Construction of Linguistic Resources for Mental Disorders -- Interdisciplinary Research in Linguistics, Cognitive Neuroscience and Artificial Intelligence

Hongwei Ding

The 2nd International Symposium on Language Resources and intelligence

上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

Construction of Linguistic Resources for Mental Disorders - Interdisciplinary Research in Linguistics, Cognitive Neuroscience and Artificial Intelligence

1. Motivation
2. Overview
3. Teams
4. Tasks
5. Database
6. Future Work

1. Motivation (SLP)

Newborn Hearing Screening

Cochlear Implantation

Speech Rehabilitation

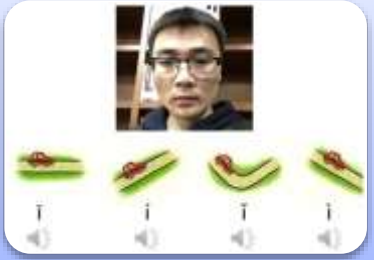
Mainstream Primary School

Deaf School

Important for typical language development

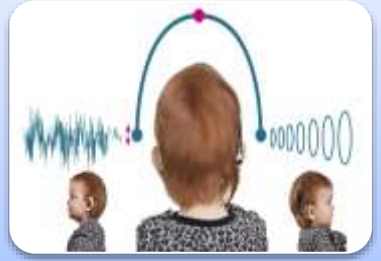
Optimal rehabilitation strategy !

1. Motivation (SLP)



Efficacy of Multi-Talker Phonetic Training

- Improvement of tonal perceptual performance through training
- Neural signatures of the training effect



Benefit of Bimodal Fitting (Cochlear Implant + Hearing Aid)

- Bimodal benefit in lexical tone perception
- Bimodal benefit in adverse listening conditions

Cognitive Predictors for Better Speech Recognition

- Top-down processing in natural and degraded speech recognition
- Cognitive factors that correlate with speech recognition outcome

1. **Motivation** (Neuro-Cognitive)

- **Inspired by**
  - LREC 2018：Workshop on Linguistic and Neuro-Cognitive Resources
- **Literature research on**
  - TalkBank (Aphasia Bank …), The Cambridge Cookie-Theft Corpus, OPTIMA, GREECAD, CoDAS, National Library of Medicine--VA EMR, Orozco-Arroyave Database, DAIC-WOZ, Health Bank, WRAP, …
  - Most on language disorders

1. **Motivation** (Example disorder)

	Corpus of Audio & Video	
Corpus Name	Cantonese Aphasia Bank (CAB)	
Participants	149/104 (patients/ controls)	
Participant characteristics	Cantonese speakers with aphasia & normal Cantonese speakers (match in age & education)	
Tasks	to describe pictures, tell stories, describe procedures of making a ham and egg sandwich, narrate an important event in their life and recount their stroke experience (for speakers with aphasia only)	
Samples	8 speeches each normal participant; 9 speeches each patients	
Media		
Transcription format	the Codes for the Human Analysis of Transcripts (CHAT)	
Annotation	Text	the Cantonese MOR tagger: part of speech (POS)
	Speech	Child Language Analyses (CLAN; MacWhinney, 2003): romanization & intonation
	Gesture	the EUDICO Linguistic Annotator (ELAN): "forms"+ "functions"
Analyzing tool		
Data Collection	demographic data (subject type, aphasia type, gender, age, education level), Action Research Arm Test (ARAT) score	

1. **Motivation** (Example Eyetracker)

Corpus of Eyetracker Data	
Ghent Eye-Tracking Corpus (GECO)	Russian Sentence Corpus (RSC)
33	96
English monolinguals and Dutch–English bilinguals	native Russian speakers
the novel The Mysterious Affair at Styles by Agatha Christie	sentences from the Russian National Corpus with an acceptability test on a scale of 1-5
5031 English sentences	144 Russian sentences, 701 single words
Eyelink 1000 Plus desktop mount eye-tracker	
the Data Viewer package (SR Research Ltd) that divides fixations and saccades	the Data Viewer package (SR Research Ltd) that divides fixations and saccades
first fixation duration (FFD), single fixation duration (SFD), gaze duration (GD), total reading time (TRT), go-past time (GPT)	first fixation duration (FFD), single fixation duration (SFD), gaze duration (GD), total reading time (TT)

1. **Motivation** (Problem-driven)

- **Why mental disorders**
  - Accelerated life rhythm
  - Intensified competition pressure
    - prominent individual psychological behavior problems and social problems
    - has aroused widespread concern of all sectors of society
  - Professional service personnel and systems are lagging behind, cannot meet the growing demand for mental illness
  - Artificial intelligence technology has shown its unique advantages in the application of assisted disease diagnosis in the early screening.

