ตอบภาษาไทย ข้อมูลที่ใช้ในการศึกษาวิเคราะห์นำมาจากบทความในคลังข้อมูลการเกษตรแห่งชาติจำนวนทั้งสิ้น 95 บทความ การวิเคราะห์ข้อมูลประยุกต์ใช้ทฤษฎีโครงสร้าง ความของ Halliday and Hasan (1985) และการเชื่อมโยง ความของ Halliday and Hasan (1976)

ผลการศึกษาพบว่าโครงสร้างของบทความด้านการเกษตรประกอบด้วย 7 องค์ประกอบ คือ (1) หัวเรื่อง (2) เกริ่นนำ (3) ส่วนประกอบ (4) วิธีทำ (5) การนำไปใช้ (6) ประโยชน์และ (7) การจบ จำนวนรูปแบบโครงสร้างมี

17 รูปแบบ การเชื่อมโยงความ ประกอบด้วยการอ้างอิงแบบปรากฏรูป แบบไม่ปรากฏรูป และการใช้คำเชื่อมส่วนชุดคำบ่งชี้เพื่อใช้ในการสกัดคำตอบมีจำนวน 5 ชุด แบ่งตามประเภทของคำถาม 5 ประเภท ประกอบด้วยคำถามประเภทถามวิธีการ เวลา สถานที่ บุคคล และการแจกแจง ทั้งนี้ผลการศึกษารูปแบบโครงสร้างการเชื่อมโยงความชุดคำบ่งชี้เหล่านี้มีความสัมพันธ์กับโครงสร้างของบทความและการเชื่อมโยงความอันจะเป็นประโยชน์ต่อผู้พัฒนาระบบถาม-ตอบอัตโนมัติ

คำสำคัญ: ข้อความ / คลังข้อมูล / การประมวลผลของระบบถาม-ตอบอัตโนมัติ

* ส่วนหนึ่งของวิทยานิพนธ์ระดับดุษฎีบัณฑิต สาขาภาษาศาสตร์ประยุกต์ มหาวิทยาลัยเกษตรศาสตร์

** นิสิตระดับปริญญาเอก สาขาภาษาศาสตร์ประยุกต์มหาวิทยาลัยเกษตรศาสตร์

***ผู้ช่วยศาสตราจารย์ อาจารย์ที่ปรึกษาหลัก ภาควิชาภาษาศาสตร์ คณะมนุษยศาสตร์ มหาวิทยาลัยเกษตรศาสตร์

**** ผู้ช่วยศาสตราจารย์ อาจารย์ที่ปรึกษาร่วม ภาควิชาภาษาศาสตร์ คณะมนุษยศาสตร์ มหาวิทยาลัยเกษตรศาสตร์

Abstract

The purposes of this research were to study the structure of agricultural academic articles, cohesion and sets of index for answering questions in the Thai-language question answering system. The samples comprised of 95 agricultural academic articles from the agricultural national corpus. The analysis is based upon the adaptation of Generic Structural Potential theory by Halliday and Hasan (1985) and Cohesion theory by Halliday and Hasan (1976).

The findings reveal that the agricultural academic articles are composed of 7 generic elements, namely 1) heading 2) introduction 3) component 4) procedure 5) implementation 6)

benefit and 7) ending. Moreover, these articles have 17 patterns. Cohesion consists of reference, ellipsis and connector. For the sets of index, there are 5 sets of index divided by 5 patterns of questions. They include questions for procedure, questions for time, questions for places, questions for persons and questions for list. The results of the research on the structure of agricultural academic articles, cohesion and sets of index are related to the structure of agricultural academic articles and cohesion. The results will be beneficial for question answering system developers in the future.

Keywords: Discourse / Corpus / Question and answering processing

1. บทนำ

ระบบประมวลภาษาด้วยคอมพิวเตอร์ คือระบบที่ทำให้คอมพิวเตอร์เข้าใจภาษามนุษย์ได้ใกล้เคียงกับมนุษย์และเป็นธรรมชาติมากที่สุด ระบบดังกล่าวเป็นส่วนหนึ่งของงานวิจัยด้านปัญญาประดิษฐ์ที่พยายามจำลองความฉลาดของมนุษย์ให้กับคอมพิวเตอร์ ดังนั้นการศึกษาเพื่อพัฒนาระบบประมวลภาษาด้วยคอมพิวเตอร์จึงมีเป้าหมายเพื่อให้คอมพิวเตอร์เข้าใจภาษามนุษย์และโต้ตอบกับผู้ใช้ระบบได้ (อัศนีย์ ก่อตระกูล และคณะ, 2541)

ระบบถาม-ตอบอัตโนมัติเป็นระบบประมวลผลภาษาด้วยคอมพิวเตอร์ประเภทหนึ่งโดยมีเป้าหมายเพื่อให้คอมพิวเตอร์สามารถโต้ตอบกับมนุษย์ได้โดยมีการทำงานของระบบสกัดคำตอบจากฐานข้อมูลทางภาษาขนาดใหญ่ อัศนีย์ ก่อตระกูลและคณะ (2541) กล่าวว่าระบบดังกล่าวเป็นเครื่องมือสำหรับค้นหาข้อมูลต่างๆ ที่ผู้ใช้ระบบต้องการทราบ โดยผู้ใช้ระบบพิมพ์คำถามเข้าสู่ระบบ และระบบจะแสดงคำตอบอัตโนมัติ ดังนั้นการพัฒนาระบบให้ตอบคำถามได้ใกล้เคียงกับการสื่อสารกับมนุษย์ได้จริงจึงเป็นสิ่งที่นักพัฒนาระบบถาม-ตอบอัตโนมัติให้ความสำคัญเป็นอย่างยิ่ง

ในปัจจุบันระบบถาม-ตอบอัตโนมัติยังคงมีปัญหาหลายประการ โดยเฉพาะอย่างยิ่งปัญหาในการสกัดคำตอบ ทั้งนี้เพราะระบบไม่มีข้อมูลตำแหน่งของประโยคคำตอบในบทความ ระบบไม่สามารถกำหนดขอบเขตของประโยคได้ หรือในภาษาไทยมีการไม่ปรากฏรูปคำจึงทำให้คำตอบที่สกัดได้อาจมีความกำกวมดังเช่นงานวิจัยของสิงห์พิย สุขสว่างโรจน์ (2557) ราชวิทย์ ทิพย์เสนา และคณะ (2557) สุพจน์ บัวเล็งและวรารัตน์ สงฆ์แป้น (2557) พบว่าระบบถาม-ตอบอัตโนมัติยังมีปัญหาการสกัดคำตอบได้ไม่ตรงกับรูปแบบคำถาม ซึ่ง Kawtrakul (2002) ได้กล่าวถึงปัญหาเหล่านี้ว่าเกิดจากโครงสร้างของเอกสารและการเชื่อมโยงความภาษาไทยมีความซับซ้อน ดังนั้นการศึกษาวเคราะห์ในประเด็นดังกล่าวอาจจะมีส่วนช่วยในการแก้ปัญหาต่างๆ เหล่านี้ได้

