



# โปรแกรมถ่ายถอดเสียงคำศัพท์อักษรโรมันเป็นอักษรไทย ตามหลักเกณฑ์ราชบัณฑิตยสถาน

ชาญชัย สุทธิวิเศษศักดิ์\* และ วิเศษ ศักดิ์ศิริ\*\*

## บทคัดย่อ

งานวิจัยนี้ได้พัฒนาโปรแกรมถ่ายถอดเสียงคำศัพท์อักษรโรมันเป็นอักษรไทยตามหลักเกณฑ์ราชบัณฑิตยสถาน โดยให้ผู้ใช้ทำการคีย์คำศัพท์อักษรโรมัน (คำศัพท์ภาษาอังกฤษ) ซึ่งมีความยาวไม่เกิน 7 ตัวอักษร โปรแกรมจะถ่ายถอดเสียงคำศัพท์อักษรโรมันให้เป็นคำอ่านอักษรไทย โดยใช้อัลกอริทึมกำหนดการถ่ายถอดเสียง ซึ่งมีใช้เป็นลักษณะของการเก็บคำอ่านไว้ในฐานข้อมูลก่อนล่วงหน้า เริ่มจากศึกษาและวิเคราะห์โครงสร้างของอักษรโรมันว่าสามารถถ่ายถอดเสียงออกมาเป็นอักษรไทยได้อย่างไร โดยใช้หลักเกณฑ์การถ่ายถอดเสียงของราชบัณฑิตยสถาน ลักษณะการทำงานคือ การสร้างตารางเปรียบเทียบอักษรโรมันแปลงให้เป็นพยัญชนะไทย โดยตารางนั้นได้อ้างอิงมาจากราชบัณฑิตยสถาน เก็บไว้ในฐานข้อมูล ซึ่งเมื่อผู้ใช้ป้อนคำศัพท์เข้ามา โปรแกรมก็จะตรวจสอบเงื่อนไข และเปรียบเทียบตัวอักษรจากตารางแล้วแปลงเป็นพยัญชนะไทย จากนั้นจะแสดงคำอ่านเป็นอักษรไทยออกมา จากการทดลองพบว่า 62.88 % สามารถถ่ายถอดเสียงคำศัพท์อักษรโรมันให้เป็นคำอ่านภาษาไทยได้ถูกต้องตามหลักเกณฑ์ราชบัณฑิตยสถาน จากตัวอย่างที่สุ่มเลือกมาทั้งหมด 431 คำ โดยนำมาจากหนังสือคำศัพท์เฉพาะทางจำนวน 5 เล่ม ฉบับราชบัณฑิตยสถาน

**คำสำคัญ :** โปรแกรมถ่ายถอดเสียง คำอักษรโรมัน อัลกอริทึม

## 1. บทนำ

นับแต่อดีตจนถึงปัจจุบัน มนุษย์ต้องติดต่อสื่อสารระหว่างกันอยู่ตลอดเวลา ซึ่งวิธีที่ใช้ในการสื่อสารอย่างง่ายที่สุดก็คือวิธีการพูด ดังนั้นการสื่อสารกันอย่างมีประสิทธิภาพจะเกิดขึ้นได้นั้น ผู้พูดต้องมั่นใจว่าผู้ฟังจะเข้าใจความหมายไปในทิศทาง

