Knowing Chinese character grammar

James Myers
National Chung Cheng University

DRAFT - July 1, 2015

Comments welcome!

Chinese characters obey regularities that are sometimes described as constituting a kind of grammar, including patterns unrelated to meaning or pronunciation. The purely formal nature of such patterns is reminiscent of the phonologies of spoken or signed languages, but also makes it unclear whether readers actually know them. To find out, we asked Chinese readers to judge the acceptability of fake characters varying both in grammaticality (obeying or violating constraints on element reduplication) and in lexicality (of the reduplicative configurations). Lexicality improved acceptability, but grammaticality did so independently as well. Acceptability was also higher for more frequent reduplicative elements, suggesting that the reduplicative configurations were decomposed. Visual complexity had no effect. Reaction times, however, showed that grammatical configurations were accepted more quickly when also lexical: knowledge of Chinese character grammar, like spoken and signed phonology, builds on lexical knowledge.

1. Introduction

Educated Chinese readers recognize around 4,000 characters by the age of 12 (Huang & Hanley, 1994); reading the entire Academia Sinica Balanced Corpus of Modern Chinese (Huang, Chen, Chen, & Chang, 1997) would require knowing around 3,000 more. What makes such feats possible is the fact that Chinese characters obey systematic regularities. Systematic regularities in syllables, words and sentences are unhesitatingly called *grammar*, and many researchers apply this notion to Chinese characters as well (Wang, 1983; Sproat, 2000; Kordek, 2013; Ladd, 2014).

The core of Chinese character grammar is recursion: characters are usually decomposable into others, which may often be decomposable in turn. Most characters (85%, as estimated by Perfetti & Tan, 1999) are decomposed, at the first step, into a semantic component (associated, at least loosely, with the meaning of the whole character) and a phonetic component (associated with the character's pronunciation, perhaps obscured through historical sound changes). If the phonetic component is decomposable, its components do not indicate meaning or pronunciation. A typical example (from Ladd, 2014, *p*. 129) is 鐳 *léi* "radium", first decomposed into the semantic component 金 *jīn* "gold, metal" and phonetic component 雷 *léi* "thunder", which in turn is decomposed into the atomic 雨 *yǔ* "rain" and 田 *tián* "field" (this paper focuses on the traditional characters still used in Taiwan).

Knowing how characters decompose into semantic and/or phonetic elements is important for character coinage (Li & Zhou, 2007), learning to read (Chan & Nunes, 1998), and mature reading (Honorof & Feldman, 2006). For example, readers are sensitive to the combinability of character components (Hsu, Tsai, Lee, & Tzeng, 2009), their typical positions (Taft, Zhu, & Peng, 1999), and whether the phonetic components are pronounced consistently across characters (Lee, Huang, Kuo, Tsai, & Tzeng, 2010).

Given that characters are ultimately used for expressing meanings and pronunciations, it is hardly surprising that readers process the character elements linked with them. Yet characters also show purely formal regularities, in the position, shape, and organization of character elements, with no effect on meaning or pronunciation. Such regularities seem to represent the second sort of patterning in Hockett's (1960) duality of patterning, analogous to phonology in spoken and signed languages (see Ladd, 2014, for critical discussion), and it is already known that some, at least, do affect processing. Thus the overall visual appearance of characters influences similarity judgments (Yeh & Li, 2002), the Visual Word Form Area is activated when discriminating semantic-plus-phonetic pseudocharacters from random jumbles of strokes (Liu et al., 2008), and writers, who need more detailed character representations than readers, are also sensitive to the elements within phonetic components (Chen & Cherng, 2013).

