Journal of Research Methodology, Volume 14, Number 1 (January - April 2001)

วารสารวิธีวิทยาการวิจัย ปีที่ 14 ฉบับที่ 1 (มกราคม - เมษายน 2544)

Interdisciplinary Research Methodology: The Case of Research in Language Technology

Sudaporn Luksaneeyanawin

ABSTRACT

Research in language technology aims at simulating human linguistic competence and ability to perform for computers. Language technology research does not aim at answering or arguing theoretical questions but rather applying the knowledge in human and computer processing of language to answer the needs in human and computer communication. The success of the research depends on how explicit the questions in terms of language use and users, and in terms of computer technology constraints are. This research area is in the domain of Artificial Intelligence, specifically in the area of Natural Language Processing. It is interdisciplinary, in its nature, across linguistics, sciences, and engineering.

Research in language technology can be approached from different disciplines. Linguistic approach to research in this area has to start from a thorough linguistic description of the language, a clear understanding of how the system being developed is going to work in terms of usage and users. In designing the algorithm for the system linguistic knowledge of the language, psycholinguistic knowledge of human language processing, as well as knowledge in computer and signal processing are applied. After laboratory tests, the developed system has to be tested and evaluated by the target users. This step is most important for the completeness of the system defined in users' terms.

In countries where research in language technology is at its advance stage, the advancement of this kind of research is strongly supported by basic research in linguistic, psycholinguistics, and basic science research in language and signal processing. Whenever there is a new question from the needs in terms of human and computer communication or new theoretical questions for the improvement of language technology, researchers do not have to start from zero. The answers to the questions will be attempted to answer from researchers who come from different disciplines. These answers contribute to the advancement of language technology. The purpose of this article is to discuss the problems of language technology research in the Thai context. It is proposed that this area of research needs interdisciplinary collaboration, and the strength or success of Thai language technology research is based on the strength in basic research in the area of Thai linguistics, psycholinguistic aspect of Thai text and speech processing, including basic research in computer signal processing.

Journal of Research Methodology, Volume 14, Number 1 (January - April 2001)

วารสารวิธีวิทยาการวิจัย ปีที่ 14 ฉบับที่ 1 (มกราคม - เมษายน 2544)

วิธีวิทยาการวิจัยสหวิทยาการ : กรณีการวิจัยเทคโนโลยีทางภาษา

สุดาพร ลักษณียนาวิน

บทคัดย่อ

การวิจัยเทคโนโลยีทางภาษา เป็นการวิจัยที่มีวัตถุประสงค์ที่จะจำลองสมรรถนะทางภาษาของมนุษย์ให้แก่เครื่องคอมพิวเตอร์ ซึ่งมีรากฐานในเชิงประยุกต์เป็นสำคัญ ในการวิจัยเทคโนโลยีทางภาษาจะต้องมีโจทย์ในแง่ของการใช้ หรือโจทย์ที่มาจากผู้ใช้ที่ชัดเจนการวิจัยจึงจะประสบความสำเร็จ การวิจัยด้านนี้จัดอยู่ในแขนงหนึ่งของการวิจัยสาขาปัญญาประดิษฐ์ คือ แขนงการประมวลผลภาษาและวัจนะ มีลักษณะที่เป็นสหวิทยาการ ระหว่างสาขาภาษาศาสตร์ วิทยาศาสตร์ และวิศวกรรมศาสตร์

การวิจัยแขนงนี้ในแง่มุมของผู้เขียนซึ่งเป็นนักภาษาศาสตร์จะเริ่มต้นจากการวิจัยทางภาษาศาสตร์ โดยเริ่มตั้งแต่การวิจัยเชิงวรรณนาเกี่ยวกับลักษณะของภาษาที่ต้องการจะนำมาจำลองให้กับเครื่องคอมพิวเตอร์ การวิเคราะห์ความต้องการของผู้ใช้เกี่ยวกับสมรรถนะทางภาษาที่ต้องการจำลองให้กับเครื่อง การนำความรู้ความเข้าใจเกี่ยวกับกระบวนการประมวลผลภาษาของมนุษย์ (human language processing) ซึ่งอยู่ในสาขาภาษาศาสตร์จิตวิทยามาใช้ร่วมกับความรู้ความเข้าใจเกี่ยวกับกระบวนการประมวลผลสัญญาณอิเล็กทรอนิกส์ (digital signal processing) และกระบวนการประมวลผลโดยเครื่อง (computer processing) ในการวางขั้นตอนการทำงาน (algorithm) ของระบบที่จะสร้างขึ้น หลังจากสร้างระบบสำเร็จแล้ว การทดสอบระบบเป็นเรื่องสำคัญที่สุดอีกประการหนึ่ง ผู้วิจัยเทคโนโลยีทางภาษามักจะต้องนำระบบไปให้ผู้ใช้ทดลองใช้จริงหลังจากการทดสอบระบบและทดลองใช้ในห้องปฏิบัติการแล้ว เพื่อพัฒนาระบบให้สมบูรณ์สำหรับผู้ใช้มากที่สุด

ในประเทศที่การวิจัยเทคโนโลยีทางภาษาก้าวหน้าในระดับสูง การสนับสนุนงานวิจัยด้านนี้จะมุ่งพัฒนางานวิจัยพื้นฐานทางภาษาศาสตร์และภาษาศาสตร์จิตวิทยา รวมทั้งการวิจัยพื้นฐานด้านการประมวลผลสัญญาณในเชิงวิทยาศาสตร์ พร้อม ๆ ไปกับการวิจัยที่เน้นการสร้างระบบหรือสิ่งประดิษฐ์ที่จะนำไปประยุกต์ในงานต่าง ๆ โจทย์การวิจัยใหม่ ๆ จึงเกิดขึ้นทั้งในแง่ของผู้ใช้ และผู้สร้างระบบที่มาจากสหวิทยาการต่าง ๆ ทำให้การพัฒนาก้าวหน้าอย่างรวดเร็วเพราะผู้วิจัยไม่จำเป็นต้องเริ่มต้นจากศูนย์ทุกครั้งเมื่อมีคำถามใหม่ในเชิงประยุกต์ บทความนี้เขียนจากประสบการณ์กว่าสิบปีของผู้เขียนซึ่งได้ทำวิจัยเกี่ยวกับเทคโนโลยีทางภาษาด้านการประมวลผลภาษาและวัจนะ เพื่อมุ่งอธิบายกระบวนการและปัญหาของการวิจัยในแนวสหวิทยาการสาขาเทคโนโลยีภาษาในปริบทไทย โดยหวังว่าจะเป็นประโยชน์แก่ผู้ที่ต้องการจะศึกษาวิจัยด้านการประมวลผลภาษาและวัจนะในประเทศไทยต่อไป

ความนำ

ผู้เขียนเป็นนักวิจัยที่ไม่ได้เรียนวิธีการวิจัยในลักษณะที่เป็นวิทยาการ แต่เป็นนักวิจัยที่ได้ทำการวิจัยเพื่อตอบคำถามหรือสร้างข้อสรุปในเรื่องต่าง ๆมาเป็นเวลากว่า 25 ปี ตั้งแต่ปี พ.ศ. 2516 โดยเริ่มงานวิจัยในฐานะนักวิจัยประจำโครงการสัทศาสตร์เกี่ยวกับภาษาและภาษาพื้นเมืองถิ่นต่าง ๆ สถาบันศูนย์ภาษาอังกฤษ โครงการพัฒนามหาวิทยาลัย ทบวงมหาวิทยาลัย เมื่อได้รับเชิญให้เขียนเกี่ยวกับวิธีวิทยาการวิจัยสหวิทยาการ ก็เริ่มถามคำถามว่า "จะเขียนอย่างไร" ถ้าจะเขียนในเชิงวิชาการของศาสตร์แห่งการวิจัยก็ไม่ได้ร่ำเรียนมาอย่างเป็นเรื่องเป็นราว จึงได้ตัดสินใจว่าจะเขียนจากประสบการณ์ของการวิจัย 27 ปีที่ผ่านมานี้โดยเน้นการวิจัยในช่วงทศวรรษที่ผ่านมาของผู้เขียนในด้านเทคโนโลยีทางภาษาซึ่งเป็นการวิจัยในแนวสหวิทยาการเป็นสำคัญ

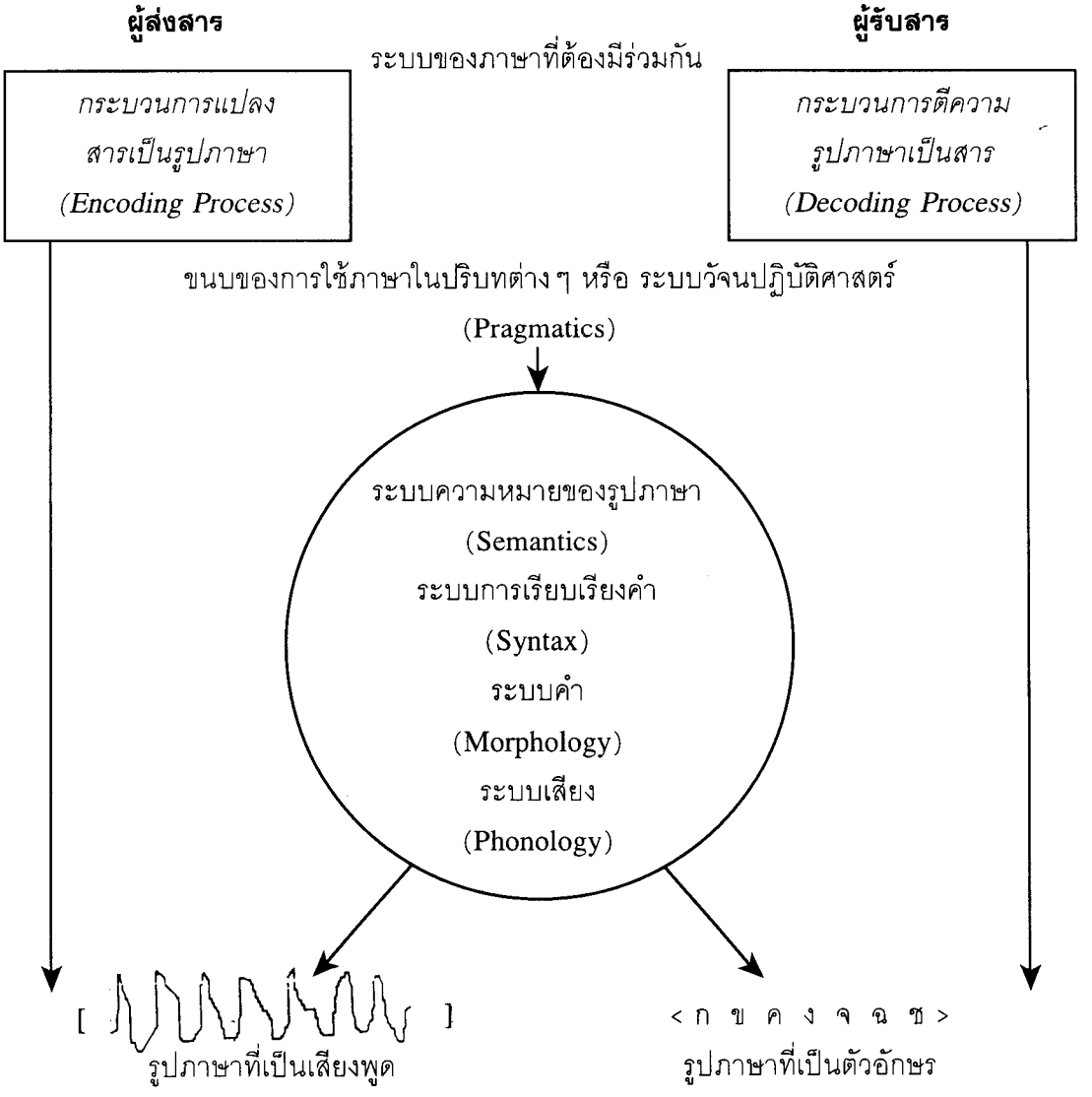
บทความนี้จะเริ่มจากการบรรยายภาพรวมของภาษาและกระบวนการสื่อสารในเชิงภาษาศาสตร์ ซึ่งเป็นหัวใจของการวิจัยด้านเทคโนโลยีทางภาษาตลอดจนแนวคิดสำคัญ ๆ ในการวิจัยภาษาศาสตร์บริสุทธิ์พอสังเขปสำหรับผู้อ่านที่อาจจะเคยศึกษาวิจัยเทคโนโลยีทางภาษาโดยเริ่มจากแง่มุมอื่นมาก่อน ก่อนที่จะกล่าวถึงประสบการณ์การวิจัยสหวิทยาการด้านเทคโนโลยีทางภาษาทั้งในเชิงกระบวนการและปัญหาในแนวสหวิทยาการต่อไป

การศึกษาวิจัยภาษาและกระบวนการสื่อสารในเชิงภาษาศาสตร์

การวิจัยทางภาษาศาสตร์บริสุทธิ์ มีลักษณะเช่นเดียวกับการวิจัยทางวิทยาศาสตร์สาขาอื่น ๆ นักภาษาศาสตร์ไม่ใช่คนที่รู้ภาษาและสามารถใช้ภาษาต่าง ๆได้ดีดังที่คนทั่วไปเข้าใจ แต่นักภาษาศาสตร์เป็นคนที่สามารถอธิบายภาษาอย่างเป็นระบบในทุกแง่มุม ตั้งแต่ ระบบเสียง ระบบคำ ระบบการเรียบเรียงคำหรือวากยสัมพันธ์ ระบบความหมาย และระบบของการใช้ภาษาในสถานการณ์ต่าง ๆ หรือที่เรียกว่าวัจนปฏิบัติศาสตร์ นักภาษาศาสตร์จิตวิทยาจะอธิบายกระบวนการต่าง ๆ ที่ทำให้มนุษย์รู้ภาษา ใช้ภาษา และเข้าใจภาษาหนึ่ง ๆ โดยใช้การวิจัยเชิงทดลองเกี่ยวกับการรับรู้วัจนะ (speech perception) การนึกรู้คำ (lexical access) หรือการตอบสนองต่าง ๆ โดยใช้ภาษา นักภาษาศาสตร์ที่สนใจความต่างของภาษาภายในภาษาหนึ่ง ๆ อยู่ในกลุ่มของนักภาษาศาสตร์ที่มีกระบวนการวิจัยเชิงเปรียบ เช่น นักภาษาศาสตร์สังคมจะอธิบายปรากฏการณ์การแปรของภาษาทั้งในแง่รูปแปรและปัจจัยต่าง ๆ ทางสังคมที่มีอิทธิพลต่อการแปรของภาษา นักวิทยาภาษาถิ่นจะอธิบายปรากฏการณ์การแปรของภาษาในเชิงภูมิศาสตร์ ส่วนนักภาษาศาสตร์เชิงประวัติจะอธิบายการเปลี่ยนแปลงของภาษาในแง่ของเวลา ตลอดจนการสัมผัสกันของภาษาอันนำมาสู่การเปลี่ยนแปลงของภาษา

ภาพข้างล่างนี้แสดงกระบวนการในการสื่อสารซึ่งผู้ส่งสารแปลงสาระที่ต้องการสื่อให้เป็นรูปภาษาทั้งที่เป็นเสียงพูดและรูปเขียน รวมทั้งกระบวนการในทางกลับกันที่ผู้รับสารตีความรูปภาษาให้เป็นสาระด้วย การสื่อสารจะบรรลุวัตถุประสงค์ได้นั้น นอกจากผู้ส่งสารและผู้รับสารต้องมีระบบต่าง ๆของภาษา (ที่อยู่ภายในวงรี) ซึ่งเป็นระบบที่เป็นนามธรรมร่วมกัน หรือ *พูดภาษาเดียวกัน* แล้ว ยังต้องมีขนบในการสื่อสารที่เรียกว่าระบบวัจนปฏิบัติร่วมกัน รวมทั้งมีการใช้รูปสัญญาณแทนระบบภาษาประเภทเดียวกันด้วยจึงจะสื่อสารกันได้อย่างราบรื่น

รูปที่ 1 ระบบของการสื่อสารและระบบของภาษา (Communicative and Linguistic System)