เดียวกับผู้พูด หากผู้พูดต้องการสื่อความหมายแบบหนึ่ง แต่ผู้ฟังกลับเข้าใจความหมายเป็นอีกแบบหนึ่ง ย่อมแสดงว่าการสื่อสารครั้งนั้นเกิดความผิดพลาดขึ้นแล้ว ซึ่งอาจทำให้เกิดปัญหาขึ้นไม่มากนักน้อย ดังนั้นการพูดหรือการอ่านออกเสียงจึงเป็นเรื่องที่สำคัญ โดยเฉพาะอย่างยิ่งกับภาษาที่ไม่คุ้นเคย ยกตัวอย่างเช่น คำในภาษาอังกฤษ ซึ่งปัจจุบันเป็นภาษาที่ใกล้ตัวมากภาษาหนึ่งและปัญหาที่พบบ่อยก็คือ คำศัพท์ที่มีตัวอักษรและสระใกล้เคียงกันแต่ความหมายแตกต่างกันทำให้การออกเสียงอ่านอาจเกิดข้อผิดพลาดได้ ดังนั้นการพูดหรือการออกเสียงควรชัดเจน และถูกต้องตามหลักเกณฑ์ด้วยสำหรับการพูดหรือการออกเสียงคำภาษาอังกฤษให้ถูกต้องตามหลักเกณฑ์ภาษาไทย ซึ่งหลักเกณฑ์ที่ใช้เป็นหลักเกณฑ์ของราชบัณฑิตยสถาน [1] ดังนั้นจึงได้ทำการพัฒนาซอฟต์แวร์นี้ขึ้นโดยมีวัตถุประสงค์เพื่อใช้ในการอ่านออกเสียงคำศัพท์อักษรโรมันให้เป็นอักษรไทยได้อย่างถูกต้องตามหลักราชบัณฑิตยสถาน โดยลักษณะการทำงานจะเป็นการสร้างเงื่อนไขเพื่อตรวจสอบตัวอักษร แล้วแสดงเป็นคำอ่าน มีใช้ลักษณะการเปรียบเทียบตัวอักษรในฐานข้อมูลแล้วดึงคำอ่านขึ้นมาแสดง ซึ่งคำศัพท์ที่ป้อนเข้ามามีขนาดไม่เกิน 7 ตัวอักษร และเมื่อโปรแกรมประมวลผลจนได้ผลลัพธ์คำอ่านออกมาแล้วก็จะเก็บบันทึกคำๆ นั้นพร้อมกับผลลัพธ์ลงในฐานข้อมูลของโปรแกรมทันที เมื่อผู้ใช้ต้องการเรียกดูคำศัพท์เก่าๆ เหล่านั้นก็สามารถทำได้โดยการเปิดฐานข้อมูลของโปรแกรมขึ้นมา โปรแกรมนี้จะสามารถนำไปติดตั้งกับเครื่องคอมพิวเตอร์เครื่องอื่นๆ ได้ เนื่องจากได้จัดทำไว้ในรูปแบบของไฟล์เซตอัพ และเครื่องคอมพิวเตอร์ที่จะนำซอฟต์แวร์นี้ไปทำการติดตั้ง ควรมีคุณสมบัติดังนี้ คือ หน่วยประมวลผลผล PENTIUM II ขึ้นไป หน่วยความจำหลัก(RAM) 128 MB ขึ้นไปและมีหน่วยความจำสำรอง (Hard disk) 10 GB ขึ้นไปสำหรับโครงการนี้ประโยชน์

\* คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ

\*\* ภาควิชาครุศาสตร์ไฟฟ้า คณะครุศาสตร์อุตสาหกรรม สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ



ที่คาดว่าจะได้รับคือ ได้ซอฟต์แวร์ที่มีประสิทธิภาพในการถ่ายเสียง และเสริมสร้างความรู้ความสามารถในการอ่านคำศัพท์ที่ถูกต้องตามหลักเกณฑ์ราชบัณฑิตยสถาน

## 2. การถ่ายเสียงอักษรโรมันเป็นอักษรไทย

การจัดทำโปรแกรมถ่ายเสียงคำศัพท์อักษรโรมันให้เป็นอักษรไทย สำหรับการออกแบบและพัฒนาตัวโปรแกรมจะต้องทำการศึกษาทฤษฎีที่จำเป็นคือภาษาวิซวลเบสิก และการประยุกต์ใช้กับระบบฐานข้อมูล เนื่องจากว่าโปรแกรมนี้จะต้องสร้างหน้าจอของโปรแกรมและเงื่อนไขที่ใช้ในการประมวลผล รวมถึงการเก็บผลลัพธ์ที่ได้ลงฐานข้อมูลโดยอัตโนมัติ ดังนั้นจึงต้องทำการศึกษาการเขียนโปรแกรม และวิธีที่จะทำให้ทั้งภาษาวิซวลเบสิก และระบบฐานข้อมูลไม่โครซอฟต์แอ็กเซสสามารถทำงานร่วมกันได้อย่างเหมาะสม รวมถึงการศึกษางานวิจัยต่าง ๆ ที่ผ่านมา ซึ่งจะสามารถนำส่วนที่มีความคล้ายคลึงเพื่อนำมาเป็นแนวทางและประยุกต์ใช้กับโปรแกรมถ่ายเสียงนี้ได้ โดยจะพบว่าการศึกษาวิเคราะห์เสียงเป็นภาษาไทย [2] - [6] จะต้องทำการวิเคราะห์โครงสร้างของคำศัพท์ ซึ่งว่าในเรื่องของเสียงของแต่ละตัวอักษรเช่น ตัวอักษรที่ทำหน้าที่เป็นพยัญชนะและทำหน้าที่เป็นสระ และในเรื่องของการแยกจำนวนพยางค์เพื่อให้สามารถออกเสียงอ่านได้อย่างถูกต้อง

