What Chinese character structure tells us about reduplicative prosody

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

SinFonIJA 12 2019/9/13 Brno

Thanks!

- Grants from Taiwan's National Science Council / Ministry of Science and Technology:
  - NSC 97-2410-H-194-067-MY3, NSC101-2410-H-194-115-MY3, MOST 103-2410-H-194-119-MY3
- Many lab assistants over the years:
  - Ko Yuguang, Yang Chentsung, Hsu Chiung-Wen, Hong Guo-Ming, Hsu Zi-Ping, Du Pei-Fen, Su Yu-Ting, Liu Yu-Jay, Chuang Wei-Chiang, Hsieh Yu-Yi, Pan Hsiao-Yin, Su Si-Qi, Wang Han-Te, …
- Jane Tsay, Chen Tsung-Ying, Niina Zhang…
- … and especially the anonymous reviewer who wrote: "I'm convinced that the argument is absolutely crazy"

The argument

- Applying "grammar" to script systems is reasonable
  - Yes, really!
- Chinese character "morphology" includes the productive operation of "reduplication"
  - Functionally and formally like reduplication in spoken and signed languages
  - Can be nontrivially analyzed in Optimality Theory
- Such findings support a broader view of the human capacity for grammar

Writing system "grammar"

- Who better to discuss writing systems than linguists (e.g., Sampson, 2015)?
- Writing systems are not entirely unnatural
  - Evolve through use, not only diktat (e.g., Crystal, 2012)
  - Learned implicitly, not only through overt teaching (Pacton *et al.*, 2001; Tsai & Nunes, 2003)
- Reading/writing depends on systematic knowledge
  - More than just phoneme-grapheme correspondences (e.g., Venetzky, 1967)
  - Writing systems even show specific grammar-like properties, including "visual prosody" (e.g., Evertz, 2018)

"Prosody" in English spelling (1/2)

- Building on McCawley (1994) & Evertz (2018):
  - FootBinarity$_{CONTENT}$: <bee> vs. <be>, <inn> vs. <in>, …
  - *<y> » *<i>: <him>, <hide>, … vs. <hymn>, <Hyde>, …
  - *<i>]$_{WD}$ » *<y>]$_{WD}$: <cry>~<cried>, <merry>~<merriment>, …
  - OCP (*<ii>, *<aa>, …): <Tarski> ~ <Tarskyite>, <cry> ~ <crying>, …
- Analysis of <die>, <dying>, <died>

<dy>	FtBin	Dep	OCP	*<i>]$_{WD}$	*<y>	Ident
<dy>	*				*	
☞ <die>		*				*
<di>	*			*		*

"Prosody" in English spelling (2/2)

- Analysis of <die>, <dying>, <died> (cont'd)

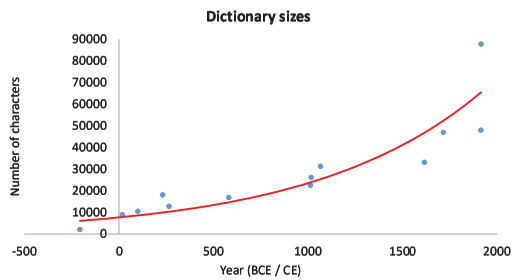
<dy>+<ing>	FtBin	Dep	OCP	*<i>]$_{WD}$	*<y>	Ident
☞ <dying>					*	
<dieing>		*				*
<diing>			*			*

<dy>+<ed>	FtBin	Dep	OCP	*<i>]$_{WD}$	*<y>	Ident
<dyed>					*	
<dieed>		*	*			*
☞ <died>						*

Chinese character "morphology"

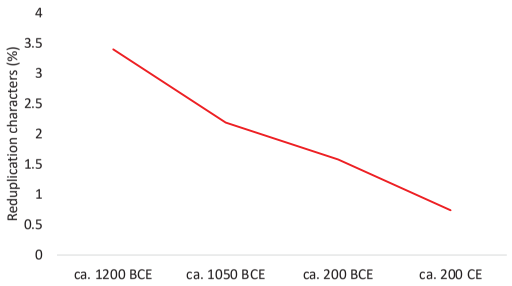
- Despite the myths, characters are not iconic
  - Xiao & Treiman (2012) found at most 15 characters with meanings guessable by non-Chinese readers
- Over 90% are semantic-phonetic characters
  - 媽 *mā* 'mother' 女 *nǚ* 'female' 馬 *mǎ* 'horse'
- Virtually all of the rest are semantic compounds
  - 明 *míng* 'bright' 日 *rì* 'sun' 月 *yuè* 'moon'
- A subset of semantic compounds involves copying
  - 木 *mù* 'wood, tree' 林 *lín* 'forest' 森 *sēn* 'full of trees'
- Claim: Copying is morphological "reduplication"

