



วิธีดำเนินการวิจัย

3.1 ที่มาของข้อมูล

ในการศึกษาการกระจายเชิงสถิติของหน่วยเสียงพยัญชนะ สระ และวรรณยุกต์ ในพยางค์ที่ปรากฏเป็นคำและส่วนประกอบของคำในภาษาไทย ผู้วิจัยใช้ฐานข้อมูลจากงานวิจัยเรื่องระบบการสังเคราะห์เสียงจากข้อความภาษาไทยของสุคาพร ลักษณะินาวิน(1989) ซึ่งฐานข้อมูลที่นำมาศึกษานั้นสร้างขึ้นมาจากคอมพิวเตอร์ โปรแกรม DBase III Plus โดยการนำเอาหน่วยเสียงที่มีจำนวนจำกัดในภาษาไทย ได้แก่ หน่วยเสียงพยัญชนะ 33 หน่วย หน่วยเสียงสระ 24 หน่วย และหน่วยเสียงวรรณยุกต์ 5 หน่วย มาจัดเรียงตามโครงสร้างพยางค์ด้วยกฎเกณฑ์การปรากฏร่วมกันของเสียง ทำให้สร้างพยางค์ที่มีศักยภาพ ที่จะเกิดใช้ในภาษาไทยได้จำนวนหนึ่ง ซึ่งจำนวนดังกล่าวเป็นจำนวนที่แน่นอนตายตัว คือ 26,928 พยางค์ โดยมีผู้ช่วยวิจัยที่มีความรู้ภาษาไทยจำนวน 4 คน ตรวจสอบและระบุว่าพยางค์แต่ละพยางค์ ที่เกิดจากการสร้างด้วยโปรแกรมคอมพิวเตอร์นั้น มีใช้ในภาษาหรือไม่ ซึ่งถ้าหากพยางค์ดังกล่าวเกิดขึ้นในภาษาก็จะระบุว่า เป็นพยางค์ประเภทที่เป็นคำ หรือเป็นส่วนประกอบของคำ มีรูปเขียนและหน้าที่อย่างไรในภาษา พยางค์ทั้ง 26,928 พยางค์ สามารถแบ่งได้ตามโครงสร้างพยางค์ประเภทต่างๆ ดังนี้คือ

ก. พยางค์เปิด ซึ่งประกอบด้วย หน่วยเสียงพยัญชนะ 33 หน่วย หน่วยเสียงสระเสียงยาว 12 หน่วย และหน่วยเสียงวรรณยุกต์ทั้ง 5 หน่วย

$$33 C_1 \text{ \& } C_2 \times 12 V_1 \text{ \& } V_2 \times 5 T = 1,980 \text{ พยางค์}$$

แต่พบว่า มีหน่วยเสียงพยัญชนะต้น 2 หน่วย คือ / kw / และ / k'w / ที่ไม่สามารถปรากฏร่วมกับสระกลมหั่ง 4 หน่วย คือ / u:, o:, @ :, u:a / ได้

$$2 C_2 \times 4 V_1 \text{ \& } V_2 \times 5 T = 40 \text{ พยางค์}$$

พยางค์ทั้ง 40 พยางค์นี้ไม่สามารถปรากฏได้ในภาษา ดังนั้น พยางค์เปิดจึงมีจำนวนทั้งสิ้น 1,940 พยางค์

ข. พยางค์เสียงสั้นที่มีเสียงพยัญชนะท้ายเป็นพยัญชนะไม่กัก ประกอบด้วยหน่วยเสียงพยัญชนะต้น 33 หน่วย หน่วยเสียงสระเสียงสั้น 12 หน่วย หน่วยเสียงพยัญชนะท้ายที่เป็นพยัญชนะไม่กัก 5 หน่วย และหน่วยเสียงวรรณยุกต์ 5 หน่วย

$$33 C_1^l \& C_2^l \times 12 V_1 \& V_2 \times 5 C_1^f \times 5 T = 9,900 \text{ พยางค์}$$

