IDEOPHONES AND THE STANDARD HISTORIES
A CASE STUDY OF MENTAL SPACE MARKERS
Thomas Van Hoey
司馬智

From Belgium

Doing PhD in Linguistics at National Taiwan University

Main research interests:

• Ideophones in (premodern) Chinese (and other languages)
• Working on the Bunun (Austronesian) part of a Formosan spoken corpus @ NTU
Main Research Question:

How did ideophones develop through the ages in Classical and Literary Chinese?
MY PREVIOUS STUDIES

Shijing
詩經
2016 (ISACG 9)
poetry

300 Tang Poems
唐詩三百首
2015 (MA thesis)
poetry

diachronic variation
2016 (CLDC8)
[2018 (CLR)]

weather
2017 (IFG)
taxonomy

intro
material
framework
method
case study
GOALS AND OUTLINE

1. Material
   Twenty-five histories: focus on the four “Three Histories”

2. Framework
   Constructions of ideophones

3. Methodology
   Corpus-driven and usage-based

4. Case study
   Emergent linguistic phenomena like mental space markers
TWENTY-FIVE HISTORIES
THE STANDARD HISTORIES

“twenty-four histories”  +  Draft History of the Qing
二十四史  清史稿

- reliable transmission
- reliable / well-known metadata
- prose
- diachronic usage

- model works
- political and moral history
- previous historical research
  (Wilkinson 2015:643)
### Overview of Standard Histories

<table>
<thead>
<tr>
<th>Year</th>
<th>Chinese Title</th>
<th>English Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>-87</td>
<td>史記</td>
<td>Records of the historian</td>
</tr>
<tr>
<td>96</td>
<td>[前]漢書</td>
<td>History of the Former Han</td>
</tr>
<tr>
<td>297</td>
<td>三國志</td>
<td>Record of the Three Kingdoms</td>
</tr>
<tr>
<td>445</td>
<td>後漢書</td>
<td>History of the Later Han</td>
</tr>
<tr>
<td>488</td>
<td>宋書</td>
<td>Song History</td>
</tr>
<tr>
<td>537</td>
<td>南齊書</td>
<td>History of Southern Qi</td>
</tr>
<tr>
<td>572</td>
<td>魏書</td>
<td>History of the Wei</td>
</tr>
<tr>
<td>636</td>
<td>梁書</td>
<td>History of the Liang</td>
</tr>
<tr>
<td>636</td>
<td>陳書</td>
<td>History of the Chen</td>
</tr>
<tr>
<td>636</td>
<td>北齊書</td>
<td>History of the Northern Qi</td>
</tr>
<tr>
<td>636</td>
<td>周書</td>
<td>History of the Zhou</td>
</tr>
<tr>
<td>636</td>
<td>隋書</td>
<td>History of the Sui</td>
</tr>
<tr>
<td>648</td>
<td>晉書</td>
<td>Jin History</td>
</tr>
<tr>
<td>659</td>
<td>南史</td>
<td>History of the Southern Dynasties</td>
</tr>
<tr>
<td>659</td>
<td>北史</td>
<td>History of the Northern Dynasties</td>
</tr>
<tr>
<td>945</td>
<td>舊唐書</td>
<td>Old history of the Tang</td>
</tr>
<tr>
<td>974</td>
<td>舊五代史</td>
<td>Old history of the Five Dynasties</td>
</tr>
<tr>
<td>1060</td>
<td>新唐書</td>
<td>New history of the Tang</td>
</tr>
<tr>
<td>1073</td>
<td>新五代史</td>
<td>New history of the Five Dynasties</td>
</tr>
<tr>
<td>1344</td>
<td>遼史</td>
<td>History of the Liao</td>
</tr>
<tr>
<td>1344</td>
<td>金史</td>
<td>History of the Jin</td>
</tr>
<tr>
<td>1345</td>
<td>宋史</td>
<td>History of the Song</td>
</tr>
<tr>
<td>1370</td>
<td>元史</td>
<td>History of the Yuan</td>
</tr>
<tr>
<td>1739</td>
<td>明史</td>
<td>History of the Ming</td>
</tr>
<tr>
<td>1928</td>
<td>清史稿</td>
<td>Draft history of the Qing</td>
</tr>
</tbody>
</table>
A TYPICAL STANDARD HISTORY