1. **Motivation** (Inquiry-based)

1. **Motivation** (Data-based)

- **Why database**
  - Accurate artificial intelligence diagnosis based on machine learning → large data for training
    - Corpus building
    - AI early screening → Artificial Intelligence research
    - Mental disease→ Cognitive Neuroscience research
    - Focused on speech and language
  - Construction of Linguistic Resources for Mental Disorders - Interdisciplinary Research in Linguistics, Cognitive Neuroscience and Artificial Intelligence

2. Overview

- **Title**
  - **精神障碍人群语料库建设及面向脑科学和人工智能的语言研究**
  - Construction of Linguistic Resources for Mental Disorders - Interdisciplinary Research in Linguistics, Cognitive Neuroscience and Artificial Intelligence

2. **Overview** (Emotional Prosody)

- **What database**
  - Language
    - Speech → Prosody, emotion
    - Transcribed text → Semantics, syntactic
  - Video
    - Facial expression
    - Gestures
  - Behavior neuro-data

3. Teams

Teams in Institute

Linguistic Database

Clinical medicine

Computer Science

Neuro science

Psychology

Linguistics, Clinical Medicine, Computer Science, Psychology

3. Teams

Nr	Subproject	Institute	Name	Function
1	Database building	SFL SJTU	Kaibao HU	Prof. Dean
2	Multimode Generation	SEIEE SJTU	Kai YU	Prof. AIspeech Principal Scientist
3	Speech Language Pathology	Minnesota Univ.	Yang ZHANG	Prof. Lab Head Visiting Prof. SJTU
4	Psychological Cognition	Inst. Psychology South China Normal Uni.	Pengmin QIN	Prof.
5	Clinical Medicine	Shanghai Mental Health Center SJTU	Chunbo LI	Prof. Deputy Director

4. Tasks

General Goal

Construction of Linguistic Resources for Mental Disorders - Interdisciplinary Research in Linguistics, Cognitive Neuroscience and Artificial Intelligence

