

Telecommunication

Application

Spoken language communication

ภาษาคืออะไร

ภาษา คือ สิ่งที่มนุษย์คิดค้นขึ้นเพื่อใช้ในการสื่อสารเป็นผลงานชั้นยอดที่เกิดจาก
มันสมองอันเลิศล้ำของมนุษย์ มนุษย์สามารถเข้าใจกันและสื่อสารกันได้
อย่างคล่องแคล่วโดยมีภาษาเป็นตัวกลาง และเมื่อรวมกับเสียงของ
มนุษย์ มนุษย์ยังสามารถสื่อสารออกมาได้ทั้งเนื้อหาที่ต้องการจะสื่อและ
น้ำเสียงที่ต้องการแสดงอารมณ์ได้อย่างละเอียดอ่อน

อุปสรรคของภาษา

เนื่องจากมนุษย์ทั่วโลกนั้นไม่ได้มีแค่ชุมชนเดียวไม่ได้มีภาษาเดียว
ดังนั้นจึงเกิดการแบ่งแยกของภาษาซึ่งทำให้เกิดช่องว่างระหว่างการสื่อสารมาก
ขึ้น ทั่วโลกตอนนี้ถึงเราจะบอกว่าภาษาอังกฤษเป็นภาษากลางแต่ก็มีคนจำนวน
ไม่น้อยที่พูดภาษาอังกฤษไม่เป็นด้วยปัญหาเหล่านี้ จึงเกิดคำถามที่ว่าเรา
สามารถทำให้คนทั้งโลกคุยกันรู้เรื่องได้อย่างไรจึงเกิดเทคโนโลยี
language Communication ขึ้น นักวิจัยต่างๆพยายามพัฒนา
เทคโนโลยีนี้ขึ้นโดยใช้ความรู้ด้านtelecommunicationต่างๆเข้ามา
ช่วยในการวิจัย

Spoken Language Communication Research



สาขาด้านการวิจัย

**นำเสนอเฉพาะสาขานี้

Department of Acoustics and Speech Research

Department of Spoken Language Research

Department of Natural Language Processing

Department of Integrated System Research

Department of Acoustics and Speech Research

- วิจัยด้านบทสนทนาและสัญญาณรบกวนรวมไปถึงสัญญาณรบกวนด้านการออกเสียง
- วิจัยบทสนทนาต่างภาษาและลักษณะสำเนียงของกลุ่มสนทนารวมไปถึงแสดง
- วิจัยกระบวนการ multi-dimensional signal processing ของ microphone array beamforming และ audio-visual information integration



-ภาพMicrophone array ที่ติดตั้งในหุ่นยนต์
เพื่อรับเสียงจุกที่ติดคือลูกศรสีแดง

Department of Spoken Language Research

- วิจัยกระบวนการทางภาษาในเชิงสถิติเพื่อการออกแบบการวิเคราะห์ภาษาพูด
- วิจัยกระบวนการการเปรียบเทียบบทสนทนากับ รูปแบบของภาษานั้นๆ
- Design and Construction of Spoken Language

