



บทที่ 2

ชาวค้เก็ท

2.1 ความหมายของชาวค้เก็ท

ชาวค้เก็ท เป็นเทคนิคของการสร้างรหัสสำหรับค้นหาข้อมูลวิธีหนึ่ง วิธีสร้างรหัสแบบนี้จะคำนึงถึงเสียงที่พูดเป็นหลัก คำหรือพยางค์ที่ออกเสียงเหมือนกัน ควรให้รหัสสำหรับการค้นหาออกมาตรงกัน ในการที่จะนำชื่อสะกดที่เป็นตัวอักษรมาเป็นรหัสค้นหา โดยตรงนั้น ย่อมจะมีการผิดพลาดได้ จากการใช้ตัวสะกดที่ผิด ดังนั้น ชาวค้เก็ท จะอาศัยการแปลง (CONVERT) ชื่อเหล่านี้ไปเป็นกลุ่มของรหัสอื่นก่อนเพื่อหลีกเลี่ยงการใช้ตัวสะกดที่ผิด กลุ่มของรหัสที่ใช้ประกอบด้วย อักษรนำหน้า 1 ตัว นอกนั้นจะเป็นตัวเลขอีกตามที่กำหนด ตัวเลขต่าง ๆ นี้ได้จากการจัดกลุ่มอักษรซึ่งมีเสียงสะกดเดียวกันไว้ด้วยกัน แต่ละอักษรในกลุ่มจะกำหนดแทนด้วยตัวเลขเหมือนกัน ซึ่งการกำหนดรหัสด้วยหลัก หรือวิธีนี้ เรียกว่า ชาวค้เก็ท

โดยทั่วไป ชาวค้เก็ทในภาษาอังกฤษ สระทุกตัวหรือพยัญชนะที่ออกเสียงสระจะถูกตัดทิ้ง แต่จะยังคงอักษรตัวแรกไว้ไม่ว่าอักษรตัวนั้นจะเป็นสระหรือพยัญชนะ ในการศึกษแปลงใช้กับภาษาไทย สระ และวรรณยุกต์ทุกตัวจะถูกตัดทิ้ง โดยไม่ยกเว้นสระซึ่งเป็นอักษรตัวแรก เช่น ในภาษาอังกฤษ สำหรับอักษรตัวแรกซึ่งใช้นำในแต่ละรหัสจะใช้พยัญชนะตัวแรกที่ปรากฏในคำเป็นตัวแทน

หมายเหตุ สำหรับการตัดสระในภาษาไทยทั้งนั้น บางกรณีอาจไม่ตัดทิ้งทุกตัว เนื่องจากสระบางตัวเมื่อออกเสียงจะให้เสียงพยัญชนะสะกดออกมา เช่น ำ เมื่อเป็นเช่นนี้สามารถใส่รหัสแทนให้กับเสียงพยัญชนะสะกดซึ่งเกิดจากสระได้ เพื่อความถูกต้องในการค้นหาข้อมูล โดยอาจนำสระรวมไว้ในกลุ่มพยัญชนะหรือจะใส่รหัสแก่สระบางตัวแยกต่างหากก็ได้

สำหรับ ชาวค้เก้กข ที่ใช้กับภาษาอังกฤษนั้น ตัวที่ถูกแปลงไปเป็นตัวเลข จะต้องเป็นพยัญชนะเท่านั้น สำหรับสระและพยัญชนะที่ออกเสียงสระจะถูกคัดทิ้ง ไม่นำมาแปลรวมเป็นรหัส ปกติชาวค้เก้กขในภาษาอังกฤษ นิยมใช้ความยาวรหัสเพียง 4 ตัว คือ จะใช้อักษรตัวแรกของชื่อ เป็นตัวนำรหัส และพยัญชนะตัวต่อไปอีก 3 ตัวจะถูกแปลงเป็นตัวเลขเป็นรหัสตาม รวมเป็นความยาวของรหัส 4 ตัวอักษร การสร้างรหัสชาวค้เก้กขภาษาอังกฤษ มีหลักเกณฑ์ที่สำคัญ ดังนี้

- 1) ตัวอักษรตัวแรกของชื่อใช้เป็นตัวนำรหัสเลย ไม่ว่าจะเป็พยัญชนะหรือ สระ
- 2) ตัวอักษรตามทุกตัวถ้าเป็นสระ หรือพยัญชนะที่ออกเสียงสระ ให้ละเว้นไม่นำมาแปลงรวมในรหัส อักษรเหล่านี้ คือ A E I O U W Y H
- 3) ตัวอักษรเหมือนกันอยู่ติดกันให้อาณาแปลงเป็นรหัส 1 ตัว
- 4) ถ้ามีอักษรไม่พอ คือแปลงเป็นรหัสแล้วไม่ครบ 4 ตัวให้ใส่ 0 จนครบ
- 5) พยัญชนะทุกตัวยกเว้นตัวแรก ให้ใช้หลักเกณฑ์การเปลี่ยนเป็นรหัสตัวเลข