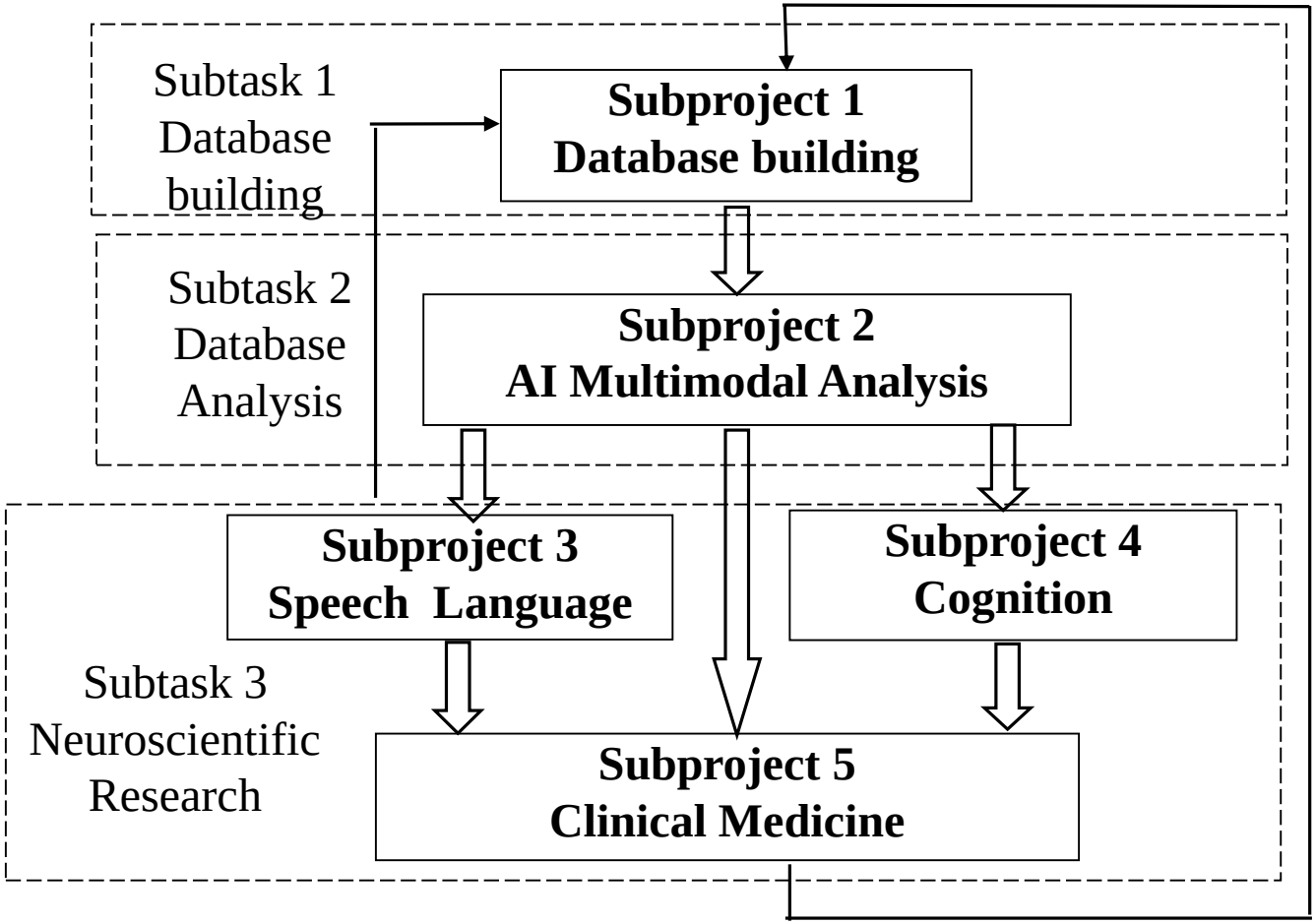
Tasks of Subprojects

Speech Language Research

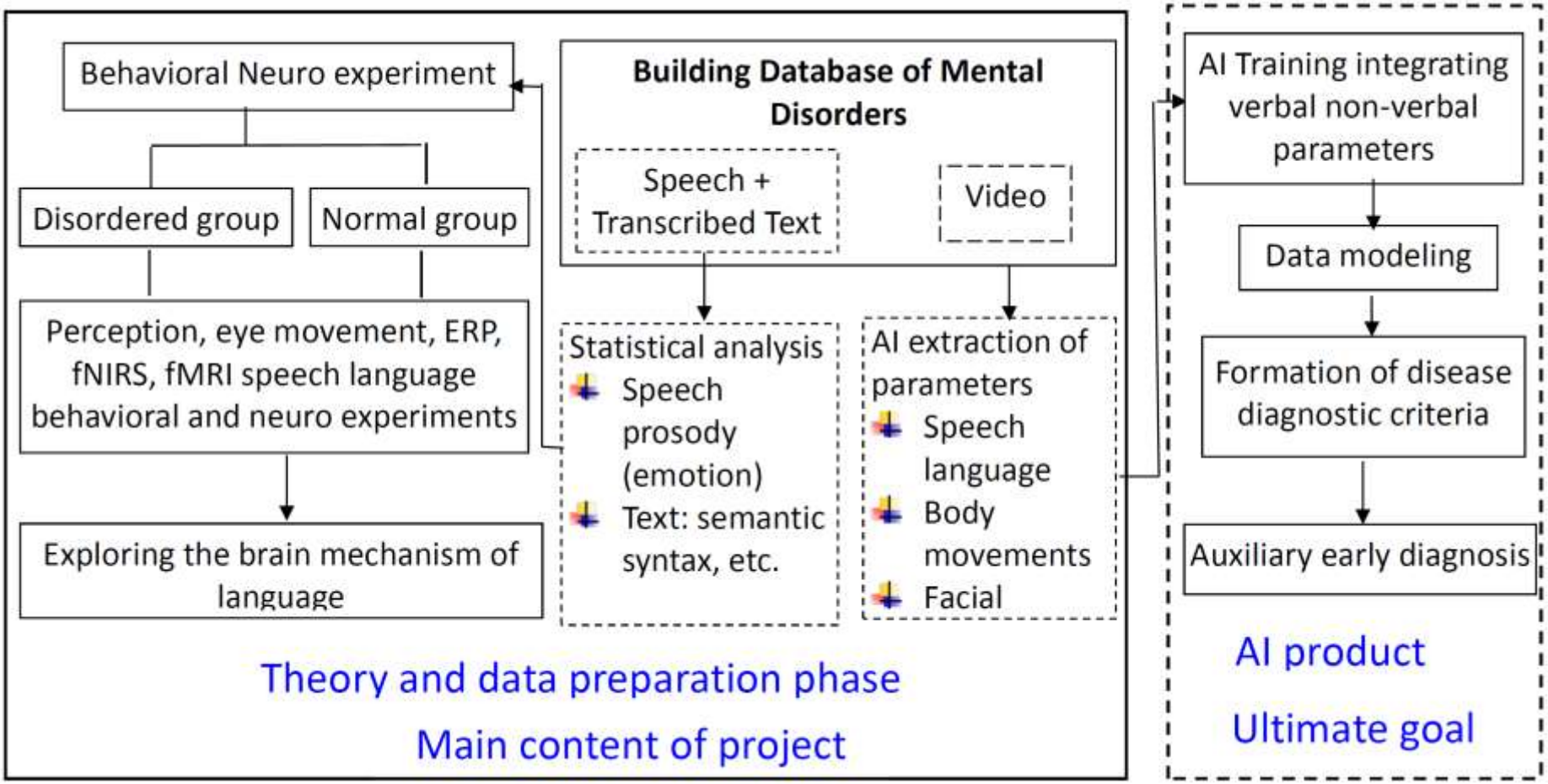
Database Building

Clinical Medicine

4. Tasks


Subtask 1
Database
building
Subproject 1
Database building
Subtask 2
Database
Analysis
Subproject 2
AI Multimodal Analysis
Subproject 3
Speech Language
Subproject 4
Cognition
Subtask 3
Neuroscientific
Research
Subproject 5
Clinical Medicine

4. Tasks


Behavioral Neuro experiment
Disordered group
Normal group
Perception, eye movement, ERP, fNIRS, fMRI speech language behavioral and neuro experiments
Exploring the brain mechanism of language
Building Database of Mental Disorders
Speech + Transcribed Text
Video