แต่มีพยางค์บางพยางค์ ที่ไม่สามารถปรากฏได้ในภาษาเพราะการไม่ปรากฏร่วมกันของหน่วยเสียงบางหน่วย ได้แก่ การไม่ปรากฏร่วมกันของหน่วยเสียงพยัญชนะต้น /kw/ และ /k'w/ กับหน่วยเสียงสระกลม / u, o, ɔ, ua / ทั้ง 4 หน่วย และหน่วยเสียงพยัญชนะท้าย / m, n, N, j / ทั้ง 4 หน่วย

$$2 C_2^l \times 4 V_1 \& V_2 \times 4 C_1^f \times 5 T = 160 \text{ พยางค์}$$

การไม่ปรากฏร่วมกันของหน่วยเสียงสระกลมทั้ง 4 หน่วย กับหน่วยเสียงพยัญชนะท้าย / w /

$$33 C_1^l \& C_2^l \times 4 V_1 \& V_2 \times 1 C_1^f \times 5 T = 660 \text{ พยางค์}$$

และการไม่ปรากฏร่วมกันของสระหน้า 4 หน่วย / i, e, x, ia / กับหน่วยเสียงพยัญชนะท้าย / j /

$$33 C_1^l \& C_2^l \times 4 V_1 \& V_2 \times 1 C_1^f \times 5 T = 660 \text{ พยางค์}$$

พยางค์ทั้ง 1,480 พยางค์นี้ไม่สามารถปรากฏใช้ในภาษาได้ ดังนั้นพยางค์เสียงสั้นที่มีพยัญชนะไม่กักท้ายพยางค์จึงมีจำนวนทั้งสิ้น 8,420 พยางค์

ค. พยางค์เสียงยาวที่มีเสียงพยัญชนะท้ายเป็นพยัญชนะไม่กัก ประกอบด้วยหน่วยเสียงพยัญชนะต้น 33 หน่วย หน่วยเสียงสระเสียงยาว 12 หน่วย หน่วยเสียงพยัญชนะท้ายที่เป็นพยัญชนะไม่กัก 5 หน่วย และหน่วยเสียงวรรณยุกต์ 5 หน่วย

$$33 C_1^l \& C_2^l \times 12 V_1 \& V_2 \times 5 C_1^f \times 5 T = 9,900 \text{ พยางค์}$$

แต่พบว่า มีพยางค์ 1,480 พยางค์ ที่ไม่สามารถปรากฏได้ในภาษา เนื่องจากการไม่ปรากฏร่วมกันของหน่วยเสียงบางหน่วย ดังที่ปรากฏในพยางค์เสียงสั้นที่มีพยัญชนะท้ายเป็นพยัญชนะไม่กัก ได้แก่ การไม่ปรากฏร่วมกันของหน่วยเสียงพยัญชนะต้น / kw / และ / kʰw / กับหน่วยเสียงสระกลม 4 หน่วย / u:, o:, @:, u:a / และหน่วยเสียงพยัญชนะท้ายที่เป็นพยัญชนะเสียงก้อง 4 หน่วย / m, n, N, j / รวมทั้งสิ้น 160 พยางค์ การไม่ปรากฏร่วมกันของหน่วยเสียงสระกลม 4 หน่วย กับหน่วยเสียงพยัญชนะท้าย / w / รวมทั้งสิ้น 660 พยางค์ และการไม่ปรากฏร่วมกันของหน่วยเสียงสระหน้า 4 หน่วย / i:, e:, x:, i:a / กับหน่วยเสียงพยัญชนะท้าย / j / รวมทั้งสิ้น 660 พยางค์ ดังนั้นพยางค์เสียงยาวที่มีพยัญชนะท้ายเป็นพยัญชนะเสียงก้อง จึงมีจำนวน 8,420 พยางค์

ง. พยางค์เสียงสั้นที่มีพยัญชนะท้ายเป็นพยัญชนะกัก ประกอบด้วยหน่วยเสียงพยัญชนะต้น 33 หน่วย หน่วยเสียงสระเสียงสั้น 12 หน่วย หน่วยเสียงที่เป็นพยัญชนะเสียงกัก 4 หน่วย คือ / p, t, k, ? / และหน่วยเสียงวรรณยุกต์ 3 หน่วย คือ วรรณยุกต์ต่ำหรือวรรณยุกต์เอก (1) วรรณยุกต์ขึ้น-ตก หรือวรรณยุกต์โท (2) และวรรณยุกต์สูง หรือวรรณยุกต์ตรี (3)

$$33 C_1' \& C_2' \times 12 V_1 \& V_2 \times 4 C_1' \times 3 T = 4,752 \text{ พยางค์}$$