ระบบภาษาที่เป็นนามธรรมจะ กล่าวถึง *หน่วย* ในแต่ละระบบ (units in a system) และการปรากฏร่วมกันหรือ *โครงสร้าง* ของหน่วยดังกล่าว (structure of the units) เช่น ภาษาไทยมีหน่วยเสียงพยัญชนะ สระ และ วรรณยุกต์อะไรบ้าง และหน่วยเสียงเหล่านี้มีกฎการปรากฏร่วมกันอย่างไรได้บ้าง นักภาษาศาสตร์จะใช้กระบวนการวิจัยแบบค้นหา (discovery procedure) ในการสร้างระบบที่เป็นนามธรรมนี้ขึ้น ในกระบวนการดังกล่าวอาจใช้กระบวนการเชิงประจักษ์นิยม (empirisism) เก็บข้อมูลเชิงประจักษ์ (empirical data) มาสร้างข้อสรุปตามวิธีการที่วางไว้อย่างมีขั้นตอน หรือในอีกแนวทฤษฎีหนึ่งจะใช้กระบวนการพิสูจน์ถูก/ผิด (verificationism or falsificationism) มาพิสูจน์สมมติฐานที่วางไว้เกี่ยวกับภาษาที่จะศึกษา วิธีการวิจัยทั้งสองประการนี้เหมือนกับวิธีการวิจัยทางวิทยาศาสตร์ทั่วไปอื่นๆ ในอดีตการวิจัยเชิงประจักษ์นิยมมักถูกโจมตีในแง่ที่ว่าเราไม่มีทางพิสูจน์ได้เลยว่าข้อสรุปที่สร้างขึ้นจากข้อมูลนั้น ถูกต้องสมบูรณ์ดีหรือไม่อย่างไร เพราะข้อสรุปที่ได้มานั้นขึ้นอยู่กับความถูกต้องสมบูรณ์หรือความบกพร่องไม่สมบูรณ์ของข้อมูล ข้อสรุปที่ได้มานั้นจึงไม่มีทางผิดเพราะเป็นการสรุปจากข้อมูลเชิงประจักษ์ คำถามสำคัญก็คือข้อมูลที่สมบูรณ์ที่สุดนั้นมีหรือไม่ อย่างไร การวิจัยแนวนี้จะอาศัยการเก็บข้อมูลภาคสนาม การฝึกอบรมนักวิจัย หรือผู้ช่วยวิจัย จะเน้นกระบวนการและระเบียบวิธีในการวิเคราะห์ข้อมูลเพื่อให้ได้ข้อมูลที่ถูกต้องสมบูรณ์ที่สุด (Bloomfield, 1933; Pike 1943, 1947) การวิจัยทางภาษาศาสตร์ในระยะแรกคือ ช่วงทศวรรษที่ 30-60 ใช้แนวทางเชิงประจักษ์นิยมในการอธิบายโครงสร้างของภาษา โดยการจำแนกหน่วยทางภาษาเป็นชนิดต่างๆ (taxonomic enterprise) และเริ่มอธิบายหน่วยทางภาษาจากระบบที่มีหน่วยย่อยที่เล็กที่สุด คือ ระบบเสียงและหน่วยเสียง มาถึงระบบคำและหน่วยคำ และระบบการเรียบเรียงคำหรือวากยสัมพันธ์และหน่วยประโยคซึ่งเป็นหน่วยที่ใหญ่ที่สุดตามลำดับ ลักษณะของภาษาที่อธิบายด้วยกระบวนการดังกล่าวนี้เรียกว่า ไวยากรณ์โครงสร้าง (Structural Grammar) ต่อมาช่วงปลายทศวรรษที่ 50 ถึงปัจจุบัน หลังจากที่ โนม ชอมสกี (Chomsky, 1957, 1965) เสนอทฤษฎีไวยากรณ์ปริวรรตแปรรูป (Generative Transformational Grammar) ซึ่งมีแนวทางในเชิงนิรนัย กล่าวคือ มีการสร้างสมมติฐานเกี่ยวกับโครงสร้างภายในของภาษาทุกภาษาอันมีลักษณะที่เป็นสากลลักษณ์ การวิจัยจะใช้กระบวนการพิสูจน์ความถูก/ผิดของข้อเสนอที่เป็นสมมติฐานเกี่ยวกับภาษามากกว่าการสร้างข้อสรุปของข้อมูลเชิงประจักษ์ การอธิบายไวยากรณ์ของภาษาจึงมักอธิบายได้เป็นส่วนย่อยๆ ในประเด็นเฉพาะที่เป็นคำถามหรือสมมติฐาน ไม่สามารถครอบคลุมรายละเอียดของภาษาทั้งภาษาได้เช่นไวยากรณ์โครงสร้าง แต่ก็มีข้อดี คือ ข้อสรุปที่ได้เป็นข้อสรุปที่แข็ง คือได้ผ่านการพิสูจน์มาแล้ว และหากจะมีข้อขัดแย้งใดๆ เกิดขึ้นในอนาคตโดยมีปรากฏการณ์ทางภาษามาสนับสนุน ข้อสรุปนั้นก็จะถูกล้มล้างไปทันที ต่างจากการสร้างข้อสรุปจากการสังเกตซึ่งเป็นข้อสรุปที่อ่อน กล่าวคือ สามารถเพิ่มเติมข้อสรุปไปได้เรื่อยๆ ตามข้อมูลใหม่ๆ

โดยข้อสรุปจะไม่ถูกล้มล้าง ข้อดีในเชิงทฤษฎีอีกประการหนึ่งของไวยากรณ์ปริวรรตแปรรูป คือ ในการอธิบายไวยากรณ์ของภาษาใดภาษาหนึ่งจะเริ่มจากโครงสร้างหลักซึ่งเป็นสากลลักษณ์ และกฎการปริวรรตและแปรรูปซึ่งเป็นกฎเกณฑ์ที่มีลำดับและขั้นตอนอย่างชัดเจนเฉพาะภาษาแต่ละภาษาในการสร้างรูปภาษาที่ถูกต้องตามหลักไวยากรณ์ ไวยากรณ์ปริวรรตและแปรรูปจึงเน้น *กระบวนการ* อันประกอบด้วยโครงสร้างลึก กฎเกณฑ์ และลำดับของกฎเกณฑ์ในการสร้างรูปภาษามากกว่ามุ่ง *อธิบายรูปภาษา* ที่ปรากฏในเชิงประจักษ์แล้ว ชอมสกีโค่นล้มแนวคิดของสกินเนอร์ (Skinner ,1957; Chomsky, 1959) เกี่ยวกับทฤษฎีการรู้ภาษา ในแง่ที่ว่ามนุษย์ไม่ได้รู้ภาษาจากการปฏิบัติตามพฤติกรรมภาษาของผู้ใหญ่โดยการเลียนแบบและเชื่อมโยงกับการกระตุ้นเร้าจากภายนอก แต่กลไกการรู้ภาษาของมนุษย์เป็นสิ่งที่เกิดมาภายในสมองของมนุษย์ (hard wired) และมีติดตัวมาตั้งแต่เกิด (innateness) ไวยากรณ์ของมนุษย์หรือสมรรถนะทางภาษาจึงเกิดขึ้นจากกลไกในการรู้ภาษา (Language Acquisition Device-LAD) ที่มนุษย์มีติดตัวมาแต่กำเนิด ข้อถกเถียงสำคัญที่ชอมสกีใช้โค่นล้มแนวคิดเกี่ยวกับการรู้ภาษาตามแนวพฤติกรรมนิยมของสกินเนอร์ คือ รูปภาษาที่เด็กเล็กๆ สร้างขึ้นและไม่ถูกต้องตามไวยากรณ์ ส่วนใหญ่เป็นรูปภาษาที่ไม่พบในภาษาของผู้ใหญ่ที่เป็นภาษาต้นแบบตามแนวคิดของนักพฤติกรรมนิยม เช่น สกินเนอร์ ชอมสกีเรียกความสามารถของมนุษย์ในการสร้างรูปภาษาใหม่ๆ ขึ้นได้ไม่รู้จบนี้ว่า ความสามารถในการสร้างสรรค์ภาษา (creativity) [1] ตัวอย่างเช่น

*I goed to school. (I went to school.)

*I sure. (I am sure.)

*หนูดูอยู่รูป (หนูดูรูปอยู่)

*พ่อหาเจอค้อน (พ่อหาค้อนเจอ)

(* คือ รูปที่ไม่ถูกต้องตามไวยากรณ์ ซึ่งเด็กสร้างขึ้นจากข้อสรุปที่เกิดขึ้นในไวยากรณ์ระยะแรกๆ มีรูปที่ต่างจากภาษาผู้ใหญ่ซึ่งให้ไว้ในวงเล็บ)

การศึกษาวิจัยภาษาในแนวไวยากรณ์ปริวรรตและแปรรูปสร้างองค์ความรู้เกี่ยวกับกระบวนการประมวลผลภาษาและวัจนะของมนุษย์เป็นอย่างมากในช่วง 4-50 ปีที่ผ่านมา โดยเฉพาะอย่างยิ่งการศึกษาวิจัยเกี่ยวกับการนึกรู้คำ (lexical access) ใช้วิธีการวิจัยเชิงทดลองสร้างองค์ความรู้เกี่ยวกับการจัดเก็บ เรียงลำดับ และกระบวนการค้นคืนคำจากคลังคำในสมอง (mental lexicon) ซึ่งเป็นประโยชน์ยิ่งต่อการนำมาประยุกต์ใช้ในการวิจัยด้านเทคโนโลยีทางภาษาซึ่งต้องอาศัยความรู้เกี่ยวกับภาษาที่มีการอธิบายอย่างมีลำดับ และแสดงในรูปกฎเกณฑ์ที่มีรูปแบบที่ชัดเจน (formalism) ดังจะได้กล่าวถึงต่อไป

ในปัจจุบันซึ่งเป็นยุคสารสนเทศ ข้อมูลภาษาในรูปอิเล็กทรอนิกส์สามารถเก็บรวบรวม จัดระเบียบ และเรียกค้นคืนเพื่อใช้ในการศึกษาทางภาษาศาสตร์อย่างมีระเบียบวิธี เรียกว่า ภาษาศาสตร์คลังข้อมูล (corpus linguistics) การสร้างคลังข้อมูลมีระเบียบวิธีในการศึกษาซึ่งอิงกับแนวคิดทางภาษาศาสตร์ สารนิเทศศาสตร์ และ วิทยาการคอมพิวเตอร์ การสร้างคลังข้อมูลโดยเก็บข้อมูลแบบเดาสุ่มจึงไม่เป็นประโยชน์ต่อการวิจัยแต่อย่างใด การศึกษาวิจัยในแนวคิดเชิงประจักษ์นิยมเริ่มคืนกลับมา แต่ทั้งนี้การสร้างข้อสรุปต่าง ๆ ตามแนวทางนี้จะต้องอาศัยคลังข้อมูลขนาดใหญ่ที่มีการออกแบบอย่างมีหลักวิชาทางภาษาศาสตร์ด้วย

การวิจัยทางภาษาศาสตร์ซึ่งนอกจากจะมุ่งอธิบายระบบของภาษาที่เป็นนามธรรม ทั้งในแนวประจักษ์นิยม และแนวเหตุผลนิยมซึ่งเน้นการตรวจสอบสมมติฐานเกี่ยวกับภาษา (rationalism) ดังกล่าวมาแล้ว ถูกจัดเป็นกลุ่มที่เรียกว่าภาษาศาสตร์รูปแบบ (Formal Linguistics) การวิจัยทางภาษาศาสตร์ยังมีอีกประเภทหนึ่งที่มุ่งศึกษาภาษาในปริบทของการสื่อสาร ภาษาศาสตร์กลุ่มนี้ถูกเรียกว่าภาษาศาสตร์หน้าที่นิยม (Functional Linguistics) ความสนใจของนักภาษาศาสตร์กลุ่มนี้อยู่ที่หน้าที่ในการสื่อสาระต่าง ๆ ของภาษาที่สำคัญ ๆ 3 ประการ ซึ่งบูฮ์เลอร์ (Buhler, 1934) ผู้นำทางความคิดเกี่ยวกับสาระของภาษาเสนอไว้และยังทันสมัยในความคิดของนักภาษาศาสตร์ในปัจจุบัน (สุดาพร ลักษณียนาวิน, 2537ก, 2537ข) ได้แก่ (1) หน้าที่ในการสื่อสาระอ้างอิงซึ่งอ้างอิงถึงความรู้ความคิดเกี่ยวกับสิ่งต่าง ๆ ทั้งที่เป็นรูปธรรมและนามธรรม (representative or referential function) เช่น รูปภาษา หรือ คำว่า <หม้อ> ใช้แทน "ภาชนะทนไฟที่ใช้ในการหุงต้ม ส่วนมากมีฝาครอบ" (2) หน้าที่ในการสื่อสาระทางทัศนคติและอารมณ์ของผู้พูด (expressive and emotive function) เช่น การใช้ลักษณะของทำนอง เสียง จังหวะ น้ำเสียง ความดังค่อยในการพูด หรือการจัดลำดับของคำในคำพูด ใช้สื่อสาระทางทัศนคติ และอารมณ์ความรู้สึกของผู้พูดทั้งสิ้น หน้าที่ในการสื่อสาระสองประการดังกล่าวแล้วเป็นหน้าที่ของภาษาในแง่มุมของผู้ส่งสาร บูฮ์เลอร์ยังมีประเด็นสำคัญเกี่ยวกับหน้าที่ของภาษาอีกประการหนึ่ง คือ (3) หน้าที่ในการสื่อสาระในผู้รับสาร (appellative, function) กล่าวคือ ผู้รับสารเกิดความรู้ ความคิด และความรู้สึกเกี่ยวกับผู้สื่อสารอย่างไร หน้าที่ประการนี้กำลังเป็นที่สนใจยิ่งในเชิงทฤษฎีของการสื่อสารปัจจุบันที่เรียกว่าวัจนปฏิบัติศาสตร์ (Levinson, 1983) เพราะไม่ว่าผู้สื่อสารต้องการจะสื่อสาระอะไรก็ตาม *สาระที่แท้จริงของภาษาที่เกิดขึ้นในปริบทของการสื่อสาร คือ สาระที่ผู้รับสารเป็นผู้ตีความเท่านั้น* ทฤษฎีวัจนปฏิบัติศาสตร์จึงให้ความสำคัญกับสาระที่เกิดกับผู้รับสารเป็นอย่างยิ่ง ทั้งในแง่ของความเชื่อเบื้องต้น (presupposed believe) รวมทั้งความรู้ร่วมระหว่างผู้ส่งสารและผู้รับสาร การศึกษาภาษาในแง่มุมของการสื่อสารของนักภาษาศาสตร์กลุ่มหน้าที่นิยมได้รับอิทธิพลทางความคิดทั้งจาก คาร์ล บูฮ์เลอร์ (1934) ของกลุ่มปราก และ เจ อาร์ เฟิร์ธ (Firth, 1957) ของกลุ่มลอนดอน ในยุโรป และใน

สหรัฐอเมริกาภายใต้แนวความคิดของ ไฮมส์ (Hymes, 1964, 1972, 1974) ซึ่งเป็นผู้เสนอว่า มนุษย์มีสมรรถนะในการสื่อสาร (Communicative Competence) เป็นตัวควบคุมการใช้ภาษาในการสื่อสาร กล่าวคือ การใช้ภาษาของมนุษย์มิได้เกิดจากสมรรถนะทางภาษา (Linguistic Competence) ซึ่งควบคุมการสร้างรูปภาษาแต่ประการเดียว ความสำเร็จในการสื่อสารเกิดจากสมรรถนะในการสื่อสารด้วยปริบทของการสื่อสารจึงเป็นเรื่องที่สำคัญยิ่งทางภาษาศาสตร์เฟิร์ธ(1957) เป็นผู้ที่ให้ความสำคัญกับเรื่องของปริบทในการสื่อสารเป็นอย่างยิ่ง เขาแสดงความสัมพันธ์ในเชิงลำดับของปริบท จากปริบทที่ใหญ่ที่สุดมายังปริบทที่เล็กที่สุดว่า การสื่อสารใดๆ ก็ตามจะถูกครอบคลุมด้วยปริบทวัฒนธรรม (cultural context) ซึ่งเป็นตัวกำหนดโครงสร้างของปริบทสถานการณ์ (context of situation) ซึ่งหมายถึงมิติของเวลาและสถานที่ในการสื่อสารเฉพาะวัฒนธรรมนั้นๆ และครอบคลุมมิติทางสังคมของผู้ร่วมสื่อสาร (context of participants) ว่ามีการจำแนกเป็นกลุ่มอย่างไร การสื่อสารยังเกิดในปริบทสัทศาสตร์ (phonetic context) ซึ่งเป็นองค์ประกอบสำคัญลำดับสุดท้ายของการสื่อสาร ขอยกตัวอย่างการสื่อสารแสดงให้เห็นทฤษฎีเกี่ยวกับปริบทในการสื่อสารของเฟิร์ธ เช่น ในวัฒนธรรมไทยมิติของเวลาและสถานที่จะถูกกำหนดในแง่มุมของ*กาละ* (เวลา) และ *เทศะ* (สถานที่) ในการแสดงออกทางภาษา คนไทยให้ความสำคัญกับ*เวลา* น้อยกว่าภาษาอังกฤษ เช่นในมิติของเวลา เช้า สาย บ่าย เย็น ค่ำ ที่ผู้พูดภาษาอังกฤษใช้ในการทักทายนั้น ในภาษาไทยไม่มีการจำแนกการทักทายตามเวลา เราอาจจะพบคำว่า อรุณสวัสดิ์ สายัณห์สวัสดิ์ เฉพาะในกรณีของการแปลตามตัวอักษรจากภาษาอังกฤษเป็นภาษาไทยเท่านั้น นอกจากนี้รูปแสดงเวลา (time) ในภาษาไทยก็มีลักษณะต่างจากภาษาอังกฤษ คือ การแสดงเวลาในภาษาไทย ใช้คำบอกเวลา หรือ ใช้ปริบทของประโยคแวดล้อมเป็นตัวบ่งชี้ และหากคำพูดถูกถอดออกจากปริบทก็เป็นเรื่องยากที่จะรู้ว่าเหตุการณ์นั้นเกิดขึ้นเมื่อใด ในขณะที่ *เวลา (time) และการณ์ลักษณะ (aspect)* ในภาษาอังกฤษจะแสดงรูปที่คำกริยาอย่างชัดเจน ตัวอย่างเช่น

He went to Australia on a business trip.

เขาไปออสเตรเลียมาด้วยเรื่องธุรกิจ

เขาไปออสเตรเลียด้วยเรื่องธุรกิจ (จะไปราวสามอาทิตย์ เห็นว่าเดินทางไปเมื่อเย็นนี้)

เวลาที่ผ่านไปแล้ว หรือ อดีตกาล ในภาษาอังกฤษจะแสดงด้วยกิริยาแสดงอดีตกาลที่ชัดเจน ในขณะที่ในภาษาไทยอาจใช้คำว่า <มา> เป็นตัวบ่งชี้ หรือไม่มีการบ่งชี้เวลาเลยดังในประโยคถัดมา การรับรู้เวลาในประโยคสุดท้ายอาศัยปริบทของคำพูดที่ตามมา

นอกจากเรื่องของการรับรู้และเข้าใจ "เวลา" ที่ต่างกันในสองวัฒนธรรมดังกล่าวแล้ว ความแตกต่างทางวัฒนธรรมที่แสดงออกทางภาษายังมีอีกมาก ทั้งในแง่ของคลังคำ เช่น ในภาษาไทยเรามีคำที่เรียกภาชนะจักสานที่ทำจากหวายหรือไม้ไผ่ (ในปัจจุบันจะพบที่ใช้วัสดุสังเคราะห์ เช่น

พลาสติก หรือ เชือกฟางด้วย) เอาไว้ใส่ของหลายชนิดจำแนกตามขนาด ประโยชน์ใช้สอย และของที่จะใช้ใส่ เช่น ตะกร้า ชะลอม กระเช้า กระบุง กระจาด กระทาย เข่ง บุ้งกี๋ หลัว ฯลฯ ในขณะที่ในภาษาอังกฤษมีคำว่า "basket" เพียงคำเดียว ปริบททางวัฒนธรรมซึ่งครอบคลุมปริบทอื่น ๆ ในการสื่อสาร ยังเข้ามาครอบคลุมมิติของปริบทของผู้ร่วมสื่อสารด้วย เช่น ในวัฒนธรรมไทยการสื่อสารจะเน้นความสัมพันธ์เชิงสัมพัทธ์ระหว่างผู้สื่อสารและผู้รับสารเป็นอย่างยิ่ง ทั้งในแง่ของ อายุ อาวุโส เพศ ระดับชั้นทางสังคม และความใกล้ชิดสนิทสนม มิติเหล่านี้แสดงออกในการเลือกใช้คำ เช่น คำสรรพนาม คำเรียกขาน คำลงท้าย ซึ่งในภาษาอื่นอาจมีมิติที่ต่างออกไป หรืออาจแสดงออกโดยใช้คำประเภทอื่น หรือ ทำนองเสียง แทนการใช้คำเหล่านี้ การศึกษาวิจัยของนักภาษาศาสตร์จึงมิได้จำกัดอยู่แต่การศึกษาระบบที่เป็นนามธรรมแต่ยังศึกษาวิจัยเกี่ยวกับการใช้ภาษาในสถานการณ์จริงด้วย ซึ่งในการนี้จะต้องใช้ระเบียบวิธีในการเก็บข้อมูลในเชิงสังคมศาสตร์ เช่นการสอบถามด้วยการสัมภาษณ์ ใช้แบบสอบถาม หรือ ใช้การสังเกตเหตุการณ์ในสถานการณ์ที่กำหนดขึ้นในการวิจัย (อมรา ประสิทธิ์รัฐสินธุ์, 2533) ทฤษฎีภาษาศาสตร์หน้าที่ให้ความสำคัญกับการสื่อสารความหมาย จึงมีบทบาทสำคัญเกี่ยวกับ ระบบการสร้างรูปภาษาโดยเครื่อง (Generation System) และระบบความเข้าใจของเครื่อง (Machine Understanding) ในการแปลอัตโนมัติโดยใช้คอมพิวเตอร์

ศาสตร์แห่งภาษาหรือภาษาศาสตร์นี้มักจะจัดอยู่ในแขนงหนึ่งของมนุษยศาสตร์ แม้นักภาษาศาสตร์จำนวนมากจะจัดศาสตร์นี้ให้เป็นแขนงหนึ่งในวิทยาศาสตร์ ที่เรียกว่าวิทยาศาสตร์พฤติกรรม (Behavioural Sciences) ในความคิดของผู้เขียนการศึกษาวิจัยทางภาษาศาสตร์ส่วนใหญ่เป็นการวิจัยแบบบูรณาการ (integrated research) คือ เรื่องที่ศึกษาวิจัยเป็นเรื่องภายในของมนุษย์ การเก็บข้อมูลมักอาศัยวิธีการทางสังคมศาสตร์และวิทยาศาสตร์ ส่วนการวิเคราะห์ข้อมูลนั้นใช้วิธีการทางวิทยาศาสตร์ การสรุปและตีความข้อมูลใช้ทั้งวิธีการทั้งทางวิทยาศาสตร์และมนุษยศาสตร์ แม้การวิจัยทางภาษาศาสตร์จะต้องใช้วิธีการแบบบูรณาการ แต่การวิจัยทางภาษาศาสตร์ไม่จำเป็นต้องเป็นการวิจัยสหวิทยาการ (interdisciplinary research) มีนักภาษาศาสตร์จำนวนมากที่ศึกษาภาษาศาสตร์เป็นศาสตร์บริสุทธิ์ภายในตนเองโดยใช้วิธีวิจัยแบบบูรณาการดังกล่าวมาแล้ว

การวิจัยเทคโนโลยีทางภาษา

เทคโนโลยีทางภาษา คือความพยายามของมนุษย์ที่จะนำภาษามาใช้ประโยชน์ในเรื่องต่าง ๆ อง (Ong, 1988) กล่าวว่าเทคโนโลยีทางภาษาชิ้นแรกที่มนุษย์ประดิษฐ์ขึ้น คือ อักขระซึ่งใช้บันทึกเรื่องราวต่าง ๆ ไว้ มีลักษณะสถิตย์ ต่างจากเสียงซึ่งมีลักษณะพลวัต เกิดขึ้นเป็นธรรมชาติของการสื่อสารของมนุษย์ กล่าวคือ เมื่อมนุษย์พัฒนามาเป็นสัตว์พูดได้ (Homoloquens) มนุษย์ก็สื่อสาร

กันด้วยการพูด มีคนจำนวนมากในโลกที่ไม่รู้หนังสือและยังสื่อสารกันด้วยการพูดเท่านั้น วอลเตอร์ องยังแสดงให้เห็นว่าเทคโนโลยีการเขียนทำให้คนที่รู้หนังสือกับคนที่ไม่รู้หนังสือมีพฤติกรรมการสื่อสารที่ต่างกัน ระบบการเขียนหรืออักขระที่มนุษย์ประดิษฐ์ขึ้นใช้ มีลักษณะต่างๆ กัน ผู้เขียนขอเปรียบระบบอักขระหรือระบบการเขียนซึ่งเป็นสิ่งประดิษฐ์หรือเทคโนโลยี ซึ่งมีหลายประเภทต่างๆ กันนี้ว่า ประเภทของระบบการเขียนที่มีต่างๆ กันนี้มีที่มาจากความต้องการของการประดิษฐ์นั้นๆ เป็นสำคัญ ลักษณะที่ต่างๆ กันของอักขระ เช่น อักขระแบบอิพิกราฟ (epigraph) ประดิษฐ์ขึ้นเพื่อใช้แทนคำพูดที่เป็นข้อความ อักขระแบบอิดิโอกราฟ (ideograph) หรือ อักษรภาพ (pictograph) ใช้แทนความหมายของคำ นักประดิษฐ์ตัวอักษรยังพยายามที่จะสร้างระบบที่ครอบคลุมคำพูดทั้งหมดในภาษาโดยประดิษฐ์สัญลักษณ์แทนเสียง เช่น อักษรแทนพยางค์ (syllabary) หรืออักษรแทนหน่วยเสียง (alpabetic) ด้วย ระบบการเขียนที่ต่างกันนี้ล้วนแสดงความต้องการที่ต่างกันของการประดิษฐ์ทั้งสิ้น การบันทึกสัญญาณเสียงพูดเป็นข้อความเพื่อใช้ในการให้ข้อมูลต่างๆ ทางโทรศัพท์ หรือ ตามสถานีขนส่ง อาจเปรียบได้กับ อิพิกราฟที่ใช้บันทึกคำพูดของชาวอียิปต์โบราณนั่นเอง เครื่องหมายที่เป็นสัญลักษณ์ (icon) ต่างๆ ที่เครื่องคอมพิวเตอร์ตระกูลแมคอินทอชนำมาใช้ในระบบปฏิบัติการแบบหน้าต่าง (window) ก่อนตระกูลพีซีซึ่งในระยะแรกใช้ภาษาเครื่องหรือดอส นั้น ก็เกิดขึ้นจากความต้องการที่จะให้ผู้ใช้สามารถสื่อสารกับเครื่องด้วยเครื่องหมายที่เป็นภาพสัญลักษณ์เช่นเดียวกับอักษรภาพนั่นเอง เมื่อคนญี่ปุ่นเริ่มมีตัวหนังสือใช้นั้นได้รับเอาอักษรจีนซึ่งเป็นอักษรภาพเข้ามาใช้แทนความหมายอ้างอิงหรือความหมายของคำต่างๆ อักษรชนิดนี้เรียกว่าตัวอักษรคันจิ (kanji) แต่เมื่อไม่สามารถแสดงความรู้ความคิดในวัฒนธรรมญี่ปุ่นหรือ ไม่สามารถแสดงความรู้สึกได้ชัดเจนด้วยคำพูดดังเช่นในภาษาพูด คนญี่ปุ่นจึงสร้างตัวอักษรแทนเสียงพยางค์ในภาษาญี่ปุ่นขึ้น 2 ชนิด เพื่อใช้เขียนถึงสิ่งต่างๆ ที่ไม่สามารถกล่าวถึงได้ด้วยตัวคันจิ คือ อักษรฮิระงะนะ (hiragana) และอักษรคะตะกะนะ (katakana) ซึ่งเป็นอักษรแทนเสียงพยางค์ในภาษาญี่ปุ่น โดยชนิดแรกใช้เขียนคำไวยากรณ์ในภาษา และ ชนิดที่สองใช้เขียนคำยืม คำแสดงความรู้สึกและอารมณ์ และเขียนหนังสือสำหรับเด็ก (Chaplin and Martin, 1991) จะเห็นได้ว่าการศึกษาวิจัยประดิษฐ์คิดค้นทางด้านเทคโนโลยีนั้นจะถูกผลักดันจากความต้องการในการใช้เป็นสำคัญ การวิจัยทางเทคโนโลยีจึงเกิดขึ้นเพื่อสนองความต้องการในการใช้ของมนุษย์เสมอ ต่างจากการวิจัยที่มุ่งตอบปัญหาเกี่ยวกับโลกภายใน โลกภายนอก ตลอดจนความสัมพันธ์ระหว่างมนุษย์กับโลกภายในและภายนอก ในแนวการวิจัยมนุษยศาสตร์ วิทยาศาสตร์ หรือ สังคมศาสตร์บริสุทธิ์ดังกล่าวมาแล้ว

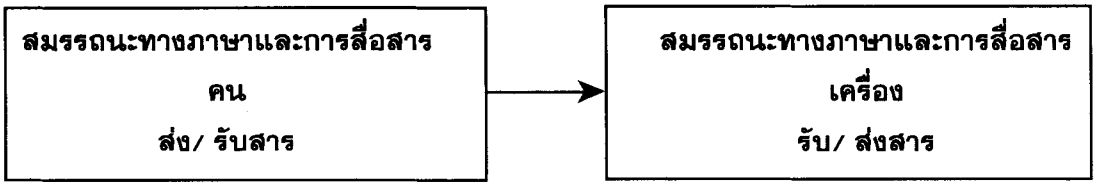
นักวิทยาการวิจัย (Research Methodologist) มีการจำแนกประเภทของการวิจัยด้วยเกณฑ์ต่างๆ เช่น จำแนกการวิจัยตามเนื้อหา-ศาสตร์หรือวิทยาการที่วิจัย หรือจำแนกตาม

เป้าประสงค์ของการวิจัย (Dane, 1990) เช่น จำแนกการวิจัยเป็นการวิจัยเพื่อแสวงหาสิ่งใหม่ๆ (Exploratory Research) บรรยายปรากฏการณ์ (Descriptive Research) ทำนายเหตุการณ์ (Predictive Research) อธิบายปรากฏการณ์ (Explanatory Research) หรือเพื่อการปฏิบัติการ (Action Research) สำหรับการวิจัยเทคโนโลยีทางภาษานั้นผู้เขียนเห็นว่าเป็นการวิจัยเชิงปฏิบัติการ แม้ว่าการวิจัยเพื่อการปฏิบัติการตามแนววิธีวิทยาการวิจัยจะหมายถึงการวิจัยที่มีจุดมุ่งหมายที่จะแก้ปัญหาทางสังคมก็ตาม (Dane, 1990) ทั้งนี้จะได้อธิบายและอภิปรายแนวคิดนี้ต่อไป

สำหรับการวิจัยเทคโนโลยีทางภาษานั้นมีลักษณะที่เป็นสหวิทยาการ คือ ต้องศึกษาเนื้อหาข้ามศาสตร์หลักๆ ทั้งสามศาสตร์ คือ มนุษยศาสตร์ วิทยาศาสตร์ และสังคมศาสตร์ ผู้เขียนเคยกล่าวว่า (สุดาพร ลักษณียนาวิน, 2526) วิศวกรรมศาสตร์ซึ่งเป็นศาสตร์ที่เน้นเรื่องการตอบสนองความต้องการของมนุษย์ด้วยเทคโนโลยีต่างๆ นั้น โดยเนื้อแท้แล้วเป็นสหวิทยาการที่ต้องการความรู้ความเข้าใจเกี่ยวกับมนุษย์และสังคมอยู่มากนอกเหนือไปจากวิทยาศาสตร์ เทคโนโลยีทางภาษาใหม่ๆ เกิดขึ้นอย่างมากมายในช่วงสองทศวรรษที่ผ่านมา พร้อมกับการเจริญเติบโตของเทคโนโลยีสารสนเทศซึ่งเกี่ยวข้องกับการจัดเก็บข้อมูลสารสนเทศและการติดต่อสื่อสารเรียกค้นคืนข้อมูล เพื่อนำมาใช้ในกิจการต่างๆ ทั้งการศึกษาวิจัย การแพทย์ การค้า ฯลฯ กระบวนการดังกล่าวต้องใช้ภาษาเป็นตัวกลางทั้งสิ้น

ขอบข่ายของงานวิจัยเทคโนโลยีทางภาษา อาจแสดงได้ด้วยรูปที่ 2 ข้างล่างนี้ รูปนี้จะมีส่วนที่ล้อกับรูปที่ 1 โดยแสดงให้เห็นขอบข่ายของระบบของการสื่อสารและระบบของภาษาในการปฏิสัมพันธ์ระหว่างคนกับเครื่อง

รูปที่ 2 ขอบข่ายของงานวิจัยเทคโนโลยีทางภาษาที่สร้างปฏิสัมพันธ์ระหว่างคนกับเครื่อง



	รับสาร	ส่งสาร
สาร (เสียงพูด/ตัวหนังสือ)	1. อ่านด้วยตาเครื่อง/เข้าใจ[2]	- พิมพ์ - ออกเสียง/พูดโต้ตอบ - กระทำตามคำสั่ง
	2. ฟังด้วยหูเครื่อง/เข้าใจ	- พิมพ์ - พูดโต้ตอบ - กระทำตามคำสั่ง
	3. อ่าน/ฟัง/เข้าใจ/แปล	- พิมพ์หรือพูดเป็นอีกภาษาหนึ่ง และอื่นๆ

จากรูปที่ 2 จะเห็นได้ว่าขอบข่ายของการวิจัยเทคโนโลยีทางภาษาก็คือการจำลองสมรรถนะทางภาษาและการสื่อสารของคนให้กับเครื่องนั่นเอง การวิจัยด้านนี้นอกจากจะช่วยอำนวยความสะดวกในการสื่อสารให้กับคนที่อาจจะมีความบกพร่องในสมรรถนะทางภาษาและการสื่อสาร เช่น คนตาบอด คนหูหนวก หรือผู้ที่มีประสาทพิการบางส่วน ทั้งในด้านการศึกษาและการดำรงชีวิตทั่วๆ ไปแล้ว ยังมีประโยชน์ต่อการจัดเก็บและบันทึกข้อมูลสารสนเทศที่มีอยู่เป็นจำนวนมหาศาลให้สามารถนำมาค้นคืนเพื่อใช้ในการศึกษาค้นคว้าได้อย่างกว้างขวางและรวดเร็ว โดยไม่มีข้อจำกัดในมิติของเวลาและสถานที่ ถ้าหากมีระบบการสื่อสารในที่ที่คนจะสื่อกับเครื่องได้ เทคโนโลยีทางภาษาที่จะเอื้อให้เกิดปฏิสัมพันธ์ระหว่างคนกับเครื่อง อาจสรุปได้เป็น 3 กลุ่มใหญ่ๆ ตามที่เราสามารถจำลองสมรรถนะทางภาษาให้กับเครื่องดังแสดงในรูปที่ 2 คือ

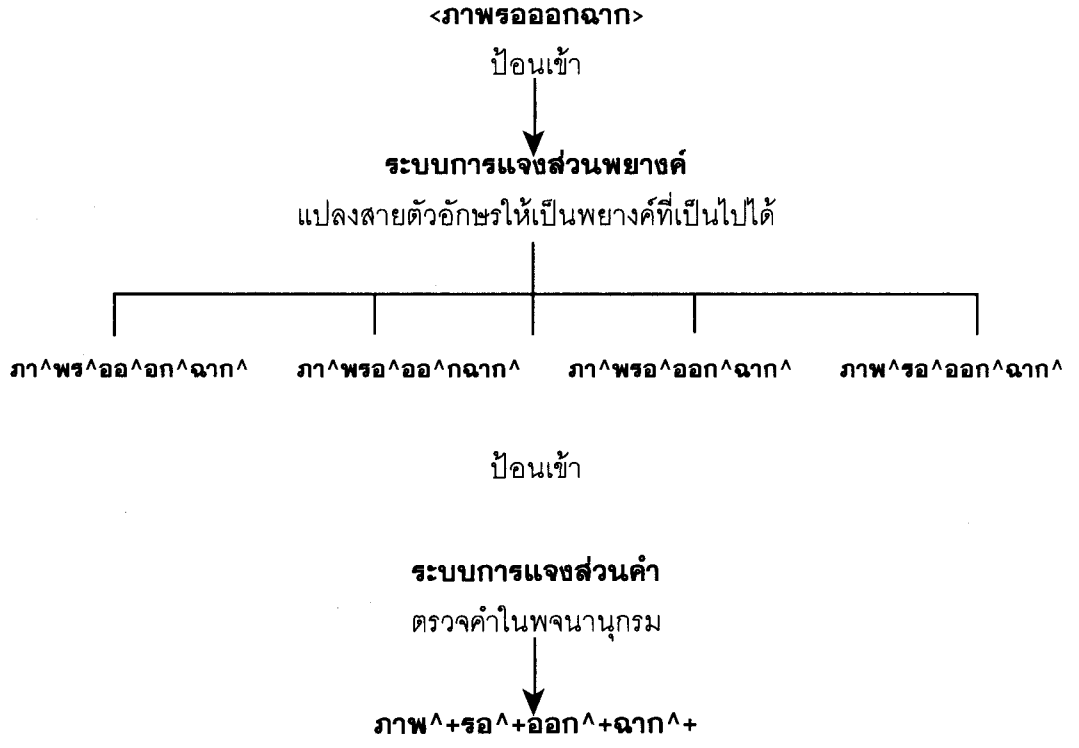
กลุ่มที่ 1 ระบบการอ่านด้วยตาเครื่อง/เข้าใจ -พิมพ์/ อ่านออกเสียง-พูดโต้ตอบ/ หรือกระทำตามคำสั่ง กลุ่มนี้จะประกอบด้วยเทคโนโลยีของการรู้จำตัวอักษรและการประมวลผลคำเพื่อการอ่านด้วยตาเครื่องและเข้าใจ เทคโนโลยีการแสดงผลเป็นตัวอักษรรูปร่างต่างๆ เพื่อเก็บและแสดงผลการอ่านบนจอหรือบนกระดาษ เทคโนโลยีการประมวลผลคำยังนำมาใช้ในระบบการค้น

คืนหรือตรวจสอบข้อมูลที่อ่านมาแล้วอีกด้วย นอกจากนี้หากต้องการให้เครื่องอ่านออกเสียง ก็ต้องอาศัยระบบการอ่านออกเสียง (text-to-speech) แปลงตัวอักษรให้เป็นรูปแทนทางเสียง แล้วใช้ระบบการสังเคราะห์เสียง สังเคราะห์เสียงพูดตามที่ระบุในรูปแทนทางเสียงนั้น

