

# Chinese Affixes in the Internet Era

## A Corpus-Based Study of X-族 *zú*, X-党 *dǎng* and X-客 *kè* Neologisms

Bianca Basciano

Università Ca' Foscari Venezia, Italia

Sofia Bareato

Università Ca' Foscari Venezia, Italia

**Abstract** In the last few decades, under the influence of foreign languages and net-speak, many word-formation patterns emerged in the Chinese lexicon. This paper proposes a corpus-based investigation of three suffixes, i.e. 族 *-zú*, 党 *-dǎng*, and 客 *-kè*, which build words indicating persons with certain characteristics or behaviour, or doing a certain activity. The paper aims at describing and comparing the three word-formation patterns based on these suffixes. It also aims at describing their evolution over time and their grammaticalisation path. In addition, it discusses the diffusion of the three patterns in Chinese and compares their productivity.

**Keywords** Derivation. Affixes. Word formation. Neologisms. Productivity.

**Summary** 1 Introduction. – 2 Derivation in Mandarin Chinese. – 3 Description of the Three Word-Formation Patterns. – 3.1 X-族 *zú* Words. – 3.2 X-党 *dǎng* Words. – 3.3 X-客 *kè* Words. – 3.4 Are X-族 *zú* and X-党 *dǎng* Words Collective Nouns? – 4 On the Development of 族 *zú*, 党 *dǎng* and 客 *kè*. – 4.1 The Evolution of 族 *zú*. – 4.2 The Evolution of 党 *dǎng*. – 4.3 The Evolution of 客 *kè*. – 4.4 A Comparison of the Three Word-Formation Patterns. – 5 A Comparison of the Productivity of 族 *-zú*, 党 *-dǎng* and 客 *-kè*. – 6 Conclusions.

## 1 Introduction<sup>1</sup>

In the last few decades, under the influence of foreign languages and netspeak, many word-formation patterns emerged in the Chinese lexicon. In this paper, we will examine three formatives, i.e. 族 *zú* (orig. ‘clan, ethnic group’), 党 *dǎng* (orig. ‘party, clique’), and 客 *kè* (orig. ‘guest’), used to form nouns indicating persons with certain characteristics or behaviour, or doing a certain activity, as in the following examples:<sup>2</sup>

1. a. 背包族  
*bèi-bāo-zú*  
back-pack-ZU  
‘backpackers’
- b. 剁手党  
*duò-shǒu-dǎng*  
chop-hand-DANG  
‘online shopaholics [those who buy things online and then regret it, wanting to cut their own hands off]’
- c. 换客  
*huàn-kè*  
exchange-KE  
‘one who sells/exchanges goods online’

Both 族 *zú* and 党 *dǎng* refer to groups of people with common characteristics or behaviour. X-族 *zú* is a quite established and widely studied word-formation pattern in Chinese. Neologisms containing the formative 族 *zú* were attested already in the Nineties and greatly increased in number over the years: between 1995 and 2006, 310 X-族 *zú* neologisms may be found in the 人民日报 *Renmin Ribao* (*People’s Daily*) (Chen, Zhu 2010; 309 according to Cao 2007). Many X-族 *zú* words are now listed in dictionaries: the words 上班族 *shàng-bān-zú* ‘go-work-ZU, office workers’ and 工薪族 *gōng-xīn-zú* ‘salary-ZU, sal-

---

<sup>1</sup> We are very grateful to two anonymous reviewers for their constructive comments and suggestions. We would also like to thank Giorgio Francesco Arcodia for carefully reading a draft of this paper.

The glosses follow the general guidelines of the Leipzig Glossing Rules, with the addition of SP = ‘structural particle’. For academic purposes, Bianca Basciano is responsible for §§ 2, 3.4, 4, 5 and 6, and Sofia Bareato is responsible for §§ 1, 3.1, 3.2 and 3.3. X-族 *zú* and X-党 *dǎng* words were collected by Sofia Bareato in her MA dissertation (Bareato 2017).

<sup>2</sup> In order to distinguish these formatives from the corresponding lexemes, we gloss them as ZU, DANG, and KE.

aried people' were included in the 现代汉语词典 *Xiandai Hanyu cidian* (The Contemporary Chinese Dictionary) in 2005 (Cao 2007).

In contrast, X-党 *dǎng* represents a quite novel pattern of word formation (Chen, Zhu 2010). Complex words containing the morpheme 党 *dǎng* as the right-hand constituent with the meaning 'group of people with common characteristics or behaviour' began to appear in the late 2000s; this pattern is mostly used on the web.

Finally, the morpheme 客 *kè* has been used as the right-hand constituent of complex words indicating 'a person doing a certain activity' or 'a person with certain characteristics', since the beginning of the twenty-first century, (again) mostly on the web.

In this paper we will examine a corpus of neologisms containing the three items at issue drawn from the following sources:

- the 新世纪新词语大词典 *Xin shiji xinciyu da cidian* (New Century Comprehensive Dictionary of Neologisms) (henceforth XCY), which collects about 5,400 neologisms coined between 2000 and 2015;
- the Leiden Weibo Corpus (henceforth: LWC),<sup>3</sup> an annotated 100-million-word corpus, consisting of 5,103,566 messages posted on Sina Weibo (China's premier Twitter-like microblogging service) in January 2012.
- the Buzzwords section of the *Shanghai Daily* (henceforth: SD).<sup>4</sup>

We collected 707 distinct words in total: specifically, 434 X-族 *zú* words, 189 X-党 *dǎng* words, and 84 X-客 *kè* words.

The aim of this paper is twofold. First, it aims at describing and comparing the three word-formation patterns at issue, highlighting their formal and semantic properties. Secondly, it aims at describing the evolution of the three formatives over time and their grammaticalisation path, as well as their diffusion in Chinese. To this end, we will also propose an analysis of productivity measures for the three word-formation patterns.

The paper is organised as follows: § 2 provides an overview of derivation in Mandarin Chinese, focusing on its status and on the charac-

<sup>3</sup> <http://lwc.daanvanesch.nl/index.php>.

<sup>4</sup> <http://buzzword.shanghaidaily.com> (2017-02-06). It is a weekly column of the *Shanghai Daily*, started in October 2005. It aims at recording and translating into English new words and phrases appearing in the press, online etc. According to the editor, the purposes of the column are "first, to provide a tentative English translation of new terms and phrases as a reference for our readers; second, to tell our readers what are the latest buzzwords used by locals in their work and daily life; and third, to invite readers to help us generate better English translations of such stylish or trendy Chinese words and phrases" (*Shanghai Daily* 2010). Unfortunately, the column is no longer available; presumably it ceased operations in 2017, when we last consulted it. The buzzwords appeared in the column up to mid-2009 have also been published as a book (*Shanghai Daily* 2010).

teristics of affixes. § 3 is devoted to the presentation of the word-formation patterns at issue, and of their formal and semantic properties. In § 4, we describe the grammaticalisation path of the three formatives, arguing for their affixal status, and we then propose a comparison of the word-formation patterns based on them. Then, in § 5 we compare their productivity in the Leiden Weibo Corpus. Lastly, in § 6 we present our conclusions.

## 2 Derivation in Mandarin Chinese

While compounding forms words made up of two or more units, be they words (Fabb 1998, 66; Katamba 1993, 54), base lexemes (Haspelmath 2002, 85), stems (Bauer 1998, 404), or roots (Katamba 1993, 54), depending on the morphological profile of the language at issue (Bauer 2006), derivation is a morphological process often involving an affix (Naumann, Vogel 2000, 933-4). Thus, in English, while a word like *zebrafish* is a compound, a word like *violinist* is a derived word. However, the distinction between compounding and derivation is not always clear-cut. In some cases, some elements have hybrid properties, which make it hard to classify them as words or as affixes (Bauer 2005, 106-7). For example, items like *monger*, *cade* or *scape* in English complex words such as *fishmonger*, *motorcade*, *seascape* are not words in Modern English but still retain some kind of full, lexical meaning. In some cases, an affix-like element co-exists with the word it originates from. For example, in Dutch the morpheme *boer* is a word meaning 'farmer'; however, it is also used as the right-hand (head) constituent in complex words with the meaning 'seller of X', as e.g. in *sigaren-boer* 'cigar-farmer, cigar seller', *kabel-boer* 'cable-farmer, provider of broadband cable services' (see Booij 2005). Therefore, we observe semantic differentiation: *boer* is considered an affixal element when, as a right-hand constituent, has the meaning of 'seller', which is not attested in its use as a free form (word).

It has been proposed by many to label these hybrid forms as pseudo-affixes or affixoids (see e.g. Naumann, Vogel 2000), a notion which has been employed in slightly different ways by different authors: as highlighted by Booij (2005), the notion of affixoid is not a theoretical notion, but a convenient descriptive label. These hybrid forms become affixes as soon as their connection with the corresponding lexeme is lost, either because of sound change or because of semantic change, following a process of grammaticalisation.

The issue of the distinction between compounding and derivation is much thornier in Chinese: the existence of derivation as a productive morphological process, distinct from compounding, is under debate (see e.g. Dong 2004). This is due to the characteristics of Mandarin Chinese, an isolating language, where words are generally formed

by the agglutination of morphemes, mostly lexical; compounding is generally regarded as the most productive means of word formation in this language (see Ceccagno, Basciano 2007, 208). In addition, the majority of lexical morphemes are bound (about 70% according to Packard 2000): this means that they cannot occupy a syntactic slot (i.e. they are not words) but have a full lexical meaning and are actively used to form complex words. However, they do not occupy a fixed position, differently from affixes: see e.g. 衣 *yī* 'clothes' (compare the corresponding free form 衣服 *yīfu*) in 大衣 *dà-yī* 'big-clothes, overcoat', 衣蛾 *yī-é* 'clothes-moth, clothes moth'. Furthermore, in Mandarin Chinese, as well as in other languages of East and South-East Asia, there is no regular formal distinction between lexical morphemes and grammatical morphemes (Bisang 1996): except for a few items, there is no phonological reduction nor other formal changes characterising grammaticalised forms.