Statistical analysis
Speech prosody (emotion)
Text: semantic syntax, etc.
AI extraction of parameters
Speech language
Body movements
Facial
Theory and data preparation phase
Main content of project
AI Training integrating verbal non-verbal parameters
Data modeling
Formation of disease diagnostic criteria
Auxiliary early diagnosis
AI product
Ultimate goal

4. Tasks

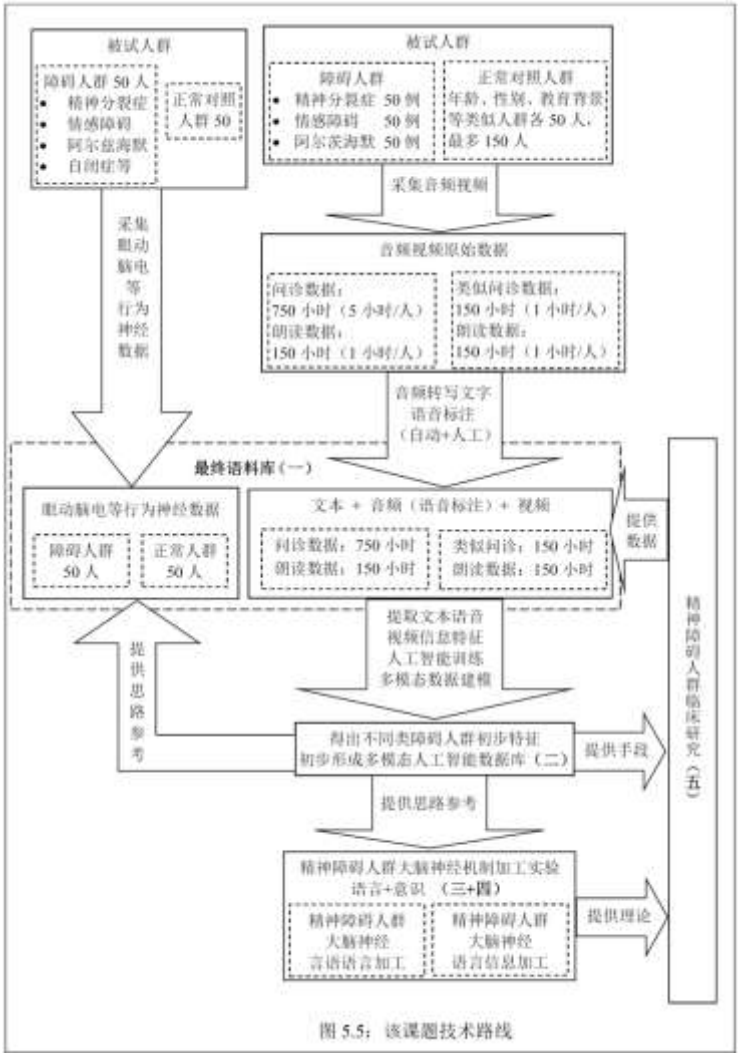


图 5.5：该课题技术路线

5. Database

- **Disordered Group**
  - Schizophrenia (50 cases)
  - Affective disorder (50 cases)
  - Alzheimer (50 cases)
- **Controlled Group (Age, sex, education)**
  - Normal vs Schizophrenia (50 cases)
  - Normal vs Affective disorder (50 cases)
  - Normal vs Alzheimer (50 cases)