More sophisticated formal regularities include those constraining the shape or configuration of character components. One such is the position-specific reduction of the semantic component, which happens readily with the component on the left or the top, but rarely when it is on the right or bottom (Sproat, 2000). For example, 忙 *máng* "busy" and 忘 *wàng* "forget" contain not only the same phonetic component 亡 *wáng* "die" but also the same semantic component 心 *xīn* "heart", which appears in its reduced form 忄 at the left in 忙 but in its full form at the bottom in 忘. These patterns may have arisen from the order in which strokes are normally written, from left to right and top to bottom, putting prominence on the stroke-final position, just as phrase-final syllables tend to be longer in speech (Beckman & Edwards, 1990) and in signed languages (Sandler, 1993). However, semantic components are closed-class (belying the common term *radical*, they are more analogous to affixes than roots), so it is difficult to test whether Chinese readers have active knowledge of these patterns.

The regularity tested in this paper is another sophisticated phonology-like pattern, involving the configuration of reduplicated constituents (Kordek, 2013). Many characters contain horizontal configurations of two copies of the same element, like 朋 *péng* "friend" (cf. 月 *yuè* "moon, month"). Other reduplicative configurations are vertical, like 炎 *yán* "inflammation" (cf. 火 *huǒ* "fire"), while still others consist of three identical elements arranged in a upward-pointing triangle, like 森 *sēn* "forest" (cf. 木 *mù* "wood"). (A rare fourth type of reduplicative configuration, not studied here, forms a two-by-two square, the only common one being 叕 *zhuó* "join together", as in 輟 *chuò* "cease"; cf. 又 *yòu* "also"). Like reduplication in speech (Hurch, 2005) and signed languages (Aronoff, Meir, & Sandler, 2005), character element reduplication is partly iconic: a forest (森) contains more trees (木) than do woods (林 *lín*). Yet iconicity is often lost in modern Chinese: 朋 and 月 are totally unrelated, and 哥 *gē* "older brother" relates to 可 *kě* "may" only in pronunciation. Iconicity is even more irrelevant when reduplicative configurations serve as phonetic components.

Reduplicative configurations obey formal constraints: triangular configurations cannot point downward (two elements over one), and horizontal and vertical configurations cannot contain three elements (with the sole exception of 㗊 *líng* "raindrops", an archaic character that now only appears in 靈 *líng* "spirit"). No such restrictions apply to mere strokes (e.g., 州 *zhōu* "prefecture", 三 *sān* "three") or to combinations of distinct components (e.g., 熒 *yíng* "glimmering" shares its upper components with 營 *yíng* "camp"). These regularities are not merely formal, but reminiscent of actual phonological constraints. In particular, reduplicative configurations obey binarity, just as stress feet are much more likely to be disyllabic than trisyllabic (Gordon, 2002). The upward-pointing triangles obey binarity both horizontally and vertically, while also making the configuration "bottom-heavy", consistent with the stroke-final prominence noted earlier.

Yet it is not obvious whether such constraints, unrelated as they are to meaning or pronunciation, are part of the active knowledge of Chinese readers. To find out, we ran an acceptability judgment experiment (Myers, 2009; Kawahara, 2011; Topolinski & Strack, 2009) using speeded binary (yes/no) responses (Weskott & Fanselow, 2011), which made it possible to analyze reaction time as well as response choice. Because judgments could also be affected by rote memorization or visual complexity, we tested fake characters that crossed grammaticality (e.g., upward-pointing vs. downward-pointing triangular configurations) with lexicality (i.e., whether the grammatical version appears in real characters), and also took into account character element frequency and the number of strokes.

2. Method

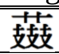
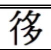
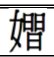
2.1 Participants

The participants were 20 university students in southern Taiwan. All were native speakers of Mandarin Chinese with normal or corrected-to-normal vision. Participants provided written consent and were paid for their participation, in accordance with the local institutional review board.

