Diachronic formation of Chinese cleft (shi-de) constructions: Mandarin and dialects

Chinese cleft constructions consist of the copula (shi or dialectal equivalent) assigning focus to its immediately adjacent element in its complement ending with an abnominaliser (de in Mandarin or dialectal g-equivalents), and while de is attested as either sentence-final (VOde) or sentence-medial (VdeO), g-forms are generally sentence-final (VOge) and not otherwise (*VgeO) (Lee and Yiu (1998), Tang (2011), Shyu (2016)):

1a) Wo shi zuotian mai piao de
   I SHI yesterday buy ticket DE
   ‘It was yesterday that I bought the ticket.’ (Mandarin, from Simpson and Wu (S&W) (2002:169))

1b) Wo shi zuotian mai de piao
   I SHI yesterday buy DE ticket
   ‘It was yesterday that I bought the ticket.’

2a) kui hai kamyat maai sue ge
   He COP yesterday buy book GE
   ‘It was yesterday that he bought the ticket.’ (Cantonese, from Lee and Yiu (1998:5))

2b) *kui hai kamyat maai ge sue
   He COP yesterday buy GE book
   ‘It was yesterday that he bought the ticket.’

It has recently been argued that Chinese cleft constructions are formed in copular constructions where the clausal complement of the copula, which is usually the nominalized relative clause is reanalyzed from being nominal to clausal (Long and Xiao (2009, 2011), Han (2012), Zhan (2012), Long (2013)), which, in formal terms, constitutes a type of ‘lateral’ reanalysis where the nominal elements (de/ge) are ‘laterally’ reanalysed from being nominal to clausal, especially when their original deictic force [+D] is weakened (Tse (2018), cf Simpson and Wu (2002)).

In the case of de (< di), there are clear instances of such weakening in both VOde and VdeO where the nominalized relative clauses (VOde/VdeO) are not necessarily referential/definite and de may hence be reanalyzed as clausal particles (C/T(past) respectively):

3a) fei shi pusa xingcang ci shi suren zuo di
   NEG COP saint behaviour this COP layman do DE
   ‘It is not the behaviour of saints; this is the doings of laymen (> it was the laymen who did this).’ (Zhenzhou liji Haizhou chanshi yulu)

3b) wukong, ni shi na shi xiu-lai de yuanfa
   Wukong you COP which life obtain-RES DE enlightenment
   ‘Wukong, as for you, of which life is your enlightenment obtained? (> Which life was it that you obtained your enlightenment?)’ (Xiyouji)

As for ge, examples of VOge are attested while *VgeO is not attested:

4) wo bian shi zhao xiang ge
   I then COP wear light GE
   ‘Then I shall be the one who is in trouble (> then I shall be in trouble.).’ (Tongshang, yierchu)

This discrepancy between de and ge can be accounted for in terms of their etymological origins, the fact that de is derived from an original nominalizer di on a par with classical Chinese zhe (Lu (1943), Aldridge (2008), Yap et al (2010)) whereas ge is widely argued to be derived from the general classifier ge (Cao (1994)), and since de being derived from nominalizer di does not necessarily denote referentiality/definiteness, it can be reanalyzed as a clausal particle when its deictic force is suppressed. Classifier ge, on the other hand, is widely attested with referentiality/definiteness in dialects (Bisang (2012)), and so its deictic force is more inherent and hence more difficult to be suppressed, which entails that it is only reanalysable as a sentence-final particle in sentence-final position (VOge, cf ex. 4)) but not as a sentence-medial particle (*VgeO). There are hence microparametric variation between Chinese dialects in the nominal hierarchy (de (n/D), ge (CL/D)) which yields different cleft constructions, since while the deictic force of de is not inherent and hence not guaranteed (VOde/VdeO), the deictic force of ge seems to be stronger and hence not attested in the sentence-medial position (VOge/*VgeO).
References:


