Using Technology to Enhance Language Learning and Research: An Omniform Approach of Learning, Teaching, and Assessment of Chinese as a Second/Foreign Language

Yao-Ting Sung
Prof., Dept. of Educational Psychology and Counseling
Dean, Office of the Aim for the Top University
National Taiwan Normal University
Using Technology for the Teaching and Learning of Languages

• The evolution of hardware in learning technology
  – Desktops
  – Laptops/Tablets
  – Handhelds
  – Wearables
Using Technology for the Teaching and Learning of Languages

• How effective are the technology-integrated learning and teaching?
• The effect of Integrating mobile devices with teaching and learning in languages: A meta-analysis
The Meta analysis: Methods

- Data resource: ERIC and ISI (SSCI) databases from year 1993 to 2013
- 51 quasi/experimental articles fit the criteria
- Experimental groups used mobile-device for learning/teaching, control groups used traditional learning/teaching (e.g., paper or desktops)
Lessons learned (1)

- Mobile-integrated teaching and learning are not necessarily effective
- Elaborated teaching/learning strategies should be included in the programs
- Teaching-oriented software is important
- Appropriate intervention time is critical
- Logistics is important for long intervention
Lessons learned (2)

We need more elaborate **strategies** for implementing CALL studies:

- To make CALL studies comprehensive
- To make CALL studies thorough
- To make CALL studies integrated
- To make CALL studies effective
Framework of Technology-enhanced Chinese learning and Teaching

• OECD-T (Omniform Empowerment for Chinese Development through Technology)
Achievement

- Papers related with Chinese and learning technology, top 1 in Taiwan, top 5 in the world
- 2014 QS University ranking top 50 in linguistics worldwide
• Important findings of the research team include
  – There are systematic differences in the cognitive processing of CSL learners from Chinese native speakers (e.g. quantifier processing, relative clauses processing)
  – Bilinguals’ native language processing doesn’t not necessarily transfer to their CFL processing.
  – The average vocabulary quantity of CSL learners is roughly equal to the average vocabulary quantity of 2nd grades of native Chinese speakers.
  – Based on the neuro-image studies, the comprehension of Chinese humorous sentences involves a three-stage process
Advanced information technology for language research and learning

- **Speaking**: Speech recognition techniques based on innovative algorithms and mechanisms of speech perception and reproduction
- **Semantics**: Defining the conceptual difficulty of vocabulary through rule-based algorithms
- **Corpus**: Developing techniques for “Chinese word segmentation”, “Keywords Capture”, “Word co-occurrence analysis”, “Speech tagging”, “Syntax parsing”, and “Semantic category tagging”
- **Application**: CSL Automatic Speech Diagnosis System, Chinese Readability Indices Explorer (CRIE), AES-Han, eM Mandarin Platform of Words and Characters (eMPOWER)
Taiwan’s largest Corpus for Chinese Learners

Collected audio data of Chinese speech, delivered by learners with Japanese, Korean, and English languages as their mother tongue
Corpus for Chinese languages-
audio/video

- Movies, Music, Words, and Meaning combined into one, the Multimedia Corpus of Movies and Music
- Providing film clips, subtitles, full scripts, and an online dictionary integrated into a single function
- Learners can see words appear on the screen as they are being spoken, replay segments, and simultaneously click on any word to have the meaning displayed
- At any time, learners could choose to repeat any part which you want to listen
Corpus for Chinese languages-speech

• Chinese Learner Spoken Corpus
• [http://140.122.110.34/mp3c/](http://140.122.110.34/mp3c/)
• The 770 thousand word corpus comprises Taiwan’s first Corpus of Speech of Chinese Learners
• The speech was transcribed into a text file so that users can read the text and listen to the recording online
華語為第二語口語語料庫

國立臺灣師範大學於2011年起接受中華民國教育部補助「邁向頂尖大學計畫」，建構華語學習者中介語語料庫。華語中介語口語語料庫現階段以生語料庫的形式進行建置，由國家華語測驗委員會提供2008年起參加「國家華語測驗(TOP)」口語文能力測驗的考生口語語音檔。華語測驗的受試者分為基礎級與進階級兩個等級，本語料庫現階段以英、日、韓三個語言為母語的學習者語料為主，且主要收錄進階級語料，日後將加入高階段語料。目前已完成450人次，約77萬3千字的語料庫。

可輸入單字或短語檢索其使用情形：

語料庫：無限制
母語：無限制
排序方式：左方 右方
送出
Corpus for Chinese languages-written texts

- A three million word database built upon the writings of learners across varying levels and over 40 different mother tongues
- A comparison between NTNU Corpora of Chinese Learners’ Compositions and Corpus of HSK Compositions
Corpus for Chinese languages-written texts