ระบบการอ่านเอกสารที่อยู่ในรูปกระดาษ ทั้งที่เป็นตัวพิมพ์ หรือตัวเขียน ประกอบด้วยระบบการรู้จำตัวอักษร (character recognition) และระบบการแจงส่วนหรือประมวลผลคำ (word parser or word processor) เมื่ออ่านด้วยตาเครื่องผ่านแสกนเนอร์แล้ว สัญญาณที่ได้จะเข้าสู่ระบบการรู้จำตัวอักษร ระบบนี้จะแปลงสัญญาณแสงที่เป็นรูปตัวอักษรนั้นให้เป็นรหัสแอสกี เพื่อนำไปเก็บบันทึกไว้เพื่อใช้แสดงบนจอ หรือพิมพ์เป็นฉบับกระดาษ เพื่อการค้นคว้า ตรวจชำระ หรือในกรณีของคนตาพิการสามารถนำไปต่อกับระบบการแปลงให้เป็นตัวอักษรพิมพ์นูนหรืออักษรเบรลล์ได้อีกด้วย หากไม่มีเทคโนโลยีทางภาษา อันได้แก่ ระบบการรู้จำตัวอักษรและระบบการประมวลผลคำแล้วเครื่องจะเก็บเอกสารในลักษณะเช่นเดียวกับภาพ ๑ ภาพ ไม่สามารถใส่คำสำคัญเพื่อค้นคว้า หรือค้นคืนได้ อีกทั้งไม่สามารถตรวจชำระหรือแก้ไข หรือนำไปแปลงรหัสให้เป็นอักษรเบรลล์ได้

รูปตัวอักษรที่บันทึกไว้แล้วด้วยการเขียนหรือการพิมพ์ลงบนกระดาษ หากต้องการจัดเก็บเป็นข้อมูลอิเล็กทรอนิกส์เพื่อเก็บรักษา จะทำอย่างไร คำถามข้อนี้อาจจะตอบง่าย ๆ ว่า ให้เก็บไว้เป็นภาพ ในหลายประเทศที่ตระหนักถึงความสำคัญของเอกสารโบราณที่อยู่ในรูปกระดาษที่ต้องอาศัยการเก็บรักษาที่สิ้นเปลืองงบประมาณมาก เริ่มเก็บหนังสือเหล่านี้ไว้ในรูปของภาพ เพราะเป็นวิธีที่ง่ายที่สุด ถ้าตอบเช่นนี้ก็ไม่ต้องอาศัยเทคโนโลยีทางภาษา แต่ถ้าถามต่อไปว่าจะเก็บไว้เพื่อนำไปใช้งานอะไร อาจจะได้คำตอบที่ทำให้เกิดการวิจัยเทคโนโลยีทางภาษาต่อไป เช่น มีเอกสารโบราณเป็นจำนวนมากในรูปของหนังสือ หรือสมุดบันทึก การศึกษาหาความรู้จากเอกสารดังกล่าวอาจจะต้องใช้เวลาในการไปยังห้องสมุดเฉพาะที่มีเอกสารเหล่านั้น ซึ่งหากเป็นห้องสมุดที่อยู่ในต่างประเทศก็จะสิ้นเปลืองมาก หากเอกสารเหล่านั้นมีเก็บไว้เป็นภาพดังกล่าวข้างต้น ก็อาจจะค้นเอกสารนั้นมาอ่านไปทีละหน้าจนกว่าจะพบข้อความหรือคำสำคัญที่จะค้นหา อย่างไรก็ตามในกรณีเช่นนี้ผู้ใช้ข้อมูลอาจต้องใช้เวลาเป็นปีหน้าจอคอมพิวเตอร์ หรือ สั่งให้เครื่องพิมพ์ออกมาเป็นฉบับบนกระดาษให้ก่อน ซึ่งก็กินทั้งเวลาและเนื้อที่ หากมีเทคโนโลยีการรู้จำตัวอักษร (Optical Character recognition-OCR) ข้อมูลเหล่านั้นจะถูกจัดเก็บเป็นอักษรต่ออักษรในรูปสัญญาณอิเล็กทรอนิกส์ และหากมีการเชื่อมต่อระบบการรู้จำตัวอักษรกับระบบการประมวลผลคำ (Word Processor) ผู้ใช้สามารถใช้ระบบการค้นคืน โดยวิธีการป้อนคำสำคัญใด ๆ ก็สามารถทำได้อย่างรวดเร็ว เราอาจจะย่นระยะทาง และระยะเวลาในการศึกษาเอกสารโบราณเหล่านั้นจากสิบปีเป็นสิบเดือน

ระบบการประมวลผลคำในภาษาไทยเป็นระบบเป็นที่สนใจเป็นอย่างมากในกลุ่มนักวิจัยด้านการประมวลผลภาษาด้วยคอมพิวเตอร์ ถึงกับกล่าวกันว่า *ภาษาไทยเป็นภาษาปราบเซียน* ทั้งนี้เพราะคำในภาษาไทยจะเรียงติดต่อกันโดยไม่มีการเว้นวรรคระหว่างคำเช่นในภาษาอังกฤษ การประมวลผลอาจใช้วิธีส่งสายตัวอักษรเข้าไปในพจนานุกรมและทำการทาบเทียบไปเรื่อย ๆ จนกว่าจะพบคำ ๆ แรกแล้วทำการทาบเทียบต่อไป สำหรับผู้เขียนซึ่งศึกษาวิจัยเกี่ยวกับการอ่านออกเสียงภาษาไทยด้วยคอมพิวเตอร์ (Luksaneeyanawin, 1989; Luksaneeyanawin and Niyompol, 1997; สุดาพร ลักษณียนาวิน, 2533, 2534, 2535) ได้ใช้ระบบการประมวลผลคำซึ่งประกอบด้วยส่วนการทำงาน 2 ส่วน (1) ระบบการแจงส่วนพยางค์ (syllable parser) ทำหน้าที่แปลงสายตัวอักษรให้เป็นพยางค์ที่เป็นไปได้ในภาษาไทย และ (2) ระบบการแจงส่วนคำ (word parser) ซึ่งทำหน้าที่ตรวจว่าพยางค์เหล่านั้นเป็นคำหรือประกอบกับพยางค์ที่ตามมาเป็นคำในภาษาไทยได้หรือไม่ ตัวอย่างเช่น สายตัวอักษร <ภาพรอออกฉาก> (ชื่อภาพเขียนของคุณจักรพันธ์ โปษยกฤต) เมื่อผ่านเข้าไปในระบบการประมวลผลคำจะมีผลดังนี้



(^ ใช้แทนเขตระหว่างพยางค์ และ + ใช้แทนเขตระหว่างคำ)

สายพยางค์สายที่ 4 **ภาพ^+รอ^+ออก^+ฉาก^+** เป็นสายเดียวที่พยางค์ทั้งหมดประกอบเป็นคำพูดได้ ด้วยวิธีนี้นอกจากจะได้สายคำที่แจงส่วนเป็นคำแล้วยังได้สายของรูปแทนทางเสียง (phonological representation) ของคำพูดซึ่งสามารถส่งไปให้ระบบการสังเคราะห์เสียงอ่านออกเสียงคำพูดนั้นออกมาได้ ในภาษาไทยมีกรณีที่การแจงส่วนคำแจงคำออกมาได้สองแบบ เรียกว่าเกิดกรณีกำกวมขึ้น เช่น <ตากลม> แจงได้เป็น ตา^+กลม ^+ หรือ ตาก^+ลม^+ <สามารถ> แจงได้เป็น สา^+มารถ^+ หรือ สา^+มา^+รถ^+ และ <เรือนรก> แจงได้เป็น เรือ^+นรก^+ หรือ เรือน^+รก^+ กรณีเช่นนี้จะไม่เกิดขึ้นในภาษาอังกฤษซึ่งมีการเว้นวรรคระหว่างคำ การตัดสินว่าสายคำที่กำกวมนั้นประกอบด้วยคำอะไรจึงต้องใช้ความรู้ที่อยู่เหนือคำ คือ วากยสัมพันธ์และความหมาย หรือ ในบางกรณีต้องใช้ปริบทของภาษาเข้ามาตีความ ในขณะนี้ยังไม่มีระบบการแจงส่วนในภาษาไทยที่แก้ปัญหานี้ได้

การสร้างระบบการแจงส่วนของภาษาไทยนี้เป็นการวิจัยสหวิทยาการ เพราะต้องนำความรู้ทางภาษาศาสตร์มาอธิบายเป็นกฎที่ชัดเจนเป็นข้อ ๆ มีวิธีการจัดลำดับเพื่อให้ประหยัดการใช้คำสั่งให้มากที่สุด ซึ่งเป็นเรื่องของวิทยาการคอมพิวเตอร์ การทำงานวิจัยร่วมกันนี้ นักภาษาศาสตร์ต้องมีความรู้พอสมควรเกี่ยวกับภาษาเครื่อง ในขณะเดียวกันนักวิทยาการคอมพิวเตอร์ก็ต้องศึกษาเนื้อหาของกฎทางภาษาเพื่อนำมาเขียนเป็นคำสั่งได้อย่างมีประสิทธิภาพ การออกแบบการทำงานของระบบจะต้องมีการเสนอให้ทุกฝ่ายที่เกี่ยวข้องได้เข้าใจ วิจารณ์ หรือเสนอให้ปรับปรุงให้เหมาะสมก่อนการลงมือวางระบบจริง ผู้เขียนขอให้รายละเอียดเกี่ยวกับการวิจัยทางภาษาศาสตร์ที่เกี่ยวข้องกับเรื่องการประมวลผลคำ ว่ามีลักษณะที่ต่างจากการวิจัยภาษาศาสตร์บริสุทธิ์อย่างไร ดังต่อไปนี้

ระบบการแจงส่วนพยางค์ (syllable parser) ระบบนี้จะจำลองสมรรถนะในการอ่านผสมตัวอักษรให้เป็นเสียง อันที่จริงการอธิบายวิธีการอ่านผสมตัวอักษรไทยนั้นได้มีผู้อธิบายไว้แล้วต่าง ๆ กัน แต่การอธิบายกฎให้ชัดเจนเป็นลำดับจากง่ายไปหายากและไม่ให้ซ้ำซ้อนนั้น ยังไม่มีผู้ใดเขียนไว้ ผู้เขียนจึงได้เขียนกฎดังกล่าวและผู้ช่วยวิจัยซึ่งเป็นวิศวกรคอมพิวเตอร์ได้ช่วยแปลงกฎเหล่านั้นให้เป็นภาษาเครื่อง ตัวอย่างเช่น

- <ก> แทนเสียง /k/ <า> แทนเสียง /a:/ ซึ่งรวมกันแล้วจะได้เป็น <กา> มีเสียงวรรณยุกต์สามัญ /0/ (เสียงกลางระดับ) ออกเสียงรวมกันว่า /ka:0/ กฎนี้ใช้กับอักษรพยัญชนะต้นที่เป็นอักษรกลาง <ก จ ด ฎ ต ฏ บ ป อ> และอักษรต่ำ <ค ฆ ช ซ ฌ ญ ฑ ฒ ณ ท ธ น พ ฟ ภ ม ย ร ล ว ฬ อ ฮ> และสระยาว คือ < า, ี, ื, ู, เ, แ, โ, _อ, เ_อ, เ_ีย, เ_ือ, _ัว>

- ถ้าอักษรพยัญชนะต้นที่เป็นอักษรสูง <ข ฉ ฐ ถ ผ ฝ ศ ษ ส ห> หรือ ชุดอักษรต่ำที่มี <ห> นำ เสียงวรรณยุกต์จะเปลี่ยนเป็น เสียงจัตวา /4/ (เสียงต่ำขึ้นสูง) ดังนี้เป็นต้น เรียกว่ากฎ

การแปลงรูปอักษรเป็นเสียง (letter to sound conversion rules) ในระบบการแปลงรูปอักษรเป็นเสียงในภาษาอังกฤษมักไม่ใช้วิธีการดังกล่าว เพราะในภาษาอังกฤษมีคำเป็นจำนวนมากที่เขียนเป็นกฎได้ยาก มีคำจำนวนน้อยที่สามารถระบุการออกเสียงในรูปของกฎ วิธีการที่ง่ายกว่าคือเก็บรูปการออกเสียงของคำแต่ละคำไว้ในรูปพจนานุกรม เนื่องจากภาษาอังกฤษมีการเว้นช่องว่างระหว่างคำและมีเครื่องหมายวรรคตอนชัดเจน การแจงคำจึงมักจะใช้การวิเคราะห์จากบนลงล่างเมื่อได้โครงสร้างทางไวยากรณ์ของคำแล้วก็สามารถบอกรูปเสียงได้ เช่น ประโยคว่า <He presented the paper.> ถ้า S => NP VP จะวิเคราะห์ประโยคได้เป็น He (NP) presented the paper (VP), วิเคราะห์ VP=> V NP จะได้ presented (V) the paper (NP) แต่ในประโยคว่า <The present is this book.> ในกรณีนี้ The present จะเป็น NP คำว่า <present> ถ้าเป็นคำกริยา (V) จะออกเสียงเป็น <pre'sent> ต่างจากเวลาที่เป็นคำนาม (N) จะออกเสียงเป็น <'present>

ในภาษาไทยกฎการสะกดคำค่อนข้างจะตรงไปตรงมา จะมีความซับซ้อนเฉพาะส่วนที่เกี่ยวกับกฎการออกเสียงวรรณยุกต์ ซึ่งมีเรื่องของลักษณะพยางค์ ชนิดของพยัญชนะต้น และเครื่องหมายวรรณยุกต์เป็นตัวกำหนด ในการเขียนกฎเหล่านี้ จะต้องแบ่งรูปตัวอักษรเป็นกลุ่มๆ ตามกฎที่เราจะระบุและใส่ไว้ในรูปของตาราง (Luksaneeyanawin, 1993a) เพื่อให้ผู้ช่วยวิจัยซึ่งทำหน้าที่แปลงกฎเหล่านั้นให้เป็นภาษาเครื่องเข้าใจได้ง่ายและทำงานได้สะดวก อย่างไรก็ตามมีกฎจำนวนมากในภาษาไทยที่ค่อนข้างจะซับซ้อนซึ่งผู้เขียนค้นพบจากการวิจัย คือ เรื่องของสระ <เอ> และ สระ <แอ> ซึ่งเป็นสระยาว ในการสะกดคำที่ไม่มีตัวสะกด สระคู่นี้จะออกเสียงยาวเสมอ แต่ในกรณีที่มีตัวสะกดในแม่ กน กง กม และมีรูปวรรณยุกต์ เอก หรือ โท สระคู่นี้มักจะเปลี่ยนเป็นเสียงสั้น เช่น

เสียงสระยาว	**เสียงสระเปลี่ยนเป็นสั้น**	
เลน	เล่น	เล้น (ทะเล้น)
แลน	แล่น	แล้น (แป้นแล้น)
เลง	เล่ง(เน่ยยี่)	เล้ง (ล้งเล้ง)
เวน	เว่น	เว้น
เซง	เซ่ง	เซ้ง
แลง	แล่ง	แล้ง

คำหลายคำที่ยกตัวอย่างมานี้ อาจจะไม่มีความหมาย เช่น เซง เว่น เซ่ง หลายคำเป็นส่วนหนึ่งของคำ เช่น เล้น เล่ง(เน่ยยี่) ฯลฯ และหลายคำมีข้อยกเว้น เช่น <เซ้ง> ออกเสียงยาวในเวลาที่ใช้ร่วมกับคำว่า <สวยเซ้ง > แต่ออกเสียงสั้นเวลาที่ใช้ร่วมกับคำว่า <ส้มเซ้ง > และ <แล้ง> ออกเสียงยาว ข้อยกเว้นต่างๆ ที่เป็นรายละเอียดปลีกย่อยเหล่านี้นักภาษาศาสตร์ทฤษฎีมักจะมองข้าม

ไป หรือไม่ก็บอกว่าเรื่องของคำที่มาจากภาษาอื่น คำพิเศษ ไม่นับรวมใน *ระบบเสียงมาตรฐาน* เมื่อผู้เขียนทำวิจัยในเชิงประยุกต์มากขึ้น ก็มักจะเกิดคำถามที่(บางครั้ง)ถูกนักวิจัยที่เน้นแต่ทฤษฎี "เลิกคิ้ว" ใส่ เช่น ผู้เขียนเขียนกฎให้พยางค์ตายทั้งสั้นและยาวมีเสียงวรรณยุกต์ได้ 3 เสียง ดังนี้

ขาด	คาด	*ค้าด
ขัด	*คั้ด	คัด

ในกฎที่เรียกว่า *ระบบเสียงมาตรฐาน* จะไม่มีเสียง *ค้าด (เสียงตรีในพยางค์ตายสระยาว) และ *คั้ด (เสียงโทในพยางค์ตายสระสั้น) ดังนี้

ขาด	คาด	–
ขัด	–	คัด

โดยจะอธิบายว่าภาษามีการเกิดช่องว่างในระบบ (systemic gap) ขึ้น ซึ่งเป็นลักษณะที่เป็นธรรมชาติ และช่องว่างเหล่านี้อาจจะมีคำยืมหรือคำแสดงความรู้สึกและอารมณ์ที่ไม่นับว่าเป็นส่วนสำคัญของภาษาเข้ามาเติมเต็มได้ ผู้เขียนซึ่งทำการศึกษาวิจัยเกี่ยวกับการแสดงความหมายในเชิงทัศนคติและอารมณ์ในภาษา รู้สึกคับข้องใจกับการตัดสินว่าส่วนใดในภาษาสำคัญกว่ากัน ระหว่างการสื่อความหมายอ้างอิงกับการสื่อความหมายทางทัศนคติและอารมณ์ของทฤษฎีมาตรฐาน ผู้เขียนมีความเห็นว่าทฤษฎีภาษาศาสตร์ที่ดีนั้นควรจะครอบคลุมปรากฏการณ์ทางภาษาใด ๆ ที่เกิดขึ้น หรือมีศักยภาพที่จะเกิดขึ้นได้ทุกกรณีมากกว่า "การกวาดข้อมูลที่ทำให้ไม่สามารถสร้างข้อสรุปชัดเจนลงใต้พรม" เช่น ในกรณีตัวอย่างจะบอกว่า ไม่มีคำว่า <ค้าด> หรือ <คั้ด> ในภาษาไทย แต่ถ้าเรานึกถึงคำว่า <โค้ก> <แซ้ด> <แว้ก> หรือ <คลั่ก> <พลั่ก> <เลิ่กลั่ก> ฯลฯ ซึ่งมีอยู่มากมายในภาษาไทยเพียงแต่ไม่ค่อยพบในภาษาเขียน เราก็คงต้องคิดใหม่ว่าเราต้องการ "ระบบเสียงมาตรฐาน" หรือ ระบบที่สามารถครอบคลุมปรากฏการณ์ในภาษาได้ครบถ้วนมากกว่ากัน การศึกษาวิจัยสหวิทยาการได้สร้างปัญหาที่ก่อให้เกิดปัญญาที่แจ่มชัดขึ้นในทางทฤษฎี

