The evolution of a topic marker in Wu (Chinese) dialects

Weifeng Han
The Hong Kong Polytechnic University
cwhan@gmail.com

John Newman
University of Alberta
john.newman@ualberta.ca
Outline

- Shanghainese
- Major functions of \( zí \) (是) in modern Shanghainese
- Corpus-based study
- The evolution of \( zí \) in Shanghainese
Shanghainese

• 上海闲话  $z\ddot{a}^{23}h\varepsilon^{34}h\varepsilon^{23}h\varepsilon^{23}h\varepsilon^{23}h\varepsilon^{23}h\varepsilon^{23}h\varepsilon^{23}h\varepsilon^{23}h\varepsilon^{23}h\varepsilon^{23}h\varepsilon^{23}h\varepsilon^{23}h\varepsilon^{23}h\varepsilon^{23}h\varepsilon^{23}$
  A dialect of the Northern Wu branch of the Sinitic language family

• Spoken in the city of Shanghai and surrounding regions.

• > 10 million native speakers
Sources of data for Shanghainese

- Written record can be traced back to the Northern Song Dynasty (960-1127)
- Historical data includes novels, tales, and dramas
- Contemporary data includes Shanghai Opera, songs, TV shows, interviews, corpus etc.
Major functions of zi in contemporary Shanghainese

- Copula (COP)
- Focus marker (FOCUS)
- Topic marker (TOPIC)
zi as a copula

The Shangahinese copula 是 typically links two nominals, in an equivalence relationship, or member-class relationship (Liang 2012).

伊 是 班长
hi35 zi pɛ35zã35
he COP class monitor
‘He is the class monitor.’
As a focus marker, Shanghainese emphasizes the importance of immediately following information (cf. Shi 2005, Huang et al. 2009: 25).

是戆呀!
zi kã³⁵ hia³⁴
FOCUS stupid SFP
‘It’s really really stupid!’
zi as a topic marker

是 is one of many topic markers in Shanghainese, most typically used with given information in the topic on the left and new information in the comment on its right.

当然，阿拉是，也支持侬个。

‘Of course, we support you!’ (SSC)
The Red Pear Blossom corpus

• *HongLiJi* 红梨记  Red Pear Blossom

• Novella (42,000 words) written by Xu Fuzuo 徐复祚 in AD1560-1630 (Later Ming Dynasty).

• Altogether 391 occurrences of 是
  • 66 copula
  • 186 focus marker
  • 65 topic marker
  • 74 other uses
拘禁

紅梨記

（凡調）

明 賴復祚

訪素 紅梨記

（六調）

明 賴復祚

陰晴無準

前腔 恨煙花戶門 恨煙花戶門 風塵仆品 豈堪與王孫貴

屏風屏幌 月色雲雨腸 自生成花狂柳狂 新詞楚楚 俏名

兒姹與秋娘抗 蘇小才貌相當 呂雙雙風流不讓 拗醉佳

人錦瑟 翠屏

前腔 郊心切 旅夢長 是何人催促餓忙

衣衫停當 坐生 作坐 與秋娘抗 蘇小才貌相當 呂雙雙

風流不讓 拗醉佳

紅梨記 訪素 1
Copula zi in the corpus

小人
是
本
县
差人
ɕiɔ³⁴ɲin²³
zɿ
pən³⁴
hyo²³
tˢʰa⁵³ɲin²³
I
COP
this
county
public servant
‘I am a public servant of this county.’ (Chapter 21)
Focus marking zi in corpus

正 是 伯畴 来 了
ʦən³⁴ zi pa⁵⁵zɤ²³ ɨ²³ ɨɔ²³
EMPHMRKR FOC Bochou come SFP

‘It is Bochou that has come.’ (Chapter 19)
Topic marking \( zi \) in the corpus

As for this country, it is full of saints and sinners.’ (Chapter 13)
Statistical analysis

• We used Hierarchical Configural Frequency Analysis (HCFA, Gries 2004) to investigate the statistically significant combinations of parts of speech and functions of zi.

• We coded all examples of copula, focus marking, and topic marking zi for
  • Part of speech of constituent to the left of zi (POSLEFT)
  • Part of speech of constituent to the right of zi (POSRIGHT)
  • Function of zi (FUNCTION: equative, contrastive etc.)
### Significant HCFA results

