

Wordlikeness judgments and the syntactic categories of Chinese compound roots

The hypothesis that words are lexically stored with their syntactic categories has been controversial (e.g., Barner & Bale, 2002, vs. Momma, 2016). It is even less clear whether the syntactic categories of compound-internal root morphemes are stored. In Chinese, Packard (2000) argues that compound structure depends on root categories, yet he admits that for some roots, category actually depends on compound structure. Zhang (2007a) extends this latter observation, arguing that a variety of phenomena in Chinese compounds are best explained if roots are underspecified for syntactic category.

This study addresses this question using wordlikeness judgments (cf. Aronoff & Schvaneveldt, 1978) in Chinese. We presented 3,000 randomly generated nonlexical two-root strings to 51 native speakers of Taiwan Mandarin, and asked them to make speeded judgments for whether they were like or unlike Mandarin. Family size (Schreuder & Baayen, 1997), a measure of lexical typicality, was quantified as the mean log number of real two-character words sharing a test item's first or second root in the same position. Syntactic category was assigned to each root based on its most common part of speech as a free word in the Academia Sinica Balanced Corpus (Chen et al., 1996). Because the two-root strings were nonwords, no word-level syntactic category could be assigned objectively, so we analyzed just the 550 NN (noun-noun) and 376 VV (verb-verb) items; Chinese strongly tends to treat NN and VV compounds as nominal and verbal, respectively (Packard, 2000).

Our key questions were how and when lexical typicality and syntactic category affected wordlikeness. In answer to the first question, a mixed-effects logistic regression on judgment choice revealed a significant positive effect of family size on acceptability, but an interaction with syntactic category showed that this effect was significantly stronger for NN compounds (Figure 1). This was not an artifact of lexical statistics because z score scaling eliminated differences in the NN and VV family size distributions, so participants were truly more sensitive to lexical typicality in NN than in VV items. Moreover, judgments for NN items, though not significantly higher or lower than those for VV items, were made significantly faster, suggesting that they were based on readily accessible lexical typicality information.

Following Baayen & Blanche (2017), we then used dynamic survival analysis to probe the temporally evolving probability of a decision "surviving" until the "Mandarin-like" button was pressed. Figure 2 shows that before leveling off, the cumulative incidence of acceptance rapidly increased as a function of family size, becoming significant (with the 95% confidence band rising above zero) by 579 milliseconds after stimulus onset. By contrast, the interaction of family size with syntactic category only became significant later, at 977 milliseconds.

Together these results are consistent with previous linguistic claims that Chinese NN compounds are lexical wholes, whereas VV compounds behave as if composed in the syntax (Packard, 2000; Zhang, 2007b). Yet the time course analysis also shows that this syntactic information is activated late: wordlikeness is initially judged in a category-free manner.

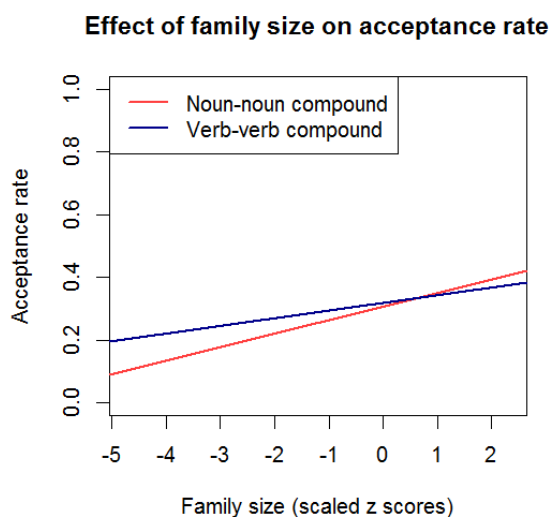


Figure 1

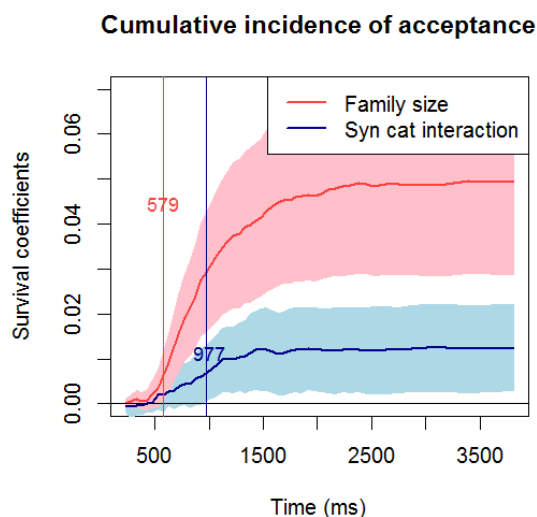


Figure 2

References

- Aronoff, M., & Schvaneveldt, R. (1978). Testing morphological productivity. *Annals of the New York Academy of Sciences*, 318(1), 106-114.
- Baayen, H., & Blanche, P. (2017). Dynamic survival analysis. Talk presented at the Workshop on New Methods in Statistics, Tübingen, Germany, January 19-20.
- Barner, D., & Bale, A. (2002). No nouns, no verbs: Psycholinguistic arguments in favor of lexical underspecification. *Lingua*, 112(10), 771-791.
- Chen, Keh-Jiann, Chu-Ren Huang, Li-Ping Chang, and Hui-Li Hsu. 1996. "SINICA CORPUS: Design methodology for balanced corpora," *Proceedings of PACLIC 11th Conference*, pp.167-176.
- Momma, S. (2016). *Parsing, generation, and grammar*. University of Maryland Ph.D. thesis.
- Packard, J. L. (2000). *The morphology of Chinese: A linguistic and cognitive approach*. Cambridge University Press.
- Schreuder, R., & Baayen, H. (1997). How complex simplex words can be. *Journal of Memory and Language*, 37, 118-139.
- Zhang, N. (2007a). Root merger in Chinese compounds. *Studia Linguistica*, 61(2), 170-184.
- Zhang, N. (2007b). A syntactic account of the Direct Object Restriction in Chinese. *Language Research*, 43(1), 53-75.