Levels of structure within Chinese character constituents

James Myers
National Chung Cheng University, Taiwan
http://personal.ccu.edu.tw/~lngmyers/

Abstract

• Character constituents are like morphemes, strokes are like segments
• In between are strokes groups, which act like syllables:
  • Target of stress-like prominence
  • Onset-nucleus-coda-like internal structure
  • Compete for space in accordance with Menzerath’s law

Thanks to Taiwan’s Ministry of Science and Technology (MOST 103-2410-H-194-119-MY3), my lab assistants, and anonymous reviewers. I absolve them of all responsibility for errors.
Levels of Chinese character structure

Constituents are psychologically real (Lee 2017; Li & Zhou 2007; Prün 1994), even if uninterpreted (Chen & Cherng 2013), like morphemes (Myers 2019)

Strokes are also psychologically real (Bohn 1998; Sze et al. 2014; Wang et al. 2020) and have distinctive features (Peng 2017; Wang 1983), like segments (Myers 2019)

But even when not separated, the stroke group may behave as a distinct level of representation

* Examples here are traditional, but simplified system works virtually the same (Myers 2019)

** History no guide to modern system: formally related 月朋服青  formerly distinct 月服朋青
Stroke groups as “prosodic” units

• Character prosody (Myers 2019)
  • Template for position-based patterns

• Reduplication of constituents
  林 lín ‘forest’ 多 duō ‘many’ 蟲 chóng ‘insects’

• Curving of strokes
  • Leftmost position, especially in tall, narrow constituents (Wang 1983)

• Prominence (“stress”)
  • Bottommost (and rightmost) constituent, stroke, and ...???

Foot-like prosodic template, with Weak vs. Strong slots
What receives prominence?

• Like syllables, stroke groups form a representation parallel to constituents (morphemes) and strokes (segments)

‘eating’ (N. American English)

\[ \text{constituents} \quad \text{prosodic templates} \quad \text{stroke groups} \quad \text{strokes} \]

'official'

\[ \text{guān} \]
Internal structure of stroke groups

• Stroke combinations favor certain interactions over others (as illustrated in two- & three-stroke characters; cf. Myers 2019)

| No contact | 二三八小川么儿刁ティ乞与凡寸又弋勺亡刃 |
| Cross      | 十乂七力九丸又又也也井寸弋子孑孓千干于才大尤丈女巾中毛土士* |
| Chain      | 了孑孑孓丫 |
| Start at contact (τ, IENTATION) | 丁刁下ティ千干于才大尤卜人久入刀刃勺万乃厂几凡亡毛工上又叉口口尸己已弓夕巾乞匕与 |
| End at contact | 上土士エ口山中口口巴己已尸么夕弓丫 |

• Contact at stroke start (its top/left point) is also seen when children copy simple line drawings (Ninio & Lieblich 1976)

• This is similar to coordination of gestures at syllable onsets (Browman & Goldstein 1988), as well as to favoring of onsets and disfavoring of codas (Prince & Smolensky 2004)

* Exceptional topmost prominence (see Myers 2019)
Structure, prominence, and curving

- Start on contact (least marked) = Onset-Nucleus: 丁卜人
  - Also complex strokes: ON: adamente 和 chains: ON+ON: 了
- No contact (most marked) = N (+ N + ...): 一二三八小川
- Cross (unmarked) = NN: 十芒
  - Unlike start contact, crossed strokes share location:
- End contact (bounded) = ... NCoda
  - Box bottom stroke is not prominent: □ = ONNC (一匂一)
- End contact (unbounded) = ambisyllabic C+N: エリ
  - Prominence shows contactee is also a nucleus:
- Curving = ambisyllabic N+O: 刃刃
  - Width effect on curving shows | in templatic slot, so it’s a nucleus
- Each stroke interaction forms a separate stroke group:
  Π = ON_{curv}+ON  十 = NN_{curv}+NN_{prom}  日 = ONNCC (一匂一)
Competing for space

• The more Xs, the simpler their mean complexity Y
  • Menzerath-Altmann law: \( y = ax^b, b < 0 \) (Altmann 1980)
  • Applies to strokes in constituents (Bohn 1998)
  • Applies to constituents in characters (Prün 1994)
  • Suggests that strokes and constituents are genuine levels

• Stroke groups seem to be genuine for the same reason

Some three-stroke characters

<table>
<thead>
<tr>
<th>Stroke groups</th>
<th>Structure</th>
<th>Mean group complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ONNC</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>ONC+NC</td>
<td>2.5</td>
</tr>
<tr>
<td>3</td>
<td>N+ON+NN</td>
<td>1.67</td>
</tr>
<tr>
<td>4</td>
<td>N+NN+ON+N</td>
<td>1.5</td>
</tr>
</tbody>
</table>

All three-stroke characters

- \( a = 3.92, b = -0.76, p < .02, D = .97, n = 50 \)
Open questions

• Can all constituents be analyzed consistently?
  • Same or different structures? 入 vs. 入 命 vs. 命
  • Scaling up? 龟 = ???

• Is any of this psychologically real?
  • Reduplication, prominence and curving are (Myers 2019)
  • For stroke groups, experimental evidence is still limited

• How far should the syllable analogy be taken?
  • Sign languages also seem to have syllables (Sandler 2008)
  • Or is sign structure more like that of segments (Channon 2002)?

• What about other writing systems?
  • Alphabetic writing also has syllables (Fuhrhop et al. 2011) and stress feet (Evertz 2018), but they directly interact with speech

• What do you think?
References