- Chinese learner Written Corpus
- [http://kitty.2y.idv.tw/~hjchen/cwrite-mtc/main.cgi](http://kitty.2y.idv.tw/~hjchen/cwrite-mtc/main.cgi)
- Conducted research on the common errors of Chinese learners as well as analyzed usage of synonyms and sentence structures
<table>
<thead>
<tr>
<th>母語</th>
<th>來源</th>
<th>文體</th>
<th>功能</th>
<th>類別</th>
<th>級分</th>
<th>CEFR LEVEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>土耳其</td>
<td>紙本</td>
<td>信件</td>
<td>表達意見與說明</td>
<td>作業</td>
<td>2</td>
<td>A2</td>
</tr>
<tr>
<td>中</td>
<td>記敘文</td>
<td>表達感謝之意或祈求</td>
<td>3</td>
<td>B1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>日本</td>
<td>論說文</td>
<td>問候近況</td>
<td>4</td>
<td>B2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>巴西</td>
<td>應用文</td>
<td>描述</td>
<td>5</td>
<td>C1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>比利時</td>
<td></td>
<td></td>
<td>6</td>
<td>C2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>巴拉圭</td>
<td></td>
<td></td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>巴拿馬</td>
<td></td>
<td></td>
<td>8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>尼加拉瓜</td>
<td></td>
<td></td>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>以色列</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>瓜地馬拉</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>布吉納法索</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

查詢詞彙 | 詞性 | 不考慮 | 排序 | 後面詞排序 | 每頁筆數 | 十筆 |
Database for orthographic rules of Chinese characters

- The Most Comprehensive and Smart Database For Rules on Chinese Character Formation
  - Including three components: radicals, character roots, and stroke order
  - Building a way of systematically displaying character formulation patterns, including both simplified and traditional characters
- Built-in categories that allow learners to compare and contrast various categories of traditional/simplified characters
  - Including nested words, character formation ratios, frequency, and differences between consistency and transparency of transliterations
- Design models for handwriting Chinese characters, as well as diagnostics and feedback, both of which are incorporated into the learning platform
### CSL Automatic Speech Diagnosis System

- Integrating empirical studies such as acoustic characteristics, speech acquisition mechanisms, assessment theories, and teaching strategies
- Compared to Nuance and MyCT

<table>
<thead>
<tr>
<th>Software background</th>
<th>Development goal(s)</th>
<th>Language unit</th>
<th>Recognition level</th>
<th>Feedback mechanism</th>
</tr>
</thead>
<tbody>
<tr>
<td>NTNU Chinese Listening &amp; Speaking</td>
<td>Combined with Speaking Automatic Diagnostics software which developed by NTNU</td>
<td>Teaching and Testing</td>
<td>character, word</td>
<td>comprehensive pronunciation ability test</td>
</tr>
<tr>
<td>MyCT Chinese communication</td>
<td>The only one developed by Taiwan</td>
<td>Teaching</td>
<td>phrase</td>
<td>pronunciation, tones, fluency &amp; volume</td>
</tr>
<tr>
<td>Nuance</td>
<td>World's most widely-used (multiple language) speech recognition software</td>
<td>Voice Transcription</td>
<td>character to paragraph</td>
<td>unclear comparison</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>pronunciation, tones, fluency &amp; volume</td>
<td>none</td>
</tr>
</tbody>
</table>
Chinese Readability Indices Explorer

• Quantitative analysis for multilevel linguistic features: words, syntax, semantics, cohesion
• Can be applied to textbook classification, assessment, compilation and editing
• Text leveling for eBooks, websites, CFL text books etc.
• Diagnosis of the linguistic features (syntax and word-difficulty levels) in texts
Chinese Readability Indices Explorer (CRIE)

- [http://www.chinesereadability.net/CRIE/](http://www.chinesereadability.net/CRIE/)
- Comparison between CRIE 2.0 and Coh-Metrix 3.0

<table>
<thead>
<tr>
<th>Feature</th>
<th>CRIE 2.0</th>
<th>Coh-Metrix 3.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of features analyzed</td>
<td>79</td>
<td>108</td>
</tr>
<tr>
<td>Features can be selected</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Diagnostic feature function</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Has multiple essay analysis</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Readability formula calculation</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Readability model types</td>
<td>non-linear, linear</td>
<td>linear</td>
</tr>
<tr>
<td>Readability analysis available for online compositions</td>
<td>yes</td>
<td>no</td>
</tr>
</tbody>
</table>
Automated Chinese Essay Scoring — AES-Han

- The composition analytics and feedback system for CFL learners

- Comparison between the composition analytics and feedback systems of AES-Han and AES-ETS

<table>
<thead>
<tr>
<th>AES-Han</th>
<th>AES-ETS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Language</strong></td>
<td><strong>Grammar is weighted most heavily</strong></td>
</tr>
<tr>
<td>Chinese</td>
<td>English</td>
</tr>
<tr>
<td>v. vocabulary, content, grammar, structure &amp; organization all equally weighted</td>
<td>numerous</td>
</tr>
<tr>
<td>few</td>
<td>high</td>
</tr>
<tr>
<td>high</td>
<td>high</td>
</tr>
<tr>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>multi-level detection</td>
<td>single-level detection</td>
</tr>
</tbody>
</table>

國立臺灣師範大學 National Taiwan Normal University
eMandarin Platform of Words and Characters (eMPower)


- Digital learning platform
  - Theory-based CFL textbook
  - Learning-strategies embedded
  - Simultaneously learning listening, speaking, reading, writing
  - Real time feedback
  - App for anytime, anywhere
第一課 這是我的家人 Lesson 1 These Are My Family Members

安娜娜：如娟，這是你媽媽嗎？
An Nàna: Rújuān, zhě shì ni māma ma?

