Form and meaning in the reading of Chinese compounds

James Myers[1], Gary Libben[2], Bruce Derwing[2]
[1]National Chung Cheng University
(Lngmyers@ccu.edu.tw)
[2]University of Alberta

Fifth International Conference on the Mental Lexicon
Oct. 11, 2006

1

Thanks to...

- Chiang Ching-Kuo Foundation grant RG001-D-02
- Assistants (and Chinese-speaking linguists) Wang Wenling, Chiu Zhenhao, Peng Yuru, Ko Yuguang

2

Factors affecting compound recognition

- Formal factors
  - Position
  - Cross-morphemic predictability
- Morpho-semantic factors
  - Semantic transparency
  - Headedness
- Reading direction…?

3

Position

- First morpheme frequency
  - English and Chinese (Taft & Forster, 1976; Zhang & Peng, 1992)
- First morpheme priming
  - Bulgarian, Greek, Polish, Chinese, and Dutch (Jarema et al., 1999; Kehayia et al., 1999; Myers et al., 2004; Sandra, 1990)
- First morpheme transparency
  - English and Chinese (Libben et al., 2003; Myers et al., 2004)

4

Cross-morphemic predictability

- Mutual information (Church & Hanks, 1990)
  - Applicable to compounds (Myers & Gong, 2002)
  Log of **ratio of** proportional **word frequency (WF)** to product of proportional **morpheme frequencies (MF)**
  [cf. morphological family size (Schreuder & Baayen, 1997), information residual (Moscoso del Prado Martín et al., 2004)]
- But mutual information is collinear with $\log(WF) - [\log(MF_1) + \log(MF_2)]$
- Solution: $\mathbf{\log(MF_1) \times \log(MF_2)}$
  - Lower value = more predictable = faster access

5

Semantic transparency

- Opaque components compete with word-level semantics (slower access)
- Yet opaque compounds are less productive, so generally more predictable (faster access)
- So transparency effects are confusing unless predictability is factored out
  - **Helps:** French, Chinese (Jarema et al., 1999; Tsai, 1994)
  - **Hurts:** Bulgarian, Chinese (Jarema et al., 1999; Su, 1998)

6

Headedness

- Headedness ≠ transparency: Hogwash!
- The modifier-noun relation can be primed
  - English and Chinese (Gagné & Spalding, 2004; Ji & Gagné, 2004)
- First-position effects have been claimed to be restricted to right-headed compounds
  - Chinese (Zhang & Peng, 1992; Zhang, 1997)

7

The time course of form and meaning processes

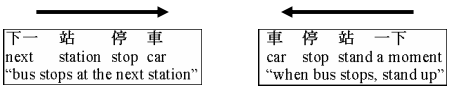
- What causes first morpheme effects?
  - Lexical representations treat it as special…
  - … or it's just the first thing you "see"?
- When do transparency and headedness come into play?
  - Partly late, after whole-word access…
  - … but could they start much earlier?
    (e.g. if morpheme access occurs early)

8

Enter an orthographic quirk…

- In Taiwan, Chinese is written three ways:
  - Top down: traditional
  - Left to right: becoming the default (e.g. computers)
  - Right to left: restricted use (e.g. headlines, old signs)

下一站停車



9

Exploiting the quirk

- The two horizontal directions allow us to test the "first thing you see" hypothesis
  - Orthographically first ≠ morphologically first
- They also help us test when transparency and headedness effects kick in
  - Direction effects must be early effects
  - So if direction modulates transparency or headedness effects, these must also start early

10

Materials: Positional transparency

- Following Libben et al. (2003), eighty compounds were divided into four types by pretested opacity (**O**) and transparency (**T**):

OO: 神經 (god-scripture) "nerve"
OT: 火車 (fire-vehicle) "train"
TO: 時光 (time-bright) "time"
TT: 白色 (white-color) "white"

11

Materials: Headedness type

- Compounds were later classified by native-speaking linguists as right-headed compounds vs. not:

	Right-headed	Not right-headed
OO	海報 (sea-report) "poster" (4)	神經 (god-scripture) "nerve" (16)
OT	火車 (fire-vehicle) "train" (16)	風景 (wind-scenery) "scenery" (4)
TO	女士 (female-scholar) "lady" (9)	時光 (time-bright) "time" (11)
TT	白色 (white-color) "white" (10)	海洋 (sea-ocean) "ocean" (10)

12

Materials: Frequency etc

- Log word frequency was matched across transparency types
- Log character frequency also matched
  - Characters ≈ morphemes
- Only nonreversible compounds
  - E.g. not used:
    蜂蜜: (bee) honey
    蜜蜂: honeybee

13

Design

- Task: Visual lexical decision
- Four groups of participants:
  - Left to right only (20)
  - Right to left only (20)
  - Both directions mixed (40)
  - Top down (20) - not discussed here