2.2 Materials and design

Sixty sets of four fake characters each were created by editing traditional characters in Microsoft MingLiU font. All characters contained reduplicated elements in a horizontal, vertical, or triangular configuration (16 sets each). The four characters in each set contained the same semantic component in its standard position, with the reduplicative configuration forming the remainder (the phonetic component in a real character) and crossing grammaticality and lexicality (see sample materials in Table 1). In grammatical configurations, reduplication obeyed the constraints discussed above. Each grammatical configuration was paired with an ungrammatical one formed of the same element, but where horizontal and vertical configurations contained three repetitions and triangular configurations formed a downward pointing triangle. Grammatical configurations were lexical if they also appear in real characters; matching ungrammatical characters were created as just described. In nonlexical configurations, the reduplicated element is never reduplicated in real characters. A three-way ANOVA on the log number of strokes showed effects of configuration shape ($F(2,180)=7.70$, $p < .001$), lexicality ($F(1,180)=4.15$, $p < .05$), and grammaticality ($F(1,180)=15.63$, $p < .001$), with an interaction between grammaticality and shape ($F(2,180)=3.85$, $p < .05$). These potential confounds with visual complexity are unavoidable because lexical reduplication, particularly vertical reduplication, favors simpler elements, and ungrammatical horizontal and vertical (but not triangular) configurations necessarily contain more strokes. Thus we included the log number of strokes as a covariate in the main analyses described below.

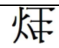
Table 1
Sample test items

	Lexical Grammatical	Lexical Ungrammatical	Nonlexical Grammatical	Nonlexical Ungrammatical
Horizontal	菻			
Reduplication	林 *lín* "woods"			
Element	木 *mù* "wood"		支 *zhī* "branch"	
Vertical				
Reduplication	多 *duō* "more"			
Element	夕 *xī* "evening"		夫 *fū* "husband"	
Triangular				
Reduplication	晶 *jīng* "crystal"			
Element	日 *rì* "sun, day"		欠 *qiàn* "owe"	

We also calculated the type frequencies of the reduplicated base elements (e.g., 木). Character components were extracted with the help of a Wikimedia resource (Chinese Characters Decomposition, 2015) that recursively decomposes 21,170 traditional and simplified characters while also providing information about reduplicative configuration shape. Type frequencies were based on just the 6,962 traditional characters in the Academia Sinica Balanced Corpus of Modern Chinese (Huang et al., 1997). Our database also showed that there are many more horizontal reduplicative configurations (450) in the lexicon than triangular (96) or vertical (110) ones.

An additional 120 fillers were created by editing other real characters (see sample materials in Table 2). The fillers, all composed of real semantic and phonetic components in their standard positions but in novel combinations, were designed to vary gradiently in acceptability: 40 had no further modification (best), 40 added or removed strokes (worse), and 40 reflected an asymmetrical element vertically or horizontally (worst).

Table 2
Sample filler items

	Novel combination	Added stroke	Removed stroke	Reflected element
Fake filler				
Real models	院 *yuàn* "court" 域 *yù* "domain"	否 *fǒu* "no" 拿 *ná* "take"	粉 *fěn* "powder" 現 *xiàn* "now"	炸 *zhà* "fry" 姓 *xìng* "surname"

Test items were divided into four lists of 60 items each in a Latin square design, so that all four participant groups saw all test item types (defined by configuration shape, grammaticality and lexicality) but never from the same matched set. All participants saw all 120 filler items.

2.3 Procedure

The experiment was run with PsychoPy v. 1.82 (Peirce, 2007, 2009). Participants were told they would see a series of characters that were not real Chinese characters. They were asked to decide if they were like or not like Chinese characters by pressing, respectively, a key on the right or left side of a computer keyboard. Trials consisted of a 500 ms display of a fixation cross at the center of the screen, followed by 500 ms of a blank screen, and finally a fake character that remained at the center of the screen for 3,000 ms or until the participant pressed one of the response keys, after which the next trial began. Characters were displayed in black on a white background, subtending approximately 4.5 degrees vertically and horizontally from a viewing distance of 50 cm. Prior to the main experiment, participants were given 10 trials of fake characters as practice; these were designed the same way as the fillers, but were not used in the main experiment. In the main experiment each participant saw 168 items (48 test items and 120 fillers) in random order. Each experimental session lasted approximately 10 minutes.