แต่มีพยางค์บางพยางค์ที่ไม่สามารถปรากฏในภาษาได้ เนื่องจากการไม่ปรากฏร่วมกันของหน่วยเสียงพยัญชนะต้น / kw / และ / kʰw / กับหน่วยเสียงสระกลม 4 หน่วย / u, o, @, ua /

$$2 C_2' \times 4 V_1 \& V_2 \times 4 C_1' \times 3 T = 96 \text{ พยางค์}$$

พยางค์ทั้ง 96 พยางค์ไม่สามารถปรากฏได้ในภาษา ดังนั้น พยางค์เสียงสั้นที่มีพยัญชนะกักท้ายพยางค์ จึงมีจำนวนทั้งสิ้น 4,656 พยางค์

จ. พยางค์เสียงยาวที่มีพยัญชนะกักท้ายพยางค์ ประกอบด้วยหน่วยเสียงพยัญชนะต้น 33 หน่วย หน่วยเสียงสระเสียงยาว 12 หน่วย หน่วยเสียงพยัญชนะท้ายที่เป็นเสียงกัก 3 หน่วย / p, t, k / และหน่วยเสียงวรรณยุกต์ 3 หน่วย คือ วรรณยุกต์ต่ำ วรรณยุกต์ขึ้น-ตก และวรรณยุกต์สูง

$$33 C_1' \& C_2' \times 12 V_1 \& V_2 \times 3 C_1' \times 3 T = 3,564 \text{ พยางค์}$$

แต่มีพยางค์บางพยางค์ที่ไม่สามารถปรากฏได้ในภาษา เนื่องจากการไม่ปรากฏร่วมกันของหน่วยเสียงพยัญชนะต้น / kw / และ / kʰw / กับหน่วยเสียงสระกลม 4 หน่วย / u:, o:, @:, u:a /

$$2 C_2' \times 4 V_1 & V_2 \times 3 C_1' \times 3 T = 72 \text{ พยางค์}$$

พยางค์ทั้ง 72 พยางค์ไม่สามารถปรากฏได้ในภาษา ดังนั้นพยางค์เสียงยาวที่มีพยัญชนะกักท้ายพยางค์จึงมีจำนวนทั้งสิ้น 3,492 พยางค์

เมื่อรวมจำนวนพยางค์ในโครงสร้างพยางค์ประเภทต่างๆ ในงานวิจัยเรื่องระบบการสังเคราะห์เสียงของสุคตาพร (Luksaneeyanawin, 1989, 1992) แล้ว พบว่ามีพยางค์ที่ไม่สามารถปรากฏได้รวม 3,168 พยางค์ มีพยางค์ที่มีศักยภาพของการปรากฏร่วมกันของหน่วยเสียงรวมทั้งสิ้น 26,928 พยางค์ และมีพยางค์ที่ใช้ในภาษาจริงๆ รวมทั้งสิ้น 5,912 พยางค์ คิดเป็น 22 % ของพยางค์ที่เกิดได้ทั้งหมด และมีพยางค์ที่ยังไม่เกิดขึ้นใช้เป็นคำในภาษาเป็นจำนวนถึง 21,016 พยางค์ คิดเป็น 78 % ของพยางค์ที่เกิดได้ทั้งหมด ซึ่งพบว่ามีจำนวนแตกต่างกันมากโดยที่พยางค์ที่เกิดเป็นคำใช้ในภาษามีจำนวนที่ปรากฏน้อยกว่าพยางค์ที่ยังไม่เกิดเป็นคำใช้ในภาษามากถึง 56 %

ฐานข้อมูลเรื่องพยางค์นี้ ได้มาจากผู้บอกภาษาจำนวน 5 คน ซึ่งรวมทั้งตัวผู้วิจัยที่เป็นหนึ่งในกลุ่มผู้ช่วยวิจัยในงานนั้น ผู้บอกภาษาจะเป็นผู้ระบุเรื่องต่างๆต่อไปนี้