ด้วยเหตุผลดังกล่าวข้างต้นผู้วิจัยจึงจะศึกษาโครงสร้างบทความ การเชื่อมโยงความและชุดคำบ่งชี้จากบทความด้านการเกษตรเพื่อประโยชน์ในการประยุกต์ใช้สกัดข้อมูลสำหรับระบบถาม-ตอบอัตโนมัติ เนื่องจากอาชีพด้านเกษตรกรรมเป็นอาชีพที่มีการสืบค้นบทความด้านการเกษตรจำนวนมากงาน วิจัยนี้จึงได้เก็บข้อมูล

จากบทความด้านการเกษตรจำนวน 95 บทความ ที่คัดเลือกมาจากบทความจาก 2,000 บทความ โดยใช้เกณฑ์การคำนวณหาจำนวนกลุ่มตัวอย่างของ Yamane (1973) การวิเคราะห์ข้อมูลใช้กรอบแนวคิดของ Halliday and Hasan (1985) และ Halliday and Hasan (1976)

2) วรรณกรรมปริทัศน์

ทฤษฎีและงานวิจัยที่เกี่ยวข้องกับการศึกษานี้ ได้แก่ การศึกษาโครงสร้างบทความ การศึกษาการเชื่อมโยงความ และการศึกษาระบบถาม-ตอบอัตโนมัติมีรายละเอียดดังต่อไปนี้

2.1 การศึกษาโครงสร้างบทความ

ผู้วิจัยเลือกศึกษาโครงสร้างบทความโดยใช้กรอบแนวคิดเชิงโครงสร้างของ Halliday and Hasan (1985) ซึ่งมีหลักทฤษฎี คือ เอกสารทุกเอกสารจะต้องมีโครงสร้าง และสามารถเขียนเป็นกฎได้โดยแสดงลำดับและส่วนประชิดขององค์ประกอบภายในเอกสาร รวมทั้งยังบอกได้ว่าองค์ประกอบใดเป็นองค์ประกอบบังคับและองค์ประกอบใดเป็นองค์ประกอบไม่บังคับ โดยทฤษฎีนี้มีมุมมองว่าการทราบข้อมูลโครงสร้างของภาษาจะมีส่วนในการพยากรณ์เหตุการณ์ที่จะเกิดขึ้นในอนาคตมีผลทำให้สามารถวางแผนจัดการกับสถานการณ์ที่จะเกิดขึ้นได้อย่างมีประสิทธิภาพ เช่น โครงสร้างการสนทนาในการซื้อขายสินค้า ทำให้ทราบว่าถ้าผู้ซื้อได้ต่อรองราคาแล้ว เหตุการณ์ต่อไปที่อาจเกิดขึ้น คือ การตกลงซื้อหรือปฏิเสธการซื้อสินค้า

งานวิจัยโดยใช้ทฤษฎีดังกล่าวในการศึกษามีหลายงานวิจัยด้วยกัน เช่น Patpong (2009) ศึกษากฎโครงสร้างนิทานพื้นบ้านภาษาตระกูลออสโตรเอเชียติก Patpong (2011) ศึกษากฎโครงสร้างนิทานพื้นบ้านไทยทรงดำ และ Sungkaman (2016) ศึกษากฎโครงสร้างเรื่องย่อละครเกาหลี ผลการศึกษาที่ได้ทำให้เห็นว่าทฤษฎีนี้ใช้ในการศึกษาวิเคราะห์กฎของเอกสารได้เป็นอย่างดีแม้ว่าเอกสารจะมีโครงสร้างซับซ้อนเป็นอย่างมาก และสามารถใช้ได้กับทั้งเอกสารที่เป็นทางการและไม่เป็นทางการ ผู้วิจัยคาดว่าการศึกษาโครงสร้างบทความจะสามารถนำไปประยุกต์ใช้ในการพัฒนาระบบถาม-ตอบอัตโนมัติและเนื่องจากข้อมูลกฎโครงสร้างของเอกสารด้านการเกษตรจะเป็นข้อมูลโครงสร้างของเอกสารทำให้สามารถจำแนกได้

ว่าเอกสารใดเป็นเอกสารที่มีเนื้อหาสมบูรณ์หรือไม่สมบูรณ์ รวมทั้งจะทำให้ทราบได้ว่าเนื้อหาในองค์ประกอบใดสามารถใช้สำหรับเป็นคำตอบของคำถามใด

2.2 การศึกษาการเชื่อมโยงความ

กรอบแนวคิดการเชื่อมโยงความของ Halliday and Hasan (1976) มีแนวคิดที่ว่าเมื่อประโยคที่เรียงต่อกันจะเชื่อมโยงกันด้วยกลไกการเชื่อมโยงความเพื่อให้ภาษาเกิดความเป็นเอกภาพ โดยการเชื่อมโยงความมี 5 ประเภท ได้แก่ 1) การอ้างอิง คือ การแสดงความสัมพันธ์ทางความหมายระหว่างคำสองคำโดยใช้รูปภาษา 2) การใช้รูปแทน คือ การเชื่อมโยงความโดยการแสดงความสัมพันธ์ของคำโดยใช้รูปแทน 3) การไม่ปรากฏรูปคำ คือ การเชื่อมโยงความโดยการแสดงความสัมพันธ์แบบไม่ใช้รูปภาษา 4) การใช้คำเชื่อม คือ การเชื่อมโยงความโดยการแสดงความสัมพันธ์โดยใช้คำเชื่อม และ 5) การใช้คำศัพท์ คือ การแสดงความสัมพันธ์โดยใช้คำศัพท์

งานวิจัยที่ศึกษาการเชื่อมโยงความจำแนกประเภทของการศึกษาเอกสารได้ดังนี้ ประเภทแรกการเชื่อมโยงความในภาษาถิ่น เช่น การศึกษาของทัศนีย์ เจริญพร (2534) ศึกษาลักษณะการเชื่อมโยงความในศิลาจารึกอีสานระหว่างพุทธศักราช 1893-2466 นิตยสุตา อภินันทภรณ์ (2543) ศึกษาหน่วยเชื่อมโยงปริจเฉทภาษาไทยตั้งแต่สมัยสุโขทัยจนถึงปัจจุบันวิยะดา ตานี (2544) ศึกษาการเชื่อมโยงความในจารึกสุโขทัย วรวรรธน์ ศรียาภัย (2544) ศึกษาการเชื่อมโยงความในภาษาถิ่นสุราษฎร์ธานี ชญานนิตย์ จินดารักษ์ (2546) ศึกษาการเชื่อมโยงความในจารึกล้านนา ฤทธิ์ดำรง มงคลสรรพ (2547) ศึกษาการเชื่อมโยงความในภาษาถิ่นเหนือ อำเภอแจ้ห่ม จังหวัดลำปาง พรศิริ ตันติจิตจารุ (2548) ศึกษาการเชื่อมโยงความในหมายรับสั่งสมัยกรุงธนบุรีและสมัยรัชกาลที่ 1 ประเภทที่สองการเชื่อมโยงความในภาษากฎหมาย เช่น การศึกษาของสุวิมล สดสะอาด (2557) ศึกษาวัจนกรรมและการเชื่อมโยงความในกฎหมายรัฐธรรมนูญ และประเภทที่สามการเชื่อมโยงความที่เกี่ยวข้องกับประเด็นทางภาษาศาสตร์สังคม เช่น การศึกษาของเทพี จรัสจรัสเกียรติ (2543) ที่ได้ศึกษาลักษณะการเชื่อมโยงในการรายงานข่าวอาชญากรรมของหนังสือพิมพ์ไทยรัฐ: การศึกษาภาษาระดับข้อความ และชลธิชา บำรุงรักษ์ (2539) ศึกษาการวิเคราะห์ภาษาระดับ