ตารางที่ 1 เปรียบเทียบอักษรโรมันเป็นพยัญชนะไทย

พยัญชนะ	อักษรโรมัน		ตัวอย่าง
	ตัวต้น	ตัวสะกด	
ก	k	K	กา = ka นก = nok
ข ข ค ค ฃ	kh	K	ขอ = kho สุข = suk โค = kho ยुक = yuk ฆ้อง = khong เมฆ = mek
ง	ng	ng	งาม = ngem สงฆ์ = song

ขั้นตอนนี้ได้ศึกษาวิธีการที่จะใช้ในการถ่ายเสียงคำศัพท์ภาษาอังกฤษ โดยผู้วิจัยได้ค้นคว้าหาข้อมูลต่างๆ ซึ่งเป็นของทางราชบัณฑิตยสถานโดยเฉพาะ เพื่อหาข้อมูลสำหรับที่จะ

นำมาวิเคราะห์และใช้สร้างอัลกอริทึมในการถ่ายเสียง สำหรับการดำเนินงานวิจัยครั้งนี้ได้นำเอาฟอร์แมต (format) การถ่ายเสียงซึ่งจะอยู่ในรูปแบบของตารางที่ทางราชบัณฑิตยสถานได้กำหนดไว้แล้วและมีลักษณะดังตารางที่ 1 โดยจะเปรียบเทียบให้เห็นว่า อักษรโรมัน ตัว  $k$  เมื่อทำหน้าที่เป็นทั้งตัวต้นและตัวสะกดแล้วจะออกเสียงเป็นพยัญชนะไทยได้คือ ก หรืออักษรโรมันที่เป็นอักษรกล้ำเช่น  $kh$  เมื่อทำหน้าที่เป็นตัวต้นแล้ว จะออกเสียงเป็นพยัญชนะไทยได้คือ ข, ค, ฃ เป็นต้นซึ่งตารางที่ได้นำเสนอเป็นเพียงส่วนหนึ่งของตารางเปรียบเทียบอักษรโรมันเป็นพยัญชนะไทยเท่านั้น

ตารางที่ 2 เปรียบเทียบอักษรโรมันเป็นสระของไทย

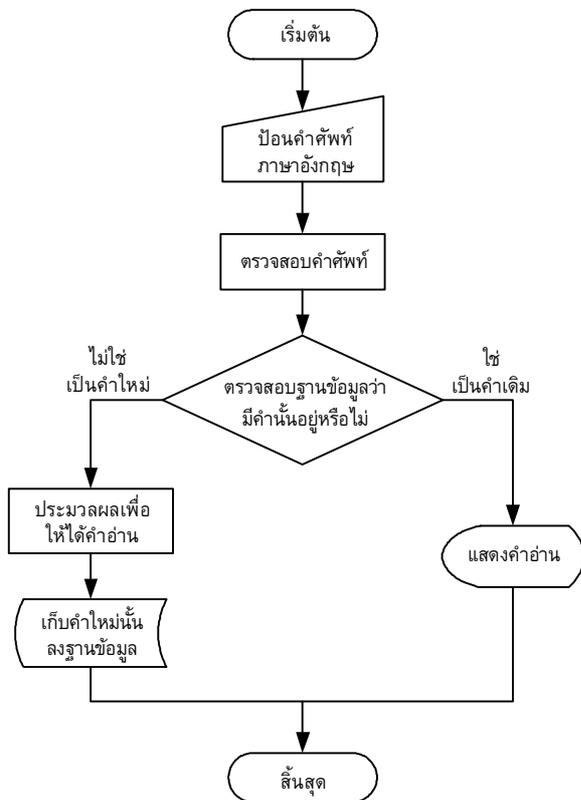
สระไทย	อักษรโรมัน	ตัวอย่าง
อะ, ะ (อะ ลดรูป), รร (มีตัวสะกด), อา	a	ปะ = pa วัน = won สรรพ = sap มา = ma
รร (มีไม่ตัวสะกด)	an	สรรหา = sanha สวรรค์ = sawan
อ่า	am	รำ = ram
อิ, อี	i	มิ = mi มีด = mit