ในระบบการแจงส่วนคำที่กล่าวมาพอสังเขปแล้วนี้ ระบบทำหน้าที่ตรวจว่าพยางค์เหล่านั้นเป็นคำหรือประกอบกับพยางค์ที่ตามมาเป็นคำในภาษาไทยหรือไม่ ในการตรวจจะต้องใช้พจนานุกรม การออกแบบพจนานุกรมอิเล็กทรอนิกส์เป็นอีกเรื่องหนึ่งที่เป็นการวิจัยในแนวสหวิทยาการ การลำดับคำในพจนานุกรมที่เป็นเล่มนั้นมีการลำดับตามตัวอักษรเป็นส่วนใหญ่ อาจมีพจนานุกรมชนิดพิเศษที่เรียงลำดับตามเสียง นักจิตวิทยาที่วิจัยด้านการนึกรู้คำ ทำการวิจัยเชิงทดลองที่แสดงให้เห็นว่า พจนานุกรมในสมองของมนุษย์นั้นมีความซับซ้อนเป็นอย่างยิ่ง (Taft, 1991) คำจะถูกเรียงด้วยวิธีการอันหลากหลาย การวิจัยใช้การทดลองกับกลุ่มตัวอย่าง ซึ่งจะได้เห็นคำ หรือได้ยินคำเพียงชั่วระยะเวลาสั้น ๆ ผู้ทดลองจะตัดสินว่า คำที่เห็นนั้นเป็นคำในภาษาหรือไม่ เช่น มีการวิจัยที่พบว่าเมื่อเปรียบเทียบ การนึกรู้คำหลัก (content word) กับคำไวยากรณ์ (grammatical word) ที่มีตัว

สะกดคล้ายคลึงกันและมีจำนวนตัวอักษรเท่ากัน การนึกรู้คำหลักจะใช้เวลานึกรู้เร็วกว่ามาก ในการทดลองว่าคำมีการจัดเก็บอย่างไร จะใช้การกระตุ้น (activate) คลังคำ โดยก่อนที่จะได้เห็นคำที่จะนึกรู้นี้ จะมีคำที่สร้างขึ้นมาเพื่อพิสูจน์สมมติฐานเกี่ยวกับการจัดลำดับ คำที่สร้างขึ้นนี้จะปรากฏให้เห็นเพียงชั่วเสี้ยววินาที เรียกว่าคำที่มาก่อน (prime) ระยะเวลาในการตอบสนอง (reaction time) จะแสดงให้เห็นว่าการจัดเก็บคำมีลักษณะที่สัมพันธ์กับคำที่มาก่อนอย่างไร การให้คำที่มาก่อนในรูปเสียง (auditory priming) แสดงให้เห็นว่าหากมีการกระตุ้นด้วยเสียง คำในพจนานุกรมในสมองนี้จะมีการจัดเรียงต่างออกไป การวิจัยของนักจิตวิทยาเกี่ยวกับการนึกรู้คำให้ความรู้ที่เป็นประโยชน์แก่การจัดทำพจนานุกรมอิเล็กทรอนิกส์ซึ่งมีการประยุกต์ใช้งานได้มาก

ความต้องการของมนุษย์มักไม่มีข้อจำกัด เทคโนโลยีใหม่ๆ ด้านการรู้จำเกิดขึ้นตลอดเวลา ในปัจจุบันมีคอมพิวเตอร์ขนาดฝ่ามือ (palm top computer) ซึ่งผู้ใช้สามารถเขียนบันทึกลงไปบนหน้าจอ ระบบรู้จำตัวอักษร และ ระบบการประมวลผลคำในคอมพิวเตอร์จะจัดเก็บข้อมูลให้ ข้อมูลนี้จะถูกจัดเก็บในรูปสัญญาณอิเล็กทรอนิกส์ เพื่อนำไปแสดงเป็นตัวอักษรบนหน้าจอหรือพิมพ์ลงบนกระดาษได้ นอกจากนี้ยังใช้ในระบบการค้นคืนได้อีกด้วย การวิจัยด้านระบบการรู้จำตัวอักษรต้องอาศัยความรู้เกี่ยวกับลักษณะสำคัญ หรือ ลักษณะบ่งความต่าง (distinctive features) ซึ่งใช้ในการจำแนกอักษรในภาษาออกจากกันอย่างมีประสิทธิภาพกว่าการที่มองตัวอักษรทั้งหมดเป็นภาพ ผู้เขียนได้ร่วมเป็นอาจารย์ที่ปรึกษาแก่งานวิจัยด้านการรู้จำตัวอักษรไทยที่เป็นลายมือเขียน(วิชาพานิช 2539; ประเสริฐ ฉอรุ่งเรืองวิวัฒน์, 2541) โดยเสนอให้นำความรู้เกี่ยวกับ ลักษณะบ่งความต่าง (distinctive features) ของอักษรไทย (Luksaneeyanawin, 1993b) ซึ่งผู้เขียนได้วิเคราะห์ไว้ไปใช้ในการวางระบบ งานวิจัยทั้งสองมีอัตราการรู้จำอยู่ในระดับสูง อย่างไรก็ตามการวิจัยดังกล่าวมุ่งที่จะรู้จำตัวอักษรเฉพาะบางส่วน โดยผู้วิจัยพยายามใช้วิธีการทางคณิตศาสตร์ในการสร้างแบบเปรียบ (templates) ของตัวอักษร และ ใช้วิธีการทางสถิติในการคำนวณความน่าจะเป็นของข้อมูลที่นำมาเปรียบเพื่อตัดสินว่าอักษรที่มีคืออักษรอะไร ระบบที่วิจัยมาได้นี้ยังต้องเชื่อมต่อกับระบบการประมวลผลคำภาษาไทย เพื่อให้ระบบสามารถรู้จำคำที่เป็นรูปเขียนเพื่อนำมาใช้ในการค้นคืน หรือจัดเรียงพิมพ์ เพื่อการใช้งานอื่นๆต่อไป

ระบบการอ่านด้วยตาเครื่อง/เข้าใจ-พิมพ์ ใช้ได้ดีกับเอกสารที่มีการพิมพ์หรือเขียนไว้แล้ว ระบบนี้หากนำไปต่อกับระบบการสังเคราะห์เสียง ก็สามารถอ่านเอกสารที่เครื่องอ่านเข้าใจแล้วนี้ให้ออกมาเป็นเสียงได้ด้วย เทคโนโลยีทางเสียงด้านการสังเคราะห์เสียงมีความก้าวหน้าไปมากภายในช่วงสองทศวรรษที่ผ่านมา เมื่อผู้เขียนเป็นนักเรียนที่มหาวิทยาลัยเอดินเบอระในระหว่างปี 1976-1983 นั้น การใช้เครื่องคอมพิวเตอร์ในการวิเคราะห์หรือสังเคราะห์เสียงยังไม่แพร่หลาย การวิเคราะห์เสียงจะต้องอาศัยการตัดต่อเทปของคำหรือคำพูดที่ต้องการจะวิเคราะห์ด้วยมือ แล้วจึง

ป้อนข้อมูลจากเครื่องบันทึกเสียงเข้าเครื่องวิเคราะห์ เครื่องจะวิเคราะห์สัญญาณเสียงพูด ด้วยการคำนวณจากองค์ประกอบของสัญญาณเสียงนั้นเป็นค่าความถี่มูลฐาน (FO) ค่าความเข้มของเสียง (Intensity) หรือ ค่าความถี่กำทอน (Formant Frequency) แล้วพิมพ์ผลการวิเคราะห์เป็นแผ่นภาพคลื่นเสียงเพื่อนำมาวัดด้วยมือและคำนวณค่าความถี่ต่าง ๆ ให้เป็นตัวเลข หากมีความผิดพลาดในการพิมพ์ผลการวิเคราะห์บนแผ่นกระดาษ จะต้องทำกระบวนการทั้งหมดใหม่ตั้งแต่การป้อนข้อมูลเข้าเครื่อง ทั้งนี้เพราะเครื่องวิเคราะห์เสียงยังไม่มีหน่วยความจำเช่นเครื่องวิเคราะห์ในปัจจุบัน การศึกษาทางสัทศาสตร์จึงต้องใช้เวลาและความอดทนในการทำงานวิเคราะห์เป็นอย่างมาก การวิเคราะห์คลื่นเสียงในลักษณะดังกล่าวนอกจากเป็นงานของนักสัทศาสตร์ ซึ่งเป็นสาขาที่ผู้เขียนทำวิจัยแล้ว ยังมีนักวิทยาศาสตร์จากสาขาวิชาอื่น ๆ มาใช้ในการศึกษาวิจัยต่าง ๆ กันเป็นที่น่าสนใจ เช่นนักสัตววิทยาที่ศึกษาพฤติกรรมการร้องของลิง หรือ นกชนิดต่าง ๆ หรือนักวิจัยทางสมุทรศาสตร์ที่นำข้อมูลเกี่ยวกับเสียงพูดที่พูดในระดับความลึกต่าง ๆ กันมาศึกษา ผู้เขียนได้มีโอกาสกลับไปศึกษาวิจัยระดับหลังปริญญาเอกที่มหาวิทยาลัยเอดินเบอระอีกครั้งในปี 1989 และ 1993 ความแตกต่างที่เห็นได้ชัดที่สุด คือ เครื่องมือวิเคราะห์เสียงเหล่านั้นถูกแทนที่ด้วยเครื่องคอมพิวเตอร์ นักเรียนจะเรียนรู้การวิเคราะห์เสียงด้วยเครื่องประมวลผลสัญญาณ (Digital Signal Processing Workstation) และต้องเรียนรู้ภาษาเครื่องเพื่อเขียนโปรแกรมเสริมที่ต้องการเพิ่มเติมนอกเหนือไปจากที่เครื่องมี ในปี 1989 นั้นมีภาควิชาปัญญาประดิษฐ์ ซึ่งมีการวิจัยหลัก ๆ 3 ด้าน คือ ด้านสมรรถนะทางภาษาเรียกว่าการประมวลผลภาษาธรรมชาติ สมรรถนะด้านการวิเคราะห์และแก้ปัญหาเรียกว่าระบบผู้เชี่ยวชาญ และสมรรถนะทางกายเรียกว่าระบบหุ่นยนต์ นักวิจัยในศูนย์วิจัยเทคโนโลยีทางภาษาซึ่งผู้เขียนได้เข้าไปทำงานวิจัยนั้น มีทั้งนักคณิตศาสตร์ นักวิทยาการคอมพิวเตอร์ วิศวกรไฟฟ้า วิศวกรเครื่องกล นักภาษาศาสตร์ รวมทั้งครูสอนภาษา นักวิจัยสหวิทยาการเหล่านี้ทำงานร่วมกันอย่างประสมประสาน การสัมมนาในศูนย์วิจัยมักจะมีคำถามทั้งจากผู้เสนอและผู้ฟังที่กระตุ้นเร้าการวิจัยเป็นอย่างยิ่ง นับเป็นเวลาของการวิจัยที่ผู้เขียนได้เรียนรู้เป็นอย่างมากในช่วงเวลาสั้น ๆ คอมพิวเตอร์เป็นเครื่องมือที่ช่วยให้การวิจัยเป็นไปอย่างรวดเร็ว การวิเคราะห์เสียงที่ผู้เขียนเคยใช้เวลาในการวิเคราะห์เป็นปีนั้น ย่นเวลาลงมาเป็นเดือนสำหรับข้อมูลปริมาณเดียวกัน นักเรียนแต่ละคนจะมีงานอดิเรกในการเขียนโปรแกรมการวิเคราะห์ต่าง ๆ โดยใช้เครื่องคอมพิวเตอร์ที่พ้นสมัยแล้ว เป็นการเล่นแข่งขันความสามารถกันที่น่าทึ่งมาก มีนักเรียนปริญญาเอกคนหนึ่งแสดงการทำงานการประมวลผลสัญญาณเสียงพูดบนเครื่องคอมพิวเตอร์แอปเปิล 8 บิท ซึ่งเป็นของเล่น เครื่องสามารถประมวลผลสัญญาณได้ดีเช่นเดียวกับเครื่อง workstation เพียงแต่ผู้แสดงต้องชงน้ำชาให้ผู้เขียนดื่มเพื่อขอบคุณสำหรับความอดทนขณะรอผลการคำนวณ มีข้อคิดจากเรื่องนี้ว่าระบบการทำงานที่พัฒนาขึ้นโดยมีข้อจำกัดนั้นในบางครั้งก็ทำให้การสร้างระบบมี

วิธีคิดที่กระชับรัดกุมและประหยัดขึ้น ผู้เขียนซึ่งมีความคิดที่จะวิจัยด้านการสังเคราะห์เสียงเมื่อกลับมาประเทศไทยแล้ว จึงเห็นความเป็นไปได้ในการวิจัยเรื่องนี้โดยใช้เครื่องคอมพิวเตอร์ขนาดเล็กในการวิเคราะห์และตัดต่อคลื่นเสียงให้เป็นคำพูด[3] โดยได้กำหนดหัวข้อการวิจัย คือ การสังเคราะห์เสียงด้วยการเชื่อมต่อสัญญาณโดยใช้หน่วยพยางค์เป็นหลัก (A syllable based waveform concatenation speech synthesis) การคิดใช้หน่วยพยางค์เป็นหลักนั้นเพราะได้ศึกษาเรื่องของพยางค์ในภาษาไทยเป็นงานวิจัยพื้นฐานทางภาษาศาสตร์มาก่อนแล้ว การที่ศึกษาเรื่องนี้เพราะเคยมีอาจารย์ทางคอมพิวเตอร์ท่านหนึ่ง[4] ถามผู้เขียนขณะที่ร่วมกันศึกษาความเป็นไปได้ของการเก็บสัญญาณเสียงพยางค์ในหน่วยความจำเพื่อใช้ในระบบการสังเคราะห์เสียง การถามคำถามในลักษณะนี้จะเกิดขึ้นในการวิจัยสหวิทยาการเสมอ ผู้เขียนไม่เคยได้ยินคำถามประเภทนี้จากนักภาษาศาสตร์บริสุทธิ์

ผู้เขียนได้เริ่มศึกษาเรื่องนี้โดยเขียนกฎการเกิดพยางค์ที่มีศักยภาพที่จะเกิดขึ้นได้ในภาษาไทยและให้ผู้ช่วยวิจัย[5] เขียนโปรแกรมสร้างพยางค์เหล่านี้ขึ้นมาโดยใช้โปรแกรม Database ง่าย ๆ โดยมีกฎเพียง 5 ข้อ ที่ใช้ในการสร้างพยางค์ในภาษาไทย (Luksaneeyanawin, 1989, 1992, 1993a) เช่น

Open Syllables = 33 C&CC * 12 V:&VV: * 5T

อ่านว่า พยางค์เปิด–สระยาว เกิดขึ้นจาก พยัญชนะต้นเดี่ยว(C) และ พยัญชนะควบกล้ำ (CC) 33 หน่วย คูณด้วย สระเดี่ยวยาว (V:) และสระผสมยาว (VV:) 12 หน่วย คูณด้วย วรรณยุกต์ (T) 5 หน่วย จากกฎโปรแกรมจะสร้างพยางค์ทั้งหมดจำนวน 1,980 พยางค์ (33*12*5) และแสดงรูปให้ในตาราง ผู้เขียนและผู้ช่วยวิจัยได้ศึกษารูปพยางค์เหล่านั้น เพื่อระบุว่าพยางค์นั้นเกิดเป็นคำ (x) หรือ คำแสดงความรู้สึกหรือคำยืม (e) หรือ ส่วนของคำ (y) ได้ หรือ ไม่ได้ (o) และถ้าเป็นคำจะมีรูปเขียน < > อย่างไร ตัวอย่างเช่น

รูปพยางค์	ไม่ใช่คำ	คำ	ส่วนของคำ	รูปเขียน
/pi:0/	–	x	y	<ปี, จำปี, ปีติ,>
/pe:0/	–	x	y	<เป, เปย์, อัปเปหิ,>
/px:0/	–	x	–	<แป>
/pv:0/	o	–	–	<ปือ>
/pq:0/	o	–	–	<เปอ>
/pa:0/	–	x	y	<ปา, จำปา, ปาฏิหาริย์,>
/pu:0/	–	x	y	< ปู, ปูชนียะ, ริปู>
/po:0/	–	x	y	< โป,เฉโป, โปเก,>

/p@:0/ - x y < ปอ, ปรเมศร์, ปรมินทร์,> ฯลฯ

จากการศึกษาพยางค์ลักษณะต่าง ๆ 5 ชนิดพบว่าพยางค์ที่มีศักยภาพที่จะเกิดได้ทั้งหมดในภาษาไทย มี 26,928 พยางค์ ในจำนวนนี้ มีเพียง 3,797 พยางค์ที่เกิดเป็นคำโดยลำพังได้ (x) และ 1,054 พยางค์ที่พบเป็นคำแสดงความรู้สึกหรือคำยืม (e) และ 1,061 พยางค์ที่พบเป็นส่วนของคำ (y) มีถึง 21,016 พยางค์ที่ไม่พบว่าใช้ในการสร้างคำ (o) ในภาษาแต่อย่างใด ฐานข้อมูลนี้นอกจากจะช่วยตอบคำถามเกี่ยวกับความเป็นไปได้ในการสังเคราะห์เสียงโดยใช้พยางค์เป็นหน่วยหลักแล้ว ยังใช้เป็นข้อมูลในระบบการประมวลผลคำ และเป็นประโยชน์ต่อการศึกษาในเชิงทฤษฎีเกี่ยวกับการเปลี่ยนแปลงของเสียงในเชิงประวัติอีกด้วย (Luksaneeyanawin, 1992)