ก. ระบุว่าพยางค์ที่ปรากฏนั้น มีชื่ออยู่ในภาษาหรือไม่ หากว่ามี ก็จะระบุต่อไปว่าปรากฏอยู่ในคำอะไรบ้าง โดยใช้สมรรถนะทางภาษาของผู้บอกภาษาเอง และอ้างอิงพจนานุกรมฉบับราชบัณฑิตยสถาน พ.ศ.2525

ข. ระบุว่าพยางค์ประเภทใด คือ พยางค์ที่เป็นคำ พยางค์ที่เป็นส่วนของคำ หรือพยางค์ที่เป็นคำพิเศษ

ค. ระบุว่าคำที่พยางค์นั้นๆปรากฏทำหน้าที่อะไรในภาษา คือ เป็นคำนาม คำสรรพนาม คำกริยา คำวิเศษณ์ ฯลฯ

จากการที่ผู้วิจัยและคณะผู้วิจัยตรวจสอบข้อมูลที่ปรากฏทั้งหมดอีกครั้งหนึ่ง เรื่องระบบการสังเคราะห์เสียงจากข้อความภาษาไทย (Luksaneeyanawin, 1992) พบว่า มีพยางค์ที่เป็นคำในภาษา 3,797 พยางค์ พยางค์ที่เป็นส่วนประกอบของคำ 1,061 พยางค์ และพยางค์ที่เป็นคำพิเศษอีก 1,054 พยางค์

จากจำนวนพยางค์ที่ใช้ในภาษา 5,912 พยางค์นี้เอง ที่ผู้วิจัยได้นำมาเป็นข้อมูล เพื่อศึกษาการกระจายของหน่วยเสียงต่างๆว่ามีการกระจายเป็นอย่างไร และมีรูปแบบการกระจายเป็นอย่างไร เนื่องจากผู้วิจัยสังเกตเห็นว่า หน่วยเสียงบางหน่วยมีการปรากฏมาก แต่บางหน่วยมีการปรากฏน้อยมาก และการปรากฏของหน่วยเสียงในพยางค์ต่างๆประเภทกันก็มีจำนวนไม่เท่ากัน ซึ่งรายละเอียดของการศึกษาวิเคราะห์การกระจายของหน่วยเสียงจะได้กล่าวถึงในบทต่อไป

### 3.2 การเก็บข้อมูล

ผู้วิจัยเก็บข้อมูลจากฐานข้อมูลดังที่ได้กล่าวไปแล้วใน 3.1 โดยนำฐานข้อมูลดังกล่าวมาตรวจสอบ เพิ่มเติมข้อมูลให้สมบูรณ์ยิ่งขึ้น ก่อนที่จะนำมาวิเคราะห์ เนื่องจากผู้บอกภาษาแต่ละคนมีคลังคำไม่เท่ากัน จึงต้องตรวจสอบข้อมูลเสียก่อน

ในการตรวจสอบและเพิ่มเติมข้อมูลให้สมบูรณ์มากขึ้น เพื่อนำข้อมูลที่ได้มาวิเคราะห์ ผู้วิจัยใช้สมรรถนะทางภาษาของผู้วิจัย และพจนานุกรม ฉบับราชบัณฑิตยสถาน พ.ศ.2525 เป็นแหล่งอ้างอิงที่มาของข้อมูล เมื่อผู้วิจัยพบพยางค์ที่สงสัย ผู้วิจัยจะถามตัวเองก่อนว่ามีพยางค์ดังกล่าวไหม ถ้ามี ผู้วิจัยจะระบุว่าปรากฏในคำอะไรบ้าง และมีรูปเขียนภาษาไทยเป็นอย่างไร นอกจากนี้ ผู้วิจัยยังค้นเพิ่มเติมจากพจนานุกรมว่า ยังมีในคำอะไรบ้าง และเป็นพยางค์ประเภทใด คำดังกล่าวมีหน้าที่อะไรในภาษา เช่น ผู้วิจัยพบว่า ในฐานข้อมูล พยางค์ / noːˈ / เป็น 0 แต่ผู้วิจัยพบว่าปรากฏในคำว่า < เสนอ > / saːˈnoːˈ / ซึ่งเป็นพยางค์ประเภท Y และมีหน้าที่เป็นคำกริยาในภาษาไทย หรือผู้วิจัยพบว่า ในฐานข้อมูล พยางค์ / haːˈ / ปรากฏในคำว่า < หา > เพียงคำเดียว เป็นพยางค์ประเภท X และมีหน้าที่เป็นคำกริยา แต่ผู้วิจัยพบว่าปรากฏในคำว่า < เวหา > / weːˈhaːˈ / ด้วย ซึ่งเป็นพยางค์ประเภท Y และมีหน้าที่เป็นคำนาม