Only a small number of items are commonly regarded as derivational affixes in the literature, in particular those items which became toneless and lost much of their meaning (and productivity), i.e. 子 *-zi* (< *zǐ* 'child'), as in 桌子 *zhuōzi* 'table', 儿 *-r* (< *ér* 'child'), as in 画儿 *huàr* 'painting', and 头 *-tou* (< *tóu* 'head'), as in 石头 *shítou* 'stone'. As a matter of fact, loss of tone and of lexical meaning seem to be the only criteria accepted by Chinese linguists to include an item among affixes (see Ma 1995).

Other formatives that are usually included among derivational affixes are 化 *-huà* 'ise, -ify' (< *huà* 'change'), as in 国际化 *guójìhuà* 'internationalise, internationalisation', and 性 *-xìng* 'nature, -ity, -ness' (< *xìng* 'inherent nature'), as in 可能性 *kěnéng-xìng* 'possible-ity, possibility'. These two suffixes began to be productively used at the beginning of the 20th century due to the influence of Japanese, where they were used to render the equivalent of European suffixes (Masini 1993). The functional correspondence with suffixes in European languages probably favoured their inclusion among derivational affixes (Pan, Ye, Han 2004, 67). However, it must be noted that these word-formation patterns already existed in Chinese; for example, words containing the suffix 化 *-huà* are found in Premodern Chinese, even though this suffix could only be attached to monosyllabic bases (Arcodia, Basciano 2012). Thus, this pattern was somehow independent from the European model, but it strongly developed starting from the 20th century, due to foreign influence. After that, it started to be used independently, creating new words by analogy, thus not only to translate foreign words (Steffen Chung 2006). Therefore, the influence of foreign languages, in this case, gave impulse to an already existent, though not very productive, word-formation pattern.

Besides the cases mentioned above, there are many ambiguous formatives: how to deal with those lexical morphemes which appear

in a fixed position in a high number of complex words, thus showing a high degree of productivity, always conveying the same meaning? Are they to be analysed as compound constituents or as derivational affixes? Recall that, as mentioned earlier, generally there is no formal distinction between lexical morphemes and grammatical morphemes in Chinese. Take, for example, the root 人 *rén* 'person', which is used both as a word and as the right-hand bound constituent in complex nouns indicating a person from a country, town etc., as e.g. 上海人 *Shànghǎi-rén* 'Shanghai-person, Shanghaiese', 意大利人 *Yìdàlì-rén* 'Italy-person, Italian'. Given its "versatility", Yip (2000, 59-60) regards it as a suffix. A similar example is that of 店 *diàn* 'shop', typically used as a constituent in complex words, indicating any kind of shop, as e.g. 书店 *shū-diàn* 'book-shop, bookstore', 布店 *bù-diàn* 'cloth-shop, cloth store', 冷饮店 *lěng-yǐn-diàn* 'cold-drink-shop, cold-drink bar/shop'. Given the high productivity of the pattern, in which 店 *diàn* has a fixed position and a stable meaning, Lü (1941, quoted in Pan, Ye, Han 2004, 468) considers it as a quasi-affix (近似词缀 *jìnsì cízhù*). However, it must be noted that the number of words built according to a morphological pattern is not normally used as a diagnostic test for affixhood, since compounding patterns too can be very productive. In addition, there is no semantic differentiation observed when these two formatives are used as right-hand constituents bearing a fixed meaning: the meaning of 人 *rén* as a bound right-hand constituent is not different from that of 人 *rén* when used as a free root (word); in the same way, 店 *diàn* retains its original meaning of 'shop', without any kind of bleaching. Also, we may remark that both formatives may be used as left-hand constituents in complex words, bearing the very same meaning: see e.g. 人产 *rén-chǎn* 'person-produce, production per person', 人堆 *rén-duī* 'person-pile, crowd', 店台 *diàn-tái* 'shop-platform, shop counter', 店员 *diàn-yuán* 'shop-member, shop assistant'.

Compare now the examples given above with the root 学 *xué*. In Modern Chinese, 学 *xué* is a free root, a verb meaning 'study'; however, it is also used as the right-hand bound constituent in complex nouns indicating a field of study, as e.g. 语言学 *yǔyán-xué* 'language-study, linguistics', 财政学 *cáizhèng-xué* 'finance-study, finance', 测地学 *cè-dì-xué* 'survey-hearth-study, geodesy'. This formative has two main characteristics: it can be used to build any word indicating a field of study, and it displays some semantic difference from the verb 学 *xué*. For these reasons, some authors have defined items like 学 *xué* as affixes (词缀 *cízhù*) or affixoids (类词缀 *lèicízhù* or 准词缀 *zhǔncízhù*); however, in the literature on the topic, the criteria for the definition of affixes and affixoids, and thus what items should be included in these categories, vary greatly from author to author (see Pan, Ye, Han 2004). Ma (1995), for example, states that in Chinese it is possible to distinguish roots from affixes: affixes are never free, and they appear in a fixed position in complex words. Affixes may be

further divided in ‘true affixes’ (真词缀 *zhēn cízhù*) and ‘quasi-affixes’ (准词缀 *zhǔn cízhù*): the former are semantically empty, always bound, and are characterised by some sort of phonological reduction, typically loss of tone, as the above mentioned 子 *-zi*, 儿 *-r*, and 头 *-tou*, while the latter have some sort of categorial meaning, i.e. they assign the complex word to a lexical category and/or a semantic class (a taxonomical category), like in the case of 学 *-xué* mentioned above. In short, to be categorised as an affixoid, an item must be bound (independently from the fact that it has a corresponding free form) and must convey a meaning which is not its core meaning.

As highlighted by Arcodia (2011), if we were to regard as affixes only those items undergoing some sort of phonological reduction, we would ignore the features of grammaticalisation processes in the languages of East and South-East Asia, which, as we mentioned, generally do not display co-evolution of form and meaning. In addition, considering as affixes only items devoid of meaning would result in a definition of derivation which is cross-linguistically inconsistent, since typically derivational affixes carry some sort of meaning.

According to Sun (2000), the distinction between affixes and affixoids is not relevant in Chinese, since the system of derivational affixes in this language is still developing: those morphemes which behave as affixes but are phonologically (and orthographically) identical to their lexematic counterparts should be regarded as not fully grammaticalised. Thus, she holds a view according to which grammaticalisation necessarily involves some formal change. Arcodia (2011) too proposes to abandon the distinction between affixes and affixoids in Chinese but holds a different view: since in Mandarin we may have grammaticalisation of a sign without sound change, then the distinction between affixes and affixoids, which may be useful for European languages, is not relevant in Chinese. However, differently from Sun, Arcodia posits that the fundamental criterion to label a morpheme as a derivational affix in Mandarin Chinese is meaning differentiation. He claims that derivational affixes in Mandarin are the evolution of compound constituents, appearing in a fixed position, with a certain meaning, in a number of complex words. In order to become an affix, a lexeme must undergo a shift in meaning, which can either be more general than the meaning it has in other uses or be the extension of one of the possible non-core meanings of the lexeme. One example provided by Arcodia (2011) is 性 *-xìng*, whose development, as mentioned above, was favoured by the influence of European languages. This item was a word (a free form) in Classical Chinese, but it has turned into a bound root in the modern language, where it can be used to form complex words, as e.g. 性能 *xìng-néng* ‘nature-capacity, natural capacity/function (of machine etc.)/property’, 个性 *gè-xìng* ‘personal-character, character/personality’. However, it also developed an affixal meaning, i.e. ‘nature, -ity, -ness’, as in 毒性 *dú-xìng*

'poison-nature, toxicity', 塑性 *sù-xìng* 'plastic-nature, plasticity' (see above). This meaning developed from the original meanings 'quality, intrinsic properties or characteristics of something' and 'inherent properties of the human being': through a process of generalising abstraction, 性 *-xìng* turned into a nominal suffix indicating just any property (not only intrinsic and everlasting properties), forming abstract nouns.<sup>5</sup>

In this paper, following Arcodia (2011), we dismiss the distinction between affixes and affixoids, since, as we have seen, in Mandarin grammaticalised signs often do not undergo any sound change. If a bound item used to build complex words appears in a fixed position with a fixed meaning, which (partially) departs from its original/core one and is more general or abstract than the meaning of the corresponding lexeme, then it can be regarded as a derivational affix, even if it is not formally different from the corresponding lexeme. Therefore, a formative like 学 *xué* 'field of study' above may be considered as a suffix; in other words, it is a grammaticalised item.

New affixes may emerge not only as a consequence of a grammaticalisation process inner to the language, but also due to the need of translating words containing affixes from foreign languages (see Shen 2015), as mentioned above. A possible example is 控 *kòng* 'buff, enthusiast, devotee', in words like 猫控 *māo-kòng* 'cat-enthusiast, cat lover', 长发控 *cháng-fà-kòng* 'long-hair-enthusiast, person extremely fond of long hair' (Ma 2016). This item is a phonetic adaptation of the Japanese suffix コン *-con*, which in turn originates from English *complex*, i.e. 'a group of attitudes and feelings that influence a person's behaviour, often in a negative way' (Cao, Mo 2012). In order to render affixes with no equivalent in Chinese, mostly lexical morphemes whose meaning is roughly similar or which are (quasi-)homophonous are chosen: when the number of words created by means of these morphemes increases, they gradually begin to assume a more general meaning.