5. Database

- **Audio visual**
  - **Disordered Group**
    - Doctoral Inquiry (5 hours/case) = (750 hours)
    - Reading + Description (1 hour/case) = (150 hours)
  - **Controlled Group (Age, sex, education)**
    - Similar Inquiry (1 hour/case) = (150 hours)
    - Reading + Description (1 hour/case) = (150 hours)

5. Database

- **Speech and Video + Neurobehavioral Database**
  - **Disordered Group (50 cases)**
    - Schizophrenia
    - Affective disorder
    - Alzheimer
    - Autism
  - **Controlled Group (Age, sex, education)**
    - Normal vs Schizophrenia
    - Normal vs Affective disorder
    - Normal vs Alzheimer
    - Normal vs Autism

- Home
- Program
- Important Dates
- Sponsors
- People
- Author Instructions
- Call for Papers
- Call for Grand Challenges
- Proceedings
- Registration
- Keynotes
- Tutorials
- Workshops
- Challenges
- Doctoral Consortium

21th ACM International Conference on Multimodal Interaction

Suzhou, Jiangsu, China. October 14-18, 2019

The 21st ACM International Conference on Multimodal Interaction (ICMI 2019) will be held in Suzhou, Jiangsu, China. ICMI is the premier international forum for multidisciplinary research on multimodal human-human and human-computer interaction, interfaces, and system development. The conference focuses on theoretical and empirical foundations, component technologies, and combined multimodal processing techniques that define the field of multimodal interaction analysis, interface design, and system development. ICMI 2019 will feature a single-track main conference which includes: keynote speakers, technical full and short papers (including oral and poster presentations), special sessions, demonstrations, exhibits and doctoral spotlight papers.

News and Updates

under construction

The proceedings of ICMI 2019 will be published by ACM as part of their series of International Conference Proceedings. ICMI 2019 will follow the ACM Policy Against Discrimination and Harassment.

https://icmi.acm.org/2019

Organizing Committee

Please submit questions/requests here

General Chairs

- Wen Gao (PKU)
- Matthew Turk (UCSB)
- Helen Mei Ling Meng (CUHK)

Program Chairs

- Kai Yu (SJTU)
- Björn Schuller (Imperial College)
- Yale Song (Microsoft, Redmond, USA)
- Susan R. Fussell (Cornell, USA)

Workshop Chairs

- Hongwei Ding (SJTU, China)
- Carlos Busso (UT Dallas, USA)
- Tadas Baltrusaitis (Microsoft, UK)

https://icmi.acm.org/2019/index.php?id=peopl

5. Database

◆**Transcribed Text + Speech （Transcription） + Video**

- Disordered group
  - Doctoral Inquiry：750 hours
  - Reading + Description：150 hours
- Controlled group
  - Similar Inquiry：150 hours
  - Reading + Description：150 hours

◆**Neurobehavioral data + Experimental design**

- Disordered group
  - Perception
  - ERP, etc.
- Controlled group
  - Perception
  - ERP, etc.

5. Presentation of Database

Classification		Schizophrenia		Aff. Disorder		Alzheimer	
		Schizophrenia	Normal	Aff. Disorder	Normal	Alzheimer	Normal
Number		50	50	50	50	50	50
Content		Inquiry + Reading	Description + Reading	Inquiry + Reading	Description + Reading	Inquiry + Reading	Description + Reading
Media		audio/video					
Annotation	audio	Audio wav（44.1kHz，16bit） Praat annotation TextGrid(Initial+Final+Tone）					
	video	Video mpg or mov （resolution ≥ 704×576）， ELAN annotation eaf					
	text	Text txt， POSAnnotatop					
Search method	word	Key word、POS、Pinyin					
	subject	Age, sex, education, subject ID, Disorder type					
	lexicon	Text ID +Frequency					
	video	Simultaneously with sounds, text, annottion					
Format	excel	Subject ID, Task ID, Text, Audio, Video					

6. Future work

- Maintenance of the platform
- Development of AI product
- More speech and language diagnosis criteria for mental disorders

Thank you for your attention

谢 谢！