Chinese character productivity



(based on data in Yip, 2000, p. 19)

Reduplication productivity (1)



(derived from Liu 2008, p. 31, Table 2.2)

Reduplication semantics

- Reduplication itself has meanings…
  - … the same ones as in spoken languages (Behr, 2006)
- Plurality and abundance
  - 多 *duō* 'many' 品 *pǐn* 'all sorts' 蟲 *chóng* 'insects'
- Intensity
  - 晶 *jīng* 'glittering' 炎 *yán* 'blazing'
- Attenuation
  - 弱 *ruò* 'weak'
- Missing is anything related to verbal aspect
  - No marking of rapidity (Behr, 2006)
  - Nor of event structure (e.g., sign languages; Wilbur, 2009)

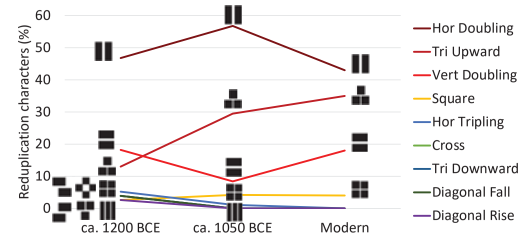
10

Reduplication shapes

- Only three shapes dominate
  - Horizontal doubling: 比 林 朋 弱 …
  - Vertical doubling: 多 哥 昌 炎 …
  - Upward triangle: 品 森 蟲 晶 …
- Other shapes are very rare
  - Square (only first is used much): 叕 茻 燚 朤
  - Horizontal tripling (only as part of constituent): 龠 壼
- Formally similar to spoken/signed reduplication
  - Abstract "templates"
  - Binary (albeit along both dimensions, with symmetry)

11

Reduplication productivity (2)



(data from Behr, 2006, pp. 88-96; Liu, 2008, pp. 16-19; Tsai, 2006)

- Simplified PRC characters may use dummy fillers
  - 双 (< 雙) 聂 (< 聶) 轰 (< 轟)

12

Reduplication productivity (3a)

Sample test items.

	Lexical grammatical	Lexical ungrammatical	Nonlexical grammatical	Nonlexical ungrammatical
Horizontal	林	[illegible]	[illegible]	[illegible]
Reduplication Element	林 *lín* 'woods' 木 *mù* 'wood'		NA 支 *zhī* 'branch'	
Vertical	侈	[illegible]	[illegible]	[illegible]
Reduplication Element	多 *duō* 'more' 夕 *xì* 'evening'		NA 夫 *fū* 'husband'	
Triangular	媪	[illegible]	[illegible]	[illegible]
Reduplication Element	晶 *jīng* 'crystal' 日 *rì* 'sun, day'		NA 欠 *qiàn* 'owe'	

(from Myers, 2016, p. 129, Table 1)

13

Reduplication productivity (3b)

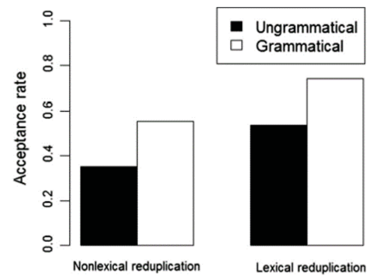
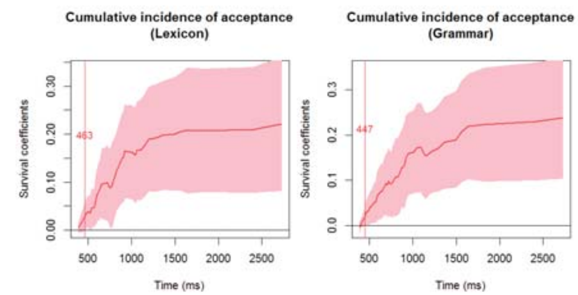


Fig. 1. Acceptance rates for fake characters containing lexical/nonlexical and grammatical/ungrammatical reduplicative configurations.