ดังนี้

กลุ่มที่แทนด้วย	1	มี	B	F	P	V
กลุ่มที่แทนด้วย	2	มี	G	C	J	K
			Q	S	X	Z
กลุ่มที่แทนด้วย	3	มี	D	T		
กลุ่มที่แทนด้วย	4	มี	L			
กลุ่มที่แทนด้วย	5	มี	M	N		
กลุ่มที่แทนด้วย	6	มี	R			
กลุ่มที่แทนด้วย	0		สำหรับกรณีที่มีตัวอักษรไม่พอ			

ตัวอย่างการสร้างรหัส ชาวค้เค็ทท์ภาษาอังกฤษจากชื่อ แสทงไวคังนี้

ชื่อ	รหัส
MILLER	M460
MUELLER	M460
MULLER	M460
LO	L000
LOEW	L000
LOWE	L000
KORFHAGE	K612
KORVAGE	K612
ROGERS	R262
RODGERS	R326
HERRING	H652
HERRON	H650

จากตัวอย่างข้างบนจะเห็นว่า ชื่อ MILLER MUELLER และ MULLER
เมื่อใช้วิธีของชาวค้เค็ทท์ สร้างรหัส จะให้รหัสออกมาเหมือนกัน ซึ่งการสะกดผิดบาง
ลักษณะ ชาวค้เค็ทท์สามารถช่วยได้

2.2 การศึกษาเปรียบเทียบไวยากรณ์ภาษาไทย และปัญหา

สิ่งที่ได้กล่าวมาข้างต้นว่า ลักษณะของภาษาไทยต่างจากภาษาอื่นหลายด้าน จึงไม่อาจที่จะนำหลักเกณฑ์และวิธีการต่าง ๆ ของไวยากรณ์ในภาษาอังกฤษทั้งหมด มาใช้กับภาษาไทย จากการพิจารณาหลักเกณฑ์ในภาษาอังกฤษ สามารถนำหลักเกณฑ์บางข้อมาใช้ บางข้ออาจต้องคัดออกหรือศึกษาเปลี่ยนแปลงตามความเหมาะสม หลักเกณฑ์สำคัญที่จะพิจารณาใช้ในภาษาไทย มีดังนี้คือ

- 1) ใช้พยัญชนะตัวแรกที่ปรากฏในชื่อ เป็นตัวนำรหัส
- 2) พยัญชนะอื่นที่ตามมาจะเปลี่ยนเป็นตัวเลขตามหลักเกณฑ์การจัดกลุ่ม และกำหนดรหัส ใช้เป็นรหัสตัวตามต่อไป
- 3) โดยปกติสระ และวรรณยุกต์ ให้ตัดทั้งหมดไม่นำมาเป็นตัวพิจารณาสร้างรหัสมากเว้นสระบางตัวเท่านั้นที่อาจนำมาพิจารณา ถ้าสระนั้นให้เสียงสะกดที่เป็นพยัญชนะออกมา
- 4) ถ้าชื่อมีพยัญชนะจำนวนที่ไม่เท่ากับความยาวของรหัสให้เติมด้วย 0 จนได้ความยาวของรหัสที่กำหนดไว้

นอกจากหลักเกณฑ์ที่กล่าวข้างต้นแล้ว ยังมีปัญหาจากการสะกดคำในภาษาไทยอีกมาก ซึ่งอาจต้องนำมาพิจารณาประกอบ แต่ปัญหาส่วนใหญ่เป็นปัญหาที่ยุ่งยาก ดังนั้นในการแก้ไขให้เกิดความถูกต้องเกือบ 100 % นั้น จะยุ่งยากซับซ้อนและเปลืองมาก ปัญหาต่าง ๆ ที่คิดว่าจะพบในกรณีที่ใช้ชื่อเป็นรหัส ซึ่งได้มาจากการพิจารณาศึกษาจากตำราไวยากรณ์ไทย มีดังนี้คือ

2.2.1 กรณีที่ใช้ ร

ชื่อที่มีพยัญชนะ ร และเขียนเหมือนกัน สามารถออกเสียงได้ 2 ลักษณะ คือ ออกเสียงเป็นพยัญชนะต้นพยางค์ หรือออกเสียงเป็นพยัญชนะสะกด

- 1) กรณีที่ออกเสียงเป็นพยัญชนะต้นพยางค์ เช่น อรทัย อรชุน อรชาติ วรชาติ อรชร

2) กรณีที่ออกเสียงเป็นพยัญชนะสะกด ซึ่งแทนด้วย ออน ไ้ เช่น
 นือร สมร ชวร อรอนงค์
 ทั้ง 2 กรณีที่กล่าวมานี้เครื่องไม่สามารถแยกไ้ได้ว่า เมื่อไร ร เป็นตัว
 สะกด และเมื่อไร ร เป็นพยัญชนะต้นพยางค์