3. Results

Seventeen trials (including two test trials) were dropped for lack of a response within the time limit, leaving 958 experimental data points (3,343 including fillers). Response choices and reaction times were analyzed with mixed-effects logistic regression and mixed-effects linear regression respectively (Bates, Mächler, Bolker, & Walker, in press), with grammaticality and lexicality in effect coding and configuration shape in dummy coding with the horizontal shape as base. Unless otherwise noted, likelihood ratio tests showed that models with random participant and item intercepts but no random slopes provided sufficient fit (cf. Barr, Levy, Scheepers, & Tily, 2013). Significance was tested with the Wald test (treating t as z for the linear regressions yielded p values very close to the bootstrapped values, wherever the latter were practical to calculate).

The overall acceptance rate was .53 (virtually identical for test items and fillers). Figure 1 shows the acceptance rate of test items as a function of lexicality and grammaticality. An analysis of response choices crossing these factors with configuration shape and the log number of strokes (visual complexity) as additive covariate showed no effect of strokes ($p > .5$), but acceptability was improved both by lexicality ($B = 0.62$, $SE = 0.19$, $z = 3.35$, $p < .001$) and by grammaticality ($B = 0.66$, $SE = 0.19$, $z = 3.45$, $p < .001$). The horizontal configuration shape was significantly more acceptable than both the triangular ($B = -0.64$, SE

= 0.26, $z = 2.44$, $p < .05$) and the vertical ($B = -0.69$, $SE = 0.27$, $z = -2.56$, $p < 0.05$), consistent with their type frequencies. There were no interactions (*p*s > .3). Separate analyses for each configuration shape showed the same patterns, with no effect of strokes (*p*s > .1), but positive effects of lexicality (respectively $p < .001$, $p < .001$, $p < .05$ for horizontal, triangular, vertical) and grammaticality (respectively $p < .01$, $p < .001$, $p < .001$) and no interactions (*p*s > .1).

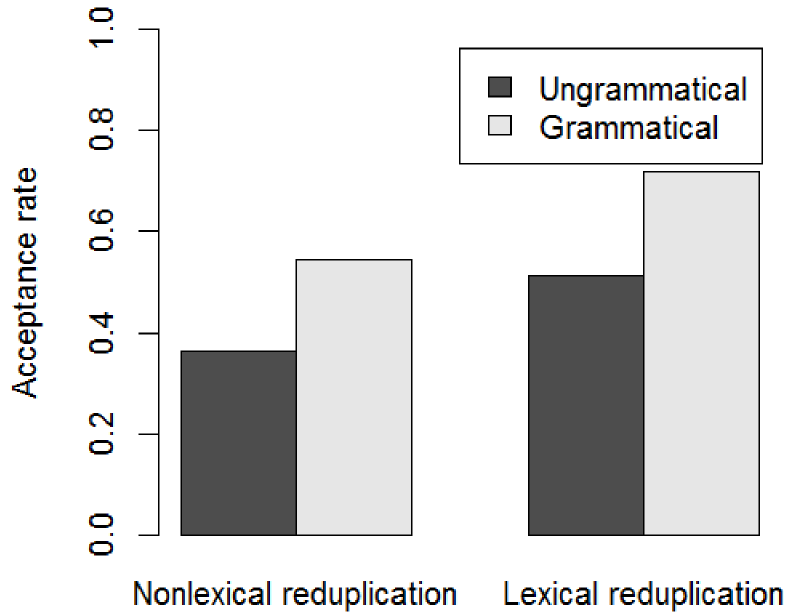


Figure 1. Acceptance rates for fake characters containing lexical/nonlexical and grammatical/ungrammatical reduplicative configurations.