Department of Natural Language Processing

&

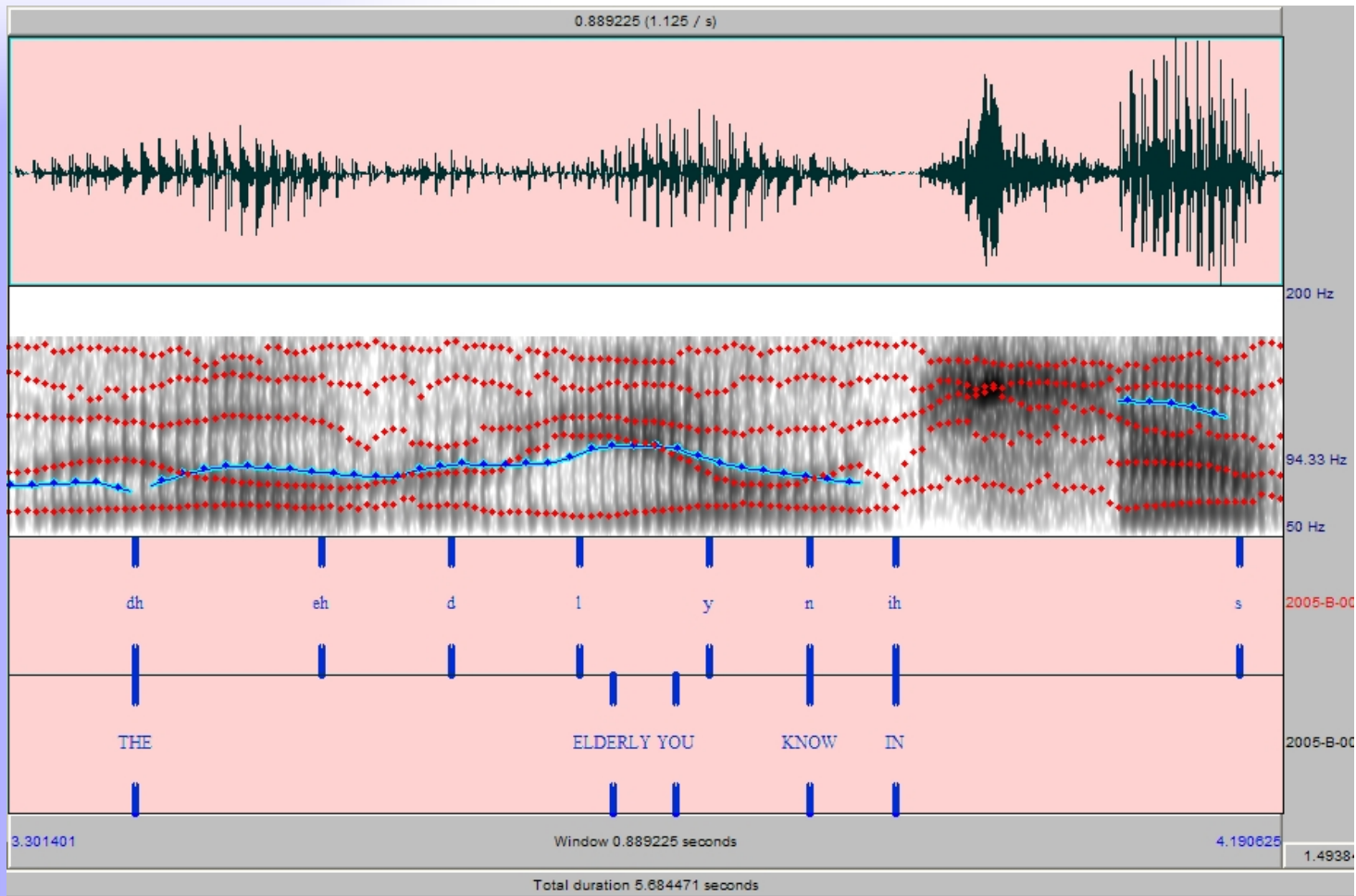
Department of Integrated System Research

-วิจัยฐานข้อมูลของภาษาต่างๆและสร้างเทคโนโลยีทางการแปล

-สร้างฐานข้อมูลด้านภาษาขนาดใหญ่ ระหว่างสองภาษา

-วิจัยและพัฒนาระบบรวมและปรับปรุงการใช้งาน

-สร้าง **speech corpora** เพื่อสังเกตการทำงานและประเมินผลวงจรรวม





Department of Acoustics and Speech Research

ASR
Department of Acoustics and Speech Research

音声音響處理研究室

วัตถุประสงค์การวิจัย

ในสาขานี้จะรับผิดชอบในด้านการออกแบบระบบทางเสียงที่สามารถได้ข้อมูลทางเสียงที่มีคุณภาพที่ดี รวมไปถึงการศึกษาสำเนียงการพูดต่างๆ กระบวนการวิเคราะห์ทางเสียงและการนำข้อมูลดังกล่าวไปใช้อย่างมีประสิทธิภาพโดยขจัดสัญญาณรบกวนต่างๆในสิ่งแวดล้อม และเป้าหมายสุดท้ายของสาขาการวิจัยนี้คือ การสื่อสารไร้สายระหว่างคู่สนทนาและระบบการแปลภาษาซึ่งระบบสามารถระบุแหล่งของเสียงที่มาหรือคู่สนทนาได้ โดยการนำเทคโนโลยี **Beamforming** เพื่อระบุบุคคลผู้ที่จะมาสนทนากับเราในสถานการณ์ต่างๆที่มีแหล่งสัญญาณรบกวน และบทสนทนาหลายภาษา