ในตารางที่ 2 เป็นการเปรียบเทียบให้เห็นว่า ตัวอักษร a เมื่อเปรียบเทียบให้เป็นสระของไทยแล้วจะออกเสียงเป็น สระ อะ, สระ อะ (ลดรูป) หรือ สระ อา หรือ ตัวอักษร am เมื่อเปรียบเทียบให้เป็นสระของไทยคือเสียง สระ อ่า และตัวอย่างสุดท้ายคือ ตัวอักษร i แสดงให้เห็นว่าอักษรโรมันหนึ่งตัวที่ทำหน้าที่เป็นสระ สามารถเป็นได้สองสระในภาษาไทย คือ สระ อิ และ สระ อี เป็นต้น

เมื่อนำข้อมูลของทางราชบัณฑิตยสถานซึ่งได้แสดงเป็นตัวอย่างดังตารางที่ 1 และตารางที่ 2 มาสร้างเป็นตารางเพื่อใช้ในการเปรียบเทียบจากอักษรโรมันให้เป็นพยัญชนะและสระของไทยและเก็บไว้ในฐานข้อมูลของโปรแกรม การที่จะถ่ายเสียงคำศัพท์อักษรโรมันให้เป็นเสียงอักษรไทยให้มีความถูกต้องนั้นจะต้องทำการศึกษาก่อนว่าตัวอักษรโรมันหรืออักษรอังกฤษเหล่านั้นอ่านออกเสียงเป็นภาษาไทยที่ถูกต้องได้อย่างไร เพราะหากว่าการอ่านออกเสียงคำศัพท์เหล่านั้นเป็น



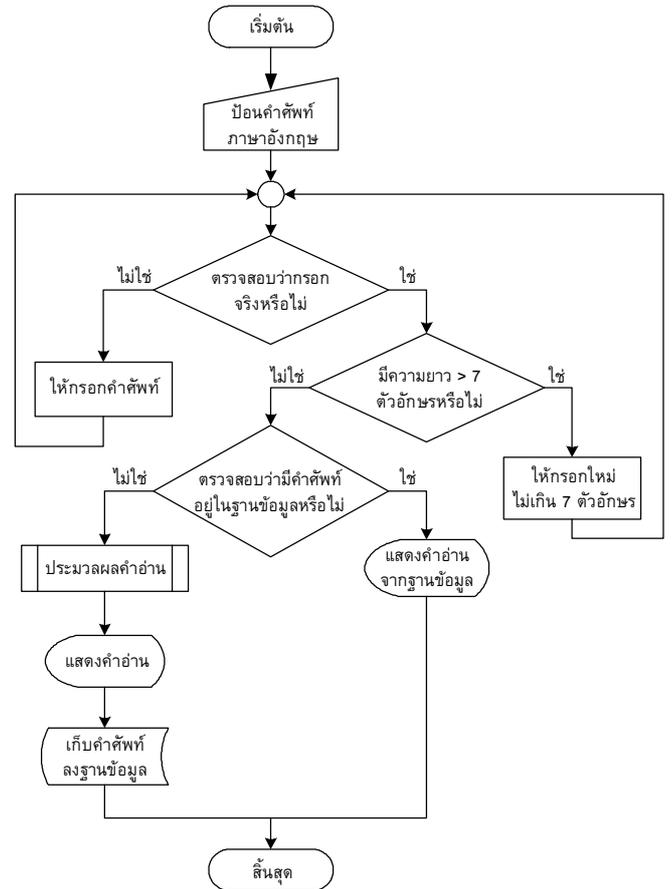
การออกเสียงที่ผิดรูปแบบ ก็จะทำให้การเขียนโปรแกรมเพื่อให้ถ่ายทอดเสียงเป็นภาษาไทยนั้นผิดไปด้วย เช่น ตัวอักษร x หากทำหน้าที่เป็นตัวอักษรตัวต้นต้องออกเสียงเป็นพยัญชนะ ข แต่หากทำหน้าที่เป็นตัวอักษรตัวสะกดก็ต้องออกเสียงเป็นพยัญชนะ ก ยกตัวอย่างเช่น คำว่า xerox อ่านว่า ซี-รอก ด้วยสาเหตุที่แต่ละคำหรือแต่ละตัวอักษรสามารถเป็นได้หลายแบบหลายกรณี ทำให้ผู้วิจัยต้องศึกษาในเรื่องของอักษรโรมันและพยัญชนะไทยเพื่อศึกษาและเก็บข้อมูลว่า แต่ละตัวอักษรโรมันออกเสียงเป็นเสียงพยัญชนะตัวใดในภาษาไทยได้และต้องให้เป็นไปตามหลักเกณฑ์ของราชบัณฑิตยสถาน และเมื่อศึกษางานวิจัยที่ผ่านมาและข้อมูลอื่นๆ แล้วจึงได้วิเคราะห์และออกแบบการทำงานของโปรแกรมโดยสังเขป ซึ่งเป็นลำดับขั้นตอนดังแผนภาพที่ 1