In other cases, affixes may develop from a phonetic adaptation of a foreign word: (part of) a loanword may undergo a grammaticalisation process leading to an affix. One such example is 吧 *bā*, phonetic part of the hybrid 酒吧 *jiǔ-bā* 'alcohol-bar, bar', defined by *The Contemporary Chinese Dictionary* (2002) as "bar; counter at which alcoholic beverages are served in a Western-style restaurant or hotel". After the acceptance of this loanword, many complex words containing 吧 *bā* as the right-hand constituent have been created, as e.g 水

<sup>5</sup> According to Arcodia (2011), the grammaticalisation process undergone by 性 *-xìng* is not fundamentally different from that leading to the English suffix *-hood* (< Old English *-hād*), as in e.g. *childhood, falsehood*. Originally a Germanic name meaning 'person, sex, condition, rank, quality', it has become a suffix forming nouns of condition or quality, or indicating a collection or group, from nouns and adjectives.



吧 *shuǐ-bā* ‘water-bar’ (a place where mostly soft drinks are served), 氧吧 *yǎng-bā* ‘oxygen-bar’ (a place where oxygen masks are available for customer usage), 网吧 *wǎng-bā* ‘internet-bar, internet café’ (see Arcodia 2011, 125-7). In the 新华新词语词典 *Xinhua xinciyu cidian* (Xinhua Dictionary of Neologisms, 2003), 吧 *bā* is listed with the following meaning: “broadly indicates an entertainment place with a particular function or supplied with some special equipment” (Arcodia 2011, 121). According to Arcodia (2011), 吧 *bā* underwent a further generalisation of meaning and does not indicate specifically an entertainment place, as e.g. 创意吧 *chuàngyì-bā* ‘creativity-bar’, a kind of enterprise in the field of business consulting, or 话吧 *huà-bā* ‘talk-bar’, basically a call shop. Thus, the starting point is a process of analogy, and then 吧 *bā* begins to be associated with more lexemes: drinks and food, other services (see e.g. 氧吧 *yǎng-bā* ‘oxygen-bar’ above), and then all sorts of meeting places (including virtual ones), where one can play games (e.g. 游戏吧 *yóuxì-bā* ‘game-bar, amusement arcade’), exchange information on a topic (e.g. 贴吧 *tiē-bā* ‘paste-bar, webpage where fans publish posts related to their idols’, lit. ‘post bar’), or even provide consulting or information for a charge (e.g. the above-mentioned 创意吧 *chuàngyì-bā* ‘business consulting service’). According to Arcodia (2011, 126), metaphor is at work here: the meaning of ‘bar’ is extended to include any place which can be associated with the defining features of a bar. He also stresses the fact that this does not mean that 吧 *bā* has become a suffix with a pure locative meaning, since the connection with the original lexical meaning is always present somehow.<sup>6</sup>

In short, affixes in Chinese may develop through a grammaticalisation process inner to the language, due to the influence of foreign languages, or due to a combination of both: word-formation patterns already attested in the language may become productive due to the necessity to introduce foreign words; or, also, loanwords may develop an affixal use over time.

In what follows, after describing the three patterns, including their formal and semantic properties, we will focus on their development, and we will argue for their affixal status.

<sup>6</sup> For further details on the development and meanings of 吧 *bā*, the reader is referred to Arcodia 2011.

### 3 Description of the Three Word-Formation Patterns

#### 3.1 X-族 zú Words

As mentioned in the introduction, X-族 zú is a quite established and widely studied word-formation pattern in Chinese. The original meaning of 族 zú is ‘clan, tribe, ethnic group’, and it is still used with this meaning in compound words, as e.g. 大族 dà-zú ‘big-clan, famous and influential clan’, 白族 bái-zú ‘Bai-group, Bai minority’. In the last decades, this root has also developed a more generic meaning, i.e. ‘a category/group of people with common characteristics or behaviour’ (see Zhao 2009), appearing in a fixed position (right-hand constituent) in complex words, as e.g. 星空族 xīng-kōng-zú ‘star-sky-ZU, night workers’ (XCY, LWC), 网购族 wǎng-gòu-zú ‘net-purchase-ZU, those who love purchasing goods online’ (XCY, LWC), 候鸟族 hòuniǎo-zú ‘migratory.bird-ZU, the commuters’ (LWC).

This use of 族 zú originates as a loan from Japanese 族 zoku ‘a group of people with similar feelings or passions’; it was introduced to mainland China through Taiwan and Hong Kong (Cao 2007; Xiao 2009; Zhao 2009; Chen, Zhu 2010; Li 2013). According to Cao (2007), in Chinese it was originally used to indicate ‘a category of things with shared characteristics or properties’, as in 水族 shuǐ-zú ‘water-ZU, aquatic animals’, and later developed the above-mentioned meaning ‘a category/group of people with common characteristics or behavior’ (Cao 2007; Lu 2010). The first words containing 族 zú with this broader meaning, coined in the early Nineties, are 上班族 shàng-bān-zú ‘go-work-ZU, office workers’, 追星族 zhuī-xīng-zú ‘follow-star-ZU, groupies’, and 打工族 dǎgōng-zú ‘have.a temporary.job-ZU, those having a temporary or casual job’ (Cao 2007; Yang, Chen 2012). Due to their use on the web and in the media, these words became widespread and increased in number over the years, as can be seen from the number of distinct X-族 zú words found in the newspaper 人民日报 *Renmin ribao* (*People’s Daily*) between 1995 and 2005, shown in table 1 (Cao 2007, 151).

**Table 1** X-族 zú words in the 人民日报 *Renmin ribao* between 1995 and 2005

1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
19	15	23	33	15	33	23	27	41	36	44

Thus, in short, X-族 zú developed into a word-formation pattern indicating different groups of people who have something in common: fans/people who love something (much like 控 kòng seen in § 2 above), as 朋克族 péngkè-zú ‘punk-ZU, punk lovers’ (LWC), 哈韩族 hā-Hán-

zú 'adore-<sup>7</sup>Korea-ZU, those who love Korean music, TV, clothes etc.' (LWC), or 爱车族 ài-chē-zú 'love-car-ZU, car lovers' (XCY, LWC); those who are addicted to something, as e.g. 爱邦族 ài-bāng-zú 'love-Lianbang.syrup-ZU, those who are addicted to the cough syrup Lianbang (联邦止咳露 Liánbāng zhǐkē lù)' (XCY), 偷菜族 tōu-cài-zú 'steal-vegetables-ZU, those addicted to online games like Happy farm (开心农场 Kāixīn nóngchǎng) or Farmville',<sup>8</sup> 点赞族 diǎn-zàn-zú 'like-<sup>9</sup>ZU, like-clicking addicted', i.e. those who always click the 'like' button, e.g. on Facebook (XCY); users of various means of transport, such as 单车族 dānchē-zú 'bicycle-ZU, cyclists' (LWC), 地铁族 dìtiě-zú 'subway-ZU, those who use the subway' (LWC); workers, such as 办公族 bàngōng-zú 'work (in an office)-ZU, people who work in an office (LWC), 星空族 xīng-kōng-zú 'star-sky-ZU, night workers' (LWC), 陪逛族 péi-guàng-zú 'accompany-stroll-ZU, personal shoppers' (SD); those who share ideals/views/lifestyles etc., as e.g. 养生族 yǎngshēng-zú 'keep.in.good.health-ZU, health-conscious people' (LWC), 慢活族 màn-huó-zú 'slow-live-ZU, those who follow a slow living lifestyle' (XCY, SD), 素食族 sù-shí-zú 'vegetarian-food-ZU, the vegetarians' (XCY); people with a particular behaviour in common or engaged in certain activities (people who often do something or who like to do something), as 手机夜游族 shǒujī-yè-yóu-zú 'mobile.phone-night-travel-ZU, those who use mobile phones in bed before sleeping' (SD), 蹭网族 cèng-wǎng-zú 'freeload-net-ZU, Wi-Fi squatters' (SD),<sup>10</sup> 晒密族 shài-mì-zú 'show-secret-ZU, those who reveal their secrets on the web' (SD); people with some characteristics in common, as e.g. 肥腿族 féi-tuǐ-zú 'fat-leg-ZU, girls with fat legs' (LWC), 榴莲族 liúlián-zú 'durian-ZU, ill-tempered co-workers who have been working for many years and are hard to get along with, just like the smelly fruit with thick thorny skin' (SD),<sup>11</sup> 向日葵族 xiàngrikuí-zú 'sunflower-ZU, people who, just like a sunflower, always look on the bright side of life and are resilient to pressure as they easily forget about unhappiness' (XCY, SD, LWC).<sup>12</sup>

7 The character 哈 *hā* is often used in Taiwan with the meaning of 'worship, adore': <https://bit.ly/39kDjSa>.

8 They are virtual farms, where you play the role of a farmer who plants and harvests crops. Players can sneak into their friends' farms and steal vegetables.

9 点赞 *diǎn-zàn*, lit. 'click-praise', in the Internet slang indicates the 'like' button, used by the users to express that they like, enjoy, or support something.