1. Basic annals *benji* 本紀 (alt. *ji* 紀 or *diji* 帝紀)
2. Biographies *liezhuan* 列傳
3. Treatises *shuzhi* 書志 (alt. *shu* 書 or *zhi* 志 or *kao* 考)
4. Chronological tables *biao* 表
5. Annals of non-legitimate regimes *zaiji* 載紀
6. Hereditary houses *shijia* 世家
THE SAN SHI 三史 CANON

• Originally:
  ❖ Shiji 史記
  ❖ Hanshu 漢書
  ❖ Dongguan Hanji 東觀漢記

• From Tang dynasty onward
  ❖ Shiji 史記
  ❖ Hanshu 漢書
  ❖ Hou Hanshu 後漢書
  ❖ Sanguo zhi 三國志

<table>
<thead>
<tr>
<th></th>
<th>Shiji 史記</th>
<th>Hanshu 漢書</th>
<th>Hou Hanshu 後漢書</th>
<th>San guo zhi 三國志</th>
</tr>
</thead>
<tbody>
<tr>
<td>本紀 annals</td>
<td>0.01</td>
<td>0.09</td>
<td>0.1</td>
<td>0.03</td>
</tr>
<tr>
<td>列傳 treatises</td>
<td>0.2</td>
<td>0.14</td>
<td>0.18</td>
<td>0.78</td>
</tr>
<tr>
<td>志 other</td>
<td>0.2</td>
<td>0.19</td>
<td>0.03</td>
<td>0.19</td>
</tr>
<tr>
<td>表 houses</td>
<td>0.09</td>
<td>0.5</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>世家 biographies</td>
<td>0.35</td>
<td>0.08</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>世家 tables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
THE *SAN SHI* 三史 CANON

**Standard history**
- **Shiji 史記**: 845-95 BC
- **Hanshu 漢書**: 206 BC - 24 AD
- **Hou Hanshu 後漢書**: 25-220
- **Sanguo zhi 三國志**: 221-280

**Covered period**
- **Finished date**
- **Author**
- **Life**

**Historical details**

Suitable as material

Evolution from 1st cent BC until 5th cent AD

Linguistic style

1. **Shiji 史記**
   - 845-95 BC
   - Sima Qian 司馬遷
   - 145-86 BC

2. **Hanshu 漢書**
   - 206 BC - 24 AD
   - Ban Gu 班固
   - 32-92

3. **Hou Hanshu 後漢書**
   - 25-220
   - Fan Ye 范曄
   - 398-445

4. **Sanguo zhi 三國志**
   - 221-280
   - Chen Shou 陳壽
   - 233-297
FRAMEWORK AND THEORETICAL BACKGROUND: CONSTRUCTIONS

A few different flavours of construction grammars (Hoffmann & Trousdale 2013), e.g.

- Fillmore’s Berkeley Construction Grammar (1988 etc.)
- Croft’s Radical Construction Grammar (2001)

Broadly mutually compatible (Croft 2001), yet still with their own differences (Langacker 2005).
ADVANTAGE OF THESE FRAMEWORKS: COGNITIVE-FUNCTIONALIST

This idea of ‘constructions’ ties in with both FORM and MEANING and the SYMBOLIC link between these two (Langacker 1987)

A recent definition:

“ideophones = marked words that depict sensory images” (Dingemanse 2011; 2012)

Older definition:

“A vivid representation of an idea in sound. A word, often onomatopoetic, which describes a predicate, qualificative or adverb in respect to manner, colour, sound, smell, action, state or intensity.” (Doke 1935)
ADVANTAGE OF THESE FRAMEWORKS: MORPHOLOGY

Identify
- Usage-based constructions
- From the bottom-up
- ...

Lu Chiarung 呂佳蓉 (2006:97) on Japanese onomatopoeia
ADVANTAGE OF THESE FRAMEWORKS: MORPHOLOGY

Similar distributions effects are observed for ideophones in Tang poetry (Van Hoey 2015)
ADVANTAGE OF THESE FRAMEWORKS: USAGE CONTEXTS

Constructions identified in *300 Tang poems* 唐詩三百首 (Van Hoey 2015:76-80)

1. PREDICATE construction 119 = 59%
2. ADVERBIAL construction 34 = 17%
3. ATTRIBUTIVE construction 30 = 15%
4. NOMINALISATION construction 6 = 3%
5. BACKGROUND construction 7 = 3%
6. UTTERANCE construction 6 = 3%