ภาพที่ 1 แผนผังขั้นตอนวิธีของระบบโดยสังเขป

จากภาพที่ 1 เป็นแผนผังแสดงลำดับการทำงานโดยสังเขปของโปรแกรมถ่ายทอดเสียง ซึ่งเริ่มต้นด้วยการให้ผู้ใช้งานป้อนคำศัพท์ภาษาอังกฤษลงไป โปรแกรมจะทำการตรวจสอบเงื่อนไขต่าง ๆ โดยเงื่อนไขที่ว่า มีคำศัพท์ค่านี้อยู่ในฐานข้อมูลแล้วหรือไม่ เป็นเพียงเงื่อนไขหนึ่งที่โปรแกรมจะต้องทำการตรวจสอบก่อนเพื่อให้ทราบว่า คำ ๆ นั้นได้เคยถูกประมวลผล

เพื่อให้ผลลัพธ์ไปแล้วครั้งหนึ่ง และในครั้งต่อไปจะไม่ประมวลผลอีกแต่จะเป็นการไปดึงค่า ๆ นั้นในฐานข้อมูลขึ้นมาแสดงทันทีเพื่อความรวดเร็ว หากโปรแกรมตรวจสอบแล้วว่าไม่มีค่า ๆ นั้นในฐานข้อมูล ก็จะประมวลผลเพื่อแสดงผลลัพธ์ที่หน้าจอโปรแกรมทันที และหลังจากนั้นก็จะเป็นค่า ๆ นั้น พร้อมทั้งผลลัพธ์ เก็บลงฐานข้อมูลของโปรแกรมโดยอัตโนมัติ หลังจากที่ผ่านขั้นตอนการวิเคราะห์และออกแบบแล้ว ขั้นตอนต่อไปคือการสร้างอัลกอริทึมเพื่อให้เห็นถึงลักษณะการทำงานของทั้งโปรแกรม ซึ่งมีรายละเอียดคือ โปรแกรมรับคำศัพท์เข้ามาทางช่องอินพุต และตรวจสอบเงื่อนไขว่ามีคำศัพท์ที่ถูกต้องหรือไม่ มีขนาดไม่เกิน 7 ตัวอักษรจริง และไม่เคยมีอยู่ในฐานข้อมูลจริง โปรแกรมประมวลผลและแสดงผลลัพธ์ที่ช่องเอาต์พุต และเก็บบันทึกคำศัพท์ที่ประมวลผลแล้วเข้าฐานข้อมูลและสิ้นสุดการทำงานของโปรแกรม



ภาพที่ 2 แผนผังการทำงานแบบสมบูรณธ์

เมื่อได้อัลกอริทึมแล้วสามารถนำไปเขียนแผนภาพเพื่อแสดงการทำงานที่สมบูรณ์ของทั้งโปรแกรมดังภาพที่ 2 หลังจากนั้นจึงได้เริ่มทำการสร้างหน้าจอเพื่อใช้ติดต่อกับผู้ใช้ (User Interface) ซึ่งหน้าจอต่าง ๆ มีดังภาพที่ 3 และภาพที่ 4



ภาพที่ 3 หน้าจอหลักของโปรแกรม



ภาพที่ 4 หน้าจอฐานข้อมูลของโปรแกรม

จากการวิเคราะห์โครงสร้างของโปรแกรมแล้วนั้น โปรแกรมถ่ายเสียงมีลักษณะการทำงานที่ต้องใช้การเปรียบเทียบอักษรโรมันเพื่อถ่ายเสียงให้เป็นอักษรไทย และได้คำอ่านซึ่งเป็นผลลัพธ์ออกมาแสดง และทำการเก็บผลลัพธ์ที่ได้นั้นลงฐานข้อมูลด้วย ดังนั้น โปรแกรมต้องมีการสร้างตารางเปรียบเทียบให้ชื่อว่าตาราง MAPP และตารางเก็บผลลัพธ์ให้ชื่อว่า ตาราง DIC เก็บเอาไว้ในฐานข้อมูลก่อน ซึ่งแสดงดังตารางที่ 3 และตารางที่ 4