ข้อความประเภทต่างๆ ในภาษาไทยโดยได้ศึกษาประเด็นทางภาษาศาสตร์สังคมที่เกี่ยวข้องกับการเชื่อมโยงความในภาษาไทย

การศึกษาข้างต้นทำให้เห็นว่าทฤษฎีการเชื่อมโยงความของ Halliday and Hasan (1976) สามารถนำไปใช้ศึกษาวิเคราะห์การเชื่อมโยงความได้ในเอกสารหลากหลายประเภท อาทิการศึกษาเอกสารภาษาถิ่น เอกสารภาษากฎหมาย รวมถึงการศึกษาการเชื่อมโยงความในเชิงภาษาศาสตร์สังคม ด้วยเหตุนี้ผู้วิจัยจึงเลือกใช้ทฤษฎีของ Halliday and Hasan (1976) มาศึกษาวิเคราะห์การเชื่อมโยงความในเอกสารด้านการเกษตร เนื่องจากเอกสารด้านการเกษตรเป็นเอกสารอีกประเภทหนึ่งที่มีรูปแบบการเขียนเฉพาะตัวและมีการเชื่อมโยงความที่เกี่ยวข้องกับประเด็นทางสังคมด้วยเช่นกัน

2.3 การศึกษาระบบถาม-ตอบอัตโนมัติ

Kawtrakul (2002) กล่าวว่าระบบถาม-ตอบอัตโนมัติภาษาไทย คือเครื่องมือหนึ่งที่ใช้ในการประมวลผลภาษาไทยด้วยคอมพิวเตอร์มีวัตถุประสงค์เพื่อตอบคำถามที่ป้อนเข้าสู่ระบบ หลักการพื้นฐานของระบบถามตอบอัตโนมัติ คือการสกัดข้อความประโยคที่เป็นคำตอบออกมาจากฐานข้อมูลขนาดใหญ่ ยุทธศักดิ์ วุฒิกุล และคณะ (2553) กล่าวถึงคำจำกัดความของระบบถามตอบอัตโนมัติว่าเป็นระบบประเภทหนึ่งที่สามารถสนทนาโต้ตอบกับมนุษย์ได้โดยผ่านการพิมพ์ข้อความ การพัฒนาระบบถาม-ตอบอัตโนมัติในภาษาไทยในปัจจุบันศึกษาโดยประยุกต์ใช้ทั้งทฤษฎีด้านโครงสร้างและความหมายเพื่อสร้างคู่คำถามคำตอบอัตโนมัติซึ่งพัฒนามาจากการตอบคำถามในยุคแรกที่จะพัฒนารูปแบบคำตอบของคำถามประเภทถามข้อเท็จจริงเท่านั้น โดยขั้นตอนส่วนหนึ่งของการศึกษา คือการตัดคำ การกำกับชนิดของคำ การกำกับวลีและการกำกับมโนทัศน์ทางความหมาย

ปัจจุบันระบบถาม-ตอบภาษาไทยยังคงมีปัญหาในการประมวลผลข้อมูลแบ่งเป็น 3 ประเด็นหลัก กล่าวคือ ประเด็นแรก คลังข้อมูลมีข้อมูลไม่เพียงพอสำหรับตอบคำถาม Deepa Yogishet al (2016) ได้อธิบายวิธีการเตรียมข้อมูลเพื่อเป็นคำตอบของระบบว่า ผู้พัฒนาระบบจะต้องสร้างรูปแบบของประโยคคำตอบขึ้นมาและเรียงลำดับความเป็นไปได้ในการใช้ประโยคเหล่านั้นในการ

ตอบคำถาม รูปแบบประโยคจะสร้างจากข้อมูลในเอกสารที่ใช้เป็นคลังข้อมูล

ประเด็นที่สอง คือ ความหมายของคำตอบคำถามและการระบุขอบเขตของประโยคคำตอบไม่ถูกต้อง เช่น การศึกษาของ Castagnola (2002) ที่ได้ศึกษาระบบการอ้างอิงเพื่อระบบถาม-ตอบอัตโนมัติกล่าวว่าความกำกวมของข้อความที่ระบบสกัดออกมาได้เกิดจากการอ้างอิงที่ไม่มียู่ในกฎการประมวลผลของระบบ และความกำกวมที่เกิดจากการอ้างอิงเพื่อแสดงบุคคลในรูปประธานและกรรม จึงสร้างกฎด้านการอ้างอิงเพิ่มเติมให้แก่ระบบโดยนำข้อมูลจากเอกสารมาวิเคราะห์ งานวิจัยดังกล่าวนี้เป็นการศึกษาการอ้างอิงในภาษาอังกฤษซึ่งคำสรรพนามที่ใช้อ้างอิงจะเปลี่ยนรูปเมื่อเป็นประธานและเป็นกรรมของประโยค อย่างไรก็ตามในภาษาไทยนั้นคำสรรพนามที่ใช้อ้างอิงไม่สามารถเปลี่ยนรูปไปตามหน้าที่เมื่ออยู่ในประโยคแต่จะเกิดการไม่ปรากฏรูปของคำ ดังนั้นการวิเคราะห์การเชื่อมโยงความจะเป็นข้อมูลให้ผู้พัฒนาระบบนำไปสร้างกฎสำหรับแสดงผลคำที่ไม่ปรากฏรูปและการอ้างอิงได้ต่อไป ส่วนการระบุขอบเขตของประโยคคำตอบจำเป็นต้องศึกษาคำบ่งชี้จากในเอกสารให้แน่ชัด ดังเช่นที่จิรวรรณ เจริญสุข (2549) ศึกษาคำบ่งชี้ขอบเขตของประโยคเพื่อการประมวลผลภาษาไทยด้วยคอมพิวเตอร์ในงานวิจัยเรื่องการแบ่งขอบเขตอนุภาคปริจเฉท (Elementary discourse unit) ในภาษาไทยโดยใช้คำระบุหน่วย (Discourse segmentation cues) และข้อสนเทศเชิงวากยสัมพันธ์ (Syntactic information)

ประเด็นที่สาม คือ ระบบตอบไม่ตรงคำถาม ดังจะเห็นได้จากที่ กมลวรรณ โพธิ์สาย (2552) ที่ศึกษาวิเคราะห์หาความหมายแฝงและการเรียนรู้เครื่องจักรสำหรับระบบถามตอบอัตโนมัติภาษาไทยงานวิจัยนี้ผลการศึกษาพบว่าประโยคคำถามในภาษาไทยมีการใช้คำที่หลากหลาย เช่น ประโยคคำถามประเภท เท่าไหร่ (How many) พบการใช้คำว่า เท่าไหร่ ก็ เหตุผลดังกล่าวทำให้ระบบสกัดคำตอบได้ผิดพลาดวัตถุประสงค์ของคำถามสอดคล้องกับงานของฉวีวรรณ เพ็ชรศิริ (2553) ที่กล่าวว่ารูปแบบคำถามของผู้ใช้ระบบจะถามว่า อะไร ที่ไหน ทำไม อย่างไร เมื่อไหร่ เท่าไหร่ สิ่งไหน แต่ในภาษาไทยอาจเกิดปัญหาคำถามมีความหมายกำกวมได้ เช่น การถามด้วย