But what about non-poetry? ➔ see material: 4 of the 25 histories

And what about the lower level patterns, which are presumed to be more entrenched? (Tuggy 2003)
## METHODOLOGY: CORPUS LINGUISTICS

<table>
<thead>
<tr>
<th>Evidence for ideophone analysis</th>
<th>OC &amp; MC</th>
<th>comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Intuition (as a native speaker)</td>
<td>X</td>
<td>not a native speaker</td>
</tr>
<tr>
<td>2. Experiments</td>
<td>X</td>
<td>no native speakers left</td>
</tr>
<tr>
<td>3. Elicitation (narratives etc.)</td>
<td>X</td>
<td>no native speakers left</td>
</tr>
<tr>
<td>4. Dictionaries</td>
<td>O</td>
<td>historical dictionaries (Geeraerts 2006)</td>
</tr>
<tr>
<td>5. Corpus-based</td>
<td>O</td>
<td>good if you already know what you want to look for</td>
</tr>
<tr>
<td>6. Corpus-driven</td>
<td>O</td>
<td>emergent phenomena</td>
</tr>
</tbody>
</table>
# METHODOLOGY: CORPUS LINGUISTICS

Evidence for ideophone analysis

<table>
<thead>
<tr>
<th>Method</th>
<th>OC &amp; MC</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Intuition (as a native speaker)</td>
<td>X</td>
<td>not a native speaker</td>
</tr>
<tr>
<td>2. Experiments</td>
<td>X</td>
<td>no native speakers left</td>
</tr>
<tr>
<td>3. Elicitation (narratives etc.)</td>
<td>X</td>
<td>no native speakers left</td>
</tr>
<tr>
<td>4. Dictionaries</td>
<td>O</td>
<td>historical dictionaries (Geeraerts 2006)</td>
</tr>
<tr>
<td>5. Corpus-based</td>
<td>O</td>
<td>good if you already know what you want to look for</td>
</tr>
<tr>
<td>6. Corpus-driven</td>
<td>O</td>
<td>emergent phenomena</td>
</tr>
</tbody>
</table>
FROM RAW MATERIAL TO DATA

raw material

data theory
FROM RAW MATERIAL TO DATA

raw material

data theory

concordances

method

framework

material

intro

case study

raw
theory
FROM RAW MATERIAL TO DATA

• Unfortunately raw data does not come in the format you want, you need to mold it into data that you can analyze and over which you can theorize.

• In the process, I used different technologies. Reasons for this include:
  – Full control over every stage of the data processing
  – Establishing a processual methodology for bigger sets of data
  – Discovering the potential of different technologies

• From raw data to data fit for theorizing in eight steps
raw material

structured material

R-script

ideophone database

MARKUS

annotated file

python

file for linguistic use

wordsketch

concordances
STEP 1: RAW MATERIAL

**Wikisource.org**
+ easy access
+ verifiable
+ structured

- differences in character sets (trad. vs. simp.)
- inconsistent inclusion of commentary
- mark-up that needs to be stripped
- some characters are missing

[https://zh.wikisource.org/wiki/後漢書](https://zh.wikisource.org/wiki/後漢書)

**Alternatives**

- Retype everything myself
- Repositories like [Kanseki](https://kanseki.org)
  - No punctuation
  - More a reflection of (cultural) printing conventions than of linguistic materials (?)
  - Maybe in the future (?)
- Available corpora like [Academia Sinica, Beijing](https://www.sinica.edu.tw)
  - Sometimes different tag-schemas (doesn’t capture all my ideophones)
  - Not all texts that I want included
STEP 1: RAW MATERIAL