### 3. ผลการดำเนินงาน

เมื่อพัฒนาโปรแกรมเป็นที่เรียบร้อยแล้วผลลัพธ์ที่ได้คือโปรแกรมถ่ายเสียงซึ่งมีหน้าจอต่าง ๆ ดังภาพที่ 5-7

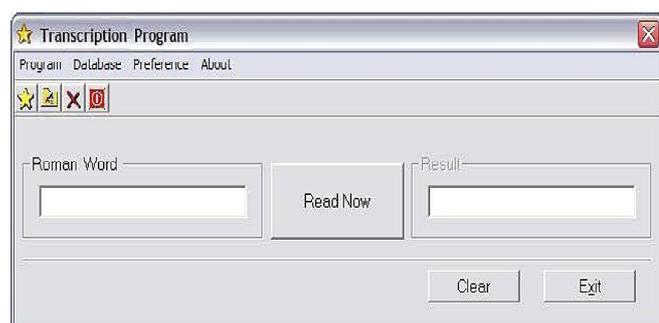
ภาพที่ 5 คือ หน้าจอหลักของโปรแกรมซึ่งประกอบไปด้วยเมนูต่าง ๆ ช่องใส่คำศัพท์ ปุ่มประมวลผล และช่องแสดงผลลัพธ์ เมื่อผู้ใช้ใส่คำศัพท์ลงไปแล้วกดปุ่ม Read Now โปรแกรมก็จะประมวลผลและแสดงผลลัพธ์ออกมา แสดงดังภาพที่ 6 และเก็บคำศัพท์เหล่านั้นลงฐานข้อมูล เมื่อเรียกเปิดฐานข้อมูลก็จะได้ผลลัพธ์แสดงดังภาพที่ 7

ตารางที่ 3 เปรียบเทียบอักษรโรมันเป็นพยัญชนะและสระไทย

ID	e_alpha	e_pos	t_alpha	e_type
22	.	N		A
5	a	1	อะ	A
6	a	N	แ	V
25	a	N	า	V
7	b	1	บ	A
8	b	N	ม	A
9	c	1	ห	A
10	c	N	ก	A
24	d	N	ด	A
23	d	1	ต	V
18	e	N	อี	V
41	e	N	เ	V
17	e	1	อ	A

ตารางที่ 4 ลักษณะการเก็บผลลัพธ์ของโปรแกรม

ID	e_word	t_word	save_date	open_time
212	bet	บีต, เบต,	29/1/2006	0
213	ni	นี, นิ,	29/1/2006	0
214	bun	บัน, บุน,	29/1/2006	1
215	VAT	แวท, วาท,	30/1/2006	0
216	CAT	แคท, คาท,	30/1/2006	1
*	มายเดซัดโนมิติ		31/1/2006	0



ภาพที่ 5 หน้าจอหลักของโปรแกรม

จากการทดสอบเพื่อประเมินความถูกต้องของโปรแกรม ได้ใช้วิธีการสุ่มเลือกคำศัพท์ภาษาอังกฤษมาจำนวน 431 คำ จำนวน 26 หมวดตัวอักษร ซึ่งเป็นคำศัพท์ที่ได้มาจากหนังสือคำศัพท์เฉพาะทาง ซึ่งเป็นฉบับของราชบัณฑิตยสถาน จำนวน 5 เล่ม คือ ศัพท์คอมพิวเตอร์และเทคโนโลยีสารสนเทศ



ภาพที่ 6 หน้าจอขณะแสดงผลศัพท์



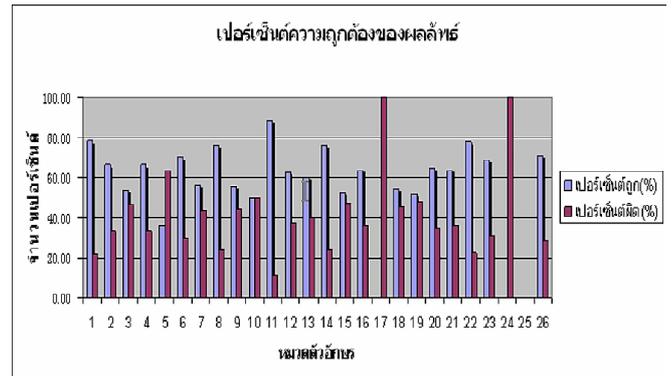
ภาพที่ 7 หน้าจอแสดงฐานข้อมูลของโปรแกรม

ตารางที่ 5 ตัวอย่างผลลัพธ์หมวดอักษร A

คำศัพท์โรมัน หมวด A	คำอ่าน (ราชบัณฑิต)	คำอ่าน (โปรแกรมถ่ายเสียง)
aztec	แอซ-เตก	แอซ-เตค
atlas	แอต-ลาส	แอต-ลาส
atlan (atlantes)	แอต-แลน	แอต-แลน
anglo	แอง-โกล	แอง-โกล
ampere	แอม-แปร์	แอม-พี-เร