(from Myers, 2016, p. 129)

14

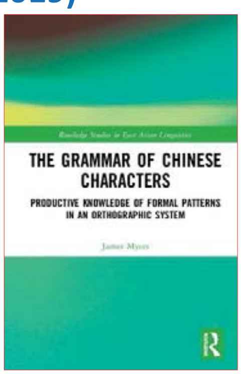
Reduplication productivity (3c)



(dynamic survival analysis [Martinussen & Scheike, 2006]; plot after Myers, 2019, p. 189, Figure 5.6)

15

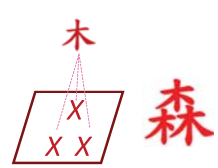
(Myers, 2019)



16

Template filling

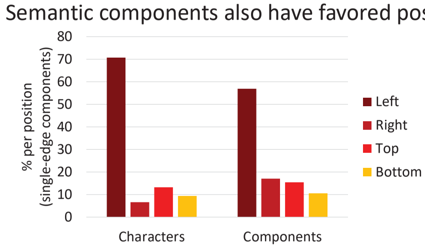
- Kinande (Mutaka & Hyman, 1990)
  - o-ku-gulu 'leg'   o-ku-**gulu**-gulu 'a real leg'
  - e-n-dwa 'wedding'   e-**n-dwa-n-dwa**-n-dwa 'a real wedding'
- Child speech (Dressler *et al.*, 2005)
  - **bebe** 'bear' (< *Bär* [German])
- Sign languages (Wilbur, 2009)
  - **SIT-SIT** 'chair' (ASL)
- Chinese characters:
  - "Prosodically" constrained morphological reduplication (cf. Inkelas, 2008)



17

Morphology and prosody

- The semantic component in a semantic-phonetic character has several affix-like properties
  - Bound, closed class, abstract function, … (Myers, 2019)
- Semantic components also have favored positions



18

Regular reduction

- Semantic components often reduce
  - Shrinking at left and top: 馬 in 駐 vs. 媽 大 in 奇 vs. 奕
  - Diagonalization/dotting: 土 in 地 vs. 型 木 in 村 vs. 李
- But all constituent types show these regular changes
  - E.g. phonetic components: 且 in 助 vs. 宜 禾 in 和 vs. 銖
- Reduplication also shows regular reduction
  - 生 in 甡 木 in 林
  - 金 in 鑫 木 in 森
- Motivated by stroke order, but not reducible to it
  - Reduces stroke distance when writing
  - Even in mechanical printing: 土 ~ 地 土 ~ 地 土 ~ 地

19

Prosodic structure

- Reduction is avoided at right and bottom
  - That's at end of left-to-right and top-to-bottom gestures
  - Gesture-final emphasis (e.g., Wann & Nimmo-Smith, 1991)
- Proposal
  - Right and bottom are "strong"
  - All other positions are "weak"

20

Deriving reduplication templates (1/4)

- Binarity
  - $*XXX$, $*\begin{matrix}X\\X\\X\end{matrix}$
- Headedness
  - Head(GestureEnd) » Head(GestureStart): $*Xx$, $*\begin{matrix}X\\x\end{matrix}$
- Symmetry
  - Symmetry(LeftRight) » Symmetry(TopBottom)
  - May also explain $\left\{XX, \begin{matrix}X\\XX\end{matrix}\right\} > \begin{matrix}X\\X\end{matrix}$
  - Common across writing systems (Morin, 2018)… but also found in sign languages (Eccarius & Brentari, 2007)

21

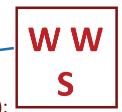
Deriving reduplication templates (2/4)

- Binarity
  - $*XXX$, $*\begin{matrix}X\\X\\X\end{matrix}$
- Headedness
  - Head(GestureEnd) » Head(GestureStart): $*Xx$, $*\begin{matrix}X\\x\end{matrix}$
- Symmetry
  - Symmetry(LeftRight) » Symmetry(TopBottom)
  - May also explain $\left\{XX, \begin{matrix}X\\XX\end{matrix}\right\} > \begin{matrix}X\\X\end{matrix}$
  - Common across writing systems (Morin, 2018)… but also found in sign languages (Eccarius & Brentari, 2007)

22

Deriving reduplication templates (3/4)

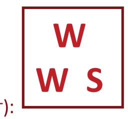
- Binarity
  - $*XXX$, $*\begin{matrix}X\\X\\X\end{matrix}$
- Headedness
  - Head(GestureEnd) » Head(GestureStart): $*Xx$, $*\begin{matrix}X\\x\end{matrix}$
- Symmetry
  - Symmetry(LeftRight) » Symmetry(TopBottom)
  - May also explain $\left\{XX, \begin{matrix}X\\XX\end{matrix}\right\} > \begin{matrix}X\\X\end{matrix}$
  - Common across writing systems (Morin, 2018)… but also found in sign languages (Eccarius & Brentari, 2007)



23

Deriving reduplication templates (4/4)