**Wikisource.org**

+ easy access
+ verifiable
+ structured

- differences in character sets  
  (trad. vs. simp.)
- inconsistent inclusion of commentary
- mark-up that needs to be stripped
- some characters are missing

https://zh.wikisource.org/wiki/《後漢書》

---

**《後漢書》**

<table>
<thead>
<tr>
<th>纪</th>
<th>[编辑]</th>
</tr>
</thead>
<tbody>
<tr>
<td>卷一</td>
<td>光武帝紀 第一上</td>
</tr>
<tr>
<td>卷一下</td>
<td>光武帝紀 第一下</td>
</tr>
<tr>
<td>卷二</td>
<td>显宗孝明帝紀 第二</td>
</tr>
<tr>
<td>卷三</td>
<td>肃宗孝章帝紀 第三</td>
</tr>
<tr>
<td>卷四</td>
<td>孝和孝殇帝紀 第四</td>
</tr>
<tr>
<td>卷五</td>
<td>孝安帝紀 第五</td>
</tr>
<tr>
<td>卷六</td>
<td>孝順孝沖孝質帝紀 第六</td>
</tr>
<tr>
<td>卷七</td>
<td>孝桓帝紀 第七</td>
</tr>
<tr>
<td>卷八</td>
<td>孝靈帝紀 第八</td>
</tr>
<tr>
<td>卷九</td>
<td>孝獻帝紀 第九</td>
</tr>
<tr>
<td>卷十上</td>
<td>皇后紀 第十上</td>
</tr>
<tr>
<td>卷十下</td>
<td>皇后紀 第十下</td>
</tr>
</tbody>
</table>
STEP 2: STRUCTURED MATERIAL

Wikisource: 《後漢書》

Structured text 《後漢書·紀》
**STEP 3: R-SCRIPT AND DATABASE**

From my **database** (Libreoffice Base v. 5.3.2.2) I select all ideophones encountered so far, both in simplified and traditional, and put them in a ‘**masterlist.txt**’ file.

Using an **R-script** to help me identify ideophones of the **AA-type (full reduplication)**, e.g. `zhengzheng` 丁丁 (not dingding!)

This I then look up, and add to my database if it wasn’t there yet. This way the database can grow throughout my research.

**Current statistics of the database:** 784 types
STEP 3: R-SCRIPT AND DATABASE

Ideophone Edit form

ideophoneID
14

pinyin
zheng-zheng

pinyin-tone
zheng-zheng

MC
tseng-tseng

OC
*Fen - *Fen

traditional
丁丁

simplified
丁丁

Form Letter
AA

Formal mechanism
reduplication

Ricei

Zdic

Kroll

same as ID

First Record Previous Record Next Record Last Record

Note
STEP 3: R-SCRIPT AND DATABASE

**Ideophone Edit form**

IDEOPHONE

- ideophoneID
- pinyin
- traditional
- simplified
- MC
- OC
- same_seq_ID

FORMS

- Form Letter
- Formal mechanism

LEXICOGRAPHIC MEANINGS

A Student’s Dictionary of Classical and Medieval Chinese

Dictionnaire Ricci Chinois-Français
STEP 3: R-SCRIPT AND DATABASE

A relational database molded after the Chinese Biographical Data Base (CBDB)
- easier to avoid errors when copying
- structural queries
- ever-growing collection of ideophones encountered
STEP 3: R-SCRIPT AND DATABASE

Currently: 784 types
If meaning in language is encyclopedic, it is very important to know about real world references:

- esp. in historiography and
- esp. in Premodern Chinese, where most markers are dropped consistently or are optional in the best circumstances.

Names = proper nouns = profiles THING (cf. Langacker 2008:316-318)

Places and Temporal references also have special semantic import: their basic domains are the SPATIAL and the TEMPORAL.

And of course the MIMETICS/IDEOPHONES we are interested in
MARKUS is designed for close reading i.e., is looks quite appealing in the screen.
STEP 4: MARKUS ANNOTATED FILE
(IN XML)

Unreadable for human eyes
STEP 5: TURNING UNREADABLE XML INTO PROCESSABLE TXT

Regular expressions in python

- All the xml `<...XXX...>` tags have been replaced by a shortened “`_XXX`” version
- Punctuation ( 。 、 ） has received a tag (if in the future I would want to delete it, it is very easy to do so)
- Commentaries have been deleted
- Segmented after every character (`'monosyllabicity of premodern Chinese' (it's not perfect I know)
raw material
wikisource

structured material
TXT

R-script
ideophone database
base

MARKUS

annotated file
NAME, IDEOPHONE, PLACE, TIME
XML

python

file for linguistic use
punctuation, segmentation, light tagging

sketch engine

concordances

ideophone database
base

R-script

data processing
XLS

-python

sketchengine

words

-method

material

frame work

-case study

intro
Concordances were the main goal from which we wanted to do research. With the previous steps, we are now able to make them.

