The lexical "morphology" and "phonology" of Chinese character form
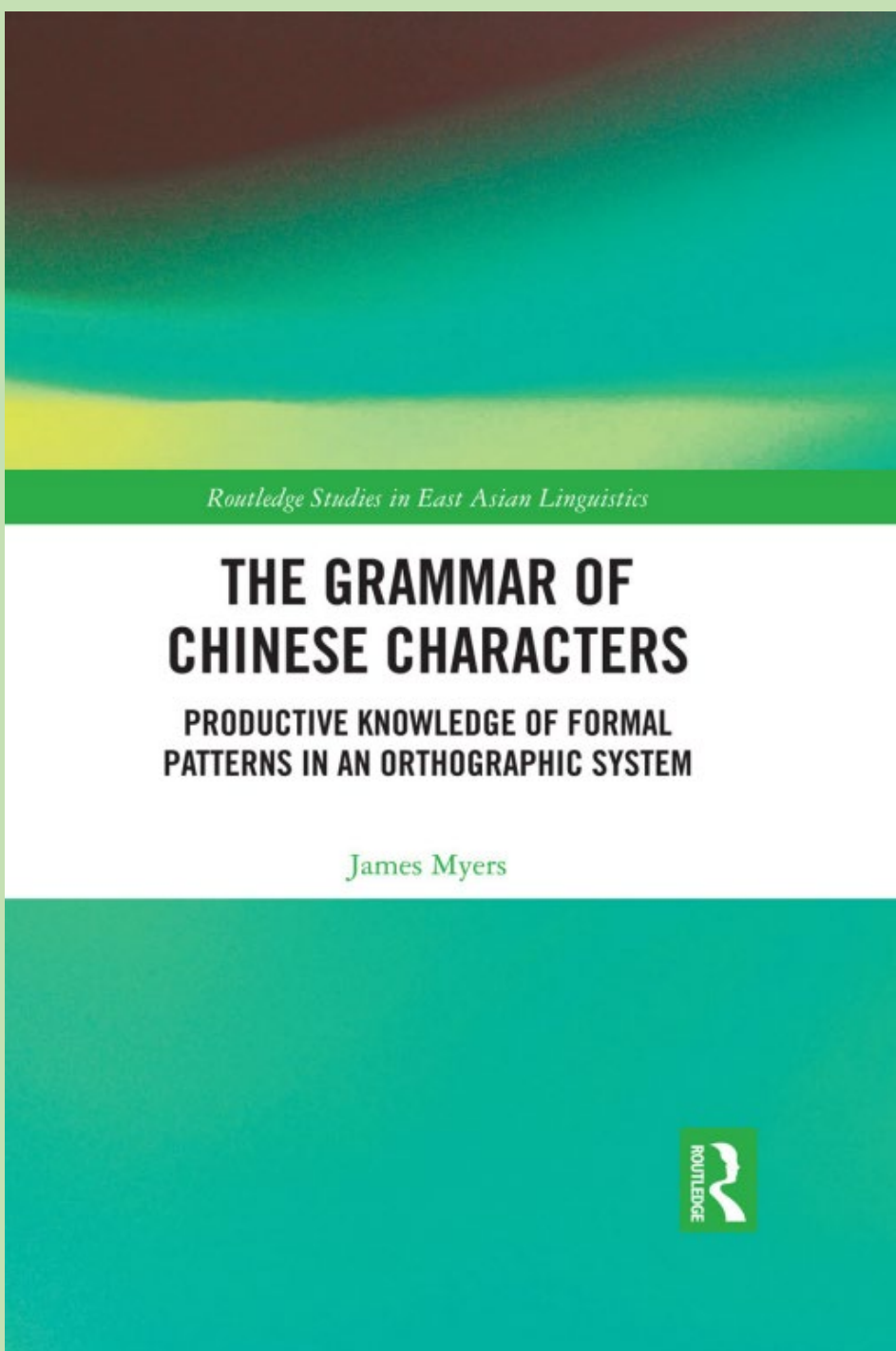
James Myers