คำถาม "อย่างไร" ในบางกรณีหมายถึง "อะไร" ดังในประโยค "ใบโหระพามีสรรพคุณทางยาอย่างไร" แท้จริงแล้วผู้สืบค้นข้อมูลต้องการคำตอบว่า "สรรพคุณทางยาของใบโหระพาคืออะไร" ดังนั้นการสร้างคำบ่งชี้สำหรับคำตอบแต่ละประเภทจึงมีความสำคัญต่อการเลือกนำคำตอบไปตอบให้ตรงกับคำถาม

การศึกษาข้างต้นทำให้เห็นว่าในบทความต่างๆ จะมีการเชื่อมโยงความกันอยู่ซึ่งการเชื่อมโยงความเหล่านี้น่าจะเป็นข้อมูลที่เป็นประโยชน์สำหรับนำไปประยุกต์ใช้กับระบบถาม-ตอบอัตโนมัติและทำให้มองเห็นว่ายังมีปัญหาที่อาจแก้ไขได้โดยการนำแนวคิดทางภาษาศาสตร์เข้าไปช่วยในการวิเคราะห์ข้อมูล

3) ผลการศึกษา

ผลการศึกษาประกอบด้วย 3 หัวข้อ ได้แก่ 3.1 โครงสร้างของบทความด้านการเกษตร 3.2 การเชื่อมโยงความในบทความด้านการเกษตร และ 3.3 ชุดคำบ่งชี้ในบทความด้านการเกษตรรายละเอียดมีดังนี้

3.1 โครงสร้างของบทความด้านการเกษตร

จากการวิเคราะห์ข้อมูลพบว่ามีองค์ประกอบ 7 องค์ประกอบ ได้แก่ หัวเรื่อง เกริ่นนำ ส่วนประกอบ วิธีทำ การนำไปใช้ ประโยชน์และการจบ ซึ่งสามารถเขียนกฎโครงสร้างบทความได้ ดังนี้

หัวเรื่อง $\wedge$ (เกริ่นนำ) $\wedge$ (ส่วนประกอบ) $\wedge$ วิธีทำ $\wedge$ [(การนำไปใช้) $\circ$ (ประโยชน์)] $\wedge$ (การจบ)

จากกฎโครงสร้างข้างต้นจะเห็นว่าปรากฏความหมายและสัญลักษณ์ที่ใช้ในกฎโครงสร้าง ดังนี้

1. องค์ประกอบที่ไม่มีวงเล็บครอบกำกับ หมายถึง องค์ประกอบบังคับ
2. () หมายถึง องค์ประกอบไม่บังคับ
3. $\circ$ หมายถึง การสลับที่ได้
4. [] หมายถึง ขอบเขตของการสลับที่
5. $\wedge$ หมายถึง การต่อเนื่องกัน

กฎโครงสร้างข้างต้นสรุปจากกฎโครงสร้างจำนวน 17 โครงสร้างดังนี้

1. หัวเรื่อง $\wedge$ ส่วนประกอบ $\wedge$ วิธีทำ $\wedge$ การนำไปใช้ $\wedge$ ประโยชน์

2. หัวเรื่อง $\wedge$ เกริ่นนำ $\wedge$ ส่วนประกอบ $\wedge$ วิธีทำ $\wedge$ การนำไปใช้ $\wedge$ ประโยชน์

3. หัวเรื่อง $\wedge$ วิธีทำ

4. หัวเรื่อง $\wedge$ เกริ่นนำ $\wedge$ ส่วนประกอบ $\wedge$ วิธีทำ $\wedge$ การนำไปใช้

5. หัวเรื่อง $\wedge$ ส่วนประกอบ $\wedge$ วิธีทำ $\wedge$ ประโยชน์

6. หัวเรื่อง $\wedge$ เกริ่นนำ $\wedge$ วิธีทำ

7. หัวเรื่อง $\wedge$ เกริ่นนำ $\wedge$ วิธีทำ $\wedge$ การนำไปใช้

8. หัวเรื่อง $\wedge$ ส่วนประกอบ $\wedge$ วิธีทำ $\wedge$ การนำไปใช้

9. หัวเรื่อง $\wedge$ เกริ่นนำ $\wedge$ ส่วนประกอบ $\wedge$ วิธีทำ

10. หัวเรื่อง $\wedge$ เกริ่นนำ $\wedge$ วิธีทำ $\wedge$ การนำไปใช้ $\wedge$ ประโยชน์

11. หัวเรื่อง $\wedge$ วิธีทำ $\wedge$ การนำไปใช้ $\wedge$ ประโยชน์

12. หัวเรื่อง $\wedge$ เกริ่นนำ $\wedge$ วิธีทำ $\wedge$ ประโยชน์

13. หัวเรื่อง $\wedge$ ส่วนประกอบ $\wedge$ วิธีทำ

14. หัวเรื่อง $\wedge$ เกริ่นนำ $\wedge$ ส่วนประกอบ $\wedge$ วิธีทำ $\wedge$ ประโยชน์

$\wedge$ การจบ

15. หัวเรื่อง $\wedge$ เกริ่นนำ $\wedge$ ส่วนประกอบ $\wedge$ วิธีทำ $\wedge$ ประโยชน์ $\wedge$ การนำไปใช้

16. หัวเรื่อง $\wedge$ เกริ่นนำ $\wedge$ วิธีทำ $\wedge$ การจบ

17. หัวเรื่อง $\wedge$ วิธีทำ $\wedge$ การนำไปใช้

จากรูปแบบบทความทั้ง 17 รูปแบบข้างต้นนี้ จะเห็นได้ว่าการปรากฏองค์ประกอบ 2 ประเภท ได้แก่ องค์ประกอบบังคับและองค์ประกอบไม่บังคับ โดยองค์ประกอบบังคับหมายถึงองค์ประกอบที่ต้องปรากฏในบทความเสมอ จากข้อมูลพบว่าองค์ประกอบบังคับมี 2 องค์ประกอบ ได้แก่ หัวเรื่อง และวิธีทำ ส่วนองค์ประกอบไม่บังคับหมายถึงองค์ประกอบที่อาจปรากฏหรือไม่ปรากฏในบทความก็ได้ ผู้เขียนมีเกณฑ์ในการพิจารณา คือ องค์ประกอบไม่บังคับอาจจะปรากฏในบทความที่ศึกษาบางบทความเท่านั้น ไม่ได้ปรากฏในบทความที่ศึกษาทั้งหมด องค์ประกอบไม่บังคับมี 5 องค์ประกอบ ได้แก่ เกริ่นนำ ส่วนประกอบ การนำไปใช้ ประโยชน์และการจบ

นอกจากเรื่ององค์ประกอบบังคับและองค์ประกอบไม่บังคับแล้ว ยังมีองค์ประกอบในลักษณะอื่นๆ ลักษณะแรกคือการสลับที่ขององค์ประกอบ ดังจะเห็นได้จากโครงสร้างรูปแบบที่ 10 และ 15 กล่าวคือ โครงสร้างรูปแบบที่ 10 องค์ประกอบ การนำไปใช้ อยู่หน้าองค์ประกอบ ประโยชน์ แต่โครงสร้างรูปแบบที่ 15 องค์ประกอบ ประโยชน์ อยู่หน้าองค์ประกอบ การนำไปใช้ ดังนั้นในการสรุป