<table>
<thead>
<tr>
<th>FUNCTION</th>
<th>POSLEFT</th>
<th>POSRIGHT</th>
<th>Freq</th>
<th>Exp</th>
<th>P.adj.Holm</th>
<th>Dec</th>
<th>Q</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUBORD</td>
<td>PRON</td>
<td>COMMON</td>
<td>15</td>
<td>1.7241</td>
<td>2.09E-07</td>
<td>***</td>
<td>0.042</td>
</tr>
<tr>
<td>EQUAT</td>
<td>PRON</td>
<td>COMMON</td>
<td>18</td>
<td>2.8212</td>
<td>4.95E-07</td>
<td>***</td>
<td>0.048</td>
</tr>
<tr>
<td>EQUAT</td>
<td>PRON</td>
<td>PROPER</td>
<td>10</td>
<td>0.8125</td>
<td>7.36E-06</td>
<td>***</td>
<td>0.029</td>
</tr>
<tr>
<td>EMPHRIGHT</td>
<td>EMPHMRKR</td>
<td>CLAUSE</td>
<td>69</td>
<td>35.6417</td>
<td>3.05E-05</td>
<td>***</td>
<td>0.119</td>
</tr>
<tr>
<td>SUBORD</td>
<td>PROPER</td>
<td>COMMON</td>
<td>7</td>
<td>0.3831</td>
<td>8.11E-05</td>
<td>***</td>
<td>0.021</td>
</tr>
<tr>
<td>COUNTER</td>
<td>ADV</td>
<td>PROPER</td>
<td>4</td>
<td>0.1236</td>
<td>0.004301176</td>
<td>**</td>
<td>0.012</td>
</tr>
<tr>
<td>INTRORIGHT</td>
<td>COMMON</td>
<td>CLAUSE</td>
<td>6</td>
<td>0.5364</td>
<td>0.009981963</td>
<td>**</td>
<td>0.017</td>
</tr>
<tr>
<td>CONTLEFT</td>
<td>ADJ</td>
<td>ADJ</td>
<td>2</td>
<td>0.0075</td>
<td>0.013894591</td>
<td>*</td>
<td>0.006</td>
</tr>
</tbody>
</table>
Significant HCFA results: copula uses

<table>
<thead>
<tr>
<th>FUNCTION</th>
<th>POSLEFT</th>
<th>POSRIGHT</th>
<th>Freq</th>
<th>Exp</th>
<th>P.adj.Holm</th>
<th>Dec</th>
<th>Q</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUBORD</td>
<td>PRON</td>
<td>COMMON</td>
<td>15</td>
<td>1.7241</td>
<td>2.09E-07</td>
<td>***</td>
<td>0.042</td>
</tr>
<tr>
<td>EQUAT</td>
<td>PRON</td>
<td>COMMON</td>
<td>18</td>
<td>2.8212</td>
<td>4.95E-07</td>
<td>***</td>
<td>0.048</td>
</tr>
<tr>
<td>EQUAT</td>
<td>PRON</td>
<td>PROPER</td>
<td>10</td>
<td>0.8125</td>
<td>7.36E-06</td>
<td>***</td>
<td>0.029</td>
</tr>
<tr>
<td>EMPHRIGHT</td>
<td>EMPHMRKR</td>
<td>CLAUSE</td>
<td>69</td>
<td>35.6417</td>
<td>3.05E-05</td>
<td>***</td>
<td>0.119</td>
</tr>
<tr>
<td>SUBORD</td>
<td>PROPER</td>
<td>COMMON</td>
<td>7</td>
<td>0.3831</td>
<td>8.11E-05</td>
<td>***</td>
<td>0.021</td>
</tr>
<tr>
<td>COUNTER</td>
<td>ADV</td>
<td>PROPER</td>
<td>4</td>
<td>0.1236</td>
<td>0.004301176</td>
<td>**</td>
<td>0.012</td>
</tr>
<tr>
<td>INTRORIGHT</td>
<td>COMMON</td>
<td>CLAUSE</td>
<td>6</td>
<td>0.5364</td>
<td>0.009981963</td>
<td>**</td>
<td>0.017</td>
</tr>
<tr>
<td>CONTLEFT</td>
<td>ADJ</td>
<td>ADJ</td>
<td>2</td>
<td>0.0075</td>
<td>0.013894591</td>
<td>*</td>
<td>0.006</td>
</tr>
</tbody>
</table>
Significant HCFA results: focus marking

<table>
<thead>
<tr>
<th>FUNCTION</th>
<th>POSLEFT</th>
<th>POSRIGHT</th>
<th>Freq</th>
<th>Exp</th>
<th>P.adj.Holm</th>
<th>Dec</th>
<th>Q</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUBORD</td>
<td>PRON</td>
<td>COMMON</td>
<td>15</td>
<td>1.7241</td>
<td>2.09E-07</td>
<td>***</td>
<td>0.042</td>
</tr>
<tr>
<td>EQUAT</td>
<td>PRON</td>
<td>COMMON</td>
<td>18</td>
<td>2.8212</td>
<td>4.95E-07</td>
<td>***</td>
<td>0.048</td>
</tr>
<tr>
<td>EQUAT</td>
<td>PRON</td>
<td>PROPER</td>
<td>10</td>
<td>0.8125</td>
<td>7.36E-06</td>
<td>***</td>
<td>0.029</td>
</tr>
<tr>
<td>EMPHRIGHT</td>
<td>EMPHMRKR</td>
<td>CLAUSE</td>
<td>69</td>
<td>35.6417</td>
<td>3.05E-05</td>
<td>***</td>
<td>0.119</td>
</tr>
<tr>
<td>SUBORD</td>
<td>PROPER</td>
<td>COMMON</td>
<td>7</td>
<td>0.3831</td>
<td>8.11E-05</td>
<td>***</td>
<td>0.021</td>
</tr>
<tr>
<td>COUNTER</td>
<td>ADV</td>
<td>PROPER</td>
<td>4</td>
<td>0.1236</td>
<td>0.004301176</td>
<td>**</td>
<td>0.012</td>
</tr>
<tr>
<td>INTRORIGHT</td>
<td>COMMON</td>
<td>CLAUSE</td>
<td>6</td>
<td>0.5364</td>
<td>0.009981963</td>
<td>**</td>
<td>0.017</td>
</tr>
<tr>
<td>CONTLEFT</td>
<td>ADJ</td>
<td>ADJ</td>
<td>2</td>
<td>0.0075</td>
<td>0.013894591</td>
<td>*</td>
<td>0.006</td>
</tr>
</tbody>
</table>
Significant HCFA results: topic marking