เมื่อผู้วิจัยตรวจสอบพยางค์แต่ละพยางค์ ผู้วิจัยจะพิจารณาว่า นอกจากที่ปรากฏในฐานข้อมูลแล้ว ยังมีปรากฏในคำใดได้อีก และเป็นพยางค์ประเภทใด คำดังกล่าวมีหน้าที่อย่างไรในภาษา โดยอ้างอิงจากพจนานุกรม ซึ่งผู้วิจัยใช้เวลาในการเก็บและตรวจสอบข้อมูลนานประมาณ 3 เดือน

เมื่อตรวจสอบและเพิ่มเติมข้อมูลแล้ว ผู้วิจัยจึงเขียนโปรแกรมคอมพิวเตอร์ ซึ่งเป็นโปรแกรมดีเบส เพื่อให้คอมพิวเตอร์แยกเฉพาะข้อมูลที่ผู้วิจัยต้องการจะศึกษาออกมา คือ พยางค์ประเภท X พยางค์ประเภท Y และพยางค์ประเภท XY ซึ่งเกณฑ์ที่ผู้วิจัยใช้ในการแยกลักษณะของข้อมูลแตกต่างจากสุคดาพร (2533) เนื่องจากผู้วิจัยจะเลือกเก็บเฉพาะพยางค์ประเภท X ที่

ไม่สามารถปรากฏเป็นพยางค์ประเภทอื่นได้ พยางค์ประเภท Y ที่ไม่สามารถปรากฏเป็นพยางค์ประเภทอื่นได้ และพยางค์ประเภท XY ที่สามารถปรากฏเป็นทั้ง X และ Y ได้ในพยางค์เดียวกัน เพื่อศึกษารูปแบบการกระจายของหน่วยเสียงในพยางค์ทั้ง 3 ประเภทดังกล่าว

เพื่อเป็นการกำหนดและจำกัดขอบเขตของข้อมูลที่จะศึกษา จากจำนวนพยางค์ที่เกิดใช้ในภาษาไทยทั้งสิ้น 5,912 พยางค์ ผู้วิจัยเลือกเก็บเฉพาะพยางค์ 3 ประเภท ซึ่งมีจำนวนแตกต่างกัน ดังนี้ คือ

- 1) พยางค์ที่สามารถปรากฏเป็นคำโดยลำพังได้ (พยางค์ประเภท X) ซึ่งมีจำนวนทั้งสิ้น 1,817 พยางค์
- 2) พยางค์ที่ไม่สามารถปรากฏเป็นคำโดยลำพังได้ ต้องปรากฏร่วมกับพยางค์อื่นเสมอ คือเป็นพยางค์ที่เป็นส่วนประกอบของคำ (พยางค์ประเภท Y) ซึ่งมีจำนวนทั้งสิ้น 539 พยางค์
- 3) พยางค์ที่สามารถจะปรากฏเป็นคำโดยลำพัง และสามารถปรากฏเป็นส่วนประกอบของคำได้ (พยางค์ประเภท XY) มีจำนวนรวมทั้งสิ้น 375 พยางค์

พยางค์ที่นำมาศึกษาเรื่องการกระจายของหน่วยเสียงต่างๆ ในพยางค์ที่เป็นคำและส่วนประกอบของคำในภาษาไทย ในวิทยานิพนธ์ฉบับนี้ จำกัดเฉพาะพยางค์จำนวน 2,731 พยางค์นี้เท่านั้น ส่วนพยางค์อีก 3,181 พยางค์ซึ่งเกิดขึ้นในรูปพยางค์ประเภทอื่นๆไม่ได้นำมาศึกษา

Phonological หน่วยเสียง	Orthographical รูปเขียน	Morphological ประเภทของพยางค์	Grammatical Cat. หน้าที่ของคำ
1) co: <sup>0</sup>	เจอ	X	V
2) ni: <sup>1</sup>	ตระ <sup>1</sup> หนึ	_Y	V AJ
3) wa: <sup>0</sup>	วา	X	N CF
	<u>วา</u> จา		N
	นา <u>วา</u>	_Y	N