หัวข้อการวิจัย

— Hands-free recognition of distant-talking speech

— Recognition of noisy and reverberated speech

— Recognition robust to speaking style and variation

— Multi-modal speech recognition based on audio and image information

— Recognition of multiple languages

— Spoken dialogue system

Hands-free recognition of distant-talking speech

เป็นการกระระยะทางของกลุ่มสนทนาว่าคนนั้นพูดกับเราหรือป่าว โดยที่
จะต้องเจอปัญหาด้าน สัญญาณรบกวนต่างๆ รวมไปถึงการสะท้อนของเสียง
ซึ่งถือว่าเป็นสัญญาณรบกวนและหลังจากหาและมั่นใจว่าเป็นกลุ่มสนทนาแล้ว
ต้องมีการใช้ **Beamforming microphone array** ช่วยใน
การตัดเสียงรบกวนโดยจับแต่เสียงของกลุ่มสนทนา

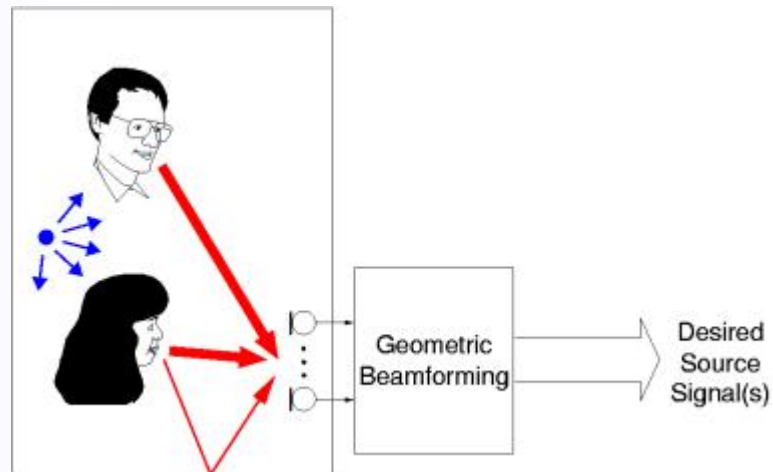
ในหัวข้อนี้หากให้วิเคราะห์ก็ต้องใช้ ในส่วนของฟิลเตอร์ในการ
กรองนอยส์ออก และมีตรรกะหาจุดที่กลุ่มสนทนาขึ้นอยู่กับคำพูด
และอาจใช้คลื่นสัญญาณในการตรวจระยะห่างจากกลุ่มสนทนาและทำการส่งใน
Beam microphone โฟกัสที่จุดดังกล่าว ในการส่งสัญญาณต่างๆ
และรับสัญญาณต่างๆจำเป็นที่จะต้องการ