When log reaction times for the test items were analyzed in terms of response choice, lexicality, and grammaticality (earlier analyses showed no effects of configuration shape or the number of strokes), the only significant results from the best-fitting model (with random participant slopes) were interactions of response choice with lexicality ($B = -0.036$, $SE = 0.01$, $t = -3.32$, $p < .001$) and with grammaticality ($B = -0.031$, $SE = 0.01$, $t = -2.73$, $p < .01$). As shown in Figure 2, these interactions involved stronger lexicality and grammaticality effects for acceptances than rejections. Analyzing just the acceptances, lexicality significantly sped up responses ($B = -0.03$, $SE = 0.01$, $t = -2.31$, $p < .05$) but grammaticality did not ($p > .1$). Moreover, unlike the case with response choices, the two factors did interact in reaction time ($B = -0.03$, $SE = 0.01$, $t = -2.36$, $p < .05$): grammatical items were responded to more quickly than ungrammatical ones if lexical configurations were involved, but the reverse was true for nonlexical configurations.

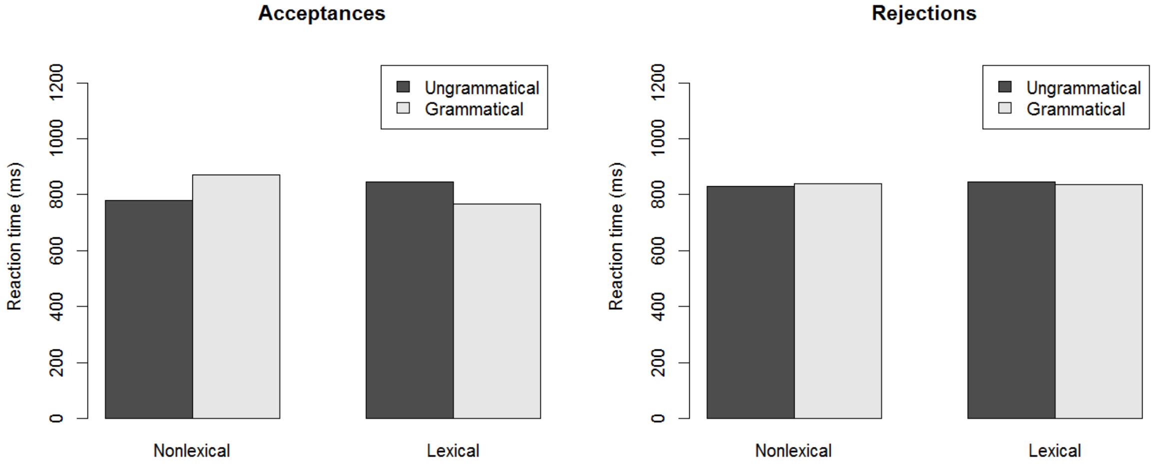


Figure 2. Reaction times for fake characters as a function of response choice, lexicality and grammaticality.