ตารางที่ 6 ตัวอย่างข้อมูลที่ใช้ในงานวิจัย พยางค์แต่ละประเภทที่ใช้ในงานวิจัย

- หมายเหตุ X คือ พยางค์ประเภทที่เป็นคำโดยลำพังได้  
 \_Y คือ พยางค์ที่เป็นส่วนประกอบอยู่ท้ายคำ  
 Y\_ คือ พยางค์ที่เป็นส่วนประกอบอยู่หน้าคำ  
 V คือ คำกริยา  
 AJ คือ คำคุณศัพท์  
 N คือ คำนาม  
 CF คือ คำลักษณะนาม

ในการเก็บข้อมูล ผู้วิจัยเลือกเก็บข้อมูลเฉพาะพยางค์ที่เป็นคำและส่วนประกอบของคำ คือ พยางค์ประเภท X Y และ XY โดยผู้วิจัยไม่ได้เก็บข้อมูลที่มีลักษณะอื่นร่วมด้วย เช่น /ke:<sup>0</sup>/ เป็นพยางค์ที่มีหลายลักษณะ คือ เป็นพยางค์ประเภท X ในคำว่า เก เป็นพยางค์ประเภท Y ในคำว่า เกเร และ ลิเก และเป็นพยางค์ประเภท Q ในคำว่า เกย ถ้าพยางค์มีลักษณะพยางค์ประเภท Q ร่วมอยู่ด้วย ผู้วิจัยจะไม่นำมาศึกษา เนื่องจากไม่ได้เป็นไปตามเกณฑ์ที่ผู้วิจัยกำหนดไว้

### 3.3 ลักษณะของข้อมูล

ลักษณะของข้อมูลที่น่ามาใช้ศึกษา ได้มาจากการสำรวจคลังคำของผู้พูดภาษาไทย มาตรฐาน รวมทั้งตัวผู้วิจัยด้วย จำนวน 5 คน ทุกคนเป็นผู้ที่มีการศึกษาระดับสูงกว่าปริญญาตรี ทั้งหมด ซึ่งแต่ละคนก็จะใช้สมรรถนะทางภาษาของตนในการระบุว่า พยางค์ดังกล่าวปรากฏในคำ ใดบ้าง และพยางค์ดังกล่าวเป็นพยางค์ประเภทใด และมีหน้าที่อย่างไรในภาษา โดยการอ้างอิง จากคำที่ปรากฏในพจนานุกรม ฉบับราชบัณฑิตยสถาน พุทธศักราช 2525 ด้วย ซึ่งคลังคำที่ได้ อาจจะไม่แตกต่างจากคลังคำของกลุ่มตัวอย่างอื่นๆ อย่างไรก็ตาม อาจถือได้ว่า คลังคำที่น่าใช้ใน การศึกษานี้เป็นคลังคำของผู้พูดที่มีการศึกษา

ลักษณะโครงสร้างของข้อมูลที่น่ามาใช้ศึกษามี 3 ลักษณะ ดังนี้ คือ

ก. พยางค์ประเภท X หรือ พยางค์ที่เป็นคำ เช่น

/ruːa / เรือ X N

ข. พยางค์ประเภท Y หรือ พยางค์ที่เป็นส่วนประกอบของคำ เช่น

/praː / ปรการ Y N

ปรการ Y V

/proː / ปรนเปรอ Y V

ค. พยางค์ประเภท XY หรือพยางค์ที่เป็นคำและส่วนประกอบของคำ เช่น

/roː / เรอ X V

เรอ Y V

เรอ Y V

### 3.4 ข้อจำกัดของข้อมูล

จากฐานข้อมูลที่ได้จากการนำเอาจำนวนหน่วยเสียงที่มีจำนวนจำกัดมาสร้างพยางค์โดย อาศัยกฎการสร้างพยางค์ ทำให้ได้พยางค์ที่มีจำนวนจำกัดที่แน่นอนตายตัว อันง่ายและสะดวก ในการศึกษาและหาข้อสรุป แต่ถ้าใช้จำนวนที่ไม่จำกัดมาศึกษา ก็อาจทำให้มีปัญหาและยากแก่การ สรุป