- Binarity
  - $*XXX$, $*\begin{matrix}X\\X\\X\end{matrix}$
- Headedness
  - Head(GestureEnd) » Head(GestureStart): $*Xx$, $*\begin{matrix}X\\x\end{matrix}$
- Symmetry
  - Symmetry(LeftRight) » Symmetry(TopBottom)
  - May also explain $\left\{XX, \begin{matrix}X\\XX\end{matrix}\right\} > \begin{matrix}X\\X\end{matrix}$
  - Common across writing systems (Morin, 2018)… but also found in sign languages (Eccarius & Brentari, 2007)



24

Copying beyond the template

- Square-shaped copying is due to compounding
  - Rare: essentially only 叕
  - Can be split left/right: 爽 爾 傘 㗊 蕊
  - Can be split top/bottom: 齒 繼 器 罝
  - Non-identical compounding: 琵 盥 蠶 蓏 棘
- No splitting of vertical & triangular redup: $*\frac{X}{X}$ $*\frac{X}{XX}$
- Splitting of horizontal copying is attested, but…
  - Only half of the examples show identity: 弼 㸚 韡 曲
  - The rest show idiosyncratic allomorphy:
    - 獄 (犭~犬) 班 班 (王~玉) 辮 (糸~糸)
  - Or curving: 辦 辯 辮 瓣 (辛)
    - … which can also be idiosyncratic allomorphy: 辛 in 辣 vs. 辟

25

Idiosyncratic allomorphy

- Common in semantic-phonetic characters
  - 水 'water' in 清 *qīng* 'clear' (青 *qīng*)
  - 火 'fire' in 照 *zhào* 'shine' (昭 *zhāo*)
- Also in compounds, but mostly at left (default position)
  - 水 'water' in 沙 *shā* 'sand' (少 'lacking')
  - 火 'fire' in 焚 *fén* 'burn' (林 'forest')
- Does not occur in reduplication at all
  - 水 'water' in 林 *zhuǐ* 'confluence of rivers'
  - 火 'fire' in 炎 *yán* 'blazing'
- Its association with splitting thus shows that true templatic reduplication cannot be split

26

Putting it all together

- Using familiar OT devices…
  - $\text{Reduction}_{\text{Weak}}$: Favor smaller size, fewer strokes, shorter distance
  - Allomorphy (parochial): *[水 » *[氵 *[犬 » *[犭
  - Ident-BR: Assuming strong (full) form is the base
  - Reduction » Ident-BR » Allomorphy: Underapplication of allomorphy
  - {Dep-IO, Max-IO} » Bin: Doubling obligatory only in reduplication

RED + 水	Dep-IO	Max-IO	Bin	Reduction	Ident-BR	Allomorphy
沝			*			*
氺水				*		*
☞ 林					*	*
沐					**	

27

Allomorphy with splitting

- Splitting is affixation or compounding, not reduplication
  - So idiosyncratic allomorphy is allowed
  - And binarity is not required

犬+言+犬	Dep-IO	Max-IO	Bin	Reduction	Ident-BR	Allomorphy
㹴			*	*		*
㹴			*			*
☞ 獄			*			
狀		*				

Conclusions

- Reduplication in Chinese characters is functionally and formally like that in spoken and signed languages
- Explaining this and other aspects of orthographic grammar seems like a worthwhile job for linguistic theory and cognitive science

Appendix 1: Reduplication is special

- Reduplication of novel forms cannot be handled by superficial analogy (Berent *et al.*, 2014)
  - {AA, BB} does not imply CC at a merely segmental level
- Reduplicative templates may help non-syntactic grammar remain "regular" (in mathematical sense)
  - Phonology (Bird & Ellison, 1994) and writing (Sproat, 2000) can be handled by finite state machines (FSM)
  - Sproat (p. 36): 鱗 *lín* 'fish scale' 魚 → [米 ⁺ [夕 → 牛]]
  - But FSMs can't count: 林 *木 → 木 → 木
  - "Templatic" operation as solution: 木|[ws]
- Thus finding it in script is especially intriguing

Appendix 2: Explaining character grammar

- Modality-free grammar (e.g., Primus, 2003)?
  - Same as in spoken and signed languages
- General cognitive constraints?
  - E.g. prosodic/segmental divide derives from coarse- vs. fine-grained processing (cf. Yamaguchi *et al.*, 2000)

- Emergent "grammar" (e.g., Abler, 2005)
  - Also seen in chemistry, genetics, etc…
  - E.g. binarity in electric charge, wave oscillations, …

References (1/4)

Abler, W. L. (2005). *Structure of matter, structure of mind*. Sofia: Pensoft.

