บทที่ 1

บทนำ

ที่มาของปัญหา

"ในปี ค.ศ. 2021 มนุษย์จะเลิกใช้แป้นพิมพ์อย่างเด็ดขาด คีย์บอร์ดจะเหลือไว้ให้ลูกหลานดูเท่านั้น ระบบสั่งงานด้วยคำพูดจะใช้กันอย่างแพร่หลาย คอมพิวเตอร์สามารถเข้าใจภาษามนุษย์ เพียงเราอ่านหนังสือให้เครื่องฟัง เครื่องก็จะพิมพ์ตามเราได้" (ยืน ภู่วรวรรณ, 2536)

นิยายเรื่องนี้ เป็นเพียงนิยายวิทยาศาสตร์หรือจะเป็นเรื่องที่เกิดขึ้นในเวลานั้น ยังไม่มีผู้ใดจะทราบได้ ซึ่งถ้าเป็นเรื่องจริงจะเป็นสิ่งที่มีประโยชน์อย่างมหาศาล

ความต้องการให้เครื่องคอมพิวเตอร์สามารถเรียนรู้ถึงวิธีการในการรู้จำคำพูดได้ ทำให้เกิดศาสตร์แขนงหนึ่งขึ้นเรียกว่า Speech Recognition หรือ Speech-to-text conversion ซึ่งเกิดขึ้นมา 40 กว่าปีมาแล้ว (Vaissiere, 1983) และได้มีการพัฒนามาเรื่อย ๆ แต่การพัฒนาค่อนข้างล่าช้า เนื่องจากปัญหาทางด้านต่าง ๆ เช่น คำพูดที่เปล่งเสียงออกมาจะมีลักษณะเฉพาะตัวเอง และแปรไปตามปริบทของการพูดเสมอ แม้ว่าเราจะพยายามเปล่งคำพูดให้มีเสียงใกล้เคียงกันเท่าไหร่ก็ตาม ก็จะพบถึงความแตกต่างที่เกิดขึ้นได้อย่างชัดเจน (Bunge, 1977)

เราสามารถแบ่งระบบการรู้จำคำพูดออกได้เป็น 2 วิธี คือ (วีระ ริ้วพิทักษ์, 2529 อ้างถึงใน ไพศาล ธรรมโพธิทอง, 2533)

1. การรู้จำโดยใช้ลักษณะทางเสียงของคำที่ประกอบเป็นคำพูดนั้น ๆ (Word Recognition) หน่วยภาษาที่พิจารณาในวิธีนี้ อาจจะเป็นพยางค์ คำ กลุ่มคำ วลี หรือประโยค ก็ได้ วิธีนี้มีผลดี คือ จะหลีกเลี่ยงผลกระทบอันเกิดจากฐานของเสียงภายในคำ หรือกลุ่มคำนั้น

2. การรู้จำโดยแยกแยะรายละเอียดทางสัทศาสตร์ของเสียงที่เป็นคำพูด (Phonetic Recognition) ในวิธีนี้จะพิจารณาลักษณะของหน่วยเสียงที่มีขนาดเล็กลงไป เช่น ระดับพยัญชนะ สระ และวรรณยุกต์ เป็นต้น โดยจะใช้หน่วยเสียงย่อยเหล่านี้เป็นหลักในการเปรียบเทียบ วิธีนี้จะมีความสะดวกสำหรับระบบที่จะรู้จักคำพูดที่มีจำนวนคำมาก ๆ

จากงานวิจัยที่ผ่านมา จะเป็นลักษณะการรู้จำโดยพิจารณาหน่วยภาษาที่เปล่งออกมาทั้งชุด มีทั้งระบบการรู้จำคำเดี่ยว (Isolated word recognition) และระบบรู้จำคำพูดต่อเนื่อง (Continuous word recognition) ระบบเหล่านี้ต่างก็มีประโยชน์ทั้งสิ้น เช่น นำไปใช้ในการสั่งงานหุ่นยนต์ เครื่องมือเครื่องใช้ต่าง ๆ ด้วยเสียง ข้อดี คือ ง่าย และไม่จำกัดภาษาที่พูด (เป็นภาษาอะไรก็ได้ แต่จำเป็นต้องมีการสอนให้ระบบ จำคำพูดเหล่านั้นเสียก่อน) แต่ข้อเสียคือ จำนวนคำที่