Software choice:

- **Antconc**: very slow
- **Natural Language Toolkit (NLTK)**: not what I want (but maybe I need more time to discover this)
- **Sketch Engine**: + free
  + does the job very well
  + extra functions
  - limited in #characters (1 million)
STEP 6: CONCORDANCES WITH SKETCH ENGINE

Ideophones in the 《後漢書》

Tag: IDE

- 987 tags in the 後漢書
- ‘IDE’ and ‘_’ are seen as separate words
- Some words tagged as IDE (and others) are reified as compounds
  - no solution?
  - but also no problem?
- Overgeneralization: some things tagged as IDE that shouldn’t, like 中元 元年、永元 元年
STEP 7: DATA PROCESSING IN EXCEL (XLS/TXT)

- Exporting the concordances from Sketch Engine, running them through a python script I wrote to get the concordances I want.
STEP 8: DATA AND THEORY

Now I have converted my raw textual data into readable concordances.

Why the long way?

• I had control over the process
• I can go back for more context should I need it
• Scripts for future research
### Sketch Engine: Token Frequency

<table>
<thead>
<tr>
<th>Document</th>
<th>Number of ideophone tokens</th>
<th>Normalized frequency (per million)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Shiji</em> 史記</td>
<td>310</td>
<td>458.26</td>
</tr>
<tr>
<td><em>Hanshu</em> 漢書</td>
<td>757</td>
<td>814.49</td>
</tr>
<tr>
<td><em>Houhanshu</em> 後漢書</td>
<td>987</td>
<td>773.18</td>
</tr>
<tr>
<td><em>Sanguo zhi</em> 三國志</td>
<td>230</td>
<td>581.13</td>
</tr>
</tbody>
</table>

There is a small error margin in these numbers. However, the mimetics do occur in prose texts.
# SKETCH ENGINE: WORDSKETCH

## IDE

### Houhanshu freq = 987 (773.18 per million)

<table>
<thead>
<tr>
<th>words before &quot;IDE&quot;</th>
<th>1,472</th>
<th>149.14</th>
</tr>
</thead>
<tbody>
<tr>
<td>元</td>
<td>58</td>
<td>9.39</td>
</tr>
<tr>
<td>中元 元 _ IDE 年</td>
<td></td>
<td></td>
</tr>
<tr>
<td>巴 applicants _ IDE</td>
<td>18</td>
<td>8.63</td>
</tr>
<tr>
<td>_ PUN 巴 applicants _ IDE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>善善 _ IDE 及 子孫</td>
<td>19</td>
<td>8.46</td>
</tr>
<tr>
<td>羽翼</td>
<td>16</td>
<td>8.46</td>
</tr>
<tr>
<td>羽翼 _ IDE ， _ PUN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>中元</td>
<td>16</td>
<td>8.41</td>
</tr>
<tr>
<td>* _ PUN 中元 元 _ IDE 年 ， _</td>
<td></td>
<td></td>
</tr>
<tr>
<td>戾蝦</td>
<td>14</td>
<td>8.26</td>
</tr>
<tr>
<td>_ PUN 戾蝦 _ IDE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>舅</td>
<td>14</td>
<td>8.22</td>
</tr>
<tr>
<td>恂恂 _ IDE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>永元元 _ IDE 年 · _</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>words just after &quot;IDE&quot;</th>
<th>246</th>
<th>24.92</th>
</tr>
</thead>
<tbody>
<tr>
<td>巴 applicants _ IDE</td>
<td>7</td>
<td>9.82</td>
</tr>
<tr>
<td>業業</td>
<td>6</td>
<td>9.61</td>
</tr>
<tr>
<td>多士</td>
<td>5</td>
<td>9.34</td>
</tr>
<tr>
<td>之氣</td>
<td>4</td>
<td>9.03</td>
</tr>
<tr>
<td>冀冀</td>
<td>4</td>
<td>8.99</td>
</tr>
<tr>
<td>嘯嬌</td>
<td>4</td>
<td>8.98</td>
</tr>
<tr>
<td>然</td>
<td>3</td>
<td>8.63</td>
</tr>
<tr>
<td>白駒</td>
<td>3</td>
<td>8.62</td>
</tr>
<tr>
<td>不得意</td>
<td>3</td>
<td>8.62</td>
</tr>
<tr>
<td>之志</td>
<td>3</td>
<td>8.62</td>
</tr>
<tr>
<td>戾蝦</td>
<td>3</td>
<td>8.58</td>
</tr>
<tr>
<td>巴 applicants _ IDE</td>
<td>3</td>
<td>8.57</td>
</tr>
<tr>
<td>中元 元 _ IDE 年 · _ PUN</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>words after &quot;IDE&quot;</th>
<th>1,039</th>
<th>105.27</th>
</tr>
</thead>
<tbody>
<tr>
<td>巴 applicants _ IDE</td>
<td>7</td>
<td>7.75</td>
</tr>
<tr>
<td>子孫</td>
<td>7</td>
<td>7.67</td>
</tr>
<tr>
<td>齊齊</td>
<td>6</td>
<td>7.54</td>
</tr>
<tr>
<td>多士</td>
<td>5</td>
<td>7.28</td>
</tr>
<tr>
<td>冀冀</td>
<td>5</td>
<td>7.26</td>
</tr>
<tr>
<td>然</td>
<td>6</td>
<td>7.24</td>
</tr>
<tr>
<td>按</td>
<td>11</td>
<td>7.20</td>
</tr>
<tr>
<td>若</td>
<td>8</td>
<td>7.18</td>
</tr>
<tr>
<td>齊齊</td>
<td>4</td>
<td>6.97</td>
</tr>
<tr>
<td>之氣</td>
<td>4</td>
<td>6.96</td>
</tr>
<tr>
<td>年</td>
<td>33</td>
<td>6.95</td>
</tr>
<tr>
<td>嘯嬌</td>
<td>4</td>
<td>6.95</td>
</tr>
<tr>
<td>知</td>
<td>6</td>
<td>6.78</td>
</tr>
</tbody>
</table>
For e.g. Hou Hanshu we can see the following constructions:

- Ideophone + 而 ér
- Ideophone + 者 zhě
- Ideophone + 也 yě

These seem to indicate PREDICATE constructions.

- Ideophone + 之 zhī

This seems to indicate a MODIFICATION construction.

- Ideophone + 如 rú
- Ideophone + 然 rán
- Ideophone + 若 ruò

These are interesting, because in general you would not expect a space-builder (cf. Langacker 2008:272; Fauconnier 1994) after an ideophone — they are usually depictive enough. E.g. in Hanshu there are 337 ideophones in clause-final position without any particles after them.
# IDEOPHONES FOLLOWED BY SPACEBUILDERS

<table>
<thead>
<tr>
<th></th>
<th>Shiji</th>
<th>Hanshu</th>
<th>Sanguozhi</th>
<th>HouHanshu</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>IDE如Compl.</strong></td>
<td>1</td>
<td>1</td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>IDE如O</strong></td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td><strong>IDE如O也</strong></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td><strong>IDE如也</strong></td>
<td>4</td>
<td>3</td>
<td>3</td>
<td></td>
<td>10</td>
</tr>
<tr>
<td><strong>IDE然</strong></td>
<td>4</td>
<td>4</td>
<td></td>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td><strong>IDE若Compl.</strong></td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>IDE若O</strong></td>
<td>1</td>
<td>3</td>
<td></td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td><strong>IDE若O也</strong></td>
<td></td>
<td>2</td>
<td></td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td><strong>IDE若此</strong></td>
<td></td>
<td>2</td>
<td></td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td><strong>IDE焉</strong></td>
<td>5</td>
<td></td>
<td></td>
<td>1</td>
<td>6</td>
</tr>
</tbody>
</table>

**Total**

<table>
<thead>
<tr>
<th>Shiji</th>
<th>Hanshu</th>
<th>Sanguozhi</th>
<th>HouHanshu</th>
<th><strong>Total</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>16</td>
<td>2</td>
<td>16</td>
<td>43</td>
</tr>
</tbody>
</table>
ABOUT THESE PARTICLES (PULLEYBLANK 1995)

2nd person pronoun

'be like'
'be thus'

'if'

'among X'

intro
material
framework
method
case study
IDEOPHONE + 若/如/然

RUO RU RAN

• Suprising: 2nd person vs. ‘if’ and ’be like’
• Less surprising: ‘if’ and ’be like’

如欲平治天下，
如若鲁德鲁然

If [Heaven] wished to bring peace and good order to the world, in the present generation, who is there except me (to bring it about)?