วิธีแก้ไข

ในการแก้ไขอาจทำได้โดยใช้อักษรพิเศษกำกับ ขึ้นในแต่ละพยางค์ที่มี ร
 เพื่อจะไ้รู้ว่า ร ตัวนี้เป็นพยัญชนะต้นพยางค์ หรือพยัญชนะสะกด ตัวอย่างเช่นใช้
 เป็นตัวคั่นบอก ถ้ามี , อยู่หน้า ก็ไ้ถือว่า ร นั้นเป็นพยัญชนะต้นพยางค์ เช่น อ,รชุน
 ว,รชาติ แต่ถ้าเป็นพยัญชนะสะกดไม่ต้องมีอะไรคั่น เช่น สมร อรอนงค์ นอกจากการ
 ใช้อักษรพิเศษกำกับแล้ว อีกวิธีหนึ่งสามารถทำได้โดยการตรวจสอบสระต่าง ๆ ซึ่งจะบอกให้
 รู้ว่าตัวพยัญชนะใดเป็นพยัญชนะต้นพยางค์ หรือเป็นพยัญชนะสะกด

2.2.2 กรณีที่ไ้ ไ้ ไ้-ย และ -ย

ในกรณีที่ไ้ ไ้ (ไม่มลาย) และ ไ้ (ไม่ม้วน) จะทำให้ไ้ไ้ส
 ต่างจากการไ้ ไ้-ย หรือ -ย เนื่องจากว่า ไ้ ไ้ สามารถแทนด้วย ไ้-ย
 หรือ -ย ไ้ไ้จึงอาจทำให้เกิดปัญหาขึ้นไ้ในกรณีที่ไ้ไ้ต่างกัน

คำที่ไ้ ไ้	เช่น	ไพโรจน์	ไพรี	ไพศาล
		โมตรี	ไพจิตร	ไ้กรสร
คำที่ไ้ ไ้-ย	เช่น	ไชย	ภูวไนย	
คำที่ไ้ -ย	เช่น	หทัย	กตัญญู	นัยนา
		วินัย	อุทัย	

วิธีแก้ไข

เนื่องจากว่าค่าที่ใช้ \bar{y} สามารถแทนด้วย \bar{y} ได้ทุกตัว ดังนั้นการแก้ไข อาจจะทำให้โดยการเปลี่ยน \bar{y} ให้เป็น y ได้คือถ้ามี \bar{y} หรือ \bar{y} ในตำแหน่งที่ g ให้ตัด \bar{y} ทิ้งแล้วใส่ y ไปในตำแหน่งที่ g บวก 2 (แทรกหลังพจน์ขณะที่ตามหลัง \bar{y} หรือ \bar{y}) แต่ถ้าในตำแหน่งที่ g บวก 2 เป็น y อยู่แล้วก็ไม่ต้องใส่ y อีกและเมื่อใส่ y ไปในตำแหน่งที่ g บวก 2 แล้ว ตำแหน่งที่ g บวก 2 เติมก็เลื่อนไปเป็นตำแหน่งที่ g บวก 3 และตัวถัดไปก็เลื่อนไปอีกเรื่อย ๆ ยกเว้นตำแหน่งที่ g บวก 1 ที่ยังคงเติม

ปัญหาที่ตามมาอีกก็คือ ทำไมไม่เปลี่ยน y ให้เป็น \bar{y} ซึ่งก็น่าจะใช้ได้เหมือนกัน เหตุที่ไม่เปลี่ยน y ให้เป็น \bar{y} ก็เพราะว่า y ไม่ใช้ค่าหลัง ทุกกรณีซึ่งเมื่อพบ y จะต้องเสียเวลาตรวจสอบย้อนหลังทุกตัว และถ้าจะเปลี่ยนจาก \bar{y} ให้เป็น \bar{y} หรือถ้าพบ \bar{y} ให้ตัดทิ้งเลยก็จะเกิดปัญหาในกรณีที่ \bar{y} เป็น \bar{y} และ \bar{y} จะได้รหัสต่างกัน จากเหตุผลเหล่านี้ จึงไม่ควรใช้วิธีเปลี่ยนจาก \bar{y} เป็น \bar{y} หรือ \bar{y}

ตัวอย่างการตรวจสอบ \bar{y} \bar{y} \bar{y} และ \bar{y} เช่น

กฎใน

ตำแหน่งที่ g คือ \bar{y}

ตำแหน่งที่ g บวก 1 คือ n

ตำแหน่งที่ g บวก 2 คือ y

เพราะฉะนั้นไม่ต้องเพิ่ม y เข้าไป นั่นคือเขียนได้เหมือนเดิม แล้วค่อยเปลี่ยนไปเป็นรหัสตามการจับกลุ่มต่าง ๆ ที่ได้จับไว้

