Keep Dialects Alive: Chinese Experience in Building Dialect Archives

YUJIAO DU (YUQIAO DU)
School of Information Management
Sun Yat-Sen University
E-mail: duyj7@mail2.sysu.edu.cn
CONTENTS

Background

Case of Zhejiang Province

Experience of Chinese Dialect Archiving

Summary
Background
China is a multi-ethnic, multi-regional and multilingual country.

Chinese language resources

- Ethnic minority language resource
- Dialect language resource
- National common language resource
In China, each region has more than one dialect and until now there is no authoritative statistics on the number of dialects. However, general speaking, there are 10 dialect distributions:

<table>
<thead>
<tr>
<th>Mandarin Dialect Distribution</th>
<th>Cantonese Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jin Dialect Distribution</td>
<td>Xiang Dialect Distribution</td>
</tr>
<tr>
<td>Wu Dialect Distribution</td>
<td>Gan Dialect Distribution</td>
</tr>
<tr>
<td>Min Dialect Distribution</td>
<td>Hui Dialect Distribution</td>
</tr>
<tr>
<td>Hakka Distribution</td>
<td>Ping and Local Dialect Distribution</td>
</tr>
</tbody>
</table>
Background

Population of Each Dialect Distribution

- Mandarin dialect: 66%
- Southwestern: 22%
- Central Plains: 15%
- Jialing: 1%
- Northeast: 8%
- Jianghuai: 7%
- Jilu: 7%
- Beijing: 2%
- Lanyin: 1%
- Jiaoliao: 3%
- Hui: 0%
- Gan: 4%
- Xiang: 3%
- Cantonese: 5%
- Wu: 6%
- Jin: 5%
- Hakka: 3%
- Ping and Local: 1%
- Northeast: 8%
- Beijing: 2%
- Jialing: 1%
- Lanyin: 1%
- Jianghuai: 7%
- Jilu: 7%
- Central Plains: 15%
- Southwestern: 22%
Number of Administrative Districts Covered By Each Dialect Distribution

- Mandarin dialect: 546
- Southwestern: 546
- Northeast: 198
- Jiaoliao: 44
- Central Plains: 397
- Jianghuai: 108
- Jilu: 162
- Beijing: 52
- Lanyin: 70
- Min: 154
- Wu: 160
- Xiang: 64
- Gan: 102
- Hui: 19
- Ping and Local: 60
- Jin: 194
- Cantonese: 141
- Hakka: 110
- Beijing: 52
- Jilu: 162
- Central Plains: 397
- Southwestern: 546
- Northeast: 198
- Jiaoliao: 44
- Lanyin: 70
- Mandarin dialect: 546
Why should we build the dialect archives?

- With the popularity of Mandarin, the speaking language in China has been gradually normalized and unified. However, some dialects which a small number of people speak or of which distributions exists sporadically may be gradually faded out.

- It is necessary to archive these dialect resources to keep social memory.
Background

Chinese dialect preserving projects

Ministry of Education & Chinese Language Council
The Chinese language audio database resource Project (2008)

National Archives of Provinces and Cities
eg: Zhejiang Province, Jiangsu Province, Shanxi Province

Non-government individuals or organizations
Case of Zhejiang Province
Why do I choose this case?

- National Archives of Zhejiang Province is the first provincial archives to start dialect archiving work.
- Zhejiang province belongs to Wu dialect distribution. Wu dialect is the one of the oldest languages in China. Compared with mandarin, Wu dialect retains more ancient pronunciation and has special words and grammars.
- Many cities of Zhejiang province have taken part in the work and much progresses have been made. Until 2015, 96 kinds of languages belonging to Wu dialect across 11 cities in Zhejiang province have been recorded and the Zhejiang dialect audio database has been built.
There are mainly 3 stages in the Zhejiang dialect archiving project:

1. Research & Plan
2. Pilot Project
3. Popularization
Case of Zhejiang province

1. Objective
To establish the Zhejiang Dialect Audio Archival Database by arranging the pronunciations, vocabularies, grammars, voice stories, folk songs, and operas of Zhejiang dialects.

4. Tasks
- Conduct survey & investigation about Zhejiang dialects
- Develop texts about what should be preserved
- Make the audios
- Archiving

2. Steps
- Pilot Step
- Popularization Step
About Voice Providers...

Voice Providers refer to those citizens who have native accent and join the dialect archival audios project by reading out the dialect words or other dialect works given. Their voices would be recorded and made into audios.

Persons who sign up would be selected after a series of strict examinations, for example they may be required to read out a dictionary or an article in dialect.
### What kind of person can be Voice Provider?

<table>
<thead>
<tr>
<th>Gender</th>
<th>Both males and females</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Middle aged and elderly</td>
</tr>
<tr>
<td>Education</td>
<td>Primary school - High school</td>
</tr>
<tr>
<td>Experience</td>
<td>Have not been away from hometown for long term</td>
</tr>
<tr>
<td>Family</td>
<td>Not only himself/ herself but also parents and spouse are native local persons.</td>
</tr>
<tr>
<td>Health</td>
<td>With complete front teeth</td>
</tr>
</tbody>
</table>
In 2013, *Call of the Local Accent* was finished by National Archives in Zhejiang based on Zhejiang dialect archival database. It includes 12 DVDs exploring the relationships between dialect and culture, dialect and Zhejiang merchants, dialect and archive.
Experience of Chinese Dialect Archiving
Experience of Chinese Dialect Archiving

Stage 1
Collection & Preservation

Stage 2
Protection

Stage 3
Exploitation & Utilization

Each stage could be led by the Archives and involved with some other organizations.
Experience of Chinese Dialect Archiving

Stage 1
Collection & Preservation

What should be preserved?

Linguistic content
pronunciation rule, words, grammars...

Extralinguistic content
historical background, dialect art works...

How to preserve?
both audio-video and digital text

Language Department collaborate
Stage 2: Protection

- Archives could collaborate with intangible cultural heritage protection association and education department.

Experience of Chinese Dialect Archiving

Record dialect resources → preservation → protection
Maintain dialect resources
Experience of Chinese Dialect Archiving

Stage 3: Exploitation & Utilization

- Language department could give some advice in making dialect archive more available because linguists would be important users of dialect archives. The way of using these archives is supposed to be multicultural and it should aim at meeting needs of users and enhancing the vitality of dialects.
Summary
Language resource is an important part of social memory. To build dialect archives is not just maintaining dialect resources integrally like antiques displayed in the museum. It aims at keeping dialects alive and making them more valuable.
THANK YOU!

YUJIAO DU (YUQIAO DU)
School of Information Management
Sun Yat-Sen University
E-mail: duyj7@mail2.sysu.edu.cn