การทำงานของ Microphone Array

ตัวอย่าง Microphone Array →



การใช้งาน →



Recognition of noisy and reverberated speech

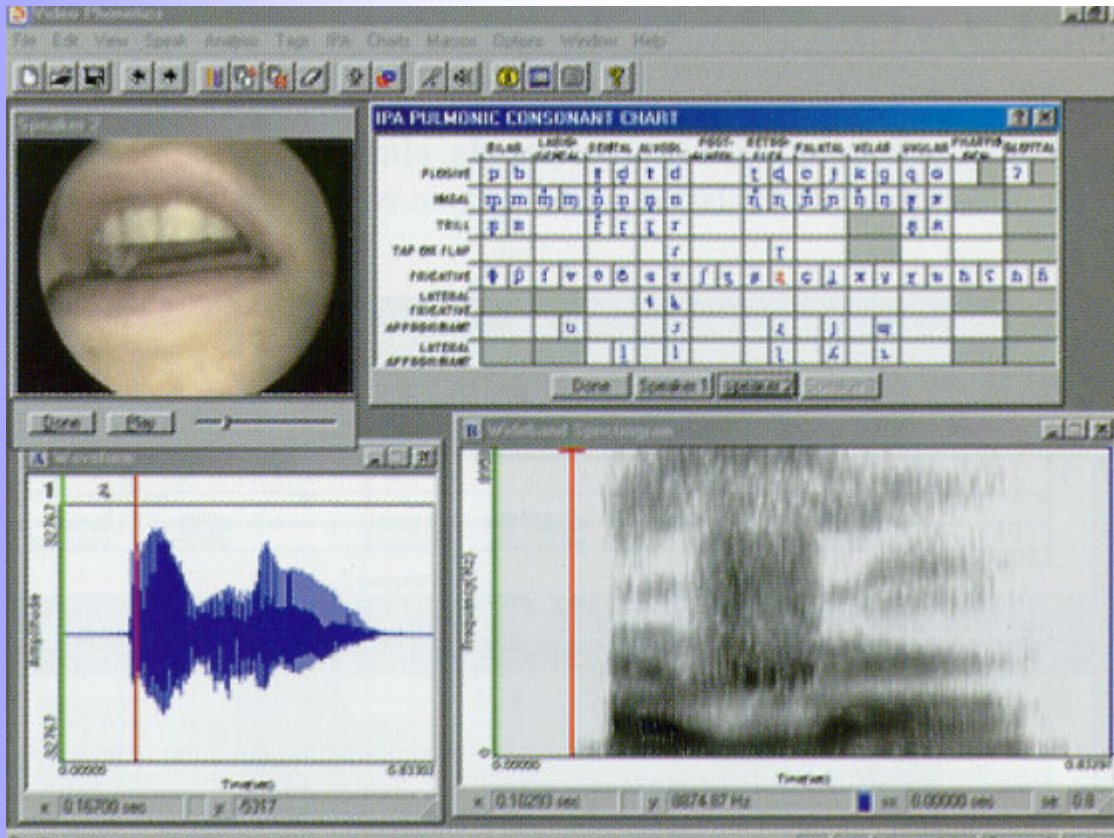
& Recognition robust to speaking style and variation

การจัดการกับสัญญาณรบกวนในบทสนทนา ซึ่งสัญญาณรบกวนนี้ก็มาจากสิ่งแวดล้อมต่างๆ รวมไปถึงเสียงที่ก้องหรือเสียงสะท้อนต่างๆของบทสนทนาในส่วนนี้คิดว่าน่าจะมีการนำฟิลเตอร์นำมากรองนอยส์ออกโดยการกรองนี้อาจต้องเป็นการกรองที่อาศัย **AI** เข้ามาช่วยในการจำเสียงนอยส์ต่างๆ และหลังจากการได้ข้อมูลจากการกรองในขั้นนี้ยังต้องดูไปถึงการแปรค่า พารามิเตอร์ต่างๆของเสียงเพื่อนำไปวิเคราะห์ในขั้นต่อไป

Multi-modal speech recognition based on audio and image information

มนุษย์นั้นไม่ได้ใช้แค่เสียงในการสนทนาซึ่งรวมไปถึงท่าทางต่างๆและการขยับริมฝีปากด้วย การนำ **multi-model speech** มาใช้นั้นจะช่วยเพิ่มสมรรถภาพในการรับข้อมูลให้มากยิ่งขึ้น โดยใช้เทคโนโลยีที่สามารถรับรู้ข้อมูลทางเสียงกับทางภาพได้ และนำไปวิเคราะห์รูปแบบของภาพต่างๆเหล่านั้นพร้อมกับเสียง ในส่วนนี้อาจต้องใช้ **image processing** เข้ามาช่วยด้วย