เนื่องจากในช่วงต้นของทศวรรษที่ 90 เทคโนโลยีหน่วยความจำยังแพงมาก ผู้เขียนได้เคยใช้วิธีการสังเคราะห์เสียงด้วยเทคนิคการเชื่อมต่อคลื่นเสียงโดยใช้พยางค์เป็นหน่วยหลัก (Luksaneeyanawin, 1989) ซึ่งเปลืองเนื้อที่ในหน่วยความจำค่อนข้างมากถ้าต้องการเก็บพยางค์ที่เป็นไปได้ทั้งหมดในภาษาไทย จึงได้เกิดความคิดในเชิงทฤษฎีใหม่ คือใช้หน่วยทางเสียงที่เรียกว่าหน่วยเริ่ม (onsets) และหน่วยตาม (rhymes) ทั้งนี้เพราะระลึกได้ถึงปัญหาที่ได้เคยอภิปรายเกี่ยวกับการสังเคราะห์ภาษาไทยซึ่งเป็นภาษาที่มีวรรณยุกต์ กับ ดร.สตีฟ อิสาร์ด [6] นักเทคโนโลยีทางเสียงที่เสนอทฤษฎีหน่วยคู่เสียง (diphone) ดร.อิสาร์ดเสนอให้ใช้เทคนิคการเชื่อมต่อคลื่นเสียงโดยใช้หน่วยคู่เสียงซึ่งเหมาะสำหรับภาษาที่มีเสียงพยัญชนะเรียงติดต่อกันเป็นจำนวนมาก เช่น <strength> การใช้หน่วยคู่เสียง /st/ /tr/ /re/ /eng/ and /ngth/ ในภาษาอังกฤษทำให้สามารถสังเคราะห์เสียงคำที่เกิดจากการเติมวิภัติปัจจัย (affix) ซึ่งเป็นลักษณะที่สำคัญของภาษาอังกฤษได้เป็นอย่างดี แต่หากมาใช้กับภาษาที่มีโครงสร้างพยางค์ง่าย ๆ แต่มีเสียงวรรณยุกต์ที่มีลักษณะทางสัทศาสตร์ที่แปรไปตามโครงสร้างพยางค์จะทำให้ดีได้ยาก การเสนอทฤษฎีหน่วยเริ่มและหน่วยตามสำหรับการสังเคราะห์เสียงในภาษาไทยที่ผู้เขียนเสนอใหม่ สามารถลดจำนวนหน่วยหลักในการสังเคราะห์เสียงที่เป็นพยางค์จำนวน 26,928 พยางค์ดังกล่าวแล้ว มาเป็นหน่วยเริ่มและหน่วยตามเพียง 1,200 หน่วยเท่านั้น ด้วยแนวคิดใหม่นี้ทำให้ผู้เขียนสามารถสร้างระบบการอ่านออกเสียงภาษาไทยด้วยคอมพิวเตอร์ด้วยซอฟท์แวร์ โดยปฏิบัติงานบนเครื่องคอมพิวเตอร์พีซี 386 ได้อย่างประหยัดเนื้อที่ของหน่วยความจำนอกจากนี้ระบบยังสามารถอ่านออกเสียงคำได้โดยไม่มีข้อจำกัดเช่น อ่านคำสแลงที่พึ่งเกิดในปีที่สร้างระบบเสร็จ เช่น หน่อมแน้ม และ สยึมกึ๋ย ได้เป็นอย่างดี นอกจากนี้ยังสามารถลดจำนวนหน่วยตามลงได้อีกด้วยการสังเคราะห์ซ้ำ (resynthesis) (Luksaneeyanawin and Conkie, 1993; Luksaneeyanawin, 1995) หลังจากการวิจัยเรื่องการสังเคราะห์ซ้ำโดยใช้หน่วยตามแล้ว ผู้เขียนยังได้แนะนำให้นิสิตใช้หน่วยเสียงอนุภาค (microphoneme) และหน่วย

เสียงอนุภาคในช่วงคลื่นเสียงคงที่กับช่วงเชื่อมต่อในการสังเคราะห์ซ้ำอีก (เอกพล อนุสุเรนทร์, 2541; ชัชวาลย์ หาญสกุลบรรเทิง, 2542) ได้ผลการวิจัยเป็นที่น่าพอใจยิ่ง งานวิจัยในลักษณะดังกล่าวคงจะไม่เกิดขึ้นหากนิสิตขาดความรู้เกี่ยวกับลักษณะของคลื่นเสียงในเชิงสัทศาสตร์ทั่วไป และความรู้เกี่ยวกับระบบเสียงของภาษาไทย แม้การวิจัยด้านการสังเคราะห์เสียงภาษาไทยจะก้าวหน้าไปมากแล้วก็ตาม เทคนิคการเชื่อมต่อคลื่นเสียงโดยใช้พยางค์เป็นหน่วยหลักก็ยังใช้ได้ดีในการสังเคราะห์คำพูดสั้น ๆ ที่ประกอบด้วยคำพูดจำนวนไม่มากนัก เช่น ระบบที่ธนาคารใช้ในการประกาศคิวของลูกค้าตามลำดับหมายเลขบัตรของลูกค้า หรือระบบเอทีเอ็มจ๊ะจ๋า ทั้งนี้เพราะเป็นระบบที่พัฒนาขึ้นได้โดยง่ายและไม่จำเป็นต้องใช้เทคโนโลยีทางภาษาแต่อย่างใด

เมื่อระบบการสังเคราะห์เสียงในภาษาไทยเริ่มทำงานได้แล้วนั้น ก็เริ่มมีความต้องการจากผู้ใช้คือกลุ่มคนตาบอดที่ต้องการระบบการอ่านออกเสียงภาษาไทยด้วยคอมพิวเตอร์เพื่อใช้ในการอ่านหรือแก้ไขข้อความในเอกสารอิเล็กทรอนิกส์ ผู้เขียนได้เริ่มศึกษาระบบการอ่านจอภาพโดยมีผู้ช่วยวิจัยตาบอดเป็นผู้ทดลองใช้และวิเคราะห์วิธีการทำงานของระบบในการสื่อสารระหว่างคนตาบอดกับเครื่องโดยใช้ระบบการอ่านจอภาพภาษาอังกฤษเป็นตัวอย่าง คนตาบอดยังแสดงความต้องการระบบที่สามารถพกพาไปเชื่อมต่อกับคอมพิวเตอร์ใด ๆ ที่มีอยู่ในสถานที่ที่เขาจะไปทำงานได้อีกด้วย แนวความคิดในการพัฒนาระบบให้เป็นฮาร์ดแวร์จึงเกิดขึ้น ทั้งนี้ผู้เขียนได้รับความร่วมมือเป็นอย่างดีจากหน่วยวิจัยเชิงเลข คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย โดยการนำของ ดร.เอกชัย ลีลารัศมี ในการพัฒนาระบบดังกล่าวเป็นฮาร์ดแวร์ที่พกพาไปได้ เรียกว่า ซียูทอล์ค และมีการพัฒนาให้ระบบการใช้งานให้สามารถเชื่อมต่อกับระบบการอ่านออกเสียงภาษาอังกฤษด้วยคอมพิวเตอร์ที่มีการผลิตจำหน่ายทั่วไป เพื่อให้สามารถอ่านออกเสียงได้ทั้งสองภาษา (Luksaneeyanawin and Niyompol, 1997)

การเก็บข้อมูลในรูปสัญญาณอิเล็กทรอนิกส์ในปัจจุบันใช้การพิมพ์ผ่านแป้นพิมพ์เพื่อแปลงให้เป็นสัญญาณแอสกี เพื่อแสดงเป็นตัวอักษรบนหน้าจอและบันทึกไว้ ภาษาไทยเป็นภาษาหนึ่งที่ก่อปัญหาให้กับการแสดงสัญญาณบนหน้าจอสำหรับภาษาไทย เพราะรหัสแอสกีมีลักษณะเรียง (linear) จากซ้ายไปขวา เช่นเดียวกับอักษรโรมันและกรีก ในขณะที่ในภาษาไทยตัวอักษรแม้จะเรียงจากซ้ายไปขวาแต่มีระดับการเรียงตามแนวตั้งได้ 4 ระดับ คือ นอกจากรูปพยัญชนะและสระบางตัวที่วางอยู่บนบรรทัดแล้ว ยังมีรูปสระที่วางอยู่เหนือหรืออยู่ใต้รูปพยัญชนะได้ด้วย นอกจากนี้ยังมีรูปอักษรวรรณยุกต์ที่วางเหนือรูปพยัญชนะหรือเหนือรูปสระในกรณีที่มีรูปสระวางเหนือพยัญชนะนั้นอีกด้วย การออกแบบตัวอักษร (font design) จึงเป็นเทคโนโลยีทางภาษาที่สำคัญอีกชนิดหนึ่ง และเคยก่อปัญหาให้กับการเรียงพิมพ์หรือจัดหน้ากระดาษด้วยคอมพิวเตอร์ (Desk Top Publishing-DTP) เป็นอย่างมากในระยะแรกของการนำคอมพิวเตอร์เข้ามาใช้ในการจัดหน้า

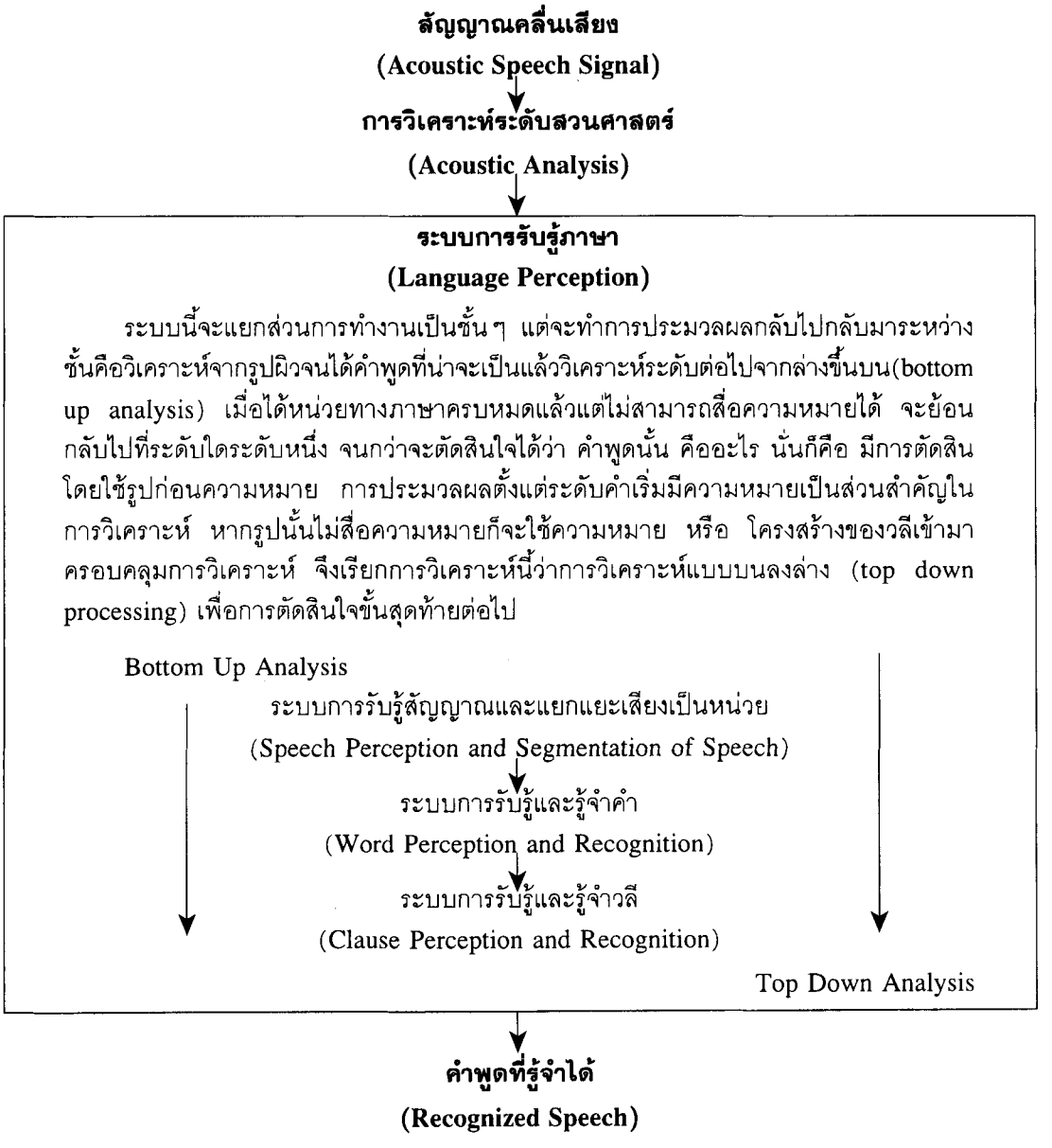
ต้นฉบับเพื่อการพิมพ์ในภาษาไทย อันที่จริงปัญหานี้ก็ยังไม่หมดไป โดยเฉพาะการจัดช่องไฟในกรณีที่มีการจัดกั้นหน้ากั้นหลังให้เสมอกัน แต่เดิมนั้นช่างเรียงพิมพ์ที่มีฝีมือจะจัดช่องไฟของการวรรคเล็ก และ วรรคใหญ่ โดยใช้ตาชั่งน้ำหนักในการจัดเรียง สมรรถนะในด้านการชั่งน้ำหนักของช่องไฟด้วยตามนุษย์นั้นยังไม่มีระบบการจัดหน้าต้นฉบับเพื่อการพิมพ์ในภาษาไทยใดๆ ทำได้ดีเท่าเรื่องนี้เป็นการวิจัยสหวิทยาการในสาขาวิทยาศาสตร์การพิมพ์ ซึ่งในประเทศไทยยังขาดผู้สนใจการวิจัยด้านนี้

กลุ่มที่ 2 ระบบการฟังด้วยหูเครื่อง/เข้าใจ -พิมพ์/ อ่านออกเสียง-พูดโต้ตอบ/ หรือกระทำตามคำสั่ง ด้วยเหตุที่การป้อนข้อมูลเข้าเครื่องคอมพิวเตอร์ต้องอาศัยการพิมพ์ลงบนแป้นพิมพ์ มีคนเป็นจำนวนมากที่ไม่มีทักษะในการพิมพ์เลย หรือในบางครั้งมีความต้องการติดต่อสื่อสารกับเครื่องโดยใช้การพูดเพราะจะต้องใช้สายตาหรือมือในการทำงานอย่างอื่น ไม่สามารถติดต่อกับเครื่องด้วยการพิมพ์หรือเขียนในขณะนั้นได้ เพื่อแก้ปัญหาในการสื่อสารกับเครื่องผ่านแป้นพิมพ์หรือแผ่นเขียน เทคโนโลยีทางภาษาอีกประเภทหนึ่งจึงเกิดขึ้น ได้แก่ การฟังด้วยหูเครื่อง/เข้าใจ เทคโนโลยีนี้ได้เข้ามาแทนการมีเลขานุการส่วนตัวที่คอยจดเชาวเลขจากข้อมูลที่เจ้านายบอกเป็นคำบอก แล้วนำไปพิมพ์ลงบนกระดาษ ปัจจุบันมีการคิดค้นระบบการรู้จำคำพูด (speech recognition system) ซึ่งประกอบด้วยการรู้จำเสียงและคำพูดในภาษาในระดับที่ผลิตออกจำหน่ายได้แล้วในภาษาอังกฤษ[7] ระบบดังกล่าวสามารถรู้จำคำพูดภาษาอังกฤษได้โดยไม่มีข้อจำกัดทั้งในแง่ของผู้พูดและคำพูด ในกรณีที่เป็นคำแปลกๆ เช่น ชื่อเฉพาะ ระบบมักจะรู้จำผิด กล่าวคือระบบไวยากรณ์และพจนานุกรมจะถูกบังคับให้ตัดสินว่าคำพูดนั้นเป็นคำพูดอะไรในภาษาอังกฤษ อย่างไรก็ตามผู้ใช้ก็อาจจะฝึกระบบให้รู้จำชื่อเฉพาะนี้ได้โดยสร้างพจนานุกรมคำเฉพาะขึ้น โดยอาจใช้อักษรย่อแทนชื่อเฉพาะนั้นได้ บริษัทผู้จำหน่ายระบุว่าหากมีการฝึกระบบให้รู้จำสำเนียงของผู้พูดแล้วระบบจะทำงานได้ดีมาก โดยเฉพาะภาษาเฉพาะแบบ เช่น ภาษากฎหมาย หรือภาษาทางการแพทย์ โดยเฉพาะการพิมพ์ตามคำบอกของแพทย์ทางรังสีที่อ่านภาพเอ็กซเรย์ แพทย์สามารถอ่านฟิล์มเอ็กซเรย์เป็นคำบอกในขณะที่ตาดูฟิล์ม เครื่องจะรู้จำคำพูดนั้นและแสดงคำบอกของแพทย์บนจอเพื่ออ่านทบทวนอีกครั้งหนึ่ง ก่อนส่งผลการวินิจฉัยนั้นต่อไป

สำหรับระบบการรู้จำคำพูดภาษาไทยนั้นยังอยู่ระหว่างการวิจัยพื้นฐาน ในระยะแรกคือช่วงทศวรรษที่ 80 มีความพยายามที่จะสร้างระบบการรู้จำคำพูดจำนวนจำกัดในภาษาไทย ส่วนใหญ่จะเป็นการรู้จำตัวเลขที่พูดเป็นคำๆ ระบบการรู้จำคำพูดยังใช้วิธีการรู้จำคำเหมือนการรู้จำสัญญาณภาพ ๆ หนึ่ง เช่นเดียวกับการรู้จำตัวอักษรที่ไม่มีระบบการประมวลผลภาษาเข้ามาจำแนกหน่วยในภาษา ระบบดังกล่าวจึงมักมีระบบการทำงานง่ายๆ ไม่ซับซ้อน ถ้าเทียบความสามารถทางภาษาก็คงได้เท่ากับลิงชิมแปนซีที่ถูกฝึกให้จำความหมายโดยการเทียบกับรูปแบบ (pattern matching)

ของแผ่นพลาสติกสีต่าง ๆ นั่นเอง ผู้เขียนพยายามที่จะให้นิสิตที่สนใจการวิจัยด้านการรู้จำคำพูดโดยมุ่งหมายการรู้จำคำพูดในระบบภาษาไทยเข้าใจความซับซ้อนของระบบการรู้จำที่มีระบบของภาษาเข้ามากำกับ (Language Modelling) โดยสร้างรูปแสดงระบบการรับรู้และรู้จำของมนุษย์ที่นักภาษาศาสตร์อธิบายไว้ในเชิงทฤษฎี (Flores d'Arcais, 1988) ด้วยแผนผังง่าย ๆ ดังต่อไปนี้