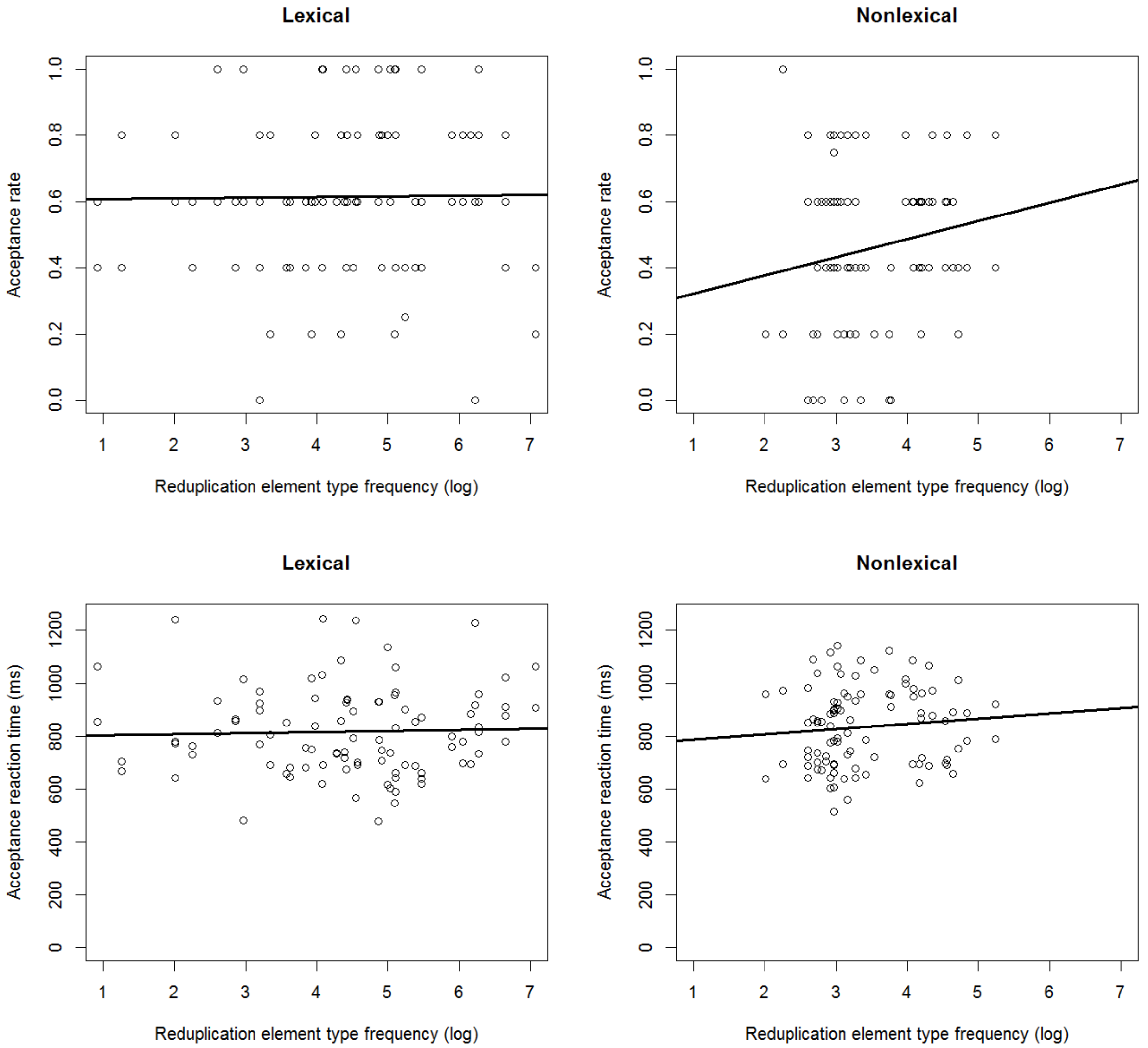


Figure 3. The effects of lexicality and element type frequency on acceptability rates (upper plots) and acceptance reaction time (lower plots).

In separate analyses, we crossed lexicality and grammaticality with the log type frequency of the test items' reduplicated elements. Frequency significantly improved acceptability (B = 0.22, SE = 0.11, z = 2.02, $p < .05$) while leaving the usual noninteracting ($p > .4$) effects of lexicality (B = 1.28, SE = 0.43, z = 2.97, $p < .01$) and grammaticality (B = 1.39, SE = 0.43, z = 3.22, $p < .01$). However, there was also a marginal interaction between frequency and lexicality (B = -0.20, SE = 0.11, z = -1.74, p = .08): the frequency effect tended to be stronger for nonlexical reduplicative configurations (see top of Figure 3). In an analysis of acceptance reaction time with the same variables, responses to grammatical forms were marginally faster (B = -0.11, SE = 0.06, t = -1.88, p = .06); a marginal interaction between lexicality and frequency also suggested that frequency slowed responses more in lexical configurations (B = -0.03, SE = 0.02, t = -1.81, p = .07) (see bottom of Figure 3).

4. Discussion

Grammaticality improved acceptability for both lexical and nonlexical reduplicative configurations of all three shapes, indeed to the same degree regardless of shape and lexical status (and visual complexity had no effect at all). Moreover, element type frequency independently improved acceptability, suggesting that readers not only know the reduplication constraints, but actively decompose reduplicative configurations into their elements, particularly (in a marginally significant interaction) the nonlexical ones.