ใช้ต้องมีจำนวนจำกัด ใช้เนื้อที่มากในการจัดเก็บแบบอ้างอิง และใช้เวลามากในการคำนวณและเปรียบเทียบ

จากข้อจำกัดเหล่านี้ ผู้วิจัยจึงมีแนวความคิดที่จะหันมาศึกษาวิจัยการรู้จำโดยแยกแยะรายละเอียดของหน่วยเสียง เพื่อลดเนื้อที่ในการจัดเก็บแบบอ้างอิง ลดเวลาที่ใช้ในการคำนวณเปรียบเทียบ อันจะช่วยให้สามารถพัฒนาไปสู่การรู้จำคำพูดระบบเปิด (Open system) ซึ่งไม่จำกัดจำนวนคำต่อไป ระบบดังกล่าวเป็นเรื่องที่ทำได้ยาก ต้องอาศัยความรู้ทางภาษาศาสตร์มาเกี่ยวข้องอย่างมาก และระบบจะสร้างขึ้นเฉพาะสำหรับภาษาใดภาษาหนึ่ง (ถ้าจะนำไปใช้กับภาษาอื่นจะต้องมีการปรับปรุงแก้ไข)

วัตถุประสงค์

1. เพื่อพัฒนาระบบการรู้จำคำพูดภาษาไทยที่เป็นคำเดี่ยว โดยใช้หน่วยเสียงเป็นหลัก
2. เพื่อแก้ปัญหาข้อจำกัดต่าง ๆ ของระบบการรู้จำในอดีต (การรู้จำโดยพิจารณาหน่วยเสียงที่เปล่งออกมาทั้งชุด) เช่น ลดเนื้อที่หน่วยความจำในการเก็บแบบอ้างอิง และลดเวลาในการเปรียบเทียบกับแบบอ้างอิง เป็นต้น
3. เพื่อนำความรู้ทางด้านภาษาศาสตร์มาใช้กับงานด้านการรู้จำคำพูด
4. เพื่อเป็นแนวทางในการพัฒนาไปสู่ระบบการรู้จำคำพูดภาษาไทยได้ทั้งภาษา

