

การแยกพยางค์ของภาษาไทยเป็นกระบวนการหนึ่งเพื่อจัดเตรียมหน่วยเสียงที่เหมาะสมสำหรับการสังเคราะห์เสียงพูดจากข้อความ ในงานวิจัยนี้ผู้วิจัยได้เสนอการแยกพยางค์จากภาษาไทย โดยใช้กฎเพื่อเปรียบเทียบหาโครงสร้างพยางค์ที่เหมาะสม (Suited-Syllable-Structure-Mapping: 3S-Mapping) วิธีการนี้มีลักษณะที่สำคัญคือจะใช้โครงสร้างของพยางค์มาสร้างเป็นกฎ โดยกฎเหล่านี้จะเริ่มต้นจากการกำหนดโครงสร้างให้อยู่ในรูปแบบอย่างง่ายที่ประกอบด้วย C V และ T ซึ่งแทนพยัญชนะ สระ และวรรณยุกต์ตามลำดับ เรียกว่า สัญลักษณ์อักขระ (Alphabet-Symbols) โครงสร้างเหล่านี้จะต้องทำงานร่วมกับการตรวจสอบกลุ่มอักขระไทย (Enhanced-Thai-Character-Cluster: ETCC) การพัฒนาด้วยวิธีการดังกล่าวจะเริ่มจากการแทนอักขระแต่ละตัวที่นำเข้ามาด้วยสัญลักษณ์อักขระ สิ่งที่ได้จากกระบวนการนี้จะถูกนำไปเปรียบเทียบกับกฎที่สร้างไว้และตรวจสอบโครงสร้างที่เหมาะสมด้วยการตรวจสอบกลุ่มอักขระไทย นอกจากนี้งานวิจัยนี้ยังได้ใช้การแยกพยางค์ด้วยพจนานุกรมด้วยวิธีการเปรียบเทียบคำตามความยาวที่เหมาะสม (Suited-Length-Word Mapping: SL-word Mapping) เพื่อแยกพยางค์ที่อยู่นอกเหนือจากกฎที่ออกแบบไว้ จากวิธีดังกล่าวในการทดลองแสดงผลให้เห็นว่าการแยกพยางค์ด้วยวิธีนี้สามารถแยกพยางค์ได้ความถูกต้องถึง 99.85 % จากตัวอย่างทั้งหมด 4,000 ข้อความ