

# A Corpus-Based Behavioral Profile Approach on Action Verbs Bào and Pù in Mandarin News Headlines

Shishan Liu  
National Taiwan Normal University

**Keywords:** Mandarin news headlines, behavioral profiles, corpus linguistics

News headlines act as the first impression for readers on any event in the society. In order to cater to market or political interest, the choice of words within the headlines becomes significant because it leaves a space to manipulate readers' ideology underlyingly. The usage of verbs is one way of realizing manipulation. The usages of action verbs Bào and Pù are common examples of Mandarin news headlines nowadays in Taiwan. Based on the definition from the Chinese WordNet, these two action verbs Bào and Pù express similar word senses, i.e., to expose something confidential to the public. In this study, we collect data from the Internet and use the corpus-based Behavioral Profile approach to explore the morphosyntactic contexts of Bào and Pù in the news headlines. Specifically, we examine their usage patterns in three structural conditions: active, passive and middle voice. Passive voice is defined as an overt co-occurrence with passive markers (e.g., Bèi or Zāo); middle voice is defined as a patient-as-subject use without overt passive markers (e.g., Example (1) and Example (2)). We annotate our data by referring to the relevant features discussed in the previous studies at multiple levels and analyze their usage differences using the behavioral profile approach. The distinctive features of these two verbs are summarized in a snake plot as shown in Figure 1. Our results show that Bào often (a) co-occurs with human subjects, (b) takes complements connected to negative events, (c) and carries the passive voice. On the other hand, Pù usually (a) co-occurs with subjects that are technology-related products (e.g., videos or images), and (b) carries the active or middle voice. The contextual features of these two verbs are also supported by the post-hoc analyses. We adopt Cramer's V to examine the effectiveness of our selected features. The statistics suggests that COMPLEMENT, VOICE, and SUBJECT are the top three distinctive features between Bào and Pù. Our comprehensive analysis of these frequently used action verbs can serve as a steppingstone for the future investigation on language and ideology in Mandarin news reports.

(1) Three structural conditions of Bào

a. Táiběi chéng bó huì méi rénqì yiyuán bào shì fǔ shī yā míntuán qiǎngpò cānguān (active voice)

“TAIPEI EXPO lacks popularity, legislator claimed that Taipei City Government pressured civil organizations to visit by force.”

b. Dà māoxióng tuántuán jīng bào diānxián! Dòngwùyuán zhèngshí `nǎo bù júbù yèhuà huàisǐ` (middle voice)

“Giant Panda TuanTuan shocked with seizures! Zoo confirmed "partial liquefaction necrosis of the brain"”

c. Táizhōng gōngchē sījī bèi bào `tíngchē mǎi biàndāng` bǎ chéngkè liàng chē shàng 10 fēnzhōng (passive voice)

“Taichung bus driver is exposed for "parking aside to buy lunch", leaving passengers stranded on the bus for 10 minutes.”

(2) Three structural conditions of Pù

a. Zàn hánjù gǎnyú wājué tícái wēn shēng háo pù 4 yuè pāi hánguó tóuzī yǐngjí (active voice)

“Praise Korean dramas for daring to explore themes, Wen ShengHao reveals that he will shoot a Korean investment album in April.”

b. Zhōurì běidōng shī dā dā! Zhuǎn qián liáng shíjiān pù xià zhōu èr kǒng diē pò 20 dù (middle voice)

“Rainy and humid in the northeast on Sunday! It is revealed that it may fall below 20 degrees next Tuesday when it turns dry and cool.”

c. Huáwèi bèi pù zài hélán diànxīn shāng wǎngluò cáng yǒu `hòumén`

“Huawei is exposed to have "backdoors" in the network of Dutch telecommunication companies.”

## References

- Chu-Ren Huang and Shu-Kai Hsieh. (2010). *Infrastructure for Cross-lingual Knowledge Representation — Towards Multilingualism in Linguistic Studies*. Taiwan NSC-granted Research Project (NSC 96-2411-H-003-061-MY3)
- Divjak, D., & Gries, S. T. (2006). *Ways of trying in Russian: Clustering behavioral profiles*.
- Divjak, D., & Gries, S. T. (2009). *Corpus-based cognitive semantics: A contrastive study of phasal verbs in English and Russian*. *Studies in cognitive corpus linguistics*, 273-296.
- Gries, S. T. (2006). *Corpus-based methods and cognitive semantics: The many senses of to run*. *Trends in linguistics studies and monographs*, 172, 57.
- Gries, S. T. (2010). *Behavioral profiles: A fine-grained and quantitative approach in corpus-based lexical semantics*. *The mental lexicon*, 5(3), 323-346.
- Gries, S. T., & Otani, N. (2010). *Behavioral profiles: A corpus-based perspective on synonymy and antonymy*. *ICAME journal*, 34(1), 121-150.
- Huang, Yi-Ru (2022). *A Quantitative Corpus-based Behavioral Profile Study of Xiǎo and Wéi Modification in Mandarin*.
- Huang, P. W., & Chen, A. C. H. (2022). *Degree adverbs in spoken Mandarin: A behavioral profile corpus-based approach to language alternatives*. *Concentric*, 48(2), 285-322.
- Kuo, S. H., & Nakamura, M. (2005). *Translation or transformation? A case study of language and ideology in the Taiwanese press*. *Discourse & Society*, 16(3), 393-417.
- Levshina, N. (2015). *How to do linguistics with R: Data exploration and statistical analysis*. John Benjamins Publishing Company.
- Zaenen, A., Carletta, J., Garretson, G., Bresnan, J., Koontz-Garboden, A., Nikitina, T., ... & Wasow, T. (2004, July). *Animacy encoding in English: Why and how*. In *Proceedings of the workshop on discourse annotation* (pp. 118-125).