

Microvariation in Chinese nominal domain: mapping formal and dialectal cartography using historical corpora and social media:

Chinese adnominalisers constitute a relatively recent category which came to existence in the late Medieval Period where phrase-final nominaliser *di* (底) gave rise to modern Mandarin *de* (的) and general classifier *ge* (個) was reanalysed as the various *g*-forms (嘅/个) in central-southern Chinese dialects (Yap et al (2011)). Despite broad similarities in denoting modification and possession below numerals and quantifiers (del Gobbo (2006), Sio (2011), Paul (2015)), microvariation exists in that *de* but not *ge* can occur as a verbal suffix in clefts (1) and head Activity nominals in subject position (2) (Tang (2011)):

- 1a) 他 是 昨天 買 票 的 / 他 是 昨天 買 的 票
he COP yesterday buy ticket DE he COP yesterday buy DE ticket
'It was yesterday that he bought tickets.' (Mandarin, in Simpson and Wu (2002:169))
- 1b) 佢 係 琴日 買 非 嘅 / *佢 係 琴日 買 嘅 非
he COP yesterday buy ticket GE he COP yesterday buy GE ticket
'It was yesterday that he bought tickets.' (Cantonese, in Tang (2008:73))
- 2a) 他 當 他 的 老師 / 他 的 老師 當 得 好
he act.as he DE teacher he DE teacher act.as COMP good
'He is acting as a teacher'/'his acting as a teacher is done well.' (Mandarin, in Huang (2008:231))
- 2b) 佢 做 佢 嘅 老師 / *佢 嘅 老師 做 得 好
he act.as he GE teacher he GE teacher act.as COMP good
'He is acting as a teacher'/'his acting as a teacher is done well.' (Cantonese, in Tang (2009:241))

Such microvariation can be traced from their diachronic roots, since *de* is widely argued to be a linker derived from phrase-final *di* in appositive constructions (den Dikken (2006), Jiang (1999), Aldridge (2009)) and *ge* may have retained some of its original quantificational and referential properties in line with East Asian classifiers (Bisang and Xu (2012), Zhang (2013)). This is supported by electronic historical corpora where KWIC searches show that *ge* [+referential] tends to denote finite and referential nouns (3b, 4b) which, in contrast to *de* (3a, 4a), does seem to pre-empt its reanalysis in clefts (3a-b) and Activity nominals (4a-b):

- 3a) 娘 原 是 氣惱-上 起 的 病
mother originally COP angry-LOC contract DE illness
'As for our mother, it was originally illness contracted via anger' > 'It was via anger that our mother originally contracted illness.' (金瓶梅詞話)
- 3b) 則 是 茶坊-裏 見 個... 官人
then COP tea.lounge-LOC see GE official
'Then it was in a tea lounge that he saw an official' (南宋話本選集)
- 4a) 太守 見 他 的 工 完 得 甚 遲
official see he DE work finish COMP quite late
'The official saw that he finished his work quite late.' (醒世姻緣傳)
- 4b) 好 個 聰明 人 相全
good GE wise people complete
'Good wise men are complete.' (敦煌變文集)

In line with dialect levelling (Trudgill (1986)), modern Chinese dialects display ambiguity in these uses of *de* and *ge*, since while the discrepancies as seen in 1-2) largely hold between extreme northern (northern Mandarin 北方話) and extreme southern dialects like Yue (粵) (Cantonese), Southern Min (閩南) (Taiwanese) and Hakka (客家) (Tang (2009)), central dialects such as Wu (吳) and Xiang (湘) display subtle divergent behaviour in (marginally) permitting *ge* in cleft constructions, though, interestingly, not as much in Activity nominals, as shown in the author's collection of dialectal data on WeChat:

- 5a) 吾 是 昨日 買 個 票 / *吾 個 飯 吃-了 蠻 久
I COP yesterday buy GE ticket I GE meal eat-ASP very long
'It was yesterday that I bought tickets'/'My eating took a long time.' (Shanghainese (Wu), WeChat)
- 5b) 俚 是 昨日 買 個 票 / *俚 個 水 游-了 蠻 久
you COP yesterday buy GE ticket you GE swimming swim-ASP very long
'It was yesterday that you bought tickets'/'Your swimming took a long time.' (Suzhou dialect (Wu), WeChat)
- 5c) 我 是 前年 到 個 北京 / ?我個 籃球 打 得 蠻 好
I COP two.years.ago arrive GE Beijing I GE basketball play COMP very good
'It was two years ago that I arrived at Beijing'/'I play basketball very well' (Xinhua dialect (Xiang), in Wu (2005:279))

The dialectal data, whose collection and consultation are permitted by powerful social media tools like WeChat in mainland China, shows an ongoing change in the Chinese nominal domain where *ge*, especially in dialects in contact with the north, is being reanalysed as a verbal suffix in clefts but not as the head of Activity nominals (5a-c), which may be explained by the fact that Activity nominals tend to be definite while postverbal objects can be indefinite (3b), and in line with Feature Interpretability (Tsimplici and Dimitrakopoulou (2007)) this definiteness feature (+D) seems to pre-empt the dialectal diffusion of *ge* in Activity nominals.

(Select) References:

- Bisang, W. and Li, X-P. (2012): 'Classifiers in Sinitic languages'. *Lingua* 122:335-355; Tang, S-W. (2011): 'On gerundive nominalization in Mandarin and Cantonese', in Yap, F-H, Grunow-Harsta, K. and Wrona, J. (eds), *Nominalization in Asian Languages: diachronic and typological perspectives*, Amsterdam/Philadelphia: John Benjamins, pp. 147-161; Tsimplici, I. and Dimitrakopoulou, M. (2007): 'The Interpretability Hypothesis: evidence from wh-interrogatives in second language acquisition'. *Second Language Research* 23(2):215-242; Wu, Y-J. (2005): *A Synchronic and Diachronic Study of the Grammar of the Chinese Xiang Dialects*. Berlin/New York: Mouton de Gruyter; Zhang, N. (2013): *Classifier Structures in Mandarin Chinese*. Berlin/New York: Mouton de Gruyter.