กฎใน

ตำแหน่งที่ g คือ \bar{y}

ตำแหน่งที่ g บวก 1 คือ n

ตำแหน่งที่ g บวก 2 คือ ที่ว่างซึ่งไม่ใช่ y

เพราะฉะนั้นเพิ่ม y เข้าไปในตำแหน่งที่ g บวก 2

และเลื่อนตำแหน่งที่ ก บวก 2 ไปเป็นตำแหน่งที่ ก บวก 3 และตัวอื่น ๆ ก็เลื่อนตามออกไป ซึ่งในกรณีนี้เป็นที่ว่างหมด เมื่อเปลี่ยนแล้วจะไ้เป็น กวไนย แล้วค่อยเปลี่ยนเป็นรหัสตามการจัดกลุ่มต่าง ๆ ที่ไ้จัดไว้

2.2.3 กรณีที่ใช้ -ัม ำ -ำม และ -รรม

กรณีที่ใช้ -ัม ำ -ำม -รรม ซึ่งมีเสียงเดียวกัน สำหรับกรณี -รรม จะกล่าวถึงในภายหลังรวมกับการแก้ปัญหากรณี รร การใช้ -ัม ำ -ำม เช่น อัมพร อำพร อมฤต อำมฤต คัมภีร์ คำภีร์ อมรินทร์ และ อำมรินทร์ จะเห็นไ้ว่า ถ้าไม่มีการแก้ไขแล้ว การใช้ -ัม และ -ำม จะไ้รหัสเหมือนกัน แต่ต่างจากการใช้ -ำ ซึ่งตามปกติแล้วถ้าใช้ -ำ และ -ัม ควรจะไ้รหัสเหมือนกันเพราะมีเสียงเหมือนกัน

วิธีแก้ไข

เพื่อที่จะให้ทั้ง 3 กรณีไ้รหัสเหมือนกัน จะแก้ไขโดย ทดสอบดูว่า ตัวที่ตามหลัง ำ เป็น ม หรือไม่ ถ้าเป็น ม ก็ให้ตัด ำ ทิ้ง ถ้าหลัง ำ ไม่ใช่ ม ให้เปลี่ยน ำ เป็น ม ถ้าหลัง ำ เป็น ม และหลัง ม เป็นสระก็ให้เปลี่ยน ำ เป็น ม เช่นเดียวกัน

2.2.4 กรณีที่ใช้ รร ำ และ -ัน

เนื่องจากว่า รร ใช้แทน -ัน ในกรณีที่ไม่มีความสะกดตามหลัง แต่ถ้าหลัง รร มีความสะกดก็ใช้แทนด้วย ำ ไ้ เช่น จรรยา วรรณา พรรณที่ ริงสรรค์ ขรรมร์คน

วิธีแก้ไข

ในการเปลี่ยนจะทำได้โดยการพิจารณาว่ามีตัวสะกดหลัง รร หรือไม่ซึ่งพิจารณากังนี้

- 1) ถ้าหลัง รร เป็นพยัญชนะและหลังพยัญชนะนั้นเป็น สระ หรือวรรณยุกต์ ให้เปลี่ยน รร เป็น ัน เช่น จรรยา รังสรรค์ วรรณ
- 2) ถ้าหลัง รร เป็นสระพวก เ แ ใ โ ก็ให้เปลี่ยน รร เป็น ัน เช่น สรรเพชร
- 3) ถ้าหลัง รร เป็นพยัญชนะและหลังพยัญชนะนั้นเป็นพยัญชนะอีกหรือเป็นสระพวก เ แ ใ โ ก็ให้เปลี่ยน รร เป็น ัน เช่น พรรณทิชรรมรคัน พรรณวไล
- 4) ถ้าหลัง รร ไม่เข้ากฎเกณฑ์ทั้ง 3 กรณีข้างต้นให้เปลี่ยน รร เป็น ัน

2.2.5 กรณีที่มีตัวการันต์

กรณีที่มีตัวการันต์ ถึงแม้ว่าเราจะคัดพยัญชนะที่มีตัวการันต์กำกับทั้ง แต่ก็ยังมีปัญหา เรื่องตัวควบกล้ำ เช่น พระลักษณ และ พระลักษมณ ถึงแม้ว่าจะคัดตัวที่มีการันต์กำกับออก คือ ณ ก็ยังคงได้รหัสที่ต่างกัน หรือ จันทรโศ และ จันโต ถึงแม้ว่าจะคัด ร์ ในตัวแรกออกก็ยังคงได้รหัสต่างจาก จันโต และยังมีชื่ออื่นอีก เช่น กนกกาญจน์ และ กนกกาญญ กรณี และ กร

วิธีแก้ไข

- 1) โดยการตรวจสอบว่าตัวที่มีการันต์กำกับ เป็น น ณ ท ร หรือเปล่า
 - ถ้าเป็น น และนำหน้าด้วย จ ให้คัด จน ึ่ง
 - ถ้าเป็น ณ และนำหน้าด้วย ม ให้คัด มณ ึ่ง

ถ้าเป็น ท และนำหน้าด้วย ณ ให้ตัด ณท ทั้ง
 ถ้าเป็น ร และนำหน้าด้วย ท ให้ตัด ทร ทั้ง
 ถ้าเป็น ร และนำหน้าด้วย ศ ให้ตัด ตร ทั้ง