National Chung Cheng University
Lngmyers@ccu.edu.tw

Thanks

- Many colleagues and lab assistants over the years
  - Particularly Wolfgang Behr, Zev Handel, Jane Tsay
- Funding
  - Ministry of Science and Technology (MOST) grant 109-2410-H-194-096-MY3
- Databases
  - Wenlin (文林) (https://wenlin.com) Chinese word learning software
  - Wikimedia
    - https://commons.wikimedia.org/wiki/Commons:Chinese_characters_decomposition
    - https://en.wiktionary.org/

Overview

- The story so far:
  - Chinese characters have a lexical grammar analogous to those of spoken and signed languages
  - Components and their combination are analogous to morphology
  - Regularities in stroke form are analogous to phonology

- In this episode:
  - Is this consistent with formal theories of lexical morphology and phonology?
  - Yes, because "morphological" operations in characters are ordered in strata…
  - … and because patterns in character "phonology" interact opaquely only when in different strata
  - (Future episodes: Diachronic and psycholinguistic evidence…?)

Routledge Studies in East Asian Linguistics

THE GRAMMAR OF CHINESE CHARACTERS

PRODUCTIVE KNOWLEDGE OF FORMAL PATTERNS IN AN ORTHOGRAPHIC SYSTEM

James Myers

Routledge, 2019

(also includes analyses of simplified characters and seal script)
(2021 paperback edition fixes some typos)

See also:

Myers (2016) in *Cognition*

Myers (2021a) in Haralambous (ed.) *Grapholinguistics in the 21st Century,* Fluxus Editions

Myers (2021b) "Areal script form patterns with Chinese characteristics", in *Written Language and Literacy*

And:

Wang (1983) *Toward a generative grammar of Chinese character structure and stroke order,* U. Wisconsin-Madison Ph.D. thesis

4

"Morphological" operations in characters

- Affixation (形聲)

  根 *gēn* 'root' = <u>木 'tree'</u> + 艮 *gèn*

  affix

- Compounding (會意)

  相 *xiàng* 'see' = 木 'tree' + 目 'eye'

- Reduplication (疊體)

  林 *lín* 'forest' = 木 'tree' + [XX]

  template

Affixation

- Closed class of affixes
- Abstract semantics

嗎 *ma* (sentence-final particle) = 口 'mouth' / 'function morpheme' + 馬 *mǎ*

賬 *zhàng* 'account' = 貝 'shell' / 'money' + 長 *cháng*

嫌 *xián* 'fed up with' = 女 'female' / 'bad things' + 兼 *jiān*

- Bound

Fully bound: 冫 宀 宀 癶 夊 彳

Bound allomorphs: 忄 扌 氵 艹 …

Bound when used with abstract semantics: 木 貝 口 女 …

- Relatively fixed positions

Most affixes appear in the same position in most of their characters

Compounding

- Concrete semantics

  吠 *fèi* 'bark of a dog' = 口 'mouth' (literally) + 犬 'dog'

- Fewer positional restrictions

  杲 *gǎo* 'bright' (sun over tree) vs. 杳 *yǎo* 'dark' (sun below tree)

- … as in spoken/signed compounds

  snowman 'man made of snow'   vs.   *mailman* 'man who delivers mail'
  māoxióng / xióngmāo 'panda'

Reduplication

- Iconic semantics
  … as in spoken/signed reduplication (Behr 2006 in Bottéro & Djamouri *Écriture chinoise*, Centre de recherches linguistiques sur l'Asie orientale)
  - Plurality/abundance: 多 *duō* 'many', 品 *pǐn* 'all sorts', 蟲 *chóng* 'insects'
  - Intensity: 晶 *jīng* 'glittering', 炎 *yán* 'blazing'
  - Attenuation: 弱 *ruò* 'weak'
- Fixed templates
  … as in spoken/signed reduplication (syllables, feet)
  - Standard: 叩　吕　品
  - Rare: 叕　靁