10 It refers to those who linger in a public location to use its Wi-Fi internet connection, or who use such a connection without authorisation. Definition from *China Daily*: [https://language.chinadaily.com.cn/trans/2012-11/22/content\\_15951634.htm](https://language.chinadaily.com.cn/trans/2012-11/22/content_15951634.htm).

11 Definition from *China Daily*: [http://www.chinadaily.com.cn/dfpd/2011-08/22/content\\_13162619.htm](http://www.chinadaily.com.cn/dfpd/2011-08/22/content_13162619.htm).

12 Definition from *China Daily*: [https://language.chinadaily.com.cn/trans/2011-06/28/content\\_12793756.htm](https://language.chinadaily.com.cn/trans/2011-06/28/content_12793756.htm).

A few base words for X-族 *zú* neologisms are phonetic adaptations, like 辣奢族 *lāshē-zú* 'luxury-zu, fans of luxury goods' (XCY, SD), 飞特族 *fēitè-zú* 'freeter-<sup>13</sup>zu, those who work only when they feel they need some money (having a work schedule more flexible than freelancers)' (SD). The base word may be also a phonetic adaptation of an acronym, as e.g. 丁克族 *dīngkè-zú* 'DINK-<sup>14</sup>zu, young couples in big cities without children' (LWC). Sometimes the written form of the phonetic adaptation contains meaningful rather than neutral characters, as e.g. in 乐活族 *lè-huó-zú* 'happy-live-zu, those following LOHAS',<sup>15</sup> where the characters chosen somehow convey the meaning of the acronym the phonetic adaptation refers to. In some cases, the base of such neologisms are direct loans, as e.g. *Emo*族 *zú* 'Emo people', including acronyms and initialisms, such as DIY族 *zú* 'DIY (do it yourself)-zu, DIY lovers'. There are also a couple of words whose base is a single Latin letter, which stands for an acronym/initialism, as e.g. T族 *zú*, which refers to Chinese students who want to study abroad and must pass the TOEFL (Test of English as a Foreign Language). Sometimes X-族 *zú* words are calques from English, as e.g. 食男族 *shí-nán-zú* 'eat-male-zu, maneaters' and 游族 *yóu-zú* 'game-<sup>16</sup>zu, gamers' (LWC). There are also graphic loans from Japanese, as e.g. (御)宅族 (*yù*)*zhái-zú* 'nerd (< Jap. *otaku*)-zu, nerds, geeks' (XCY, SD, LWC).

In addition, the variant 一族 *yīzú* (lit. 'one group') is attested as well, as e.g. 微博一族 *Wēibó-yīzú* 'Weibo users': we found 31 types in our corpus. Out of 31 neologisms ending in 一族 *yīzú*, 20 do not display the corresponding X-族 *zú* form, as e.g. 哈哈一族 *hāhā-yīzú* 'Harry Potter lovers' (XCY), while 11 appear both in the 一族 *yīzú* and in the 族 *zú* form, as e.g. 拇指族 *mǔzhǐ-zú* and 拇指一族 *mǔzhǐ-yīzú* 'thumb-(YI)ZU, young people who use text messages as main means of communication' (XCY, SD, LWC), or 上网族 / 上网一族 *shàng-wǎng-zú* / *shàng-wǎng-yīzú* 'go-web-zu, web users', apparently with the same meaning (but see fn. 38).

### 3.2 X-党 *dǎng* Words

Complex words containing 党 *dǎng* as the right-hand constituent, indicating groups of people with common characteristics or behaviour, started to appear in the late 2000s, thus X-党 *dǎng* is a quite novel pattern of word formation. This pattern is typical of the web but is occasionally attested in other media as well (Chen, Zhu 2010).

13 From English *free* and German *Arbeiter* 'worker'.

14 DINK is the acronym of *dual income, no kids*.

15 LOHAS is the acronym of *Lifestyles of Health and Sustainability*.

16 游 *yóu* stands for 游戏 *yóuxì* 'game'.

The original meaning of 党 *dǎng* is ‘political party, clique’, and with this meaning it is found in complex words such as 党员 *dǎng-yuán* ‘party-member, party member’, 黑手党 *hēi-shǒu-dǎng* ‘black-hand-clique, mafia, gang’. As the right-hand constituent in complex words, it also developed the meaning ‘a group/category of people with common interests and characteristics or behaviour’, much like the morpheme 族 *zú*: e.g. 剧透党 *jùtòu-dǎng* ‘spoiler-DANG, people who like to spoil (films etc.)’ (LWC). Chen and Zhu (2010) argue that this use of 党 *dǎng* derives from the meaning ‘clique’, which has a strong derogatory sense. However, they point out that after the morpheme acquired the meaning of ‘political party’ (from Japanese 党 *tō*, e.g. in 国民党 *kōkumintō* ‘Chinese Nationalist Party’), especially after the foundation of the Chinese Communist Party, it started to have a positive connotation: this new meaning contributed to ‘lighten’ the derogatory sense connected to ‘clique’.<sup>17</sup>

Among neologisms containing 党 *dǎng*, we find words indicating different groups of people with something in common, as e.g.: people with a particular behaviour or habit in common, such as 自拍党 *zìpāi-dǎng* ‘selfie-DANG, people who take a lot of selfies’ (LWC), 游戏党 *yóuxì-dǎng* ‘game-DANG, those who play online videogames’ (LWC), 睡衣党 *shuìyī-dǎng* ‘pyjamas-DANG, those who go out in pyjamas’ (LWC), 早起党 *zǎo-qǐ-dǎng* ‘early-wake.up-DANG, the early risers’ (LWC), 格格党 *gégé-dǎng* ‘princess(a loan from Manchu)-DANG, Chinese girls born after 1985 who do not take their work seriously, do not obey their superiors, are arrogant, pay too much attention to their own needs without understanding those of other people, thus being incompatible with traditional jobs’ (XCY, SD, LWC); people addicted to something or who like something very much, be it a videogame, a sport, a musical genre, a dressing style, an instrument, or a brand, as e.g. 手机党 *shǒujī-dǎng* ‘mobile.phone-DANG, mobile phone addicted’ (LWC), 剁手党 *duò-shǒu-dǎng* ‘chop-hand-DANG, online shopaholics (example (1b)), 甘党 *gān-dǎng* ‘sweet-DANG, sweet lovers’, 爱凤党 *àifèng-dǎng* ‘Iphone-DANG, Iphone lovers’ (LWC), where 爱凤 *àifèng* is a phonetic adaptation; people sharing some particular characteristics, such as 白意党 *bái-yì-dǎng* ‘pure-intention-DANG, the sentimental’ (LWC), 无聊党 *wúliáo-dǎng* ‘bored-DANG, the bored’ (LWC), 一见钟情党 *yī-jàn-zhōngqíng-dǎng*, one-see-fall.in.love-DANG, those who fall

<sup>17</sup> Chen and Zhu (2010) highlight that in Japanese 党 *tō* has also the meaning ‘clique’, just like in Chinese, as e.g. in 凶党 *kyō-tō* ‘gang of partners in crime’ (lit. ‘evil/villain-clique’). Furthermore, it has also the meaning of ‘a group/category of people with common interests and characteristics’, much like in Chinese, but it has very low productivity. Some of the few examples that can be found are 烟党 *kemuri-tō* ‘smoke-TO, smokers’ (compare Chinese 抽烟党 *chōuyān-dǎng* ‘smoke-DANG’), and 甘党 *ama-tō* ‘sweet-TO, sweet lovers’ (compare Chinese 甘党 *gān-dǎng* ‘sweet-DANG’; 甜食党 *tián-shí-dǎng* ‘sweet-eat-DANG’).

in love at first sight' (LWC), 美丽党 *měilì-dǎng* 'beautiful-DANG, beautiful people' (LWC), or 苍白党 *cāngbái-dǎng* 'pale-DANG, people with little vitality and energy' (LWC).

Among these words indicating different types of people, there are also words originating from online buzzwords, such as 寂寞党 *jìmò-dǎng* 'lonely-DANG', i.e. web users who often use the buzzword (哥)... 的不是..., 是寂寞 (*gē...de bù shì..., shì jìmò*) 'what X is Y-ing is not Z, it is loneliness' (XCY, SD, LWC).<sup>18</sup>

In addition, just like 族 *zú*, 党 *dǎng* too can form neologisms which indicate certain types of workers, such as 上班党 *shàng-bān-dǎng* 'go-work-DANG, office workers' (LWC; compare the above-mentioned 上班族 *shàng-bān-zú* 'go-work-ZU, office workers'), and 配音党 *pèiyīn-dǎng* 'dub-DANG, dubbers' (LWC).