2) เนื่องจากสระต่าง ๆ เราไม่นำมาใช้ในการลงรหัส เป็นส่วนใหญ่
 ดังนั้น ตัวสระที่มีอาร์บ์ นั้นเราจะต้องตัดพยัญชนะที่อยู่หน้าสระนั้น ๆ ทั้งไปด้วย เช่น
 จักรวรรค์ จะต้องตัด กิ ทั้งไป นั่นคือ ไม่นำพยัญชนะและสระที่มีตัวอาร์บ์กำกับ
 มาคิดในการกำหนดรหัส จะตัดทิ้งไปเลย

2.3 การจํกกลุ่มพยัญชนะไทย

การจํกกลุ่มพยัญชนะไทยนั้น ถือหลักในการพิจารณาครั้งนี้ คือ

- 1) พยายามจํกกลุ่มพยัญชนะเพื่อให้โอกาสที่เกิกรหัสซ้ำซ้อนกันน้อยที่สุด หรือ
 ไม่เกิดขึ้นเลย นั่นคือ ถ้าชื่อที่ออกเสียงต่างกันแล้ว ไม่ควรที่จะไกรหัสเหมือนกัน
- 2) พยายามจํกกลุ่มเพื่อให้ชื่อต่าง ๆ ที่ออกเสียงเหมือนกันไกรหัสเดียวกันมาก
 ที่สุด ถึงแม้ว่าจะเขียนต่างกันก็ตาม

จากการพิจารณาจํกกลุ่มพยัญชนะไทย จะแบ่งการจํกกลุ่มไค้ดังนี้

2.3.1 การจํกกลุ่มที่ 1

เป็นการจํกกลุ่มโดยยึดคามพยัญชนะซึ่งเป็นตัวสะกดในแม่เดียวกันของพยัญชนะไทย
 เป็นหลักในการจํกเป็นส่วนใหญ่ (แต่ไม่ไค้จํกแบ่งตามตัวสะกดในแม่ต่าง ๆ ทั้งสิ้น) ซึ่งตัว
 สะกดในแม่ต่าง ๆ ที่นำมาใช้ในการพิจารณามีดังนี้

แม่กง มีพยัญชนะสะกด คือ ง
 แม่กน มีพยัญชนะสะกด คือ ณ ฌ น ร ล ท
 แม่กม มีพยัญชนะสะกด คือ ม
 แม่เกย มีพยัญชนะสะกด คือ ย

แม่กข = ข ฃ ฅ ฆ ๗ ๘ ๙

๑ กอ

๒

กค

ก ฃ ค ฅ

กด

ค ฃ ฅ ๗ ๘ ๙ ๗ ๘ ๙ ๗ ๘ ๙
 ท. ๗ ๘ ๙ ๗ ๘ ๙

แม่เกอว	มีพยัญชนะสะกด	คือ	ว				
แมกก	มีพยัญชนะสะกด	คือ	ก	ข	ค	ฌ	
แมกค	มีพยัญชนะสะกด	คือ	จ	ช	ฌ	ฉ	ฉ
			ฎ	ฏ	ฐ	ฑ	ฒ
			ด	ต	ถ	ด	ด
แมกบ	มีพยัญชนะสะกด	คือ	บ	ป	พ	ฟ	ภ

คงได้กล่าวข้างต้นแล้วว่า ในการจัดกลุ่มเราไม่ได้จัดแบ่งตามตัวสะกดในแม่ต่าง ๆ ทั้งสิ้น ก็เพราะว่าตัวสะกดบางแม่ เช่น แมกก มีขนาดของกลุ่มใหญ่เกินไป ถ้าจัดไว้ในกลุ่มเดียวกันแล้ว จะทำให้ไทรหัสซ้ำซ้อนกันมากขึ้น ดังนั้นจึงได้แบ่งตัวสะกดในแมกก ออกเป็น 3 กลุ่ม แต่ในบางแม่มีขนาดไม่ใหญ่เกินไป ก็สามารถยกมาเป็นกลุ่มพยัญชนะกลุ่มหนึ่งได้เลย เช่น แมกค สำหรับในบางแม่ เช่น แม่กง กม เกย เกอว มีขนาดเล็กเกินไป ถ้าจัดไว้เป็นแต่ละกลุ่ม ก็จะให้มีกลุ่มพยัญชนะมากขึ้น ทำให้ความยาวในการลงรหัสของคำในภาษาไทยเพิ่มขึ้น เพราะแทนที่จะใช้ตัวเลขหลักเดียวแทนพยัญชนะแต่ละตัว อาจจะต้องใช้ 2 หลักจึงจะพอ ทั้งนี้เป็นต้น ซึ่งจะทำให้สิ้นเปลืองค่าใช้จ่าย จึงต้องนำมารวมกัน เพื่อจะได้ใช้กลุ่มพยัญชนะน้อยกลุ่มลง ซึ่งในการพิจารณาตามที่กล่าวข้างต้น จะจัดกลุ่มได้ดังนี้