กฎจึงใช้เครื่องหมาย [] เพื่อแสดงขอบเขตการสลับที่ และใช้เครื่องหมาย o เพื่อแสดงว่าองค์ประกอบทั้งสองดังกล่าวข้างต้นสามารถสลับที่กันได้ ลักษณะที่สอง คือการประชิดกันขององค์ประกอบ ซึ่งสามารถพิจารณาได้จากโครงสร้างรูปแบบที่ 1, 2, 10, และ 11 กล่าวคือองค์ประกอบหัวเรื่องปรากฏข้างหน้าองค์ประกอบเกริ่นนำเสมอ องค์ประกอบเกริ่นนำปรากฏข้างหน้าองค์ประกอบส่วนประกอบเสมอ องค์ประกอบส่วนประกอบปรากฏข้างหน้าองค์ประกอบวิธีทำเสมอ องค์ประกอบวิธีทำปรากฏข้างหน้าองค์ประกอบการนำไปใช้หรือองค์ประกอบประโยชน์ก็ได้ เนื่องจากองค์ประกอบทั้งสองนี้สามารถสลับที่กันได้ และองค์ประกอบการจบจะเป็นองค์ประกอบสุดท้ายเสมอ ดังนั้นองค์ประกอบที่ประชิดกันเหล่านี้จึงต้องกำกับด้วยเครื่องหมาย ^

3.2 การเชื่อมโยงความในบทความด้านการเกษตร

จากข้อมูลพบการเชื่อมโยงความ 2 ประเภท ได้แก่ การเชื่อมโยงความด้วยการอ้างอิงโดยใช้รูปแทนและการใช้คำเชื่อม ทั้งนี้การเชื่อมโยงความด้วยการอ้างอิงสามารถแบ่งเป็นการอ้างอิงปรากฏรูปและการอ้างอิงไม่ปรากฏรูป โดยการอ้างอิงทั้งสองประเภทสามารถปรากฏได้ทั้งตำแหน่งประธานและกรรม การอ้างอิงโดยใช้รูปแทนปรากฏรูปแบ่งเป็น 6 ประเภท ได้แก่ การอ้างอิงโดยใช้คำบอกจำนวนการอ้างอิงโดยใช้คำยกย่องให้เกียรติ การอ้างอิงโดยใช้คำบอกเพศการอ้างอิงโดยใช้คำเรียกญาติ การอ้างอิงโดยใช้คำสรรพนาม และการอ้างอิงโดยใช้คำสรรพนามแทนตนเอง

การอ้างอิงโดยใช้คำเชื่อมหมายถึงการนำคำเชื่อมมาแสดงความสัมพันธ์ของคำทั้งสองคำ แบ่งเป็นการอ้างอิงปรากฏรูปคำเชื่อมและการอ้างอิงไม่ปรากฏรูปคำเชื่อม การอ้างอิงทั้งสองรูปแบบแบ่งเป็นสองลักษณะตามตำแหน่งที่ปรากฏ ได้แก่ การปรากฏรูปและการไม่ปรากฏรูปคำเชื่อมที่ใช้เชื่อมระหว่างประโยคและการปรากฏรูปและไม่ปรากฏรูปคำเชื่อมที่ใช้เชื่อมระหว่างประโยคกับประโยคย่อย โดยการอ้างอิงปรากฏรูปคำเชื่อมแบ่งเป็น 9 ประเภท ได้แก่ การเชื่อมแบบคล้ายตาม ได้แก่ และ แล้ว จากนั้น ก็ การเชื่อมแบบขัดแย้ง ได้แก่ แต่การเชื่อมแบบแสดงทางเลือก ได้แก่ หรือ การเชื่อมแบบแสดงเหตุผล ได้แก่ เพราะ เนื่องจาก การเชื่อมแบบแสดงเงื่อนไข ได้แก่

ถ้า หาก เมื่อ การเชื่อมแบบขยายความ ได้แก่ โดย ที่ ลำคัญ ซึ่ง การเชื่อมแบบแจกแจง ได้แก่ ดังนี้ คือ การเชื่อมแบบตัดตอนหรือแยกส่วน ได้แก่ ส่วนและการเชื่อมแบบแสดงวัตถุประสงค์ ได้แก่ เพื่อ

การศึกษาการเชื่อมโยงความในงานวิจัยนี้มีจุดประสงค์เพื่อศึกษาแต่เพียงว่าบทความด้านการเกษตรมีการเชื่อมโยงความอย่างไร ส่วนการนำผลการศึกษาไปประยุกต์ใช้ ผู้พัฒนาระบบอาจจะนำข้อมูลไปประยุกต์ใช้ให้เหมาะสมกับระบบของตนต่อไป

3.3 ชุดคำบ่งชี้ในบทความด้านการเกษตร

ชุดคำบ่งชี้ หมายถึง ชุดคำที่ใช้เป็นคำบ่งบอกสำหรับการสกัดคำตอบให้ถูกต้องตรงกับรูปแบบคำถามแบ่งเป็น 6 ประเภท ได้แก่ คำถามเกี่ยวกับ วิธีการ เวลา สถานที่ บุคคล และการแจกแจง โดยการพิจารณาคำบ่งชี้จะพิจารณาว่าคำใดเป็นชุดคำบ่งชี้สำหรับตอบคำถามประเภทดังกล่าวผลการศึกษารายงานหนึ่งมีดังต่อไปนี้

1. ชุดคำบ่งชี้ในการตอบคำถามเกี่ยวกับวิธีการ เป็นคำบ่งชี้ที่ใช้เป็นคำบ่งบอกสำหรับการสกัดคำตอบให้ถูกต้องและตรงกับคำถามที่ถามเกี่ยวกับวิธีการมีจำนวนทั้งสิ้น 3 คำ ได้แก่ วิธีทำ วิธีการ และขั้นตอน โดยคำตอบเกี่ยวกับวิธีการจะสัมพันธ์กับองค์ประกอบวิธีทำ
2. ชุดคำบ่งชี้ในการตอบคำถามเกี่ยวกับเวลาเป็นคำบ่งชี้ที่ใช้เป็นคำบ่งบอกสำหรับการสกัดคำตอบให้ถูกต้องและตรงกับคำถามที่ถามเกี่ยวกับเวลาที่มีจำนวนทั้งสิ้น 6 คำ ได้แก่ ช่วง ตอน ใน จนกระทั่ง ถึง ตั้งแต่ ประมาณ
3. ชุดคำบ่งชี้ในการตอบคำถามเกี่ยวกับสถานที่ เป็นคำบ่งชี้ที่ใช้เป็นคำบ่งบอกสำหรับการสกัดคำตอบให้ถูกต้องและตรงกับคำถามที่ถามเกี่ยวกับสถานที่มี 7 คำ ได้แก่ ใน ภายใน จาก รอบ อำเภอบ ตำบล และจังหวัด
4. ชุดคำบ่งชี้ในการตอบคำถามเกี่ยวกับบุคคล เป็นคำบ่งชี้ที่ใช้เป็นคำบ่งบอกสำหรับการสกัดคำตอบให้ถูกต้องและตรงกับคำถามที่ถามเกี่ยวกับบุคคลมี 7 คำ ได้แก่ คุณ ลุง ป้า ผู้ใหญ่ นาย รศ.ดร. พี่
5. ชุดคำบ่งชี้ในการตอบคำถามเกี่ยวกับการแจกแจง เป็นคำบ่งชี้ที่ใช้เป็นคำบ่งบอกสำหรับการสกัดคำตอบให้ถูกต้องและตรงกับคำถามที่ถามเกี่ยวกับการแจกแจงประกอบด้วย การเว้นวรรคและตัวเลข
6. การแจกแจงข้อเท็จจริงเป็นคำบ่งชี้ที่ใช้เป็น