Chen and Zhu point out that 族 *zú* and 党 *dǎng* as the right-hand constituents of complex words indicating people with common characteristics or behaviour are actually interchangeable, i.e. they can attach to the same base without any apparent change in meaning: see e.g. 熬夜族 *áo yè-zú* 'stay.up.late-ZU' / 熬夜党 *áo yè-dǎng* 'stay.up.late-DANG', both indicating 'those who stay up late or all night'. However, Chen and Zhu (2010) observe that the oldest words containing 族 *zú* are generally not found in the corresponding X-党 *dǎng* form. In addition, after becoming an established pattern, X-族 *zú* words lost their novelty; at the same time, X-党 *dǎng* words started to appear on forums, becoming more and more widespread and replacing X-族 *zú* words as the most popular way to indicate groups of people with common interests, characteristics, or behaviour. Through a Baidu search, Chen and Zhu show that between 2008 and 2009 党 *dǎng* was the most used formative for words referring to groups of people: they considered the frequency of X-党 *dǎng* and X-族 *zú* words formed with the same base, showing that the X-党 *dǎng* pattern is the most frequently used for recent neologisms, while for older ('typical') words it is rarely used (Chen, Zhu 2010, 67). Therefore, apparently the difference between the two items is that 族 *zú* is more established, while 党 *dǎng* is more recent, popular and fashionable, and

<sup>18</sup> This buzzword emerged in 2009 in the Chinese BBS community Baidu World of Warcraft forum: an user posted a low-resolution webcam image of a man eating noodles accompanied by the sentence 哥吃的不是面, 是寂寞 *gē chī de bù shì miàn, shì jìmò* 'what this brother is eating aren't noodles, but loneliness!'. Shortly after, other users on the forum began repeating this sentence with slight variations, giving rise to the template illustrated above, creating a series of parody images centred around the theme of loneliness, as e.g. 我呼吸的不是空气, 是寂寞 *wǒ hūxī de bù shì kōngqì, shì jìmò* 'what I am breathing is not air, is loneliness', 哥灌的不是水, 是寂寞 *gē guàn de bù shì shuǐ, shì jìmò* 'what (this brother) is pouring is not water, is loneliness', 我用的不是手机, 是寂寞 *wǒ yòng de bù shì shǒujī, shì jìmò* 'what I am using is not a mobile phone, is loneliness'. <https://baike.baidu.com/item/%E5%AF%82%E5%AF%9E%E5%85%9A>; <https://knowyourmeme.com/memes/loneliness-party-%E5%AF%82%E5%AF%9E%E5%85%9A#fnr1>.

it is mainly used on the web (we will return on this issue in §§ 4.4 and 5). A hint of the fact that 党 *dǎng* is perceived as more popular and fashionable is the significant presence in our corpus (53 out of 189 words, 28%) of X-党 *dǎng* words indicating fans of actors, singers, characters, books, TV series, comics etc., as e.g. 天使党 *tiānshǐ-dǎng* 'angel-DANG, fans of the Japanese anime television series *Angel beats!*' (LWC), 松井党 *Sōngjǐng-dǎng* 'Rena Matsui-DANG, fans of Rena Matsui (松井玲奈, Japanese actress and singer)'. In our corpus we did not find any X-族 *zú* words of this kind, with the exception of 哈哈一族 *hāhā-yīzú* 'Harry Potter lovers' (XCY), containing the variant 一族 *yīzú* (see § 3.1). Rather, among X-族 *zú* words we found some examples of fans/enthusiasts of a particular genre or category, like 朋克族 *péngkè-zú* 'punk-zu, punk lovers', 哈韩族 *hā-Hán-zú* 'adore-Korea-zu, those who love Korean music, TV, clothes etc.' (§ 3.1).

Furthermore, it must be noted that, among X-党 *dǎng* words, we find words referring to a series of illegal activities, which cannot be found among X-族 *zú* words, as e.g. 拎包党 *lǐnbāo-dǎng* 'bag-DANG, pickpockets' (SD), 撞车党 *zhuàng-chē-dǎng* 'collide-car-DANG, people who wilfully get hit by other cars to extort money from drivers' (LWC), 敲墙党 *qiāo-qióng-dǎng* 'knock-wall-DANG, a mafia-style group that forces people to rely on their companies when they need to renovate their properties' (LWC), and 黄牛党 *huánɡniú-dǎng* 'scalper-DANG, scalpers'. This negative nuance is apparent in the word 摩托党 *mótuō-dǎng* 'motorcycle-DANG, the motorcyclists' as well, which usually refers to gangs of motorcyclist disturbing public security etc.,<sup>19</sup> and not simply to people who ride a motorcycle. Therefore, in some cases, 党 *dǎng* retains to an extent the negative nuance of its original meaning 'clique' (see Chen, Zhu 2010; we will return to this issue in § 4.2). We believe that this is the source of the ambiguity displayed by some neologisms, which can have two different meanings: e.g. 狗党 *gǒu-dǎng* 'dog-DANG' can refer either to 'close friends' or to 'spies' (LWC). The latter meaning retains the negative nuance of the term 'clique'.

### 3.3 X-客 *kè* Words

The original meaning of the morpheme 客 *kè* is 'guest, traveller', and with this meaning it is found in compound words as e.g. 旅客 *lǚ-kè* 'travel-guest, hotel guest/traveller', 请客 *qǐng-kè* 'invite-guest, invite/entertain guests', 客车 *kè-chē* 'guest-vehicle, passenger train'.

However, in recent years it started to appear as the right-hand constituent of complex words indicating 'a person doing a certain activity' or 'a person with certain characteristics'. Arguably, the most pop-

19 <https://baike.baidu.com/item/%E6%91%A9%E6%89%98%E5%85%9A/18818561>.



ular of these complex words is 黑客 *hēi-kè* 'black-KE, hacker', which entered the Chinese lexicon in the late Nineties, as a phonetic-semantic adaptation of the English word *hacker*: the Chinese word approximately recalls the pronunciation of the source word; in addition, the left-hand constituent, 黑 *hēi* 'black, shady, illegal', conveys the negative meaning of the word (compare 黑车 *hēi-chē* 'black-vehicle, illegal taxi, unlicensed motor vehicle'). This word-formation pattern has become popular starting from the beginning of the twenty-first century: according to Zhang and Xu (2008), with the spread and popularity of blogs (in Chinese 博客 *bókè* 'blog', also 'blogger') at the beginning of the 2000s, more and more X-客 *kè* words appeared, which, together with words already coined, like 黑客 *hēikè* 'hacker', contributed to form a word-formation pattern typical of the web.

Along with words indicating different kinds of 'hackers', such as 白客 *bái-kè* 'white-KE, online security guard; hacker-fighter', 红客 *hóng-kè* 'red-KE, patriotic hacker, defending the security of domestic networks and fending off attacks', 灰客 *huī-kè* 'grey-KE, unskilled hacker',<sup>20</sup> we find neologisms indicating persons engaged in different kinds of activities, such as 刷书客 *shuā-shū-kè* 'scan-book-KE, a person who record extracts from a book, either in a bookstore or in a library, with an electronic mini scanner, without any intention to buy it' (XCY), 换客 *huàn-kè* 'exchange-KE, one who sells/exchanges goods online'.

As Zhang and Xu (2008) point out, this word-formation pattern is typical of the web and was then extended to the media in general and to everyday language too, even though it is still mainly used by young people. Actually, many X-客 *kè* words belong to the domains of technology and the web, often indicating people doing some kind of activity online (38 out of 84 words in our corpus, almost half of the total); we will go back to this issue in § 4.3.

Among X-客 *kè* words, we find many neologisms which are phonetic adaptations: however, differently from what happens with 族 *zú* and 党 *dǎng*, generally speaking it is the whole complex word ending in 客 *kè* that is a phonetic adaptation (not just the base), as e.g. 极客 *jí-*

**20** Following Arcodia and Basciano (2018), we excluded 'hacker' words from our analysis, since they do not indicate 'a person doing a certain activity' or 'a person with certain characteristics' related to the base. Rather, they are best analysed as analogical formations (see Booij 2010) from 黑客 *hēi-kè* 'black-KE, hacker', where the modifier is invariably a colour term, which is always understood in a metaphorical rather than in a literal sense. An anonymous reviewer pointed out that the semantic mechanism at work could be similar to reductions observed in English words such as *cheeseburger* or *fish-burger*, where *burger* is the truncated form of *hamburger*, or also in Italian words like *auto-strada* 'car-road, motorway' or *auto-lavaggio* 'car-washing, car washing', where *auto* stands for *automobile* 'car'. Thus, in a word as 红客 *hóng-kè* 'red-KE, patriotic hacker' (lit. red hacker), 客 *kè* would be the truncated form of 黑客 *hēi-kè* 'hacker' (红(黑)客 *hóng-(hēi)-kè*). However, in our opinion analogy best explains these cases, since the modifier is always a colour term, which replaces 黑 *hēi* 'black' in 黑客 *hēi-kè* 'hacker', and is interpreted in a metaphorical sense, just like in 黑客 *hēi-kè*.



kè 'extremely-KE, geek', much like in the case of 黑客 *hēi-kè* 'hacker' (we will go back to this issue in § 4.4). Out of the 84 X-客 *kè* words collected from our sources, 15 (17.86%) are phonetic adaptations of this kind, i.e. the whole complex word is a phonetic adaptation. It must be noted, though, that 客 *kè* is not just a component of the phonetic adaptation, but is also the element which conveys the agentive meaning to the complex word. For example, 切客 *qiē-kè* 'cut-KE, fan of location-based services who regularly checks in to keep friends and relatives posted on her/his whereabouts' is a phonetic adaptation of English *check-in*; however, the word indicates a person, and this meaning is conveyed by the morpheme 客 *kè*. The same goes for the word 粉飞客 *fěn-fēi-kè* 'fan-<sup>21</sup>fly-KE, fanfictioner (fan who likes to write sequels or change plots of TV series to express her/his ideas, passions etc.)', which is a phonetic adaptation of English *fanfic*: besides recalling the pronunciation of the last part of the word, 客 *kè* also conveys the meaning of 'person'; as a matter of fact, the whole word means *fanfictioner*, not *fanfic*. Therefore, in these cases the X-客 *kè* word indicates a person involved in an activity connected to the meaning of the phonetic adaptation as a whole ('a person doing an activity connected to X-客 *kè*', not 'a person doing an activity connected to X'), where 客 *kè* is part of the phonetic adaptation but, at the same time, contributes the meaning of 'person'.