Nevertheless, the reaction times also revealed that acceptability judgments were filtered through a processing system better at judging lexicality than grammaticality: lexical grammatical configurations were accepted more quickly than ungrammatical ones, but the reverse was true for nonlexical configurations. Decomposing nonlexical configurations even seemed to slow acceptance, judging by the marginally significant interaction between lexicality and element type frequency.

Should we dignify knowledge of such formal orthographic patterns with the term *grammar*? The influence of lexicality on the judgment process, as reflected in the reaction times, raises the possibility that *analogy* may be the more appropriate concept, but it is already known that lexical statistics influence nonword judgments in both spoken (e.g., Bailey & Hahn, 2001) and signed language (Carreiras, Gutiérrez-Sigut, Baquero, & Corina, 2008), and most linguists consider these systems to involve grammar. Distinguishing grammar from analogy is notoriously difficult anyway (Pinker & Ullman, 2002; McClelland & Patterson, 2002).

Could knowledge of Chinese character grammar merely be a form of visual learning (Chang & Knowlton, 2004; Stobbe, Westphal-Fitch, Aust & Fitch, 2012)? This is an equally difficult question, but discussion of it should take into account the observations that characters are constructed recursively, employ closed-class bound elements reduced in non-final positions, and obey mentally active constraints on reduplication that reflect binarity and final prominence.

Acknowledgments

This study was supported by grants from Taiwan's National Science Council (now Ministry of Science and Technology) NSC 101-2410-H-194-115-MY3 and MOST 103-2410-H-194-119-MY3. [Further thanks TBA.]

References

Aronoff, M., Meir, I., & Sandler, W. (2005). The paradox of sign language morphology.

Language, 81(2), 301-344.
Bailey, T. M., & Hahn, U. (2001). Determinants of wordlikeness: Phonotactics or lexical neighborhoods? *Journal of Memory and Language, 44*(4), 568-591.
Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255-278.
Bates, D., Mächler, M., Bolker, B., & Walker, S. (in press). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*.
Beckman, M. E., & Edwards, J. (1990). Lengthening and shortening and the nature of prosodic constituency. In J. Kingston & M. E. Beckman (Eds.), *Papers in laboratory phonology I: Between the grammar and the physics of speech* (pp. 152-178). Cambridge, UK: Cambridge University Press.
Carreiras, M., Gutiérrez-Sigut, E., Baquero, S., & Corina, D. (2008). Lexical processing in Spanish sign language (LSE). *Journal of Memory and Language, 58*(1), 100-122.
Chan, L., & Nunes, T. (1998). Children's understanding of the formal and functional characteristics of written Chinese. *Applied Psycholinguistics, 19*(01), 115-131.
Chang, G. Y., & Knowlton, *B*. J. (2004). Visual feature learning in artificial grammar classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(3), 714-722.
Chen, J.-Y., & Cherng, R.-J. (2013). The proximate unit in Chinese handwritten character production. *Frontiers in Psychology, 4*, 517. doi:10.3389/fpsyg.2013.00517
Chinese characters decomposition. (2015). In *Wikimedia Commons*. Retrieved June 9, 2015, from http://commons.wikimedia.org/wiki/Commons:Chinese_characters_decomposition.
Gordon, M. (2002). A factorial typology of quantity-insensitive stress. *Natural Language & Linguistic Theory, 20*(3), 491-552.
Hockett, C. (1960). The origin of speech. *Scientific American, 203*, 88-111.
Honorof, D. N., & Feldman, L. (2006). The Chinese character in psycholinguistic research: Form, structure, and the reader. In P. Li, L. H. Tan, E. Bates, & O. J. L. Tzeng (Eds.) *The handbook of East Asian psycholinguistics, vol. 1: Chinese* (pp. 195-208). Cambridge, UK: Cambridge University Press.
Hsu, C. H., Tsai, J. L., Lee, C. Y., & Tzeng, O. J. L. (2009). Orthographic combinability and phonological consistency effects in reading Chinese phonograms: an event-related potential study. *Brain and Language, 108*(1), 56-66.
Huang, C.-R., Chen, K.-J., Chen, F.-Y., & Chang, L.-L. (1997). Segmentation standard for Chinese natural language processing. *Computational Linguistics and Chinese Language Processing, 2* (2), 47-62.
Huang, H. S., & Hanley, J. R. (1995). Phonological awareness and visual skills in learning to read Chinese and English. *Cognition, 54*(1), 73-98.
Hurch, B. (Ed.) (2005). *Studies on reduplication*. Berlin: Walter de Gruyter.
Kawahara, S. (2011). Experimental approaches in theoretical phonology. In M. van Oostendorp, C. Ewen, E. Hume, & K. Rice (Eds.) *The Blackwell companion to phonology* (pp. 2283-2303). Oxford, UK: Wiley-Blackwell.
Kordek, N. (2013). *On some quantitative aspects of the componential structure of Chinese characters*. Wydawnictwo Rys.
Ladd, D. R. (2014). *Simultaneous structure in phonology*. Oxford, UK: Oxford University Press.
Lee, C. Y., Huang, H. W., Kuo, W. J., Tsai, J. L., & Tzeng, J. L. O. (2010). Cognitive and neural basis of the consistency and lexicality effects in reading Chinese. Journal of Neurolinguistics, 23(1), 10-27.
Li, J., & Zhou, J. (2007). Chinese character structure analysis based on complex networks.