[7] ศัพท์เรอร์โมไดนามิกส์ [8] ศัพท์ทันตแพทยศาสตร์ [9] ศัพท์สถาปัตยกรรมศาสตร์ [10] ศัพท์พฤกษศาสตร์ [11] หลังจากนั้นได้นำคำศัพท์ทั้งหมดที่ได้ไปเข้าโปรแกรมถ่ายเสียงเพื่อประมวลผลหาผลลัพธ์คำอ่านภาษาไทยออกมา ซึ่งผลลัพธ์ที่ได้ดังตัวอย่างตารางที่ 5 ผลลัพธ์ที่ได้จากการทดสอบการประมวลผลของโปรแกรมนั้น เช่นคำว่า Aztec คำอ่านของทางราชบัณฑิตอ่านว่า แอซ-เตก และโปรแกรมอ่านได้ว่า แอซ-เตค ซึ่งเป็นผลลัพธ์ที่ถูกต้องแล้ว เนื่องจากตัวอักษรตัว c สามารถที่อ่านออกเสียงเป็นพยัญชนะ ค ได้ ส่วนคำว่า anglo

คำอ่านของทางราชบัณฑิตอ่านว่า แอง-โกล แต่โปรแกรมอ่านได้ว่า แอง-โล ซึ่งเป็นผลลัพธ์ที่ผิด โดยจะทำการรอบสี่ทศวรรษ แสดงไว้เฉพาะคำที่ผิด ผลสรุปคือโปรแกรมประมวลผลได้ถูกต้องเป็นจำนวน 271 คำ และประมวลผลออกมาผิดจำนวน 160 คำ ซึ่งเมื่อคิดเป็นเปอร์เซ็นต์ความถูกต้องเท่ากับ 62.88% โดยสามารถสรุปออกมาเป็นตารางเปอร์เซ็นต์ความถูกต้องได้ดังตารางที่ 6 และแสดงเป็นกราฟได้ดังภาพที่ 8



ภาพที่ 8 กราฟแสดงเปอร์เซ็นต์ความถูกต้อง

ตารางที่ 6 สรุปผลเปอร์เซ็นต์ความถูกต้อง

หมวดอักษรโรมัน	จำนวนคำทั้งหมด	คำอ่านถูก	คำอ่านผิด	เปอร์เซ็นต์ถูก (%)	เปอร์เซ็นต์ผิด (%)
A	23	18	5	78.26	21.74
B	27	18	9	66.67	33.33
C	26	15	13	53.57	46.43
D	21	14	7	66.67	33.33
E	11	4	7	36.36	63.64
F	27	19	8	70.37	29.63
G	25	14	11	56.00	44.00
H	21	16	5	76.19	23.81
I	9	5	4	55.56	44.44
J	4	2	2	50.00	50.00
K	9	8	1	88.89	11.11
L	24	15	9	62.50	37.50
M	25	15	10	60.00	40.00
N	17	13	4	76.47	23.53
O	19	10	9	52.63	47.37
P	25	16	9	64.00	36.00
Q	2	0	2	0.00	100.00
R	22	12	10	54.55	45.45
S	27	14	13	51.85	48.15
T	20	13	7	65.00	35.00
U	11	7	4	63.64	36.36
V	9	7	2	77.78	22.22
W	16	11	5	68.75	31.25
X	2	0	2	0.00	100.00
Y	0	0	0	0.00	0.00
Z	7	5	2	71.43	28.57
รวม(ค่า)	431	271	160		
เปอร์เซ็นต์ของคำถูก (%) หารได้จาก (271/431)*100				62.88	
เปอร์เซ็นต์ของคำผิด (%) หารได้จาก (160/431)*100					37.12

#### 4. สรุป

ในงานวิจัยนี้ได้พัฒนา “โปรแกรมการถ่ายเสียงคำศัพท์อักษรโรมันเป็นอักษรไทยตามหลักราชบัณฑิตยสถาน” โดยให้ผู้ใช้ทำการป้อนคำศัพท์อักษรโรมันที่มีขนาดไม่เกิน 7