In addition to these cases, we also found 4 complex words (4.76%) the base of which is a phonetic adaptation, as e.g. 秀客 *xiù-kè* 'show-KE' (秀 *xiù* is the phonetic adaptation of *show*), which refers to those who share videos from the e-commerce platform 秀兜 *Xiūdōu* on their Weibo, among their friends (they receive a fee from the platform every time someone clicks on their sponsored links and then completes the purchase). All in all, we can observe that the proportion of phonetic adaptations among X-客 *kè* words is much higher than among X-族 *zú* (13 out of 434, about 3%) and X-党 *dǎng* words (10 out of 189, 5.29%). We will return to the possible motivations for this in § 4.3.

Besides phonetic adaptations, we also find calques and hybrid forms, as e.g.: 追客 *zhuī-kè* 'follow-KE, someone who regularly refreshes web pages to follow the latest updates of online series, TV series, bloggers, or podcasts', which looks like a calque of English *follower* (追 *zhuī* translates *follow*, and 客 *kè* is roughly equivalent to *-er*); 创客 *chuàng-kè* 'create-KE, maker', which can be regarded as a hybrid, where 创 *chuàng* translates *make*, while 客 *kè* acts as the equivalent of the suffix *-er* and, at the same time, recalls the pronunciation of the last part of the word *maker*.

However, X-客 *kè* words are not limited to loans and words connected to the Internet and new technologies; the X-客 *kè* pattern is

21 粉 *fěn* 'powder' stands for 粉丝 *fěnsī*, phonetic adaptation of the English word *fans*.

also used to coin words indicating persons involved in all sorts of different activities or having certain characteristics, as e.g. 必剩客 *bì-shèng-kè* 'certainly-remain-KE, a person above the typical marriage age but still single, considered to be doomed to remain unmarried', 代扫客 *dài-sǎo-kè* 'take.the.place.of-sweep-KE, a person who offers a service consisting in visiting tombs (sweeping and offering sacrifices) during the Qingming festival' (XCY), 排客 *pái-kè* 'line.up-KE, a person paid to stand in a queue for others', 帕客 *pà-kè* 'handkerchief-KE, a green consumer who prefers to use handkerchiefs instead of throw-away paper tissues in support of low-carbon life'<sup>22</sup> (LWC). However, even when X-客 *kè* words are not nouns connected to the Internet and new technologies, the role of the web in their creation and diffusion is apparent, at least for part of them. Take for example the word 帕客 *pà-kè* just mentioned above: it became popular after one of China's online messaging service providers launched a handkerchief design campaign in 2009 to encourage the use of handkerchiefs to protect the environment; the winner was called 帕客 *pà-kè* 'handkerchief-KE'.<sup>23</sup>

All in all, it can be stated that the morpheme 客 *kè* as the right-hand constituent of complex words has acquired a more general meaning, appearing in a fixed position, indicating various kinds of persons, with a function roughly comparable to that of English *-er* (Arcodia, Basciano 2018).

### 3.4 Are X-族 *zú* and X-党 *dǎng* Words Collective Nouns?

The three morphemes at issue, as we have seen, have apparently acquired a more general meaning, appearing in a fixed position (to the right of complex words), indicating various kinds of persons. At a first look, it would seem that 客 *kè* forms individual nouns, while 族 *zú* and 党 *dǎng* form collective nouns, thus preserving part of their original meaning, as suggested by the following examples:

2. a. 刷书客  
*shuā-shū-kè*  
scan-book-KE  
'a person who scans with a mini-scanner the content from the books in a bookstore or a library'

<sup>22</sup> [http://language.chinadaily.com.cn/trans/2010-02/21/content\\_9480739.htm](http://language.chinadaily.com.cn/trans/2010-02/21/content_9480739.htm).

<sup>23</sup> [http://language.chinadaily.com.cn/trans/2010-02/21/content\\_9480739.htm](http://language.chinadaily.com.cn/trans/2010-02/21/content_9480739.htm).

- b. 刷书族  
*shuā-shū-zú*  
scan-book-ZU  
'people who scan with a mini-scanner the content from the books  
in a bookstore or a library'

In (2), we have two words differing only for the right-hand constituent used, i.e. 客 *kè* or 族 *zú*. The only difference in meaning between the two words seems to be individual vs collective. The X-族 *zú* term, thus, apparently denotes a collective whole, a (semantic) plurality ('more than one') obtained by grouping together a number of entities, which share a part-whole relation (see Gardelle 2019). This is further suggested by the fact that 客 *kè* and 族 *zú* may combine in the same word. See the following examples:

3. a. 换客  
*huàn-kè*  
exchange-KE  
'one who sells/exchanges goods online'
- b. 换客族  
*huàn-kè-zú*  
exchange-KE-ZU  
'those who sell/exchange goods online'
4. a. 晒客  
*shài-kè*  
expose-KE  
'a person who shares his experiences and thoughts with others on  
the Internet'
- b. 晒客族  
*shài-kè-zú*  
expose-KE-ZU  
'those who share their experiences and thoughts with others on  
the Internet'

However, a closer look at the data reveals a different picture: X-族 *zú* nouns can apparently refer to members of the group rather than the group as a whole, as in the following examples.<sup>24</sup>

---

**24** In these examples, the plural classifier 些 *xiē* (i.e. the only plural classifier available in Chinese) is used. 些 *xiē* is never used in counting; it combines with the demonstratives 这 *zhè* 'this' or 那 *nà* 'that', resulting in 'these' and 'those', or with the numeral 一 *yī* 'one', leading to the indefinite meaning 'some' (cf. Eng. *a few, a couple of, a num-*

5. a. 现在,人们的工作节奏较快,对一些上班族来说,下班之后想自己做顿像样的晚餐成了一种奢望 [...]
- xiànzài rén-men de gōngzuò jíezuò jiào kuài duì yī*  
 now person-PL SP work rhythm quite fast for one  
**xiē shàng-bān-zú** *lái shuō xiàbān zhīhòu xiǎng*  
 CFL<sub>PL</sub> go-work-ZU concerning finish.work after want  
*zìjǐ zuò dùn xiàngyàng de wǎncān chéng le*  
 oneself make CFL decent SP dinner become PFV  
*yī zhǒng shē-wàng*  
 one CFL extravagant-hope  
 ‘Nowadays, the working rhythm of people is quite fast. For **some office workers**, preparing a decent meal for themselves after work has become an extravagant hope [...].’<sup>25</sup>
- b. [...]到了双休日那些爱运动的上班族都来了 [...]
- dào le shuāng-xiū-ri nà xiē ài yùndòng de*  
 arrive PFV double-rest-day that CFL<sub>PL</sub> love sports SP  
**shàng-bān-zú** *dōu lái le*  
 go-work-ZU all come PFV  
 ‘[...] In the weeks with two rest days, all **the/those office workers who love sports** came [...].’<sup>26</sup>

This is observed with X-党 *dǎng* nouns as well:

6. a. [...]全国各地都有一些睡衣党出没。
- quán-guó-gè-dì dōu yǒu yī xiē shuìyī-dǎng*  
 whole-country-each-place all have one CFL<sub>PL</sub> pyjamas-DANG  
*chūmò*  
 come.and.go  
 ‘[...] everywhere in the country there are **some people who go out in pyjamas** [...].’<sup>27</sup>
- b. 对于这些熬夜党,尤其是女性熬夜党来说,护肤尤为重要。<sup>28</sup>
- duìyú zhè xiē áo yè-dǎng yóuqí shì nǚxìng*  
 for this CFL<sub>PL</sub> stay.up.late-DANG especially be woman

*ber of*; Sybesma 2017). According to Ilijc (1994), 些 *xiē* is a collective marker, referring to wholes, rather than a plural marker.

<sup>25</sup> [http://www.peopledailynews.eu/sp/20190417\\_57656.html](http://www.peopledailynews.eu/sp/20190417_57656.html).

<sup>26</sup> [https://hznews.hangzhou.com.cn/xinzheng/quxian/content/2010-06/23/content\\_3327660.htm](https://hznews.hangzhou.com.cn/xinzheng/quxian/content/2010-06/23/content_3327660.htm).

<sup>27</sup> <https://kknews.cc/zh-my/news/ebbe42z.html>.

<sup>28</sup> [https://k.sina.com.cn/article\\_7026285403\\_1a2cc9b5b00100saxq.html?from=fashion](https://k.sina.com.cn/article_7026285403_1a2cc9b5b00100saxq.html?from=fashion).

áo-yè-dǎng      láishuō      hùfū      yóuwéi      zhòngyào  
 stay.up.late-DANG    concerning    skincare    particularly    important  
 ‘For **these people who stay up late**, especially for women, skin-care is particularly important’.