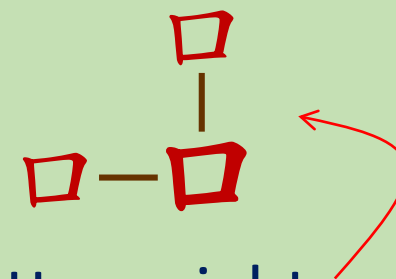# "Phonological" patterns in characters

- Idiosyncratic allomorphy

  忄（心）扌（手）氵（水）艹（艸）⋯

- Featural distributions and alternations

  川, 月　　→土 →地

  ↑　　↑
  curving　　diagonalization

- "Prosody"

  Binary branching along one or both axes, head at bottom right

Character "prosody"

- Prosodic templates
  - Same as reduplicative templates
  - E.g. for horizontal axis: [XX] = [WS] (weak-strong, strong = head)

- Prominence ("stress")
  - Components larger in right/bottom head position: 林 昌　大 in 尖 vs. 奇
  - Strokes and stroke groups larger at right/bottom: 川 三 井　由 vs. 甲

- Domain for featural regularities
  - Curving in leftmost stroke in weak position [WS]: 川 井 月 周
  - Wider = [S][S], so not necessarily curved: 門 冊 肉 同　角 vs. 甬
    (Wang 1983; Unicode 冂 vs. 刀)

10

Featural patterns

- Curving

  (see above)

- Diagonalization

  - In prosodically weak position [WS]: 土 地　　工 in 江 vs. 鴻

- Axis assimilation

  三　彡 (cf. 才 …) (see also Morin 2017 in *Cognitive Science*)

  - Only applies within components: 形

- Many others

  - Dotting (木 in 析), shrinking (雨 in 電), stretching (走 in 起), left-hooking (門), right-hooking (長), dot axis (米) …

Psychological reality

- Readers mentally activate the semantics of affixes
  Feldman & Siok (1999) in *Journal of Memory and Language*, etc

- Readers know that non-affixes (phonetic components) are open-class
  Mattingly & Hsiao (1999) in *Psychologia*

- Readers generalize reduplicative templates to novel characters
  Myers (2016) in *Cognition*

- Readers know that prominence and curving depend on position
  Myers (2019)

- Readers know that curving is sensitive to width
  Myers (2022) in NACCL-34

Theory of lexical morphology and phonology

- Kiparsky (1982) in *Linguistics in the morning calm*, Seoul

  Halle & Mohanan (1985) in *Linguistic Inquiry;* Hargus & Kaisse (eds.) (1993) *Studies in lexical phonology*, Academic Press; etc

- Morphology applies in ordered strata (levels)
  - English: Less productive operations before more productive operations

    radicalizing   [[[[[rad] ic] al]$_1$ ize]$_2$ ing]$_3$

- Strata are associated with different phonological regularities
  - English: stress on an earlier stratum, stem-final tensing on a later stratum

    órigin, oríginal$_1$, originálity$_1$ vs. *rádical$_1$, rádicalize$_2$, rádicalizing$_3$*

    happy [i] ~ *happiness$_2$* [ɪ] (dialectal) but *cities$_3$* [i]