รูปที่ 3 แผนผังแสดงการทำงานของระบบการรับรู้และรู้จำคำพูดของมนุษย์



จะเห็นจากรูปว่าการรับรู้สัญญาณและแยกแยะออกเป็นหน่วยนั้นจะต้องอาศัยการวิเคราะห์ทั้งในระดับสวนศาสตร์(acoustics) และการมีโมเดลทางภาษาเข้ามาในการประกอบเสียงที่วิเคราะห์ทางสวนศาสตร์มาแล้วให้เป็นหน่วยต่างๆ ในภาษา การวิจัยเรื่องการแบ่งคำพูดต่อเนื่องให้เป็นพยางค์ในภาษาไทย (นัฏฐชา จิตติวรางกูล, 2541) ใช้ความรู้เกี่ยวกับเสียงซ้อนในระดับพยางค์ 3 ประการ คือ ความเข้ม ความสั้น/ยาว และเสียงสูงต่ำ ในระบบการแยกพยางค์แบบอัตโนมัติ โดยเน้นการแยกพยางค์ในชุดที่มีปัญหา คือ ในคำพูดที่ประกอบด้วยเสียงเรียงชุดเดียวกัน แต่มีจุดแบ่งพยางค์ที่ต่างกัน เช่น /ja:tam/ ใน <หญ้าตำเท้า> = /ja: tam thaw/ และ <ญาติอำภา> = /ja:t am pha:/ ระบบสามารถแบ่งพยางค์ในภาษาไทยออกจากกันแบบอัตโนมัติได้ผลเป็นที่น่าพอใจ งานวิจัยนี้เป็นงานวิจัยพื้นฐานที่สำคัญยิ่งต่อการสร้างระบบการรู้จำของภาษาไทยทั้งภาษาเพราะจะช่วยย่นระยะเวลาในการวิเคราะห์แยกพยางค์ในคำพูดต่อเนื่องเพื่อนำไปใช้ในการสร้างแบบเปรียบ ซึ่งต้องการสัญญาณเสียงพยางค์ และการระบุโดยการให้นาม (labelling) แก่หน่วยทุกหน่วยที่จะนำไปใช้ในการฝึกให้เครื่องรู้จำ ตัวอย่างนี้แสดงให้เห็นว่าการวิจัยเทคโนโลยีทางภาษาในระดับพื้นฐานนั้นก็เป็นการวิจัยสหวิทยาการที่ต้องอาศัยการวิจัยจากศาสตร์อื่นเข้ามาช่วยตอบปัญหา

ในการรู้จำหน่วยในภาษานั้นหากระบบจัดเก็บรูปแทนของหน่วยในแง่ของลักษณะบ่งความต่างระหว่างหน่วยจำนวนจำกัดในภาษา เช่น รูปตัวอักษร หรือ หน่วยเสียง จะเป็นระบบที่ประหยัดและก่อให้เกิดประสิทธิภาพในการรู้จำการวิเคราะห์ลักษณะบ่งความต่างของหน่วยต่างๆ ทางภาษาศาสตร์เกิดขึ้นพร้อมกับทฤษฎีไวยากรณ์ปริวรรตแปรรูป เพราะไวยากรณ์ดังกล่าวจะแสดงรูปของภาษาเป็นองค์ประกอบที่ใช้บ่งความต่างของหน่วยนั้นออกจากหน่วยอื่นๆ ในระบบเดียวกัน ในการระบุลักษณะบ่งความต่างอาศัยทฤษฎีปฏิลักษณ์ (theory of opposition) ในเชิงปรัชญาจำแนกสรรพสิ่งออกจากกัน โดยระบุการมีคุณลักษณะนั้น (+) หรือการไม่มีคุณลักษณะนั้น (-) ซึ่งเป็นคุณลักษณะแบบสองทาง (binary feature) เช่น ในระบบเสียงภาษาไทย มีสระเดี่ยวอยู่ 18 หน่วย เป็นสระที่มีลักษณะต่างกันในด้านตำแหน่งของลิ้น (quality difference) 9 หน่วย และแต่ละหน่วยมีความแตกต่างกันในแง่ของความสั้น-ยาว (quantity difference) เป็นคู่กัน สระทั้ง 18 หน่วย ในภาษาไทยจึงจำแนกออกจากกันได้ด้วยลักษณะบ่งความต่างดังต่อไปนี้

รูปที่ 4 ตารางแสดงลักษณะบ่งความต่างของสระทั้ง 18 หน่วยในภาษาไทย

สระ	i	ɯ	u	e	ɣ	o	æ	a	ɔ	i:	ɯ:	u:	e:	ɣ:	o:	æ:	a:	ɔ:
ลักษณะบ่งความต่าง																		
สูง	+	+	+	−	−	−	−	−	−	+	+	+	−	−	−	−	−	−
ต่ำ	−	−	−	−	−	−	+	+	+	−	−	−	−	−	−	+	+	+
หน้า	+	−	−	+	−	−	+	−	−	+	−	−	+	−	−	+	−	−
หลัง	−	−	+	−	−	+	−	−	+	−	−	+	−	−	+	−	−	+
(ปากกลม)	(−)	(−)	(+)	(−)	(−)	(+)	(−)	(−)	(+)	(−)	(−)	(+)	(−)	(−)	(+)	(−)	(−)	(+)
ยาว	−	−	−	−	−	−	−	−	−	+	+	+	+	+	+	+	+	+

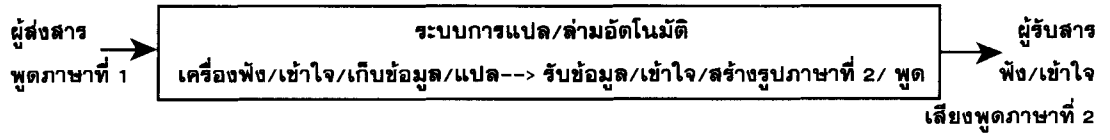
จะเห็นได้ว่าลักษณะบ่งความต่างที่ใช้ในการจำแนกสระทั้ง 18 หน่วยในภาษาไทยออกจากกัน มีเพียง 6 ลักษณะ และลักษณะ ± ปากกลม เป็นลักษณะที่เกินความจำเป็น (redundant feature) หมายความว่าเป็นลักษณะในรูปผิว แต่ไม่ใช่ลักษณะสำคัญที่ใช้ในการจำแนกสระออกจากกันเพราะเราสามารถทราบได้ว่าสระนั้นมีลักษณะปากกลมหรือไม่ จากลักษณะ ± หลัง คือสระที่มีลักษณะ + หลัง จะมีลักษณะ + ปากกลมด้วย จะเห็นได้ว่าการใช้ลักษณะบ่งความต่างในการจำแนกสรรพสิ่งออกจากกันในเชิงตรรกศาสตร์นั้นเมื่อนำมาใช้ในการจำแนกหน่วยทางภาษาออกจากกันเพื่อสร้างแบบเปรียบในระบบการรู้จำ จะทำให้สามารถวางระบบการทำงานได้อย่างประหยัดและมีประสิทธิผลมากกว่าการจำแนกหน่วยในภาษาออกจากกันโดยมองหน่วยเหล่านั้นเป็นหน่วยเดี่ยวๆ แต่ละหน่วย ผู้เขียนได้ให้ความรู้ทางตรรกศาสตร์เกี่ยวกับทฤษฎีปฏิลักษณ์เป็นรากฐานแก่นิสิตก่อนที่จะทำการวิจัยทางด้านระบบการรู้จำทั้งการรู้จำตัวอักษรและการรู้จำเสียง โดยเฉพาะในด้านการรู้จำคำพูด (Speech Recognition) ได้เกิดผลงานวิจัยที่นำความรู้ด้านนี้ไปใช้ในการวิเคราะห์รูปอักษรภาษาไทยในการวิจัยด้านการรู้จำตัวอักษรดังกล่าวมาแล้ว และในการวิเคราะห์เสียงภาษาไทยเพื่อสร้างแบบเปรียบ และวางระบบการทำงานรู้จำอย่างมีประสิทธิภาพ (วิศรุต อาชุบุตร, 2539; ชัย วุฒิวิวัฒน์ชัย, 2540; เอกฤทธิ์ มณีน้อย, 2541) การวิจัยด้านการรู้จำคำพูดในภาษาไทยในขณะนี้มีความเข้มข้นขึ้น และมีความเป็นสหวิทยาการมากขึ้นโดยการสร้างระบบที่มีนัยสำคัญทางภาษา (language modelling)เป็นส่วนสำคัญของการทำงาน

นอกจากการวิจัยด้านการรู้จำคำพูดแล้ว การรู้จำผู้พูด (speaker recognition) ก็เป็นเทคโนโลยีที่เริ่มมีการวิจัยกันมากขึ้นในประเทศไทย ในระบบการรู้จำคำพูด (speech recognition) เราจะต้องกำจัดลักษณะเฉพาะของผู้พูดแต่ละคนที่ต่างออกไป ให้มีลักษณะที่เป็นกลาง (normalization process) ตรงข้ามกับการวิจัยด้านการรู้จำผู้พูด ซึ่งระบบจะต้องสกัดลักษณะ

เฉพาะของผู้พูดแต่ละคนออกมา ความรู้เกี่ยวกับการแปรของเสียงในด้านภาษาศาสตร์สังคมเป็นเรื่องที่ต้องนำมาพิจารณา ผู้เขียนได้ให้คำแนะนำในเรื่องนี้แก่นิสิตที่ทำการรู้จำโดยกำหนดบทพูด (ศวิต กาสุริยะ, 2542) ได้ผลการวิจัยเป็นที่น่าพอใจยิ่ง โดยเฉพาะบทพูดที่ประกอบด้วยเสียง /ส/ ซึ่งงานวิจัยทางสัทศาสตร์ภาษาไทยทั้งในเชิงวรรณาและทางภาษาศาสตร์สังคมรายงานว่ามีรูปแปรที่หลากหลายในภาษาไทย แนวทฤษฎีเกี่ยวกับลักษณะสัทสัมพันธ์ในการศึกษาคุณลักษณะของเสียง (voice quality theory) ยังเป็นอีกเรื่องหนึ่งซึ่งสามารถนำมาใช้ในการศึกษาเรื่องการรู้จำผู้พูดได้

กลุ่มที่ 3 การอ่าน/ฟัง/เข้าใจ/แปล และ พิมพ์หรือพูดเป็นอีกภาษาหนึ่ง การวิจัยด้านนี้เป็นการวิจัยที่ยากที่สุดในกระบวนการวิจัยด้านเทคโนโลยีทางภาษา เพราะนอกจากจะต้องสร้างสมรรถนะทางภาษา 1 ภาษา ดังเช่นในกลุ่มที่ 1 และ กลุ่มที่ 2 ดังกล่าวมาแล้ว ยังต้องสร้างสมรรถนะทางภาษาของภาษาที่ 2 ด้วย ในศาสตร์ด้านการแปลด้วยเครื่อง (machine translation) นอกจากจะต้องสร้างสมรรถนะเกี่ยวกับระบบภาษา และการสร้างรูปภาษาที่ถูกต้องตามไวยากรณ์แล้ว ในการแปลยังมีเรื่องของขนบที่เกี่ยวข้องกับการใช้ภาษาในการสื่อสาร ซึ่งหมายถึงปริบทของการสื่อสาร โดยเฉพาะปริบทวัฒนธรรมที่เข้ามามีอิทธิพลต่อปริบทอื่นๆ ดังกล่าวมาแล้วด้วย การแปลตามตัวหนังสืออาจจำลองขึ้นได้ในระดับหนึ่ง ดังพบการใช้ในพจนานุกรมอิเล็กทรอนิกส์แบบ 2 ภาษา หรือ แบบหลายภาษา ระบบเหล่านี้มักจะเป็นผลพลอยได้ของการวิจัยด้านการแปลด้วยเครื่องทั้งสิ้น แต่การแปลในระดับที่เหนือคำทำไม่ได้ง่ายนักแม้ในการแปลโดยมนุษย์ ตัวอย่างเช่น ในคำเรียกญาติภาษาอังกฤษ <aunt> คำนี้ในภาษาอังกฤษ หมายถึงพี่หรือน้องของพ่อหรือของแม่ที่เป็นผู้หญิง ในภาษาไทยคำนี้แปลได้ 3 คำ คือ <ป้า น้า อา> ในการแปลจึงต้องทราบว่า <aunt> ที่พูดถึงนั้นเป็นพี่สาวหรือน้องสาวของพ่อหรือของแม่ ถ้าเป็นพี่สาวของพ่อหรือของแม่ เรียกว่า <ป้า> ถ้าเป็นน้องสาวของพ่อ เรียกว่า <อา> แต่ถ้าเป็นน้องสาวของแม่ เรียกว่า <น้า> ตัวอย่างนี้เป็นตัวอย่างในระดับคำง่ายๆ ที่อาจก่อให้เกิดปัญหาไม่ใช่น้อยในการแปลถ้าผู้แปลไม่ได้ศึกษาความสัมพันธ์ระหว่างตัวละครให้ดีก่อนการแปล ปัญหาในการตีความคำพูดยังอยู่ที่ปริบทเฉพาะสถานการณ์ที่คำพูดนั้นปรากฏอยู่ด้วย การจะตีความว่าคำพูดว่า <นั่น เหว นะ> เป็นคำบอกคำเตือน หรือคำขู่ อาศัยปริบทในการตีความทั้งสิ้น การแปลภาษาจึงเป็นเรื่องที่ซับซ้อนยิ่ง แม้การวิจัยพื้นฐานในด้านนี้จะมีการศึกษากันไม่ใช่น้อย

ความฝันอันสูงสุดของการวิจัยเพื่อสร้างระบบการแปลด้วยเครื่อง แสดงในรูปข้างล่างนี้



ในแง่ของความเข้าใจในภาษาธรรมชาติ โดยเฉพาะภาษาสนทนานั้นเป็นเรื่องที่ซับซ้อนมากดังกล่าวมาแล้วในเรื่องของระบบภาษาและระบบการสื่อสาร มนุษย์ใช้ความรู้ที่มีมาก่อนในการตีความและเกิดความเข้าใจ การวิจัยพื้นฐานด้านภาษาศาสตร์จิตวิทยาได้รับการนำไปจำลองเป็นโมเดลของระบบการรับรู้/รู้จำ/เข้าใจ/และตีความ การฝึกการแปลอัตโนมัติที่ใช้งานได้ดีในปัจจุบันยังห่างไกลจากความสามารถของนักแปลอาชีพมาก นอกจากการแปลที่พบในพจนานุกรมอิเล็กทรอนิกส์ชนิดพูดได้ ซึ่งเป็นการแปลระดับคำแล้ว ยังมีระบบการแปลภาษาแบบเฉพาะ เช่น การแปลเอกสารทางกฎหมาย การแปลคู่มือการใช้เครื่องมือ ซึ่งมีสำนวนและวงคำศัพท์จำนวนจำกัด ส่วนใหญ่ผู้สร้างระบบจะสร้างระบบเพื่อใช้เป็นเครื่องช่วยโดยให้คนแปลมาเติมข้อความหรือแก้ไขดัดแปลงได้ การวิจัยทั้งด้านความเข้าใจของเครื่อง (Natural Language Understanding) และระบบการสร้างรูปภาษา (Generation System) ที่จะช่วยสร้างรูปผิวที่ถูกต้องเหมาะสมนั้น ยังมีการศึกษาในระดับทฤษฎีอยู่มาก

บทสรุป

การวิจัยสหวิทยาการด้านเทคโนโลยีทางภาษาที่ได้กล่าวมาแล้วทั้งหมด ตามความเห็นของผู้เขียนเป็นการวิจัยเพื่อการปฏิบัติการ กล่าวคือ โจทย์ของการวิจัยมักจะเริ่มมาจากความต้องการทางเทคโนโลยีใหม่ๆ จากผู้ใช้ ก่อนที่ผู้เขียนจะเริ่มวิจัยด้านเทคโนโลยีทางภาษาได้มีความคิดว่าจะทำการวิจัยเรื่องลักษณะทางกลสัทศาสตร์ของเสียงในภาษาไทยให้สมบูรณ์ที่สุดเพื่อให้วิศวกรคอมพิวเตอร์นำไปใช้ในระบบการสังเคราะห์เสียงหรือระบบการรู้จำเสียง โดยได้นำเรื่องนี้ไปปรึกษากับ ศาสตราจารย์ ดร.ไพรัช ธัชยพงศ์ ผู้อำนวยการศูนย์วิจัยเทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ(เนคเทค)ในขณะนั้น ท่านให้คำแนะนำว่าหากต้องการทุนวิจัยจากเนคเทคให้วิจัยเรื่องระบบการอ่านออกเสียงภาษาไทยด้วยคอมพิวเตอร์ โจทย์นี้ค่อนข้างยากสำหรับผู้เขียนซึ่งไม่เคยวิจัยสหวิทยาการมาก่อน ผู้เขียนได้ไปศึกษาความเป็นไปได้ว่าในการวิจัยเรื่องดังกล่าวจะต้องมีนักวิจัยจากสาขาใดเข้ามาเกี่ยวข้องบ้าง ท้ายที่สุดก็ได้พบกับ ดร.วีระ ริ้วพิทักษ์ (ดูโน้ต[4]) ซึ่งเป็นผู้เดียวที่สนใจด้านเทคโนโลยีทางเสียงเพราะท่านวิจัยระดับปริญญาเอกในด้านนี้ ผู้เขียนได้นำเรื่องนี้ไปปรึกษากับ รศ.ดร.ปราณี กุลละวณิชย์ ซึ่งขณะนั้นเป็นคณบดีคณะอักษรศาตร์ เพื่อขอคำแนะนำเกี่ยวกับทุนการศึกษาระยะสั้นสำหรับการไปทำวิจัยหลังปริญญาเอก ด้าน Speech Computing and Speech Technology ซึ่งอาจารย์ท่านหนึ่งของผู้เขียน[8]มีความเชี่ยวชาญและท่านยินดีที่จะให้ไปทำงานวิจัยกับท่าน ท่านคณบดีได้แนะให้ขอทุนระยะสั้นของจุฬาฯในความต้องการของคณะซึ่งทำให้ผู้เขียนได้ไปศึกษาวิจัยภาษาศาสตร์ในเชิงสหวิทยาการเป็นครั้งแรก ทำให้เกิดความเข้าใจในธรรมชาติของการวิจัยแนวนี้ กล่าวคือ โจทย์ของการวิจัยต้องชัดเจนมาก ผู้วิจัย