(Mengzi 2B/13) (Pulleyblank 1995:150)
It is noteworthy that these particles seem to function in different degrees of schematicity as clitic-like elements to ideophones.

This may be due to the experiential nature ideophones try to convey, the perceptions these concepts attempt to recreate (Dingemanse 2016).

That these particles occur after certain words has been noted before (e.g. Pulleyblank 1995:102; Van Hoey 2015), but this connection in Classical Chinese has not been studied from a cognitive linguistics perspective before.
**IDEOPHONE + 若/如/然**

**RUO RU RAN**

**Ideophone + 若** often precedes an object or complement → there is room for an object

- 「叔度 汪汪 若 千頃陂」  
  Shu Du boundless like broad.tolerance  
  “Shu Du has broad an boundless-like tolerance.” (Houhanshu)

- 「咎-敗 灼灼 若 此」  
  misfortune-loss brilliant like this  
  “The misfortune and loss were brilliant like this.” (Hanshu)

**Ideophone + 然** does not, because ran 然 =‘如+之’ → the transitive gap has already been filled

- 「天-下 警警=然」  
  heaven-below raucous.and.rackety=like.it  
  The empire is raucous. (Hanshu)
Ideophone + 如 often has no object, only ‘pause-marker’ 也

- 「魯道衰，洙泗之間　斷斷=如也」
  Lu way decline Zhu-Si LNK between wrangle=be.like YE
  “The Way of Lu is declining, between the Zhu and Si river fights are commonplace.” (Hanshu)

- 「孔子於鄉黨，恂恂=如也」
  Confucius LOC towns-men respect.care=like YE
  “Confucius was respectful and careful toward the townsman.” (Hou Hanshu)
My hypothesis is that the spacebuilder works as a redundant clue that stresses the perceptual nature ideophones depict.

Let’s look at an example.

"The Way of Lu is declining, between the Zhu and Si river fights are commonplace.” (Hanshu)
My hypothesis is that the spacebuilder works as a redundant clue that stresses the perceptual nature ideophones depict.

Let’s look at an example.

鲁 道 衰， 洙-泗 之 間 斷斷= 如 也
Lu way decline Zhu-Si LNK between wrangle=be.like YE
“The Way of Lu is declining, between the Zhu and Si river fights are commonplace.” (Hanshu)
"The Way of Lu is declining, between the Zhu and Si river fights are commonplace."
「魯道衰，洙-泗之間，斷斷=如也」 (Hanshu)

Lu way decline  Zhu-Si LNK between wrangle=be.like  YE

“The Way of Lu is declining, between the Zhu and Si river fights are commonplace.”
“The Way of Lu is declining, between the Zhu and Si river fights are commonplace.”
The Way of Lu is declining, between the Zhu and Si river fights are commonplace.
This paper has focused mostly on the methodology and procedures for future research.

– It is possible to magnify the scope, i.e. investigate the other Standard Histories as well
– Other trends may emerge through observed time
– Construction and collocation research of ideophones
– Cognitive relevance of ideophones
– Mental space research in relation to ideophones is worth further exploration
– Digital methodology is starting to come together
CONCLUSION

intro

material

framework

method

case study

2nd person pronoun

女
nǚ
MC nrjoX
OC *nra?

汝
rú
MC nyoX
OC *ne?

爾
ěr
MC nyeX
OC *ne?

若
ruò
MC nyak
OC *nak

然
rán
MC nyen
OC *nan

焉
yān
MC jen
OC *?an

猶
yóu
MC yuw
OC *gu

由
yóu
MC yuw
OC *lu

戎
róng
MC nyuwng
OC *njŋ

乃
nǎi
MC nojX
OC *nʔə?

而
ér
MC nyi
OC *nə

然
rán
MC nyen
OC *nan

焉
yān
MC jen
OC *?an

be like
‘be thus’

‘if’

‘among X’
THANK YOU!
REFERENCES


HARVARD UNIVERSITY; ACADEMIA SINICA; and PEKING UNIVERSITY China Biographical Database. https://projects.iq.harvard.edu/cbdb.
REFERENCES


QUOTATIVES

• Japanese  
  to  と  
  quotative  
  adverbial use

• Mandarin Chinese  
  de  地  
  adverbial use

• Premodern Chinese  
  -ran/ru/ruo 然如若  
  quotative, but redundant, less frequent  
  no adverbial use