กลุ่มที่แทนควย	0	มี	ม	ว			
กลุ่มที่แทนควย	1	มี	ก	ข	ค	ฌ	
กลุ่มที่แทนควย	2	มี	ง	ย			
กลุ่มที่แทนควย	3	มี	ญ	ณ	น		
กลุ่มที่แทนควย	4	มี	ฎ	ฏ	ฐ	ฑ	ฒ
กลุ่มที่แทนควย	5	มี	ด	ต	ถ	ด	ด
กลุ่มที่แทนควย	6	มี	บ	ป	พ	ฟ	ภ
กลุ่มที่แทนควย	6	มี	ด	ด	ฟ	ห	อ

กลุ่มที่แทนควย	7	มี	จ	ฉ	ช	ซ	ฅ
กลุ่มที่แทนควย	8	มี	ฐ	ฑ	ฒ	ณ	ด
กลุ่มที่แทนควย	9	มี	ร	ฤ	ล	ฬ	พ

2.3.2 การจกกลุ่มที่ 2

เนื่องจากว่าในการจกกลุ่มที่ 1 นั้นจะเห็นว่าจกกลุ่ม เฉพาะพยัญชนะสะกหรือ
ตัวตามเท่านั้น แต่ไม่ได้จกกลุ่มพยัญชนะต้นคำ ซึ่งจากการพิจารณาตามพยัญชนะไทย 44
รูปซึ่งมี 21 เสียงดังนี้คือ

- 1) เสียง ก มีรูปพยัญชนะ ก
- 2) เสียง ค มีรูปพยัญชนะ ช ซ ฅ ฌ ฆ
- 3) เสียง ง มีรูปพยัญชนะ ง
- 4) เสียง จ มีรูปพยัญชนะ จ
- 5) เสียง ช มีรูปพยัญชนะ ช ฅ ฉ
- 6) เสียง ซ มีรูปพยัญชนะ ซ ฌ ฅ ฆ
- 7) เสียง ก มีรูปพยัญชนะ ก ฌ
- 8) เสียง ค มีรูปพยัญชนะ ค ฌ
- 9) เสียง ท มีรูปพยัญชนะ ท ฑ ฒ ฌ ณ ฐ
- 10) เสียง น มีรูปพยัญชนะ ณ น
- 11) เสียง บ มีรูปพยัญชนะ บ
- 12) เสียง ป มีรูปพยัญชนะ ป
- 13) เสียง พ มีรูปพยัญชนะ พ ฬ ฌ
- 14) เสียง ฟ มีรูปพยัญชนะ ฟ ฝ
- 15) เสียง ม มีรูปพยัญชนะ ม
- 16) เสียง ย มีรูปพยัญชนะ ฌ ย

- 17) เสียง ร มีรูปพยัญชนะ ร
 18) เสียง ล มีรูปพยัญชนะ ล ท
 19) เสียง ว มีรูปพยัญชนะ ว
 20) เสียง อ มีรูปพยัญชนะ อ
 21) เสียง ฮ มีรูปพยัญชนะ ห ฮ

ซึ่งจากการพิจารณาเสียงของพยัญชนะ 21 เสียงนั้น จะเห็นได้ว่ามีพยัญชนะ
 ๓ ๓ ซึ่งจะนำมาจัดกลุ่มในพยัญชนะต้นได้ และบางเสียงอาจจะจัดเข้าด้วยกันได้ เนื่องจาก
 จากในภาษาไทย มีพยัญชนะต้นที่ออกเสียงเหมือนกันหลายกลุ่ม ถ้าใช้ต่างกันก็จะทำให้ได้
 รหัสนิคพลาทได้ ดังนั้นจึงได้นำเสียงของพยัญชนะดังกล่าวข้างต้นมาจัดกลุ่มเป็นพยัญชนะต้น
 ควบ ใ้ด้วยกันคือ

- | | | | | | | | |
|-------------------------|---|-----|---|---|---|---|---|
| 1) พยัญชนะต้นที่แทนควย | ก | คือ | ก | | | | |
| 2) พยัญชนะต้นที่แทนควย | ข | คือ | ข | ช | ค | ฅ | ฆ |
| 3) พยัญชนะต้นที่แทนควย | ง | คือ | ง | | | | |
| 4) พยัญชนะต้นที่แทนควย | จ | คือ | จ | | | | |
| 5) พยัญชนะต้นที่แทนควย | ช | คือ | ช | ฅ | ฉ | | |
| 6) พยัญชนะต้นที่แทนควย | ส | คือ | ซ | ศ | ษ | ส | |
| 7) พยัญชนะต้นที่แทนควย | ค | คือ | ก | ฅ | | | |
| 8) พยัญชนะต้นที่แทนควย | ค | คือ | ค | ฅ | | | |
| 9) พยัญชนะต้นที่แทนควย | ท | คือ | ฐ | ฑ | ฒ | ณ | ถ |
| 10) พยัญชนะต้นที่แทนควย | น | คือ | ณ | น | | | |
| 11) พยัญชนะต้นที่แทนควย | บ | คือ | บ | | | | |
| 12) พยัญชนะต้นที่แทนควย | ป | คือ | ป | | | | |
| 13) พยัญชนะต้นที่แทนควย | พ | คือ | พ | ภ | ผ | | |