14

Analysis

- Multi-level modeling (Pinheiro & Bates, 2000)
  - Subjects and items both treated as random
- Independent variables (predicting **log(RT)**):
  - **Dir**: left to right vs. right to left
  - **Mix**: one consistent direction vs. mixed
  - **CS_1, CS_2**: character transparency types (T vs. O)
  - **Head**: right-headed compound vs. not
  - **CF_1, CF_2**: log character frequencies
  - **WF**: log word frequency

15

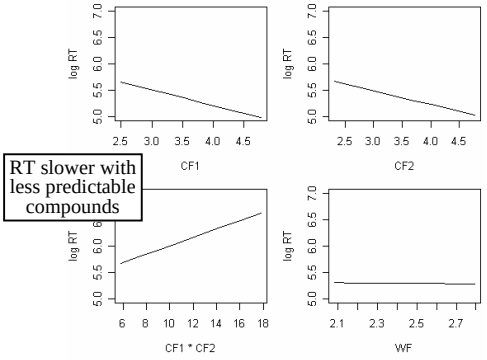
Interactions we care about

- **$CF_1 \times CF_2$**: Cross-morphemic predictability
- **$CS_1 \times CS_2$**: TT/OO vs. OT/TO
- **Dir × CS, Dir × Head**: Influence of direction
  - Do transparency & headedness effects occur early?
- Model tested:

$$\underbrace{\text{Dir} \times \text{Mix} \times CS_1 \times CS_2 \times \text{Head}}_{\text{“early” \& “late” factors}} + \underbrace{CF_1 \times CF_2 + \text{WF}}_{\text{predictability \& other nuisance}}$$

16

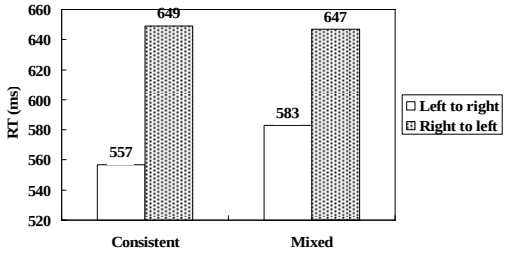
Frequency and predictability



RT slower with less predictable compounds

17

Direction and mixing

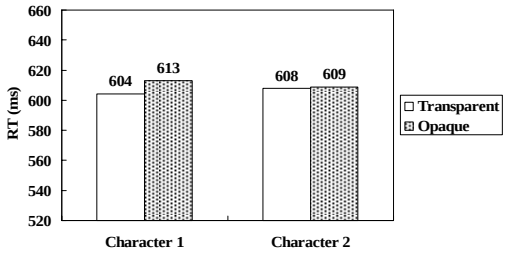


Right to left is significantly slower

(Consistency and interaction not significant)

18

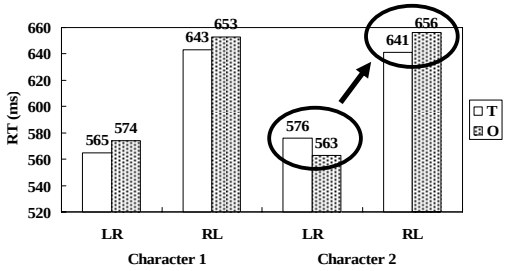
Semantic effects not robust overall



(CS_1 and interaction significant only by subject)

19

Right-to-left reading reverses C_2 transparency effect



Only when C_2 is on left does its opacity hurt access

(C_1 unaffected by direction: late process?)

20

Transparency and reading direction

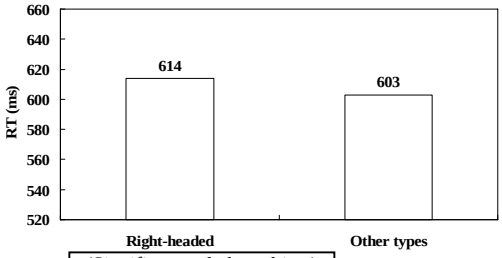
- The processor first accesses leftmost character, even if it's the "second" morpheme
- If opaque, word access is slowed
- Character-based access makes sense:

  Chinese reading requires composition, not decomposition (Myers et al. 2006)

  - Characters ≈ morphemes
  - No word boundaries

21

Headedness effects not robust overall

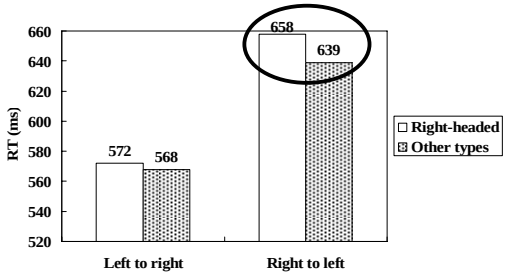


(Significant only by subject)

(Also, no interaction with transparency)

22

Right-to-left reading enhances headedness effect



Only when head is on left does headedness hurt access

23

Headedness and reading direction

- The processor expects to the head to be "second" in a left-to-right direction
- In the less familiar right-to-left direction, the head (on left) is misclassified as modifier
- Recovering from this mistake takes time

24

Summary

- Reading direction influences both positional transparency and headedness effects
  - Hence both start early
  - Their effects are distinct from each other and from predictability
- Semantic transparency effects start early because of character-by-character access
- Yet head assignment must also start early
  - Characters are accessed with their expected morphological roles in mind

25

References (1/2)

Church, K. W., & Hanks, P. (1990). Word association norms, mutual information, and lexicography. *Computational Linguistics, 16*, 22-29.