ในการศึกษาการกระจายของหน่วยเสียงแต่ละหน่วยเสียงในพยางค์ในภาษาไทย ผู้วิจัยจำกัดขอบเขตของข้อมูลที่ศึกษา โดยเลือกศึกษาเฉพาะพยางค์ที่มีลักษณะเป็นคำและส่วนประกอบของคำ ดังที่ได้กล่าวไปแล้วใน 3.2 พยางค์ลักษณะอื่นๆ ที่ไม่ได้นำมาศึกษามีลักษณะดังต่อไปนี้ (Luksaneeyanawin, 1989, 1992) คือ

ก. พยางค์ประเภท Q หรือพยางค์ที่เป็นคำพิเศษ เช่น

/t<sup>h</sup>u:a<sup>0</sup>/ ทัวร์ Q N, V

ข. พยางค์ประเภท XQ หรือพยางค์ที่เป็นคำและมีลักษณะที่เป็นคำพิเศษรวมอยู่ด้วย เช่น

/pi:a<sup>0</sup>/ เปีย X N, V

เปีย Q NP (ชื่อขนมอย่างหนึ่งเรียก ขนมเปีย

หรือขนมเปียะ...ในพจนานุกรมฉบับราชบัณฑิตยสถาน พ.ศ.2525 )

(หมายเหตุ NP คือ คำนามเฉพาะ)

ค. พยางค์ประเภท Y, Q หรือพยางค์ที่มีทั้งลักษณะที่เป็นส่วนประกอบของคำและคำพิเศษอยู่ด้วยกัน เช่น

/c<sup>h</sup>x:<sup>0</sup>/ แชร่ Q N, V

แซ่เขื่อน Y\_ V

ง. พยางค์ประเภท X, Y, Q คือ พยางค์ที่มีทั้ง 3 ลักษณะอยู่ด้วยกัน เช่น

/k<sup>h</sup>a:<sup>0</sup>/ คา X N, V

คา Q NP

คารม Y\_ N

คาร์บอน Q, Y\_ NP

จ. พยางค์ประเภท O หรือพยางค์ที่ถูกต้องตามกฎการสร้างพยางค์แต่ยังไม่เกิดเป็นคำในภาษา เช่น

/t<sup>h</sup>x:<sup>0</sup>/ แท O

### 3.5 การวิเคราะห์ข้อมูล

ผู้วิจัยนำฐานข้อมูลที่ได้จากการสำรวจคลังคำของผู้พูดภาษาไทยมาตรฐานมา และมีการตรวจสอบ และเพิ่มเติมฐานข้อมูลดังกล่าว โดยทำการแยกเฉพาะข้อมูลที่ต้องการจะศึกษา โดยเขียนโปรแกรมคอมพิวเตอร์ ด้วยโปรแกรมดีเบส (DBase III Plus) โดยให้คอมพิวเตอร์แยกเฉพาะข้อมูลที่เป็นไปตามกฎเกณฑ์ที่ผู้วิจัยกำหนดไว้ ดังที่ปรากฏใน 3.2 และให้แยกข้อมูลแต่ละประเภทตามที่กล่าวออกมาแล้วออกจากกัน และนอกจากนี้ ยังให้คอมพิวเตอร์แจ้งนับด้วยว่า หน่วยเสียงแต่ละหน่วยมีการปรากฏในพยางค์ที่เป็นคำ (X) จำนวนเท่าใดปรากฏในพยางค์ที่เป็นส่วนประกอบของคำ (Y) จำนวนเท่าใด และปรากฏในพยางค์ที่เป็นทั้งคำและส่วนประกอบของคำ (XY) จำนวนเท่าใด นอกจากนี้ให้คอมพิวเตอร์แยกข้อมูลด้วยว่า หน่วยเสียงแต่ละหน่วยปรากฏในพยางค์เสียงสั้นและพยางค์เสียงยาวเป็นจำนวนเท่าใด หน่วยเสียงแต่ละหน่วยปรากฏในพยางค์เป็นและพยางค์ตายเป็นจำนวนเท่าใด และในโครงสร้างพยางค์ต่างๆ ดังต่อไปนี้ คือ