Spoken English: Lexical morphology & phonology

	Morphology	Phonology
Stem	Less productive derivation	Stress
	More productive derivation	Stem-final tensing
Word	Inflection	…
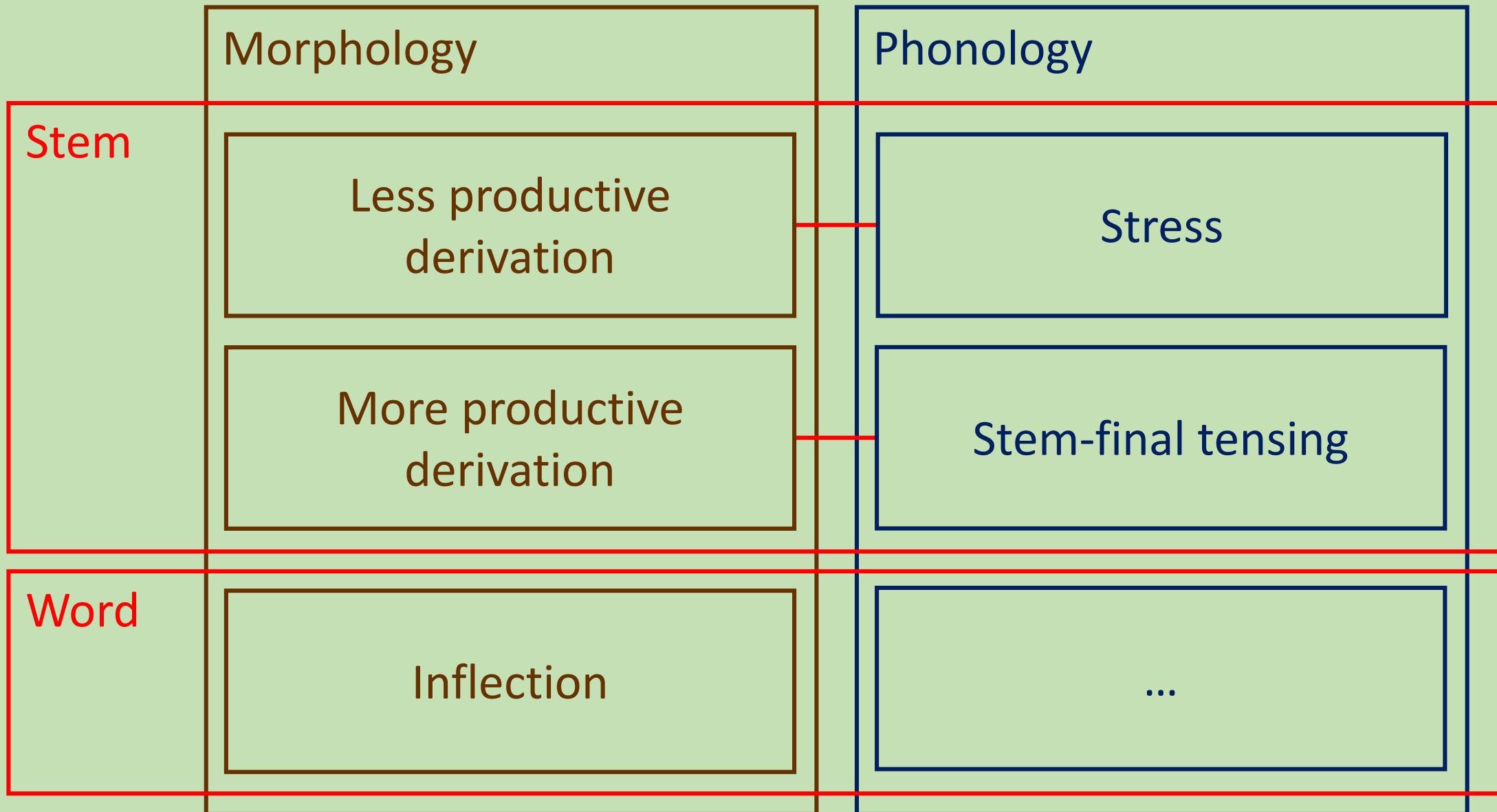
Stratal Optimality Theory

- Kiparsky (2000) *The Linguistic Review*
  - Bermúdez-Otero (2018) in Hannahs & Bosch *The Routledge handbook of phonological theory*, etc

- Universal set of ordered strata (levels): Stem, Word, Phrase

- Stem stratum
  - May consist of ordered substrata (Jaker & Kiparsky 2020 *Phonology*)

- Each stratum is an ordinary Optimality Theory (OT) grammar
  - Ranked constraints rather than ordered rules
  - Phonological interactions in ordinary OT must be transparent
  - Opaque interactions must thus apply across different strata

Stratum ordering in character morphology

- Reduplication only applies to base components (including bound)*

  多　昌　炎　圭　　　林　朋　比　弱　羽　艸　　　品　森　蟲　晶　轟

- Compounding only applies to bases (including bound) & reduplication
  - Base components: 明　相　尋
  - Reduplicated forms: 區　雙　器　替　琵
- Affixation can apply to any type of form (except bound)
  - Base components: 住　根　英
  - Reduplicated forms: 錢　琳　臨
  - Compounds: 請　想　份
  - Affixed forms: 燙　榴　眯　　(recursion)