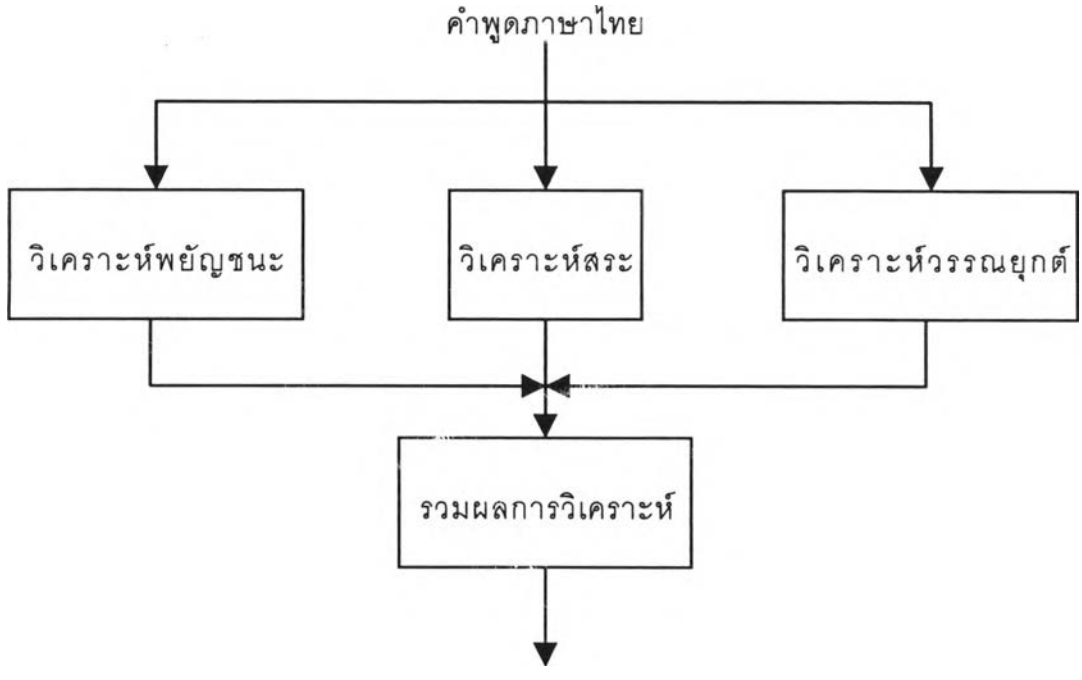
แนวความคิด

พยางค์ประกอบด้วยหน่วยเสียงพยัญชนะ หน่วยเสียงสระ และหน่วยเสียงวรรณยุกต์ การรู้จำคำพูดโดยการแยกพิจารณาแต่ละหน่วยเสียงและนำผลที่ได้มารวมกัน จะเป็นแนวทางที่จะนำไปสู่การรู้จำคำพูดได้ทั้งภาษา

การแยกหน่วยเสียงต่าง ๆ ออกจากกัน จะใช้ลักษณะบ่งความต่าง (Distinctive features) โดยพิจารณาลักษณะเฉพาะบางส่วนที่ต่างกันของหน่วยเสียงต่าง ๆ และใช้ลักษณะต่างนั้นในการแบ่งแยก

จากแนวความคิดนี้ จะมีกระบวนการประมวลผลดังรูปที่ 1.1

รูปที่ 1.1 การประมวลผลการรู้จำคำพูดโดยแยกวิเคราะห์หน่วยเสียง



ขอบเขตงานวิจัย

1. การวิจัยนี้มุ่งศึกษาพัฒนาบนเครื่องไมโครคอมพิวเตอร์ โดยมีอุปกรณ์เพิ่มเติม ได้แก่ การ์ดเสียง (Sound Card) ไมโครโฟน และลำโพง เป็นต้น

2. การวิจัยนี้ได้เลือก ศึกษาเฉพาะหน่วยเสียงในภาษาไทยบางหน่วยเสียงดังนี้

ก) หน่วยเสียงพยัญชนะ 10 หน่วยเสียง คือกลุ่มพยัญชนะกัก (Stop Consonants) 5 หน่วยเสียง ได้แก่ /p/, /ph/, /d/, /c/ และ /ch/ กลุ่มพยัญชนะไม่กัก (Non-stop Consonants) 5 หน่วยเสียง ได้แก่ /n/, /f/, /s/, /r/ และ /w/

ข) หน่วยเสียงสระ 6 หน่วยเสียง คือ /i/, /ii/, /a/, /aa/, /u/ และ /uu/ (หน่วยเสียงสระ อิ อี อะ อา อุ และ อู ตามลำดับ)

ค) หน่วยเสียงวรรณยุกต์ 5 หน่วยเสียง คือ สามัญ เอก โท ตรี และจัตวา

ขั้นตอนการวิจัย

1. ศึกษาระบบการรู้จำคำพูดในลักษณะต่าง ๆ ที่ได้มีการศึกษามาแล้ว
2. ศึกษาทฤษฎีทางด้านภาษาศาสตร์ และอัลกอริทึมในการวิเคราะห์เสียงพูด
3. กำหนดขอบเขต และเลือกกลุ่มคำทดสอบสำหรับงานวิจัย
4. บันทึกข้อมูลเสียงพูดลงบนแถบบันทึกเสียง ใช้ผู้พูดเพศละ 3 คน รวม 6 คน จำนวนคำในรายการคำ 134 คำ จำนวนตัวอย่างที่ใช้ในการรู้จำ 804 คำ
5. เก็บข้อมูล วิเคราะห์ กำหนดกฎเกณฑ์การรู้จำ และทดสอบ
6. สรุปผลการวิจัย ประเมินผล และเสนอแนะแนวทางในการวิจัยต่อไป
7. เสนอรายงานการวิจัยในรูปวิทยานิพนธ์

ประโยชน์ที่คาดว่าจะได้รับจากงานวิจัย

1. ทราบลักษณะทางสวนสัทศาสตร์ของหน่วยเสียงต่างๆ ที่กำหนดในงานวิจัยนี้
2. เป็นแนวทางการพัฒนาไปสู่ระบบการรู้จำคำพูดในภาษาไทยทั้งหมดต่อไป
3. สามารถนำไปประยุกต์ใช้ในงานด้านต่าง ๆ เพื่อเพิ่มประสิทธิภาพการติดต่อกับเครื่องมือต่าง ๆ โดยใช้เสียงพูด