ต้องวางแผนการทำงานโดยแบ่งงานเป็นส่วนๆ การกำหนดตารางเวลาการทำงานของส่วนงานแต่ละส่วนจะต้องมีรายละเอียดที่ชัดเจน หากมีการวิจัยเป็นทีมต้องมีการประชุมเพื่อตกลงกันในเรื่องของกำหนดการทำงานและการส่งต่อระหว่างส่วนงานแต่ละส่วน การวิจัยในโครงการแรกของผู้วิจัยใช้การจ้างผู้ช่วยวิจัยเต็มเวลาสำหรับส่วนงานแต่ละส่วน และมีผู้ช่วยวิจัยคนหนึ่งรับหน้าที่เป็นเลขานุการของโครงการด้วย โดยมีหน้าที่จัดการประชุมสัมมนา และติดตามงาน ดูแลเรื่องการส่งต่อระหว่างส่วนงาน ทำให้การวิจัยสามารถดำเนินไปได้เป็นอย่างดีตรงตามกำหนดเวลา อย่างไรก็ตาม การวิจัยในแนวนี้จะประสบปัญหาอย่างมากหากส่วนใดส่วนหนึ่งในระบบการทำงานไม่สามารถทำส่วนงานนั้นให้สำเร็จตามกำหนดเวลา ซึ่งจะทำให้การวิจัยทั้งโครงการไม่สำเร็จไปด้วยเพราะไม่สามารถนำส่วนงานมาประกอบกันได้ ผู้วิจัยและผู้ช่วยวิจัยจะต้องมีสปิริตของทีมสูงมาก และมีความมุ่งมั่นที่จะทำให้ส่วนการทำงานของตนเสร็จตามกำหนดเวลาเพื่อเชื่อมต่อและเพื่อประโยชน์ของส่วนการทำงานอื่น ในกรณีของการวิจัยสหวิทยาการที่เป็นส่วนหนึ่งของการศึกษาในมหาวิทยาลัย ผู้วิจัยอาจจะไม่ต้องเริ่มต้นจากศูนย์ ผู้วิจัยควรศึกษาว่ามีส่วนการทำงานใดที่พัฒนาไว้แล้วและส่วนการทำงานใดที่ตนจะต้องพัฒนาขึ้นใช้ใหม่ เป็นการวางขอบข่ายการวิจัยของตนให้ชัดเจนก่อนการลงมือศึกษา ดังที่กล่าวมาแล้วว่าการวิจัยสหวิทยาการด้านเทคโนโลยีทางภาษาเป็นการวิจัยเชิงปฏิบัติการ ซึ่งมีที่มาจากความต้องการของผู้ใช้เป็นสำคัญ การวิจัยในแนวนี้มิใช่จะเสร็จสิ้นเมื่อตอบสนองความต้องการของผู้ใช้แล้ว นวัตกรรมใหม่ๆ เกิดขึ้นได้อีกทั้งจากความต้องการใหม่ หรือจากการที่ผู้วิจัยมีแนวคิดหรือทฤษฎีใหม่ที่จะนำมาใช้แก้ปัญหาเก่าให้มีประสิทธิภาพยิ่งขึ้น จากประสบการณ์การวิจัยเชิงสหวิทยาการ ผู้เขียนมีความเห็นว่าการวิจัยในแนวนี้ผู้วิจัยควรมีพื้นฐานทางทฤษฎีที่ลึกซึ้ง แต่ไม่ยึดติดกับทฤษฎีที่มีอยู่แล้ว การได้ทำงานร่วมกับนักวิจัยจากศาสตร์อื่นจะทำให้เกิดความคิดและคำถามใหม่ๆ ในประเด็นที่อาจไม่เคยมีการศึกษาวิจัยมาก่อน ก่อให้เกิดองค์ความรู้ใหม่ ทั้งในแง่ อะไร ทำไม เป็นเรื่องที่ดีสำหรับนักวิจัยเชิงทฤษฎีที่ได้มาทำวิจัยเชิงปฏิบัติการหรือเชิงประยุกต์บ้าง เพราะการตอบคำถามที่สำคัญที่สุดของนักวิจัยสหวิทยาการเชิงปฏิบัติการคือ ***อย่างไร*** การตอบคำถาม ***อย่างไร*** มีปัจจัยเกี่ยวกับความเหมาะสม ความประหยัด และ การมีประสิทธิภาพในการทำงานสูงสุดเป็นองค์ประกอบในการตอบคำถาม ซึ่งเป็นเรื่องที่ไม่ง่ายนักหากจะตอบคำถามจากมุมมองของศาสตร์ของตนแต่ประการเดียว การช่วยกันตอบเป็นกิจกรรมที่ประเทืองปัญญาสำหรับคนที่เป็นนักทฤษฎีมาก่อนไม่ใช่น้อย หากพยายามที่จะมองปัญหาต่างออกไปจากมุมที่เคยอยู่

Note

[1] การศึกษาเรื่องพยางค์ที่มีศักยภาพจะเกิดขึ้นได้ในภาษาไทย เป็นตัวอย่างที่แสดงให้เห็นความสามารถในการสร้างสรรค์ภาษา

[2] คำว่าเข้าใจในระบบของเครื่องในปัจจุบัน (ปี ค.ศ. 2000) หมายถึงการที่เครื่องสามารถรู้ว่าคำๆ นั้น คือคำอะไร ในระบบที่ฉลาดมากๆ เช่น ระบบที่ใช้ในการแปลอาจรู้ว่า คำๆ นั้นในปริบทดังกล่าวมีความหมายว่าอะไร อย่างไรก็ตามในแง่ของความเข้าใจภาษา เทคโนโลยียังไม่ก้าวหน้าไปถึงระดับการเข้าใจความของคำพูดต่อเนื่องอย่างลึกซึ้งเช่นมนุษย์

[3] คอมพิวเตอร์ที่มีใช้อยู่ทั่วไปในประเทศไทยขณะนั้นเป็นเครื่องพีซี 386

[4] ดร.วีระ ริ้วพิทักษ์ ซึ่งเป็นอาจารย์อยู่ที่ภาควิชาวิศวกรรมคอมพิวเตอร์ จุฬาลงกรณ์มหาวิทยาลัย เป็นผู้ที่ชักชวนให้ผู้เขียนเข้ามาศึกษาวิจัยด้านเทคโนโลยีทางเสียง อาจารย์ลาออกจากราชการไปทำธุรกิจส่วนตัว เมื่อผู้เขียนกลับมาจากการศึกษาวิจัยหลังปริญญาเอก

[5] ผู้ที่มีส่วนช่วยอย่างมากกับฐานข้อมูลนี้ โดยไม่มีค่าตอบแทนเพราะยังไม่ได้เสนอขอทุนวิจัย คือ คุณเชาวณัฐ ช่งเจริญสุข เทคนิเชียนของหน่วยปฏิบัติการวิจัยทางภาษาศาสตร์ และอาจารย์วิสิทธิ์ ลีลาศิริวงศ์ จากภาควิชาฟิสิกส์ จุฬาลงกรณ์มหาวิทยาลัย ซึ่งกำลังทำวิจัยด้านการรู้จำเสียงสระอยู่กับผู้เขียนในขณะนั้น

[6] ผู้เขียนได้พบกับ ดร.สตีฟ อิสาร์ด (Isard) ในขณะที่ไปวิจัยหลังปริญญาเอกที่ศูนย์วิจัยเทคโนโลยีทางเสียง ที่มหาวิทยาลัยเอดินเบอระในปี ค.ศ. 1989 ดร.อิสาร์ด พยายามโน้มน้าวให้ผู้เขียนใช้ทฤษฎีของหน่วยเสียงคู่ (diphone) ที่ท่านกำลังทดลองใช้ในการสังเคราะห์เสียงภาษาอังกฤษได้อย่างมีประสิทธิภาพ แต่ผู้เขียนพบว่าการใช้ทฤษฎีดังกล่าวจะเป็นอุปสรรคต่อการสังเคราะห์เสียงวรรณยุกต์ซึ่งเป็นเสียงสำคัญในภาษาไทยจึงเลือกใช้พยางค์เป็นหน่วยหลักในการสังเคราะห์

[7] ผู้เขียนได้รับมอบระบบการรู้จำคำพูดจากบริษัท ไอ บี เอ็ม ประเทศไทย ซึ่งมอบให้หน่วยวิจัยของผู้เขียนเพื่อทดลองใช้ และได้รับระบบจากบริษัท แดรกอน ซิสเต็ม โดย คุณเกรแฮม แวนเดอร์สตุล ซึ่งเป็นผู้พูดอีกคนหนึ่งนอกจากผู้เขียนที่ American Chamber of Commerce เชิญไปบรรยายพิเศษในหัวข้อเทคโนโลยีทางเสียง เมื่อ 23 มีนาคม 2543

[8] Professor John Laver ซึ่งเป็นผู้อำนวยการศูนย์วิจัยเทคโนโลยีทางเสียง ที่มหาวิทยาลัยเอดินเบอระในขณะนั้น

หนังสืออ้างอิง

ภาษาไทย

ชัชวาลย์ หาญสกุลบรรเทิง. (2542). ***การสังเคราะห์เสียงพยางค์ภาษาไทยด้วยวิธีการสังเคราะห์แบบวิเคราะห์โดยใช้คู่เส้นสเปกตรัม.*** วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาวิศวกรรมไฟฟ้า จุฬาลงกรณ์มหาวิทยาลัย.

ชัย วุฒิวิวัฒน์ชัย. (2540). ***การรู้จำเสียงคำไทยหลายพยางค์แบบไม่ขึ้นกับผู้พูดโดยใช้เทคนิคแบบฟัซซีและนิวรอลเน็ตเวิร์ค.*** วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาวิศวกรรมไฟฟ้า จุฬาลงกรณ์มหาวิทยาลัย.

นัฏฐชา จิตติวรางกูล. (2541). ***ระบบการแบ่งพยางค์ในคำพูดต่อเนื่องภาษาไทย.*** วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาวิศวกรรมไฟฟ้า จุฬาลงกรณ์มหาวิทยาลัย.

ประเสริฐ ฉอรุ่งเรืองวิวัฒน์. (2541). ***การรู้จำตัวอักษรเขียนภาษาไทยโดยใช้การวิเคราะห์ลักษณะบ่งความต่าง.*** วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาวิศวกรรมไฟฟ้า จุฬาลงกรณ์มหาวิทยาลัย.

วิชา พานิช. (2539). ***ระบบการรู้จำอักษรภาษาไทยโดยใช้ลักษณะบ่งความต่างของตัวอักษรไทย.*** วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาวิศวกรรมไฟฟ้า จุฬาลงกรณ์มหาวิทยาลัย.

วิศรุต อาขุบุตร. (2539). ***ระบบการรู้จำคำไทยหลายพยางค์แบบไม่ขึ้นกับผู้พูดโดยใช้แบบจำลองฮิดเดนมาร์คอฟ.*** วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาวิศวกรรมไฟฟ้า จุฬาลงกรณ์-มหาวิทยาลัย.

ศวิต กาสุริยะ. (2542). ***ระบบการรู้จำผู้พูด: การบ่งชี้ผู้พูดแบบขึ้นกับบทคำพูด.*** วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาวิศวกรรมไฟฟ้า จุฬาลงกรณ์มหาวิทยาลัย.

สุดาพร ลักษณียนาวิน. (2526). มนุษยศาสตร์ตายแล้วหรือ. ***ศิลปวัฒนธรรม.***

สุดาพร ลักษณียนาวิน. (2537). การใช้ภาษาเพื่อการสื่อสาร. ***ประมวลสาระชุดวิชาทักษะทางภาษา. หน่วยที่ 2.*** บัณฑิตศึกษา สาขาศึกษาศาสตร์ มหาวิทยาลัยสุโขทัยธรรมาธิราช. หน้า 35-96.

สุดาพร ลักษณียนาวิน. (2537). ภาษากับความหมาย. ***ประมวลสาระชุดวิชาทักษะทางภาษา. หน่วยที่ 3*** บัณฑิตศึกษา สาขาศึกษาศาสตร์ มหาวิทยาลัยสุโขทัยธรรมาธิราช. หน้า 96-152.

สุดาพร ลักษณียนาวิน. (2533). ระบบการสังเคราะห์เสียงจากข้อความภาษาไทย. ***เอกสารสืบเนื่องในการประชุมวิชาการทางการวิจัยและพัฒนาทางเทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ ครั้งที่ 2*** เล่มที่ 1. ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และ

คอมพิวเตอร์แห่งชาติ. กระทรวงวิทยาศาสตร์เทคโนโลยีและสิ่งแวดล้อม. กรุงเทพฯ. 23 หน้า.

สุดาพร ลักษณียนาวิน. (2534). ปัญหาของระบบการสังเคราะห์เสียงจากข้อความภาษาไทย. ***เอกสารสืบเนื่องในการประชุมวิชาการทางการวิจัยและพัฒนาทางเทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ ครั้งที่ 3*** เล่มที่ 1. ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ. กระทรวงวิทยาศาสตร์เทคโนโลยีและสิ่งแวดล้อม. 20-21 ตุลาคม 2534 กรุงเทพฯ. หน้า 41-62.

สุดาพร ลักษณียนาวิน. (2535). คอมพิวเตอร์อ่านออกเสียงภาษาไทย. ***เอกสารสืบเนื่องในการประชุมวิชาการทางการวิจัยและพัฒนาทางเทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ ครั้งที่ 4*** ศูนย์เทคโนโลยีอิเลกทรอนิกส์แห่งชาติ. กระทรวงวิทยาศาสตร์เทคโนโลยีและสิ่งแวดล้อม. 21-22 ตุลาคม 2535 กรุงเทพฯ. หน้า 65-78.

อมรา ประสิทธิ์รัฐสินธุ์ (2533). ***ภาษาศาสตร์สังคม.*** กรุงเทพฯ: โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.

เอกพล อนุสุเรนทร์. (2541). ***การสังเคราะห์เสียงวรรณยุกต์และสระในพยางค์เปิดภาษาไทยโดยใช้หน่วยเสียงอนุภาค.*** วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาวิศวกรรมไฟฟ้า จุฬาลงกรณ์มหาวิทยาลัย.

เอกฤทธิ์ มณีน้อย (2541). ***การรู้จำหน่วยเสียงสระภาษาไทยโดยใช้โครงข่ายประสาทเทียม.*** วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาวิศวกรรมไฟฟ้า จุฬาลงกรณ์มหาวิทยาลัย.

ภาษาอังกฤษ

Bloomfield, L. (1933). ***Language.*** New York: Holt.

Buhler, Karl. (1934). ***Sprachtheorie.*** Jena: Fisher.

Chaplin, H. Ito and S.E. Martin. (1991). ***Japanese: A Manual of Reading and Writing.*** Tokyo: Charles E. Tuttle Co., Inc.

Chomsky, N. (1957). ***Syntactic Structures.*** The Hague: Mouton.

Chomsky, N. (1959). Review of Skinner's verbal behavior. ***Language.*** 35:26-58.

Chomsky, N. (1965). ***Aspects of the Theory of Syntax.*** Cambridge, Mass.: MIT Press.

Dane, F. C. (1990). ***Research Methods.*** California: Brooks/Cole Publishing Company.

Firth, J. R. (1957). ***Papers in Linguistics 1934-1951.*** London: Oxford University Press.

Flores d'Arcais, G.B. (1989). Language Perception. In Federick J. Newmeyer (ed). ***Linguistics: The Cambridge Survey III, Language : Psychological and Biological Aspects.*** Cambridge University Press. 97-123.

Hymes, D. (1964). ***Language in Culture and Society.*** New York: Harper and Row.

Hymes, D. (1972). Models of the interaction of language and social life. In John Gumperz and Dell Hymes (eds), ***Directions in Sociolinguistics: The Ethnography of Communication.*** New York: Holt Rinehart and Winston.

Hymes, D. (1974). ***Foundations in Sociolinguistics: An Ethnographic Approach.*** Philadelphia: University of Pennsylvania Press.

Levinson, Stephen C. (1983). ***Pragmatics.*** Cambridge: Cambridge University Press.

Luksaneeyanawin, S. (1989). A Thai text-to-speech system. ***Proceedings of the Regional Workshop an Computer Processing of Asian Languages (CPAL).*** Asian Institute of Technology. 305-315.

Luksaneeyanawin, S. (1992). Three-dimensional phonology: A historical implication. ***Proceedings of the Third International Symposium on Language and Linguistics-Pan-Asiatic Linguistics 1992.*** Chulalongkorn University, Bangkok, Thailand. 75-90.

Luksaneeyanawin, S. (1993a). Speech computing and speech technology in Thailand. ***Proceedings of the First Symposium on Natural Language Processing in Thailand - SNLP' 93.*** 17-20 March. Chulalongkorn University. 276-321.

Luksaneeyanawin, S. (1993b). Linguistics research and Thai speech technology. Paper read at ***The 5th International Conference on Thai Studies. School of Oriental and African Studies,*** University of London.

Luksaneeyanawin, S. (1995a). Tone transformation. ***Proceedings of the Second Symposium on Natural Language Processing - SNLP' 95.*** Kasetsart University. 345-353.

Luksaneeyanawin, S. (1995b). Speech, speech processing and, speech technology. Invited paper presented at ***The Second Symposium on Natural Language Processing - SNLP' 95.*** Kasetsart University.

Luksaneeyanawin, S. and Conkie, Alistair. (1993). Synthesis and Tone Transformation : Resynthesizing Thai tones using PSOLA technique. Paper read at ***The 9th Australian International Conference on Language and Speech.*** Sydney. 18-19 November.

Luksaneeyanawin, S. and Niyompol, W. (1997). The Application of the Thai-Text-to-Speech System for the Education of the Blind. Paper read at ***The Third UNESCO-ACEID International Conference : Educational Innovation for Sustainable Development.*** Jointly organized by Asia-Pacific Centre of Educational Innovation for Development (ACEID) and Asia-Pacific Programme of Education for All (APPEAL) in co-operation with the Office of the National Education Commission of Thailand. 1-4 December, Bangkok, Thailand.

Ong, W.J. (1988). ***Orality and Literacy: The Technologizing of the Word.*** London: Routledge.

Pike, K.L. (1943). ***Phonetics: A Critical Analysis of Phonetic Theory and Technique for the Practical Description of Sounds.*** Ann Arbor: University of Michigan Press.

Pike, K.L. (1947). ***Phonetics: A technique for Reducing Languages to Writing.*** Ann Arbor: University of Michigan Press.

Taft, M. (1991). ***Reading and the Mental Lexicon.*** Hove: Lawrence Erlbaum.