*Three kinds of reduplication

- Base reduplication
  (see above)

- Reduplication of fossilized complex forms
  赫 *Hè* (surname): 赤 *chì* 'red' (compound) < 大 'big' + 火 'fire'
  哥 *gē* 'older brother': 可 *kě* 'able' (affixed form) < 口 'mouth' + 丂 *qiǎo*

- Ludic reduplication (outside grammar proper?)
  囍 *xǐ* 'double happiness': 喜 *xǐ* 'happiness' (compound) = 壴 'drum' + 口 'mouth'
  靊 *duì* 'cloudy': 雲 *yún* 'cloud' (affixed form) = 雨 'rain' + 云 *yún*
  棽 *yàn* (unclear meaning): reduplication of reduplicated 林

All character morphology is stem-level

- Affixation acts somewhat like inflection, since it shows "agreement"…
  (Myers 2019; cf. Handel 2019 *Sinography,* Brill)

  葡萄 *pútáo* 'grape'   蝴蝶 *húdié* 'butterfly'

- … But it's also recursive, and recursive inflection is extremely rare
  (Arista 2009 in Butler & Arista *Deconstructing constructions*, John Benjamins)

- Recursive affixation = characters are never "finished"
  - So affixation cannot be a word-level operation
  - So affixation is stem-level
  - Since affixation comes last, reduplication and compounding must also be stem-level

Interactions in character phonology

- Most patterns in character phonology do not interact at all

  川 simultaneously obeys axis assimilation, curving, and prominence

- Most others interact in a transparent way

  川 conforms to prosodic structure [WS], which "feeds" curving and prominence

|||	Prosody: [WS]	Curving (leftmost W)	Prominence (S)
☞)||ws			
|||sw	*		
|||ws		*	
)||ws			*

Interactions with reduplicative identity

- Idiosyncratic allomorphy is restricted to affixation and compounding*

  - Affixation:　　　水 → 氵 in 清 *qīng* 'clear' = 水 'water' + 青 *qīng*

  - Compounding: 　水 → 氵 in 涉 *shè* 'wade' = 水 'water' + 步 'walk'

  - Reduplication: 　水 →水 in 㴅 *zhuǐ* 'water, rivers'

- So reduplicative identity "bleeds" (blocks) idiosyncratic allomorphy
  (transparent interaction)

水 + REDUP	Prosody: [WS]	Ident-BaseRedup	IdiosynAllomorph(水)
☞ 㴅			*
氵水		*	
水水水	*		

*Affixation: common in all positions　vs.　Compounding: only common on left

Opaque interactions in reduplication

- Reduplicative identity doesn't block diagonalization

  比　甡　珏　孖　竝　鉖　龖

- Reduplicative identity also tolerates curving

  艸　卝　辡　舙

- Both are opaque "counterbleeding" interactions
  - In rule ordering, reduplicative identity would apply first
  - But ordinary OT constraint ranking cannot handle any type of opacity

- Stratal OT: Reduplicative identity must apply in an earlier stratum

Opacity and stratum ordering

- Axis assimilation also tolerates (counterbleeds) diagonalization

    土 地　　立 站　　牛 物　　工 功

- So Stratal OT must say:
  - Diagonalization applies in a stratum after axis assimilation and reduplicative identity
  - Curving must also apply in a later stratum than reduplicative identity
- This earlier stratum must be the stem stratum
  - Reduplication applies in the stem stratum
  - Axis assimilation only applies within base components (input to stem stratum)
- So curving and diagonalization must apply on the word stratum

Chinese characters: Stratal OT

	Morphology	Phonology
Stem	Reduplication ↓ Compounding ↓ Affixation	Prosody Reduplicative identity Prominence Idiosyncratic allomorphy Axis assimilation
Word	(none)	Curving Diagonalization

23

In conclusion

- It works

- So what?

  - Formal analyses raise questions about history and psychology:
  How did these patterns arise and do they affect modern readers & writers?

  - If character grammar is "real", what does this mean for the nature and scope of the human capacity for language?

- Stay tuned….

That's it.