คำบ่งบอกสำหรับการสกัดคำตอบให้ถูกต้องและตรงกับคำถามที่ถามเกี่ยวกับข้อเท็จจริงโดยระบบถาม-ตอบอัตโนมัติด้านการศึกษาด้านวิศวกรรมใช้วิธีเรียกว่า การค้นหาความรู้ในฐานข้อมูลเอกสาร วิธีการดังกล่าวเป็นวิธีการค้นหาข้อความที่เป็นคำตอบในเอกสารจำนวนมหาศาลโดยใช้ความรู้จากแม่แบบคำถามคำตอบในระบบ โดยแม่แบบคำถามคำตอบในระบบสร้างมาจากความสัมพันธ์ทางโครงสร้างประโยคและความหมายเชิงโมโนทัศน์ของคำถามซึ่งใน

ด้านวิศวกรรมเรียกว่ากระบวนการสร้างตัวแทนเชิงความหมายในรูปแบบมโนภาพ

4. สรุปและการอภิปรายผล

จากการวิเคราะห์บทความทางเทคโนโลยี 3 ลักษณะ ได้แก่ โครงสร้างบทความ การเชื่อมโยงความ และชุดคำบ่งชี้ ผู้วิจัยจะนำเสนอว่าสามารถนำข้อมูลที่ได้วิเคราะห์ไปใช้ได้อย่างไรโดยพิจารณาจากตัวอย่างต่อไปนี้

หัวเรื่อง	สูตรน้ำหมักผักบุงบำรุงดิน
เกริ่นนำ	นายอุทัย แก่นสากล ปัจจุบันอยู่บ้านเลขที่ 110 ม.5 ต.เริงราง อ.เสาไห้ จ. สระบุรี พื้นที่การเกษตร 50 ไร่แบ่งเป็นนาข้าว 40 ไร่ บ่อน้ำและแปลงเกษตรสาธิต ประมาณ 3-4 ไร่ คุณอุทัยทำการเกษตรปลอดสารพิษ ซึ่งปุ๋ยต่างๆที่ใช้ 01 ทำเองทั้งหมด เช่น ผลิตภัณฑ์ชีวภาพเพื่อไล่แมลงหรือบำรุงให้เจริญเติบโต
ส่วนประกอบ	ส่วนประกอบ 1. ผักบุง จำนวน 9 กิโลกรัม 2. กากน้ำตาล จำนวน 3 กิโลกรัม 3. พด 7 จำนวน 1 ซอง 4. น้ำสะอาดจำนวน 10 ลิตร
วิธีทำ	02 นำผักบุงมาหั่นให้ยาวประมาณ 2-3 นิ้ว แล้วนำไปลงในถังหมัก และละลายสารเร่งพด7 ในน้ำสะอาด 04 เท 05 ลงไปในถังหมัก ตามด้วยกากน้ำตาล 06 คนให้เข้ากันแล้วหมักทิ้งไว้ประมาณ 3 เดือน ก็สามารถนำมาใช้ในการปรับปรุงบำรุงดินในนาข้าวได้
การนำไปใช้	วิธีการใช้นั้นในเรื่องการปรับปรุงบำรุงดินในนาข้าวพื้นที่เคยใช้สารเคมีมาก่อนก็ให้น้ำหมักผักบุงช่วยในการปรับปรุงบำรุงดิน โดยใช้น้ำหมักผักบุง จำนวน 5 ลิตร ต่อพื้นที่ 1 ไร่ 07 ปล่อย 08 ไปตามน้ำใช้ก่อนการไถดิน
ประโยชน์	เพื่อปรับสภาพให้ดินนั้นอ่อนนุ่มบำรุงดินให้มีจุลินทรีย์มากขึ้น

จากตัวอย่างข้างต้นจะเห็นว่าปรากฏสัญลักษณ์ ดังนี้

1. △ หมายถึง คำที่ใช้เป็นคำบ่งชี้สำหรับการสกัดคำตอบของระบบ
2. □ หมายถึง คำที่ใช้ในการเชื่อมโยงความ
3. ∅ หมายถึง การไม่ปรากฏรูปคำ
4. △ และเครื่องหมาย □ ซ้อนกัน หมายถึง คำที่เป็นทั้งคำบ่งบอกสำหรับการสกัด คำตอบของระบบและคำที่ใช้ในการเชื่อมโยงความ

จากตัวอย่างข้างต้นเขียนเป็นกฎโครงสร้างได้ดังนี้

หัวเรื่อง(เกริ่นนำ) (ส่วนประกอบ) วิธีทำ(การนำไปใช้) (ประโยชน์)

จะเห็นได้ว่าบทความมีองค์ประกอบ 6 องค์ประกอบ ประกอบด้วย "องค์ประกอบหัวเรื่อง คือ สูตรน้ำหมักผักบุงบำรุงดิน" "องค์ประกอบเกริ่นนำ คือ นายอุทัย แก่นสากล...เจริญเติบโต" "องค์ประกอบส่วนประกอบ คือ ส่วนประกอบ...น้ำสะอาดจำนวน 10 ลิตร"

"องค์ประกอบวิธีทำ คือ นำผักบุ้งมาหั่นให้ยาวประมาณ 2-3 นิ้ว...ในนาข้าวได้" "องค์ประกอบการนำไปใช้ คือ วิธี การใช้...ก่อนการตีดิน" และ "องค์ประกอบประโยชน์ คือ เพื่อปรับสภาพ...มีจุลินทรีย์มากขึ้น"

จากตัวอย่างข้างต้นพบการเชื่อมโยงความ 2 ประเภท ได้แก่ การอ้างอิงโดยใช้รูปแทนและการอ้างอิงโดยใช้คำเชื่อม ทั้งนี้การอ้างอิงโดยใช้รูปแทนปรากฏเฉพาะการอ้างอิงไม่ปรากฏรูป ส่วนการอ้างอิงโดยใช้คำเชื่อมปรากฏทั้งการอ้างอิงปรากฏรูปและการอ้างอิงไม่ปรากฏรูป การเชื่อมโยงความประเภทแรกที่พบในตัวอย่าง คือ การอ้างอิงโดยใช้รูปแทนแบบไม่ปรากฏรูปจำนวน 5 คำ คำที่หนึ่งคือ "คุณอุทัย" ในประโยค "คุณอุทัยทำการเกษตรปลอดสารพิษ ซึ่งปุ๋ยต่างๆที่ใช้[๑]ทำเองทั้งหมด" คำที่สองคือ "คุณ" ในประโยค [๒]นำผักบุ้งมาหั่นให้ยาวประมาณ 2-3 นิ้ว คำที่สามคือ "ผักบุ้ง" ในประโยค "นำผักบุ้งมาหั่นให้ยาวประมาณ 2-3 นิ้ว แล้วก็นำ[๓]ลงไปในถังหมัก" คำที่สี่คือ "น้ำสะอาด" ในประโยค "ละลายสารเร่งพด 7 ในน้ำสะอาด เท [๔]ลงไปในถังหมัก" และคำที่ห้า คือ "ปล่อย" ในประโยค "ปล่อย[๕]ไปตามน้ำใช้ก่อนการตีดิน"