<table>
<thead>
<tr>
<th>FUNCTION</th>
<th>POSLEFT</th>
<th>POSRIGHT</th>
<th>Freq</th>
<th>Exp</th>
<th>P.adj.Holm</th>
<th>Dec</th>
<th>Q</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUBORD</td>
<td>PRON</td>
<td>COMMON</td>
<td>15</td>
<td>1.7241</td>
<td>2.09E-07</td>
<td>***</td>
<td>0.042</td>
</tr>
<tr>
<td>EQUAT</td>
<td>PRON</td>
<td>COMMON</td>
<td>18</td>
<td>2.8212</td>
<td>4.95E-07</td>
<td>***</td>
<td>0.048</td>
</tr>
<tr>
<td>EQUAT</td>
<td>PRON</td>
<td>PROPER</td>
<td>10</td>
<td>0.8125</td>
<td>7.36E-06</td>
<td>***</td>
<td>0.029</td>
</tr>
<tr>
<td>EMPHRIGHT</td>
<td>EMPHMRKR</td>
<td>CLAUSE</td>
<td>69</td>
<td>35.6417</td>
<td>3.05E-05</td>
<td>***</td>
<td>0.119</td>
</tr>
<tr>
<td>SUBORD</td>
<td>PROPER</td>
<td>COMMON</td>
<td>7</td>
<td>0.3831</td>
<td>8.11E-05</td>
<td>***</td>
<td>0.021</td>
</tr>
<tr>
<td>COUNTER</td>
<td>ADV</td>
<td>PROPER</td>
<td>4</td>
<td>0.1236</td>
<td>0.004301176</td>
<td>**</td>
<td>0.012</td>
</tr>
<tr>
<td>INTRORIGHT</td>
<td>COMMON</td>
<td>CLAUSE</td>
<td>6</td>
<td>0.5364</td>
<td>0.009981963</td>
<td>**</td>
<td>0.017</td>
</tr>
<tr>
<td>CONTLEFT</td>
<td>ADJ</td>
<td>ADJ</td>
<td>2</td>
<td>0.0075</td>
<td>0.013894591</td>
<td>*</td>
<td>0.006</td>
</tr>
</tbody>
</table>
Example sentences for the significant combinations:

**EQUAT, PRON, PROPER**

我 是 王 太傅
ŋu²³  zɿ  hua̯ ³⁵ tʰa³⁴fu³⁴
I COP Wang high rank official
‘I’m Official Wang.’ (Chapter 18)
It is a fact that information has been leaked somewhere else.

(Chapter 22)
元来　是　赵相公

actually　TOPIC　Mr. Zhao

‘Actually, it is Mr. Zhao.’ (Chapter 24)
惹祸的是花容月貌。
Trouble? That's beauty in the flesh.

赚人的 是云魂雨梦。
Fascination? That's imaginary beauty.' (Chapter 4)
Shanghainese blending

COP

POSLEFT
NOM

POSRIGHT
NOM

FOCUS

POSLEFT
EMPHMRKR
CLAUSE

POSRIGHT

TOP

POSLEFT
NOM

POSRIGHT
CLAUSE
German blending  (Newman 1997)

Impersonal Construction  
*Es juckt mich*  
'I feel itchy'  

2-place GIVE Construction  
*Berge geben Trauben*  
'Mountains produce grapes'  

\[ \text{Es gibt Beulen...} \]  
'There will be lumps [on the body]'
Summary

• We draw upon attested historical data to understand earlier stages, rather than simply relying upon imagined intermediate stages as is often the case in grammaticalization studies.

• Even simple statistical methods can be very useful in identifying significant factors at work in data.

• The emergent zi topic construction inherits properties from the historically earlier copula and focus constructions and can be understood as a kind of constructional blending.
谢谢你！
zia³⁴ zia³⁴ noŋ⁵³
Thank you!
😊