1. พยางค์เปิด
2. พยางค์ปิด สระเสียงสั้น พยัญชนะไม่กักท้ายพยางค์
3. พยางค์ปิด สระเสียงยาว พยัญชนะไม่กักท้ายพยางค์
4. พยางค์ปิด สระเสียงสั้น พยัญชนะกักท้ายพยางค์
5. พยางค์ปิด สระเสียงยาว พยัญชนะกักท้ายพยางค์

มีหน่วยเสียงแต่ละหน่วยปรากฏอยู่เป็นจำนวนเท่าใด หลังจากนั้น ผู้วิจัยจึงนำจำนวนที่ได้มาคำนวณหาอัตราร้อยละของการปรากฏของหน่วยเสียงแต่ละหน่วยในพยางค์ที่เป็นคำ พยางค์ที่เป็นส่วนประกอบของคำ และพยางค์ที่เป็นได้ทั้งคำและส่วนประกอบของคำด้วยโปรแกรมโลตัส (Lotus Ver.2.2) เพื่อเปรียบเทียบรูปแบบการกระจายของหน่วยเสียงแต่ละหน่วย และรูปแบบการกระจายของหน่วยเสียงแต่ละชุด เมื่อหาอัตราร้อยละของหน่วยเสียงได้แล้ว ผู้วิจัยจึงสร้างแผนภูมิแสดงอัตราการกระจายดังกล่าว โดยใช้โปรแกรมคอมพิวเตอร์ Perspective Business and Scientific Graphics (3D Graphics) Version 1.2 และ EXCEL 4.0 เพื่อเปรียบเทียบการกระจายของหน่วยเสียงแต่ละหน่วยในพยางค์แต่ละประเภท และศึกษารูปแบบการกระจายของหน่วยเสียงในพยางค์ประเภท X กับพยางค์ประเภท Y และพยางค์ประเภท XY การวิเคราะห์ข้อมูลนั้นแบ่งหน่วยเสียงออกเป็น 3 ระบบ คือ ระบบหน่วยเสียงพยัญชนะ ระบบหน่วยเสียงสระ และระบบหน่วยเสียงวรรณยุกต์ โดยจำแนกการกระจายของหน่วยเสียงออกเป็น

การกระจายของหน่วยเสียงที่เกิดเป็นคำ ส่วนประกอบของคำและที่เกิดเป็นคำและส่วนประกอบของคำ การกระจายของหน่วยเสียงตามชุดของหน่วยเสียง การกระจายของหน่วยเสียงตามตำแหน่งของหน่วยเสียง การกระจายของหน่วยเสียงเมื่อปรากฏกับหน่วยเสียงอื่น และการกระจายของหน่วยเสียงในโครงสร้างพยางค์ประเภทต่างๆ หลังจากทีวิเคราะห์ข้อมูลแล้วก็นำมาศึกษาเปรียบเทียบกัน โดยใช้สถิติเชิงพรรณนา (Descriptive Statistics) มีการเปรียบเทียบร้อยละของการปรากฏของหน่วยเสียงแต่ละหน่วย และหน่วยเสียงแต่ละชุด และมีการจัดลำดับเพื่อหารูปแบบของการกระจาย มีการเรียงลำดับการปรากฏของหน่วยเสียงแต่ละหน่วย จากหน่วยเสียงที่มีการปรากฏใช้ในภาษาสูงสุด ไปหาหน่วยเสียงที่มีปริมาณการปรากฏใช้ในภาษาต่ำที่สุด เพื่อใช้เป็นลำดับของหน่วยเสียงในการพิจารณารูปแบบของการกระจายของหน่วยเสียง ในการตัดสินใจว่าหน่วยเสียงมีรูปแบบการกระจายเหมือนกันหรือไม่นั้น ผู้วิจัยมีเกณฑ์การตัดสินใจโดยพิจารณาจากหน่วยเสียง ที่มีการปรากฏมากที่สุดในแต่ละชุดของหน่วยเสียงในพยางค์แต่ละประเภทเป็นหลัก แล้วนำมาเปรียบเทียบกัน ซึ่งผลที่ได้จากการศึกษานี้อาจใช้เป็นแนวทางในการอธิบายการเปลี่ยนแปลง และพัฒนาการของหน่วยเสียงแต่ละหน่วย ที่ใช้ในภาษาไทยปัจจุบัน