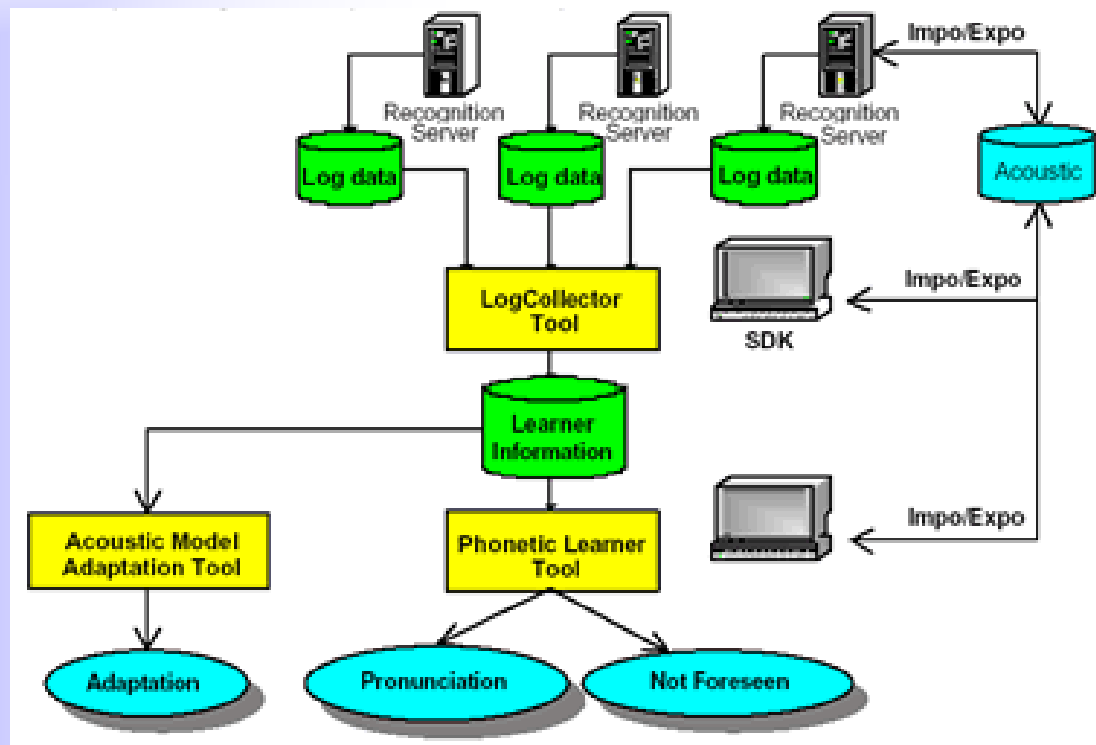
Multi-model speech & Image processing



รูปของโปรแกรมที่ใช้ในการวิเคราะห์ speech model

Recognition of multiple languages

เมื่อเราจัดการกับสิ่งต่างๆด้านข้อมูลในการรับรู้บทสนทนาได้แล้ว ปัญหาต่อมาคือเรื่องภาษาเราอาจต้องใช้หน่วยความจำที่ใหญ่มากที่จะเก็บฐานข้อมูลด้านภาษาของทุกภาษา ดังนั้นทางการวิจัยจึงคิดค้น **universal acoustic model** ซึ่งมาจากความคิดที่ว่ารูปแบบของน้ำเสียงมนุษย์ในการสนทนาจะประกอบเป็นเสียงของ คำพูดและประโยค มีความตายตัวในช่วงความถี่หนึ่งๆ เพราะฉะนั้นการที่เราสามารถรู้รูปแบบของสัญญาณของ คำพูดและประโยคเราสามารถสร้าง โมเดลของน้ำเสียงที่เป็นเสมือนรหัสเพื่อกรองบทสนทนาได้โดยจับว่าเขาพูดว่าอะไร และหลังจากนั้นค่อยเอาไปเทียบกับข้อมูลภาษาเบื้องต้นที่มีด้วยวิธีนี้เราไม่ต้องเสียฐานข้อมูลในการจำรูปแบบเสียงต่างๆของทุกภาษา



acoustic model adaptation

Spoken dialogue system

เป็นส่วนที่ช่วยให้ ส่วนของ **Recognition of multiple language** สมบูรณ์โดยเป็นส่วนที่จะจับ บทสนทนาออกมาได้จากแหล่งข้อมูล ณ เวลานั้นๆ และอาจต้องใช้ส่วยของ **AI**. ในการพัฒนาระบบการจับบทสนทนา