การเชื่อมโยงความที่พบในตัวอย่างประเภทที่สอง คือ การอ้างอิงโดยใช้คำเชื่อมแบบปรากฏรูป มีจำนวน 8 คำ ได้แก่ "และ" ประโยคตัวอย่าง เช่น "พื้นที่การเกษตร 50 ไร่ แบ่งเป็นนาข้าว 40 ไร่ บ่อน้ำ [และ] แปลงเกษตรสาธิต ประมาณ 3-4 ไร่" "ซึ่ง" ประโยคตัวอย่าง เช่น "คุณอุทัยทำการเกษตรปลอดสารพิษซึ่งปุ๋ยต่างๆ ที่ใช้ทำเองทั้งหมด" "เพื่อ" ประโยคตัวอย่าง เช่น "ผลิตจุลินทรีย์ชีวภาพ[เพื่อ]ไล่แมลง" "หรือ" ประโยคตัวอย่าง เช่น "[หรือ]บำรุงให้เจริญเติบโต" "แล้วก็" ประโยคตัวอย่าง เช่น "นำผักบุ้งมาหั่นให้ยาวประมาณ 2-3 นิ้ว [แล้วก็]นำไปในถังหมัก" "ตามด้วย" ประโยคตัวอย่าง เช่น "[ตามด้วย]กากน้ำตาลคนให้เข้ากันแล้ว" "แล้ว" ประโยคตัวอย่าง เช่น "[แล้ว]หมักทิ้งไว้" "โดย" ประโยคตัวอย่าง เช่น "น้ำหมักผักบุ้งช่วยในการปรับปรุงบำรุงดิน[โดย]ใช้น้ำหมักผักบุ้ง จำนวน 5 ลิตร ต่อพื้นที่ 1 ไร่"

การเชื่อมโยงความที่พบในตัวอย่างประเภทที่สาม คือ การอ้างอิงโดยใช้คำเชื่อมแบบไม่ปรากฏรูป จำนวน 3 คำ ได้แก่ คำที่ ๐4 ในประโยค "ละลายสารเร่งพด 7 ในน้ำสะอาด[๐4] เทลงไปในถังหมัก" คำต่อไป คือ คำที่ ๐6

ในประโยค "ตามด้วยกากน้ำตาล ๐6 คนให้เข้ากัน" และ คำสุดท้ายคือคำที่ ๐7 ในประโยค "ใช้น้ำหมักผักบุ้ง จำนวน 5 ลิตร ต่อพื้นที่ 1 ไร่ ๐7 ปล่อยไปตามน้ำใช้ก่อนการตีดิน" โดยคำเชื่อมที่พบในตัวอย่างข้างต้นทั้ง 3 คำเป็นคำเชื่อมไม่ปรากฏรูปแบบแสดงลำดับวิธีการ

ตัวอย่างข้างต้นยังพบว่ามีการใช้ชุดคำบ่งชี้ในการตอบคำถาม 6 ประเภท ดังนี้

1. การตอบคำถามประเภทถามวิธีการเช่นคำถาม "นำหมักผักบุ้งบำรุงดินอย่างไร" คำบ่งชี้ ได้แก่  "แล้ว"  "แล้ว" และ  "ก็"
2. การตอบคำถามประเภทถามเวลา เช่นคำถาม "นำหมักผักบุ้งบำรุงดินต้องหมักทิ้งไว้นานเท่าใด" คำบ่งชี้ ได้แก่  "ประมาณ" คำตอบ ได้แก่ "ประมาณ 3 เดือน"
3. การตอบคำถามประเภทสถานที่เช่นคำถาม "ควรใช้น้ำหมักผักบุ้งบำรุงดินที่ไหน" คำบ่งชี้ ได้แก่  "ใน" คำตอบ ได้แก่ "ในนาข้าวพื้นที่เคยใช้สารเคมีมาก่อนก็ให้ใช้"
4. การตอบคำถามประเภทถามบุคคลเช่นคำถาม "สูตรนำหมักผักบุ้งบำรุงดินเป็นสูตรของใคร" คำบ่งชี้ ได้แก่  "นาย" คำตอบ ได้แก่ "นายอุทัย แก่นสากล"
5. การตอบคำถามประเภทแจกแจงเช่น คำถาม "ส่วนประกอบสูตรนำหมักผักบุ้งบำรุงดินคืออะไร" คำบ่งชี้ ได้แก่  "ส่วนประกอบ" และ  "1, 2, 3, 4" คำตอบ ได้แก่ "1. ผักบุ้ง จำนวน 9 กิโลกรัม 2. กากน้ำตาล จำนวน 3 กิโลกรัม 3. พด 7 จำนวน 1 ชอง 4. น้ำสะอาดจำนวน 10 ลิตร"
6. การตอบคำถามประเภทถามหาข้อเท็จจริง เช่นคำถาม "นำหมักผักบุ้งใช้ทำอะไร" ระบบจะวิเคราะห์ประโยคคำถามทั้งด้านโครงสร้างและด้านความหมายเพื่อเทียบหาคำถามคำตอบให้ตรงกับกฎที่ระบบมีอยู่ในฐานข้อมูลและประมวลผลออกมาเป็นคำตอบ คือ "นำหมักผักบุ้งใช้บำรุงดิน"

ผลการศึกษาทำให้ทราบว่าโครงสร้างความของเอกสารด้านการเกษตรมีรูปแบบเฉพาะ ดังจะเห็นได้จากผลการศึกษาในครั้งนี้แตกต่างจากผลการศึกษาของงานวิจัยอื่นๆที่ศึกษาโครงสร้างความดังเช่นการศึกษาของ Danilo (2008) Wee (2009) Inya (2010) Olagunju (2015) ด้านการศึกษาการเชื่อมโยงความ ผลการศึกษาทำให้ทราบว่าการเชื่อมโยงความของบทความด้านการเกษตร

มีปัจจัยทางสังคมมาเกี่ยวข้อง เช่น การใช้คำเรียกญาติ และ ผลการศึกษาคำบ่งชี้อาจเป็นข้อมูลที่มีประโยชน์กับการพัฒนาระบบถาม-ตอบอัตโนมัติ เนื่องจากการศึกษาคำบ่งชี้ในการตอบคำถามเป็นการศึกษาอีกรูปแบบหนึ่งที่ไม่ได้เน้นการแปลงข้อมูลทางภาษาให้กลายเป็นกฎ

การศึกษาในอนาคตควรต่อยอดศึกษารูปแบบคำถามในภาษาไทยด้านความหมายแฝงซึ่งประเด็น

ดังกล่าวนี้สอดคล้องกับที่ กมลวรรณ โพธิ์สาย (2552) จีรวรรณ เพ็ชรศิริ (2553) สุพจน์ บัวเล็ง และวราภรณ์ สงฆ์แป้น (2557) ที่ได้กล่าวหาว่าควรมีการศึกษาวเคราะห์ รูปแบบและความหมายของคำถามอย่างละเอียดเพื่อนำมาใช้เป็นข้อมูลในการศึกษาเพื่อพัฒนาระบบถาม-ตอบอัตโนมัติต่อไป