Behr, W. (2006). Homosomatic juxtaposition and the problem of 'syssymantic' (*huiyi*) characters. In F. Bottéro & R. Djamouri (Eds.), *Écriture chinoise: données, usages et représentations* (pp. 75-114). Paris: École des hautes études en sciences sociales, Centre de recherches linguistiques sur l'Asie orientale.

Berent, I., Dupuis, A. & Brentari, D. (2014). Phonological reduplication in sign language: Rules rule. *Frontiers in Psychology, 5*, Article 560.

Bird, S. & Ellison, T.M. (1994.) One-level phonology: Autosegmental representations and rules as finite automata. *Computational Linguistics, 20*(1), 55-90.

Crystal, D. (2012). *Spell it out: The singular story of English spelling*. London: Profile Books.

Dressler, W. U., Dziubalska-Kolaczyk, K., Gagarina, N., & Kilani-Schoch, M. (2005). Reduplication in child language. B. Hurch (Ed.) *Studies on reduplication* (pp. 455-474). Berlin: De Gruyter Mouton.

References (2/4)

Eccarius, P., & Brentari, D. (2007). Symmetry and dominance: A cross-linguistic study of signs and classifier constructions. *Lingua, 117*(7), 1169-1201.

Evertz, M. (2018). *Visual prosody: The graphematic foot in English and German*. Berlin: Walter de Gruyter.

Inkelas, S. (2008). The dual theory of reduplication. *Linguistics, 46*(2), 351-401.

Liu, M. (2008). *Diétǐzì de lìshí kǎochá yǔ rènzhī bǐjiào yánjiū* [The diachronic study and cognitive comparison of stacked characters]. Ph.D. thesis, Tsinghua University, Beijing.

Martinussen, T. & Scheike, T. H. (2006). *Dynamic regression models for survival data*. Berlin: Springer.

McCawley, J. D. (1994). Some graphotactic constraints. In W.C. Watt (Ed.), *Writing systems and cognition: Perspectives from psychology, physiology, linguistics, and semiotics* (pp. 115-127). Berlin: Springer.

Morin, O. (2018). Spontaneous emergence of legibility in writing systems: the case of orientation anisotropy. *Cognitive Science, 42*(2), 664-677.

References (3/4)

Mutaka, N., & Hyman, L. M. (1990). Syllables and morpheme integrity in Kinande reduplication. *Phonology, 7*(1), 73-119.

Myers, J. (2019). *The grammar of Chinese characters: Productive knowledge of formal patterns in an orthographic system*. Abingdon, UK: Routledge.

Pacton, S., Perruchet, P., Fayol, M. & Cleeremans, A. (2001). Implicit learning out of the lab: The case of orthographic regularities. *Journal of Experimental Psychology: General, 130*(3), 401-426.

Primus, B. (2003). Zum Silbenbegriff in der Schrift-, Laut und Gebärdensprache - Versuch einer mediumübergreifenden Fundierung. *Zeitschrift für Sprachwissenschaft, 22*(1), 3-55.

Sampson, G. (2015). *Writing systems*, 2nd ed. Sheffield: Equinox.

Sproat, R. (2000). *A computational theory of writing systems*. Cambridge, UK: Cambridge University Press.

Tsai, C. H. (2006). Frequency and stroke counts of Chinese characters. http://technology.chtsai.org/charfreq/.

References (4/4)

Tsai, K. C. & Nunes, T. (2003). The role of character schema in learning novel Chinese characters. In C. McBride-Chang & H.-C. Chen (Eds.), *Reading development in Chinese children* (pp. l09-125). London: Praeger.

Venezky, R. L. (1967). English orthography: Its graphical structure and its relation to sound. *Reading Research Quarterly, 2*(3), 75-105.

Xiao, W. & Treiman, R. (2012). Iconicity of simple Chinese characters. *Behavior Research Methods, 44*(4), 954-960.

Wann, J. & Nimmo-Smith, I. (1991). The control of pen pressure in handwriting: A subtle point. *Human Movement Science, 10*(2-3), 223-246.

Wilbur, R. B. (2009). Productive reduplication in a fundamentally monosyllabic language. *Language Sciences, 31*, 325-342.

Yamaguchi, S., Yamagata, S. & Kobayashi, S. (2000). Cerebral asymmetry of the 'top-down' allocation of attention to global and local features. *Journal of Neuroscience, 20*(9), Article RC72.

Yip, P. C. (2000). *The Chinese lexicon: A comprehensive survey*. Abingdon, UK: Routledge.