ตัวอักษรลงไป โปรแกรมจะประมวลผลด้วยการตรวจสอบตัวอักษรทุกตัวของคำศัพท์แล้วเปรียบเทียบจากตาราง และแปลงให้เป็นอักษรไทยรวมกันเป็นคำอ่านออกมาแสดง โดยตารางที่ใช้ในการเปรียบเทียบแปลงจากอักษรโรมันให้เป็นอักษรไทยอ้างอิงจากทางราชบัณฑิตยสถาน ซึ่งโปรแกรมจะประมวลผลหาคำอ่านตามอัลกอริทึมที่ได้เขียนไว้ด้วยภาษาไมโครซอฟต์วิซวลเบสิกเวอร์ชัน 6.0 และเก็บผลลัพธ์ไว้ฐานข้อมูลไมโครซอฟต์แอ็กเซส ด้านการทดลองเพื่อวิเคราะห์ความถูกต้อง ได้สุ่มเลือกคำศัพท์จากหนังสือคำศัพท์เฉพาะทางฉบับราชบัณฑิตยสถาน จำนวน 5 เล่ม รวมคำศัพท์ทั้งหมด 431 คำ โดยเรียงลำดับตัวอักษรจากอักษร A-Z ซึ่งแต่ละหมวดได้ทำการสุ่มเลือกมาเป็นจำนวนไม่เท่ากัน ซึ่งเมื่อได้นำคำศัพท์เหล่านั้นไปทำการประมวลผลด้วยโปรแกรมถ่ายเสียงแล้วผลลัพธ์ที่ได้ คือ มีจำนวนคำที่ประมวลผลถูกต้องเท่ากับ 271 คำ และจำนวนคำที่ประมวลผลผิดเท่ากับ 160 คำ เมื่อคิดเป็นเปอร์เซ็นต์ความถูกต้องได้เท่ากับ 62.88 %

## 5. บรรณานุกรม

- [1] ราชบัณฑิตยสถาน. “หลักเกณฑ์การถอดอักษรไทยเป็นอักษรโรมันแบบถ่ายเสียง.” ประกาศ ณ วันที่ 12 ต.ค. 2548, ประเภท เอกสารอิเล็กทรอนิกส์ (www.royin.go.th).
- [2] กิตติพันธ์ วานิช, สาวิตรี วงศ์ทะยาน. “การพัฒนาโปรแกรมสังเคราะห์เสียงภาษาไทย.” สาขาวิทยาการคอมพิวเตอร์ประยุกต์ สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2546.
- [3] พัชรินทร์ ปราชญ์สุนัย. “โครงสร้างหลักของคลื่นเสียงพยัญชนะและสระสำหรับการสังเคราะห์เสียงภาษาไทย.” สาขาเทคโนโลยีคอมพิวเตอร์ สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2546.
- [4] เศรษฐา ศรีพล, เอกชัย แซ่จิ่ง. “โปรแกรมอ่านและแปลคำศัพท์.” สาขาวิทยาศาสตร์ประยุกต์ สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2545.
- [5] ดวงแก้ว สวามิภักดิ์. “การสร้างซอฟต์แวร์วิเคราะห์ไวยากรณ์ไทยภายใต้ระบบยูนิคซ์.” สาขาวิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยธรรมศาสตร์, 2533.
- [6] สัญญา เครือหงษ์. “การพัฒนาการแยกพยางค์จากข้อความภาษาไทยเพื่อการสังเคราะห์เสียงพูด.” สาขาวิชาเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2546.
- [7] ราชบัณฑิตยสถาน “ศัพท์คอมพิวเตอร์และเทคโนโลยีสารสนเทศ” พิมพ์ครั้งที่ 6, 2546.
- [8] ราชบัณฑิตยสถาน “ศัพท์เชอร์โมไดนามิกส์” พิมพ์ครั้งที่ 1, 2548.
- [9] ราชบัณฑิตยสถาน “ศัพท์ทันตแพทยศาสตร์” พิมพ์ครั้งที่ 2, 2541.
- [10] ราชบัณฑิตยสถาน “ศัพท์สถาปัตยกรรมศาสตร์” พิมพ์ครั้งที่ 1, 2548.
- [11] ราชบัณฑิตยสถาน “ศัพท์พฤกษศาสตร์” พิมพ์ครั้งที่ 2, 2546.