Besides, it must be noted that both X-族 *zú* words and X-党 *dǎng* words may be followed by the plural / collective suffix 们 *-men*:

7. a. 眼看着本月底地铁4号线就将推行“禁食令”,本市不少“地铁快餐族”们同样提出了自己的质疑  
 yǎnkànzhe      běn yuè-dǐ      dìtiě      sì hào      xiàn  
 watch.helplessly this month-end subway 4 number line  
 jiù      jiāng      tuīxíng      jìn      shí      lìng      běn shì  
 then will carry.out forbid eat decree this city  
 bùshǎo      **dìtiě-kuài-cān-zú-men**      tóngyàng      tíchū      le      zìjǐ  
 many subway-fast-food-ZU-PL same pose PFV oneself  
 de      zhíyí  
 SP call.into.question  
 ‘While watching helplessly that by the end of this month Line 4 of the subway will implement a ‘no eating decree’, many “**subway fast-food eaters**” in town called it into question’.
- b. [...] 酱油党们也因为在这片中露脸而找到了狂欢的理由 (LWC)  
 jiàngyóu-dǎng-men      yě      yīnwèi      zài piàn      zhōng      lùliǎn  
 soy.sauce-DANG-PL also because at film in appear  
 ér      zhǎodào      le      kuánghuān      de      lǐyóu  
 and find PFV revel SP reason  
 ‘**Those who feign ignorance**<sup>29</sup> too found a reason to revel because they appeared in the film’.

According to Li and Thompson (1981, 40), the suffix 们 *-men* is generally used only when there is some reason to emphasise the plurality of the noun. According to others (e.g. Iljic 1994; Cheng, Sybesma 1999), it is a collective rather than a plural marker. Iljic (1994, 96), for example, points out that “[t]he speaker resorts to *men* whenever he has grounds to view several persons as a group, either relative to himself or relative to a third party”. The function of this suffix, then, would be to group different units, to construct a group from several elements. According to

<sup>29</sup> From 打酱油 *dǎ jiàngyóu* ‘it’s none of my business; it has nothing to do with me’ (orig. ‘buy soy sauce’). This meaning developed from a buzzword: in 2008 the Guangzhou Broadcasting Network interviewed a local man about the Edison Chen (a celebrity from Hong Kong) photo scandal, who answered: “关我鸟事, 我出来打酱油的 *guān wǒ niǎo shì, wǒ chūlái dǎ jiàngyóu de*” (it’s none of my business / what the f\*\*k does it have to do with me? I was just out buying soy sauce). This answer then became a meme, applicable to any context: [https://chinadigitaltimes.net/space/Get\\_soy\\_sauce](https://chinadigitaltimes.net/space/Get_soy_sauce).

Cheung (2016), count nouns suffixed with 们 *-men* can be used to refer to a group of people that are known to both speakers and hearers. As a matter of fact, they are regularly used as a term of address in gatherings, as e.g. 女士们、先生们 *nǚshìmen, xiānshēngmen* 'ladies and gentlemen'.

Therefore, the co-occurrence of 族 *zú* and 党 *dǎng* with the suffix 们 *-men* would be unexpected if they were simply used to form collective nouns (which involve the gathering of a plurality of entities, specifically a group), unless 们 *-men* is seen just an emphatic marker (i.e. if it is used to emphasise collectivity). If the function of 们 *-men* is to group several entities, we should then conclude that X-族 *zú* and X-党 *dǎng* nouns in these contexts refer to members of the group, rather than to the collective whole.

In addition, individuation may be observed in yet other contexts: X-族 *zú* and X-党 *dǎng* nouns can combine with sortal classifiers<sup>30</sup> (or individual classifiers, Peyraube 1998) used for humans, and individual members can be counted. See the following examples, where X-族 *zú* nouns clearly indicate single entities, and not the collective whole:<sup>31</sup>

8. a. 背上旅行包, 带上相机, 做个背包族, 继续我的浪子情怀。(LWC)  
*bèi-shàng lǚxíng-bāo dài-shàng xiàngjī zuò ge bēi-bāo-zú*  
back-on travel-bag bring camera be CLF back-pack-ZU  
*jìxù wǒ de làngzǐ qínghuái*  
continue 1SG SP wastrel mood  
'Carrying my luggage on the shoulder, taking the camera with me,  
**being a backpacker**, carrying on my nomad spirit'.
- b. 粗略统计, 3分钟内竟出现40个“车缝族”。(XCY)  
*cūlüè tǒngjì sān fēnzhōng nèi jìng chūxiàn*  
rough statistics 3 minute inside actually appear  
*sìshí ge chē-fèng-zú*  
40 CLF vehicle-crack-ZU  
'With a rough estimate, in 3 minutes **40 jaywalkers** appeared'

<sup>30</sup> As pointed out by Croft (1994), sortals simply name the unit that is already present in the semantic denotation of the noun, while measures create a unit by which we can count or measure; they include real measures (kilo, mile), containers (cup, spoon), and collectors (group, mass). Measures carry their own, noun-independent semantics, as confirmed by the fact that they can be used with count nouns and mass nouns alike (Sybesma 2017). Chinese sortal classifiers represent a closed class, and each classifier combines with a set of nouns that can be seen to belong to one and the same class. Classifiers are compulsory with numerals, i.e. there is no counting without a classifier, so that they are often referred to as numeral classifiers (Sybesma 2017).

<sup>31</sup> In the examples, we observe the use of the classifiers 个 *ge*, used for all humans (regardless of sex, age, social status, occupation etc.), and the honorific classifier for people 位 *wèi*. Actually, 个 *ge* is also used as a generic classifier for nouns lacking more specific sortals, or even as a 'default' - speakers often use it with nouns that combine with another sortal according to prescriptive grammar (see Sybesma 2017).

- c. 我只是一个朝九晚五的上班族 (Zhao 2009, 36)  
*wǒ zhǐ shì yī gè zhāo jiǔ wǎn wǔ de*  
 1SG only be one CLF morning nine evening five SP  
**shàng-bān-zú**  
 go-work-ZU  
 'I am only a **9- to 5-er**' (translation provided by the source)
- d. 一位27岁的上班族写完一首诗后跳楼自杀了[...]<sup>32</sup>  
*yī wèi èrshíqī suì de shàng-bān-zú xiě-wán yī*  
 one CLF 27 year SP go-work-ZU write-finish one  
*shǒu shī hòu tiào-lóu zì-shā le*  
 CLF poem after jump-building self-kill PFV  
 'A **27-year-old office worker** jumped to his death from a building after finish writing a poem [...]'

X-党 *dǎng* words too are attested in numeral-classifier constructions like the ones above, as in the following examples (see also Chen, Zhu 2010):

9. a. 群里面就我一个电脑党 (LWC)  
*qún lǐmiàn jiù wǒ yī ge diànnǎo-dǎng*  
 group inside only 1SG one CLF computer-DANG  
 'I am the **only computer expert** of the group'
- b. [...] 这份调查报告研究了南京1840位“剁手党” [...]  
*zhè fèn diàochá bàogào yánjiū le Nánjīng yīqiānbābāisìshí*  
 this CLF survey report study PFV Nanjing 1840  
**wèi duò-shǒu-dǎng**  
 CLF cut-hand-DANG  
 '[...] this survey studied 1840 online shopaholics in Nanjing [...]'<sup>33</sup>

Further examples where X-族 *zú* and X-党 *dǎng* nouns are used to indicate individuals rather than groups are the following ones, where a member-class/category relationship is displayed: the X-族 *zú* and X-党 *dǎng* nouns represent a class/category indicating the nature of the individuals (see § 4.1):

10. a. 你是御宅族吗? (LWC)  
*nǐ shì yùzhái-zú ma*  
 2SG be otaku-ZU Q  
 'Are you an otaku (nerd)?'

<sup>32</sup> [https://3g.163.com/dy/article\\_cambrian/EIU2S9G10544809Y.html](https://3g.163.com/dy/article_cambrian/EIU2S9G10544809Y.html).

<sup>33</sup> [http://china.cnr.cn/qqhygbw/20160123/t20160123\\_521212278.shtml](http://china.cnr.cn/qqhygbw/20160123/t20160123_521212278.shtml).

- b. 可怜我这熬夜党每晚只睡三小时 (LWC)  
*kělián wǒ zhè áoyè-dǎng měi wǎn zhǐ*  
 poor 1SG this stay.up.late-DANG each evening only  
*shuì sān xiǎoshí*  
 sleep three hour  
 ‘Poor me, this night owl who sleeps only three hours per night’

But what about cases like those in (3) and (4), where 客 *kè* and 族 *zú* may combine in the same word, so that both the X-客 *kè* and the X-客族 *kèzú* version of a word are attested? Generally speaking, in those cases it seems that actually the X-族 *zú* word is not used to refer to individuals. As a matter of fact, X-客族 *kèzú* words, differently from X-客 *kè* words, are not generally used with a sortal classifier in numeral-classifier constructions:

11. a. 一个/位换客  
*yī ge/wèi huàn-kè*  
 one CLF exchange-KE  
 ‘an exchanger’
- b.<sup>??</sup> 一个/位换客族  
*yī ge/wèi huàn-kè-zú*  
 one CLF exchange-KE-ZU

However, both of them are apparently allowed with a measure numeral classifier, as e.g. the collector 群 *qún* ‘group’:

12. a. 在惠州,也有一群“换客”<sup>34</sup>  
*zài Huìzhōu yě yǒu yī qún huàn-kè*  
 in Huizhou also have one CLF<sub>group</sub> exchange-KE  
 ‘There is a **group of “exchangers”** in Huizhou too’
- b. [...] 并涌现出一群“换客族”<sup>35</sup>  
*bìng yǒngxiànd-chū yī qún*  
 and emerge.in.large.numbers-come.out one CLF<sub>group</sub>  
**huàn-kè-zú**  
 exchange-KE-ZU  
 ‘[...] and a large **group of “exchangers”** emerged’

<sup>34</sup> [http://news.ifeng.com/gundong/detail\\_2013\\_11/19/31364520\\_0.shtml](http://news.ifeng.com/gundong/detail_2013_11/19/31364520_0.shtml).

<sup>35</sup> <http://news.sina.com.cn/c/2011-05-16/112422472490.shtml>.