Gagné, C. L., & Spalding, T. L. (2004). Effect of relation availability on the interpretation and access of familiar noun-noun compounds. *Brain and Language, 90*, 478-486.

Jarema, G., Busson, C., Nikolova, R., Tsapkini, K., & Libben, G. (1999). Processing compounds: A cross-linguistic study. *Brain & Language, 68*, 362-369.

Ji, H., & Gagné, C. L. (2004, June-July). Lexical and relational influences on the processing of Chinese modifier-noun compounds. Poster presented at the Fourth International Conference on the Mental Lexicon, Windsor, Canada.

Kehayia, E., Jarema, G., Tsapkini, K., Perlak, D., Ralli, A., & Kadzielawa, D. (1999). The role of morphological structure in the processing of compounds. *Brain and Language, 68*, 370-377.

Libben, G., Gibson, M., Yoon, Y. B., & Sandra, D. (2003). Compound fracture. *Brain and Language, 84*, 50-64.

Moscoso del Prado Martín, F., Kostić, A., & Baayen, R. H. (2004). Putting the bits together: An information theoretical perspective on morphological processing. *Cognition, 94*, 1-18.

Myers, J., & Gong, S. (2002). Cross-morphemic predictability and the lexical access of compounds in Mandarin Chinese. *Folia Linguistica, 26* (1-2), 65-96.

Myers, J., Derwing, B., & Libben, G. (2004). The effect of priming direction on reading Chinese compounds. *Mental Lexicon Working Papers, 1*, 69-86.

26

References (2/2)

Myers, J., Huang, Y.-C., & Wang, W. (2006). Frequency effects in the processing of Chinese inflection. *Journal of Memory and Language, 54*, 300-323.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-Plus*. Berlin: Springer.

Sandra, D. (1990). On the representation and processing of compound words: automatic access to constituent morphemes does not occur. *The Quarterly Journal of Experimental Psychology, 42A* (3), 529-567.

Schreuder, R., & Baayen, H. (1997). How complex simplex words can be. *Journal of Memory and Language, 37*, 118-139.

Su, Y.-C. (1998). The representation of compounds and phrases in the mental lexicon: evidence from Chinese. *University of Maryland Working Papers in Linguistics, 6*, 179-199.

Taft, M., & Forster, K. I. (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning & Verbal Behavior, 15* (6), 607-620.

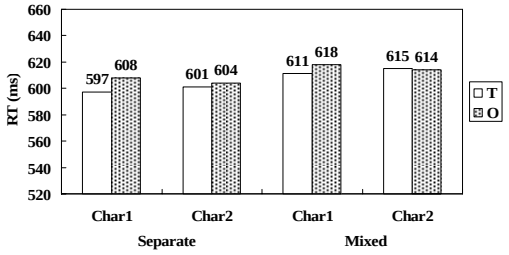
Tsai, C.-H. (1994). *Effects of semantic transparency on the recognition of Chinese two-character words: Evidence for a dual-process model*. Unpublished master's thesis, National Chung Cheng University, Chia-Yi, Taiwan.

Zhang, B. & Peng, D. (1992). Decomposed storage in the Chinese lexicon. In H.-C. Chen & O. J. L. Tzeng (Eds.), *Language processing in Chinese* (pp. 131-149). Amsterdam: North-Holland.

Zhang, B. (1997). Zhongwen shuangzici zai xinli cidian zhong de chucun moshi [Storage model of Chinese two-character words in the mental lexicon]. In R. Peng, H. Shu & H. Chen (Eds.), *Hanyu renzhi yanjiu* [Chinese cognitive research] (pp. 217-230). Jinan, China: Shandong Jiaoyu Publishers.

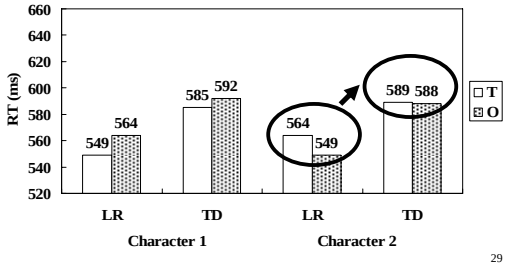
27

Appendix I: Direction consistency affects semantic consistency effects...?



28

Appendix II: Top-down direction flattens C_2 semantic transparency effects



29