- 14) พยัญชนะคนที่แทนควย ฟ คือ ฝ ฟ
 15) พยัญชนะคนที่แทนควย ม คือ ม
 16) พยัญชนะคนที่แทนควย ย คือ ญ ย
 17) พยัญชนะคนที่แทนควย ร คือ ร ล พ ฤ ฦ
 18) พยัญชนะคนที่แทนควย ว คือ ว
 19) พยัญชนะคนที่แทนควย อ คือ อ
 20) พยัญชนะคนที่แทนควย ฮ คือ ฮ

สำหรับพยัญชนะตัวตามอื่น ๆ จักกลุ่มไว้ดังนี้

กลุ่มที่แทนควย	0	มี	ม	ว	ำ				
กลุ่มที่แทนควย	1	มี	ก	ข	ช	ค	ศ	ฆ	
กลุ่มที่แทนควย	2	มี	ง	ย					
กลุ่มที่แทนควย	3	มี	ญ	ณ	น				
กลุ่มที่แทนควย	4	มี	ฎ	ฏ	ก	ค	ศ	ฆ	ส
กลุ่มที่แทนควย	5	มี	บ	ป	พ	ภ			
กลุ่มที่แทนควย	6	มี	ฝ	ฝ	ฟ	ห	อ	ฮ	
กลุ่มที่แทนควย	7	มี	จ	ฉ	ช	ซ	ฌ		
กลุ่มที่แทนควย	8	มี	ฐ	ฑ	ฒ	ณ	ท	ธ	
กลุ่มที่แทนควย	9	มี	ร	ฤ	ล	ฦ			

2.4 หลักเกณฑ์และข้อกำหนดต่าง ๆ ในการให้รหัส

ในวิธีการสร้างรหัสให้กับชื่อหรือคำพ้องเสียงในภาษาไทย ควรคำนึงถึงหลักใหญ่ ๆ 2 ประการคือ หลักการสร้างรหัส และ หลักการกำหนดขนาดของรหัส หลักการดังกล่าวนี้จะช่วยให้รหัสตามเสียงที่โครงการและมีประสิทธิภาพเหมาะสมกับงานที่จะทำ

2.4.1 หลักการสร้างรหัส

หลักเกณฑ์และข้อกำหนดส่วนใหญ่ได้มาจากการศึกษาเปลี่ยนแปลงหลักเกณฑ์ของชาวคึกซ์ในภาษาอังกฤษ มีการแก้ไขและเพิ่มเติมบางส่วนเพื่อให้เหมาะสมกับภาษาไทย หลักเกณฑ์และข้อกำหนดต่าง ๆ มีดังนี้

1) พยัญชนะตัวแรกที่ปรากฏในชื่อจะเป็นพยัญชนะตัวแรกของรหัส (ซึ่งต่างจากชาวคึกซ์ในภาษาอังกฤษ เพราะในภาษาอังกฤษจะใช้อักษรตัวแรกที่ปรากฏในชื่อเป็นตัวแรกของรหัสเลย ไม่ว่าอักษรตัวแรกนั้นจะเป็นพยัญชนะหรือสระก็ตาม) หรือบางกรณีอาจจะมีการจัดกลุ่มพยัญชนะที่มีเสียงพยัญชนะต้นเดียวกันไว้ด้วยกัน และใช้แทนด้วยพยัญชนะเพียงตัวเดียวเท่านั้นในแต่ละกลุ่ม ส่วนตัวที่ไม่สามารถเข้ากลุ่มใดก็ได้ใช้พยัญชนะตัวนั้นเลย (เช่นในการจัดกลุ่มที่ 2)

เหตุที่มีการจัดกลุ่มพยัญชนะต้นด้วยเนื่องจากว่าในภาษาไทยนั้นมีพยัญชนะที่มีเสียงเดียวกันและอาจใช้แทนกันได้ ดังนั้นถ้าเขียนต่างกันโดยพยัญชนะต้นแล้วก็อาจทำให้การค้นหาข้อมูลนั้นล้มเหลวได้เนื่องจากตัวแรกของรหัสต่างกัน เช่น ยคา และ ญคา จะเห็นว่า ยคา ไคร้รหัส ย100 ส่วน ญคา ไคร้รหัส ญ100 (สมมุติว่าใช้ขนาดของรหัสเป็น 4) ดังนั้นถ้าเขียน ญคา เป็น ยคา แล้วจะทำให้การค้นหาข้อมูลนั้นล้มเหลวได้ถึงแม้ว่าจะมีเสียงเดียวกันก็ตาม แต่ถ้ามมีการจัดกลุ่มพยัญชนะต้นด้วยก็จะหลีกเลี่ยงปัญหานี้ได้