เอกสารอ้างอิง

- กมลวรรณ โพธิ์สาย. (2552). การวิเคราะห์หาความหมายแฝงและการเรียนรู้เครื่องจักรสำหรับระบบถามตอบอัตโนมัติภาษาไทย. วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยธรรมศาสตร์.
- จีรวรรณ เจริญสุข. (2549). การแบ่งขอบเขตอนุพาทย์ปริเฉทในภาษาไทยโดยใช้คำระบุนัย และข้อสันเทศเชิงวากยสัมพันธ์. วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาวิศวกรรมคอมพิวเตอร์ มหาวิทยาลัยเกษตรศาสตร์.
- จีรวรรณ เพ็ชรศิริ. (2553). การสกัดความรู้เกี่ยวกับสรรพคุณทางยาของพืชสมุนไพรไทยจากเอกสารภาษาไทย เพื่อสนับสนุนการตอบคำถามอัตโนมัติ. รายงานการวิจัยมหาวิทยาลัยธุรกิจบัณฑิต.
- ชฎานุกตย์ จินดารักษ์. (2546). การเชื่อมโยงความในจารึกล้านนา. วิทยานิพนธ์ศิลปศาสตรมหาบัณฑิต สาขาจารึกภาษาไทย. มหาวิทยาลัยศิลปากร.
- พรศิริ ตันติจิตจากร. (2548). การเชื่อมโยงความในหมายรับสั่งสมัยกรุงธนบุรีและสมัยรัชกาลที่ 1. วิทยานิพนธ์ศิลปศาสตรมหาบัณฑิต สาขาจารึกภาษาไทย มหาวิทยาลัยศิลปากร.
- ยุทธศักดิ์ วุฒิกุล และคณะ. (2553). การขยายฐานความรู้แบบอัตโนมัติสำหรับระบบถาม-ตอบโดยใช้ความหมายของประโยค. วารสารรวมบทความการประชุมวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 6. 940-945.
- ราชวิทย์ ทิพย์เสนา และคณะ. (2557). การจำแนกกลุ่มคำถามอัตโนมัติบนกระดานสนทนา โดยใช้เทคนิคเหมืองข้อความ. วารสารวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยมหาสารคาม. 33(5), 493-502.
- ฤทธิ์ดำรง มงคลสรรพ. (2547). การเชื่อมโยงความในภาษาถิ่นเหนือ อำเภอแจ้ห่ม จังหวัดลำปาง. วิทยานิพนธ์ศิลปศาสตรมหาบัณฑิต สาขาจารึกภาษาไทย มหาวิทยาลัยศิลปากร.
- วรวรรณ ศรียาภัย. (2544). การเชื่อมโยงความในภาษาถิ่นสุราษฎร์ธานี. วิทยานิพนธ์ศิลปศาสตรมหาบัณฑิต สาขาจารึกภาษาไทย มหาวิทยาลัยศิลปากร.
- วิยะดา ตานี. (2544). การเชื่อมโยงความในจารึกสุโขทัย. วิทยานิพนธ์ศิลปศาสตรมหาบัณฑิต สาขาจารึกภาษาไทย มหาวิทยาลัยศิลปากร.
- สิงห์ทัย สุขสว่างโรจน์. (2557). ระบบถาม-ตอบภาษาไทยเพื่อสนับสนุนการตัดสินใจของนักศึกษามหาวิทยาลัยรามคำแหง. สารนิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาระบบสนับสนุนการตัดสินใจ มหาวิทยาลัยรามคำแหง.
- สุชาติา เจียพงษ์. (2553). การศึกษาเปรียบเทียบคำบุรุษสรรพนามในภาษาไทย 4 ถิ่น. วิทยานิพนธ์อักษรศาสตร์ดุสิตบัณฑิต สาขาภาษาไทย คณะอักษรศาสตร์ มหาวิทยาลัยศิลปากร.

- สุพจน์ บัวเลิง และวราวัตร สงษ์แป้น. (2557). การแยกกลุ่มคำถามเพื่อค้นหาคำตอบโดยใช้เว็บเชิงความหมายและการทำเหมืองข้อมูล. *วารสารรวมบทความการประชุมวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 10*. 101-108.
- อัศนีย์ ก่อตระกูลและคณะ. (2541). การประมวลผลภาษาธรรมชาติด้วยคอมพิวเตอร์. *วารสารมนุษยศาสตร์*. 6(1), 1-12.
- Castagnola. (2002). *Anaphora Resolution for Question Answering*. Master of Engineering in Electrical Engineering and Computer Science. Massachusetts Institute of Technology
- Cooper, R. J. and S. M. Ruger. (2000). "A Simple Question Answering System". In *Proceedings of the 9th Text Retrieval Conference (TREC'00)*. 251-257.
- Danilo, T.D. 2008. "The Discourse of Print Advertising in the Philippines: Generic Structures and Linguistic Features". In *Proceeding of 22nd Pacific Asia Conference on Language, Information and Computation (PACLIC)*. 1-15.
- Deepa Yogish et al. 2016. A Survey of Intelligent Question Answering System Using NLP and Information Retrieval Techniques. *International Journal of Advanced Research in Computer and Communication Engineering*. 5(5), 536-540
- Halliday, M.A.K and Hasan.R. (1985). *Language, Context, and Text: Aspects of Language in a Social-Semiotic Perspective*. Oxford University Press.
- _____. (1976). *Cohesion in English*. London, Longman.
- Inya, O. (2010). *Discourse Structure and Pragmatic Acts in Christian Apologetic Discourse*. Master of Arts Thesis. Federal University of Technology, Akure.
- Kawtrakul, A. et al. (2002). "A State of the Art of Thai Language Resources and Thai Language Behavior Analysis and Modeling". In *Proceedings of the International Conference on Natural Language Processing*. Post COLING 2002 Workshop. Taipei, Taiwan. 1-8.
- Olagunju, S. 2015. Generic Structure Potential of Football Matches in Newspaper Reporting. *IOSR Journal of Humanities and Social Science (IOSR-JHSS)*. 20(3), 38-50.
- Patpong, P. (2009). "An Analysis of Generic Structure Potential of Some Selected Austroasiatic Folktales: A Preliminary Report". *The 4th International Conference on Austroasiatic Linguistics (ICAAL 4)*, Research Institute for Languages and Cultures of Asia Mahidol University Thailand
- Patpong, P. (2011). "A Generic Structure Potential Analysis of A Generic Structure Potential Analysis of Thai Song Dam Folktales Thai Song Dam Folktales". *Presented at 21st Annual Conference of Southeast Asian Linguistics Society*. Kasetsart University Thailand.
- Sungkaman. (2016). "The Study of Korean Series Synopsis Structure: Discourse Analysis". In *Proceeding of the 4th International Joint Conference on Korean Studies and Thai Studies*. 184-191.
- Srihari, R. and Li, W. (2010). "A Question Answering System Supported by Information Extraction". In *Proceedings of the Sixth Conference on Applied Natural Language Processing*. 166-172.

- Wee, S.C. (2009). *Newspaper reporting of SARS in Singapore: A Systemic Functional Approach*. Ph.D.dissertation, National University of Singapore.
- Yamane, T. (1973). *Statistic: An Introductory Analysis*. 3rd ed. New York , Harper and Row.