Physica A: Statistical Mechanics and its Applications, 380, 629-638.

Liu, C., Zhang, W. T., Tang, Y. Y., Mai, X. Q., Chen, H. C., Tardif, T., & Luo, Y. J. (2008). The visual word form area: evidence from an fMRI study of implicit processing of Chinese characters. *Neuroimage, 40*(3), 1350-1361.

McClelland, J. L., & Patterson, K. (2002). Rules or connections in past-tense inflections: what does the evidence rule out? *Trends in Cognitive Sciences, 6*(11), 465-472.

Myers, J. (2009). Syntactic judgment experiments. *Language & Linguistics Compass, 3* (1), 406-423.

Peirce, J. W. (2007). PsychoPy: Psychophysics software in Python. *Journal of Neuroscience Methods, 162*(1-2), 8-13.

Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics, 2*, 10. doi:10.3389/neuro.11.010.2008

Perfetti, C. A., & Tan, L. H. (1999). The constituency model of Chinese word identification. In A. Inhoff, J. Wang, & H. C. Chen (Eds.), *Reading Chinese scripts: A cognitive analysis* (pp. 115-134). Mahweh, NJ: Erlbaum.

Pinker, S., & Ullman, M. T. (2002). The past and future of the past tense. *Trends in Cognitive Sciences, 6*(11), 456-463.

Sandler, W. (1993). A sonority cycle in American Sign Language. *Phonology, 10*(02), 243-279.

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

Stobbe, N., Westphal-Fitch, G., Aust, U., & Fitch, W. T. (2012). Visual artificial grammar learning: Comparative research on humans, kea (*Nestor notabilis*) and pigeons (*Columba livia*). *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*(1598), 1995-2006.

Taft, M., Zhu, X., & Peng, D. (1999). Positional specificity of radicals in Chinese character recognition. *Journal of Memory and Language, 40*(4), 498-519.

Topolinski, S., & Strack, F. (2009). The architecture of intuition: Fluency and affect determine intuitive judgments of semantic and visual coherence and judgments of grammaticality in artificial grammar learning. *Journal of Experimental Psychology: General, 138*(1), 39-63.

Wang, J. C. (1983). *Toward a generative grammar of Chinese character structure and stroke order*. University of Wisconsin-Madison, Ph.D. dissertation.

Weskott, T. & Fanselow, G. (2011). On the informativity of different measures of linguistic acceptability. *Language, 87*(2), 249-273.

Yeh, S. L., & Li, J. L. (2002). Role of structure and component in judgments of visual similarity of Chinese characters. *Journal of Experimental Psychology: Human Perception and Performance, 28*(4), 933-947.