2) พยัญชนะตัวต่อ ๆ มาให้ยึดหลักการให้รหัสตามการจัดกลุ่มที่จัดไว้โดยนำรหัสแต่ละตัวเข้าแทนที่พยัญชนะต่าง ๆ ของชื่อ

- 3) ให้ตัดพยัญชนะและสระที่มีตัวการันต์กำกับทิ้ง
- 4) หากมีพยัญชนะไม่พอ ให้ใช้ ๐ เติมท้ายรหัสจนได้ความยาวรหัสตาม

ที่ต้องการ

- 5) ปกติจะตัดวรรณยุกต์ทุกตัว และตัวสระทุกตัวทิ้ง นอกจากนี้บางกรณีก็อาจจะมีการนำสระมาจัดกลุ่มด้วย

สมมุติว่าถ้ากำหนดให้ขนาดของรหัสเป็น 7 แล้วจะได้รูปแบบของรหัส
ชวาค์เด็กซ์ดังนี้

$$(\text{รหัสชวาค์เด็กซ์}) = (\text{ตัวอักษร})(\text{ตัวเลข})(\text{ตัวเลข})(\text{ตัวเลข})(\text{ตัวเลข})(\text{ตัวเลข})(\text{ตัวเลข})$$

2.4.2 การกำหนดขนาดของรหัส

การกำหนดขนาดของรหัส ต้องพิจารณาว่า ความยาวรหัสเท่าใ้จะให้ความถูกต้อง และมีประสิทธิภาพมากที่สุดในการใช้งาน การพิจารณาควรคำนึงถึงหลักดังต่อไปนี้

- 1) การใช้ชื่ออย่างเกี่ยว การที่จะใช้รหัสมีความยาวเท่าไรนั้นต้องพิจารณาว่าปกติความยาวของชื่อคนไทย มีขนาดเฉลี่ยเท่าไรเมื่อไม่นำสระมาคิดด้วย (ซึ่งอาจได้ขนาดเป็นช่วง ๆ) ก็ควรใช้รหัสตามค่าประมาณในช่วงนั้น ๆ เพราะถ้าใช้รหัสสั้นไป ก็จะทำให้ไ้รหัสซ้ำกันเป็นจำนวนมาก หรืออาจจะใช้วิธีการทางสถิติอื่น ๆ เพื่อหาขนาดของรหัสที่เหมาะสมก็จะได้กล่าวในบทต่อไป

- 2) การจะใช้ชื่อ ร่วมกับ นามสกุล ให้รหัสต่อกันไปเลยซึ่งในการให้รหัสในกรณีนี้ อาจทำได้เป็น 3 กรณี คือ

- ก) ให้รหัสทั้งชื่อและนามสกุลต่อกันไปเลย
 - ข) ให้รหัสแก่ชื่อและมีช่องว่าง 1 ช่องก่อนให้รหัสแก่นามสกุลต่อ
 - ค) กำหนดขนาดของรหัสแก่ชื่อ และขนาดของรหัสแก่นามสกุลเลย
- โดยถ้าชื่อไม่พอก็ให้ใส่ ๐ จนครบขนาดรหัสของชื่อ แล้วจึงต่อกับรหัสของนามสกุล ถ้านามสกุลไม่พอก็ให้ใส่ ๐ จนครบขนาดรหัสของนามสกุล

ซึ่งในการวิจัยครั้งนี้จะใช้รหัสทั้งชื่อและนามสกุลต่อกันไปเลย

3) ค่าใช้จ่าย เพราะถ้าขนาดของรหัสยาวมาก ถึงแม้ว่าจะมีความผิดพลาดจากการซ้ำซ้อนน้อยลง แต่จะมีความผิดพลาดจากการค้นหาข้อมูลไม่พบมากขึ้น ดังนั้น ต้องหาขนาดของรหัสที่ทำให้ได้ค่าอัตราประโยชน์สูงสุด (OPTIMUM) ในทั้ง 2 กรณี ในขณะเดียวกัน ก็ต้องพิจารณาด้วยว่าค่าที่ให้อัตราประโยชน์สูงสุดนี้คุ้มกับค่าใช้จ่ายที่เพิ่มขึ้นเมื่อขนาดของรหัสยาวขึ้นหรือไม่

4) ลักษณะการใช้งานว่า รหัสยาวเท่าไร เหมาะสมกับการใช้งานแบบไหน คือพิจารณาว่างานนั้น ๆ ต้องการความถูกต้องมากน้อยแค่ไหน ถ้าต้องการความถูกต้องชนิดที่ไม่สนใจว่าข้อมูลที่จะออกมาจะมากน้อยแค่ไหน คือให้มีโอกาสที่ค้นไม่พบข้อมูลน้อย ก็ต้องให้ขนาดของรหัสสั้นลงซึ่งจะทำให้มีการซ้ำซ้อนของข้อมูลมากก็ตาม