姜如娟：不是，這是我外婆。
Jiāng Rújuān: Búshì, zhě shì wǒ wàipó.

安娜娜：那女孩是你的姑姑嗎？
An Nàna: Nà nǚhái shì ní de gūgū ma?

姜如娟：她是我媽媽的姐妹，我叫阿姨。
Jiāng Rújuān: Tā shì wǒ māma de jiěmèi, wǒ jiào āyí.

安娜娜：這是你媽媽的爸爸嗎？
An Nàna: Zhě shì ni māma de bābā ma?

姜如娟：是，他是我外公。
Jiāng Rújuān: Shì, tā shì wǒ wàigōng.
# e MPOWER — learning characters & words

## 第一課 這是我的家人

<table>
<thead>
<tr>
<th>部件</th>
<th>Radical characters</th>
</tr>
</thead>
<tbody>
<tr>
<td>女父</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>字族</th>
<th>Character-Family</th>
</tr>
</thead>
<tbody>
<tr>
<td>女: 奶奶 女</td>
<td>姨姑 姨娭 娘</td>
</tr>
<tr>
<td>女: 爸爸</td>
<td>爹</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>其它字</th>
<th>Others characters</th>
</tr>
</thead>
<tbody>
<tr>
<td>家人 是你嗎</td>
<td>不我外那孩的叫阿他公</td>
</tr>
</tbody>
</table>
第一課 這是我的家人
Lesson 1 These Are My Family Members

我是王若凡。 PLAY REC
我是________  {王若凡} PLAY REC
我是________  {王若凡} PLAY REC
我是________  {安娜的爸爸} PLAY REC
我是安娜的爸爸。 PLAY REC

她是安娜的姐姐，我叫她安娜。 PLAY REC
她是安娜的妹妹，我叫安娜。 PLAY REC
他是安娜的姐姐，我叫安娜。 PLAY REC
他是安娜的妹妹，我叫安娜。 PLAY REC

她是我爸爸的姐姐，我叫她安娜。 PLAY REC
她是安娜的爸爸，我叫安娜。 PLAY REC
他是安娜的爸爸，我叫安娜。 PLAY REC
他是安娜的爸爸，我叫安娜。 PLAY REC

我姓王，我叫王若凡。 PLAY REC
我姓________  {安娜的爸爸} PLAY REC

Learning for communication and interaction

- 3D Online Virtual Environment Second Life (SL)
Superior-multimedia, multilevel, multinational—CSL Teaching Materials

• 初級華語學習詞典: The first sentence-based book globally, the first promotion overseas

• Chinese CUBES: iF Communication Design Award – The first award in character materials globally

• Hello·華語: The maximum sales in Chinese material for children globally
Innovative hardware—EyeNTNU-180

- Natural Light Eye Tracker & Low-cost Eye Tracker
- Innovative Eye Movement Detection Technology
- The first eye tracking system that does not rely on infrared light

<table>
<thead>
<tr>
<th></th>
<th>EyeNTNU-180</th>
<th>Tobii X120</th>
<th>Tobii TX300</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price</td>
<td>less than NT 150,000 ✓</td>
<td>NT 750,000</td>
<td>NT 1,500,000</td>
</tr>
<tr>
<td>Sampling frequency</td>
<td>180Hz</td>
<td>120Hz</td>
<td>300Hz ✓</td>
</tr>
<tr>
<td>Angle of error</td>
<td>0.3 degrees ✓</td>
<td>0.5 degrees</td>
<td>0.5 degrees</td>
</tr>
<tr>
<td>Quantity of infrared LED</td>
<td>2 W/st ✓</td>
<td>60 W/st</td>
<td>60 W/st</td>
</tr>
<tr>
<td>Tester's head can move</td>
<td>no</td>
<td>yes ✓</td>
<td>yes ✓</td>
</tr>
<tr>
<td>Has hardware and software technologies</td>
<td>yes ✓</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Customized software can be developed</td>
<td>yes ✓</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>
• EyeNTNU180
  – A parity and competitive eye tracker in the world
Global Network of Chinese Learning

- Peking University
- CUHK, H.K.
- PolyU, H.K.
- NTU, Singapore
- Monash University
- Griffith University
- Canterbury
- UH, Germany
- MU, Czech
- TUFS, Japan
- Waseda, Japan
- Simon Fraser
- Illinois
- Ohio
- CMU
- PSU
- UP
- ETS
- UK, Lancaster
- UK, Leeds
- UK, Cambridge
- Marylands
- Ohio
Thanks for your attention!

Yao-Ting SUNG

sungtc@ntnu.edu.tw