- c. 福州的这群“换客”们,带来的“宝贝”都不太多 [...] <sup>36</sup>  
Fúzhōu de zhè qún huàn-kè-men dàilái de bǎobèi  
Fuzhou SP this CLF<sub>group</sub> exchange-KE-PL bring SP treasure  
dōu bù tài duō  
all not too many  
'This group of "exchangers" in Fuzhou did not bring many 'treasures [...]'
- d. 这可不是在做梦,而是一群“换客族”们在网络交换平台上发出的召唤。 <sup>37</sup>  
zhè kě bù shì zài zuòmèng ér shì  
this actually NEG be PROG dream but be  
yī qún huàn-kè-zú-men zài wǎngluò jiāohuàn  
one CLF<sub>group</sub> exchange-KE-ZU-PL at Internet exchange  
píngtái shàng fāchū de zhàohuàn  
platform on issue SP call  
'This is not a dream, but the call issued by a group of "exchangers" in an exchange platform on the web'.

This is possibly due to the fact that for X-客族 kèzú nouns a less degree of individuation is licensed, and thus they can be used to refer to the members of the group but not to indicate a single entity; accordingly, they imply plurality. This issue requires further investigation.

In a nutshell, both 族 zú and 党 dǎng have undergone further extension of meaning, departing more from their original meaning – indicating a group –, and at present they can be used to refer to individuals. <sup>38</sup> As pointed out by an anonymous reviewer, similar cases of collective > individual metonymic shift are observed in different languages, as e.g. Spanish *policia* 'police': *un policia* 'a policeman' (lit. 'a police'). We will go back to this issue in § 4.1.

#### 4 On the Development of 族 zú, 党 dǎng and 客 kè

In the preceding section, we have shown that 族 zú, 党 dǎng and 客 kè appear in a fixed position, with a fixed meaning, building families of words indicating people doing certain activities or with shared characteristics or behaviour. Can they be labelled as suffixes then? In order to answer this question, in this section we will focus on the evolution of the three items at issue.

<sup>36</sup> <http://news.sina.com.cn/s/2006-11-21/010010551237s.shtml>.

<sup>37</sup> <https://baike.baidu.com/item/%E6%8D%A2%E5%AE%A2%E6%97%8F>.

<sup>38</sup> We may remark that 一族 yīzú (see § 3.1), despite bearing the same meaning as 族 zú, cannot refer to individuals (see Cao 2007; Lu 2010).

#### 4.1 The Evolution of 族 *zú*

As we have mentioned in § 3.1, 族 *zú* as an affix-like item originates from Japanese. Its original meaning of ‘clan, tribe, group’ developed into the affixal 族 *zoku* ‘a group of people with similar feelings or passions’, which was then imported in Taiwan, Hong Kong, and later in Mainland China, as we have seen. It then acquired the more generic meaning of ‘a category/group of people with common characteristics or behaviour’.

13. 族 *zú* ‘clan/ethnic group’ > a group of people with similar feelings or passions > a category/group of people with common characteristics or behaviour

Thus, it is evident that this item underwent a process of generalising abstraction, which involves taking a lexeme to a higher taxonomical level (Heine, Claudi, Hünnemeyer 1991; Arcodia 2011). This is confirmed by the fact that 族 *zú* can be used to indicate a variety of referents (see § 3.1): fans/people who love something, workers, people with a particular behaviour in common or engaged in certain activities, people with some characteristics in common etc. This can be seen as a process of grammaticalisation through metaphorical extension, with increased lexical generality and contextual expansion (see Arcodia 2011, 126-7); we argue that the different meanings conveyed by this item may all be subsumed under the meaning ‘a category/group of people with common characteristics or behaviour’. Given these characteristics, we maintain that 族 *zú* can be classified as a proper suffix: as pointed out by Arcodia (2011, 125-6),

[s]ince the meaning expressed by a derivational affix, a grammaticalised sign, may be very general, it is not surprising that it can be used to design a huge variety of referents, provided that it is still possible to identify the commonalities among the various instances.

Recall that in Chinese grammaticalisation usually does not display co-evolution of form and meaning, i.e. affixes are generally characterised by meaning generalisation but not by phonological reduction (see § 2).

In addition, we have shown that 族 *zú* underwent further meaning extension, and it is now used to indicate single entities as well (§ 3.4). This appears to be similar to the development of the suffix 家 *jiā*, which was first used to indicate a group (‘school of thought’), as e.g. 法家 *fǎjiā* ‘Legalists’ (\*一个法家 *yī ge fǎjiā* ‘one Legalist’), and then started to form individual nouns, with the meaning of ‘expert’, as e.g. 艺术家 *yìshùjiā* ‘artist’, 语言学家 *yǔyánxuéjiā* ‘linguist’ (一个艺术家 / 语言学家 *yī ge yìshùjiā / yǔyánxuéjiā* ‘an artist / a linguist’); see Wang ([1980] 2002, 230).

As we mentioned in § 3.4, this kind of metonymical semantic shift is not uncommon in the world's languages. Specifically, this metonymical pattern can be seen as an extension of the part-whole relationship to the domain of collections, i.e. sets of roughly equal members: for example, a swarm of bees is made up only of bees, thus it is a collection, because its parts are largely identical (Peirsman, Geeraerts 2006, 302). In collections, entities are conceived as relatively independent but still closely associated. Through this kind of metonymical pattern, a collective term can be used for one entity only, as e.g. in the case of German *Imme* 'bee' (single entity), which developed from Middle High German 'swarm of bees' (collection; Peirsman, Geeraerts 2006, 304). This phenomenon, as we mentioned in § 3.4, is observable in the polysemy displayed by some nouns in synchrony as well. Gardelle (2019, 112-19) observes that in English originally collective nouns, as e.g. *crew*, may come to mean "more than one member in a group", as in *these crew* (uninflected plural), or even, for some of them, "a member in a group" (*one crew*, in the sense of 'one member of a crew'). Another example is *police* used in the sense of 'policeman', as in *those police* 'those policemen', *two police* 'two policemen'.<sup>39</sup>

For these nouns of collective origin, Gardelle argues that the mechanism at work is 'type coercion', i.e. a rather unusual use of a word as regards its grammatical features (in this case, use as uninflected plural instead of singular count) (see Audring and Booij 2016). Gardelle (2019, 115-16) hypothesises that this kind of type coercion goes through three stages: 1) the noun has collective sense and takes grammatical agreement (*this crew has...*); 2) the noun, still having a collective sense, licenses semantic override agreement (foregrounding of the individuals) outside the NP, in the verb and in pronouns (*this crew have...they...*) - non-additivity is lost, and the predicates and anaphors only apply to the individuals; 3) uninflected lexical plural use (*these crew have...*). This plural denotes units, not a collective whole, "though they are expected to belong to a group of the kind denoted by the collective sense of the noun" (Gardelle 2019, 115). This is considered a type of coercion by Gardelle, since this sense is not freely accessible with all collective nouns that denote humans (*\*these*

---

**39** Gardelle (2019, 109-10) notes that the uninflected plural, meaning "more than one member in a group", is less individuated than the noun that names the separate units: she points out that *those police* is found in cases in which the police officers act together, react together, without any differentiation, while *those policemen* may be used in the same contexts or where there is individuation, as in "[i]t was directed to those policemen who kill and mistreat Blacks". Similarly, *two police* is found only in contexts of professional activities (arrests, or to count victims) - what matters is that they belong to the same socio-professional category -, while *two policemen* are found either in the same contexts or with a higher degree of individuation.

*committee*) and does not allow for free combination with determiners. As a matter of fact, Gardelle shows that the only determiners licensed by all the uninflected plurals of collective origin are plural demonstratives (these/those), while quantities (one, two, several) are acceptable only with a few nouns (e.g. *crew, police, faculty*). Gardelle observes that, semantically, conceptualisation with a demonstrative determiner only requires a very low degree of individuation of the units, if compared e.g. with quantities. This could explain why numerals are found only with some of these nouns. She further stresses that actual numerals seem to stand one step further in the evolution of these uses, since they are available only for some of the nouns examined: as for 'one' ('one member'), it is restricted to very few nouns (*clergy, crew, faculty, police, staff*), possibly due to potential referential ambiguity. Finally, the use of the indefinite article *a* is very rare. Gardelle argues that stage 3 is reached through plural uses (*these/those*); only at this point, for some nouns and to some speakers, more individuation may be licensed, including, ultimately, the singular.

Gardelle (2019, 116) points out that type coercion is accompanied by semantic coercion, from group to members, "as the loss of the /count/ feature entails a loss of boundedness at lexical level"; the noun becomes polysemous. The shift from the collective sense to the uninflected plural sense takes place at a notional level, from the notion of group to that of members; the uninflected plural denotes a class, a socio-professional category, "albeit one in which people are expected to be members of groups".

Gardelle (2019, 117) concludes that these uninflected plural nouns are not collective: the units do not stand in a part-whole relationship with the plurality (*\*crew are composed of crew/members/members of crew*). These nouns do not denote a collective whole but a class, indicating the nature of the individuals. Thus, they are characterised by a member-class relation (e.g. *she is crew*).

Let us now go back to X-族 *zú* nouns in Chinese. Given the characteristics displayed by these nouns observed in § 3.4, we argue that 族 *zú* underwent a metonymical semantic shift from 'group' to 'members' (see examples 5a-b, 7a), and then more individuation has been licensed, as shown by the compatibility of X-族 *zú* nouns with quantities (one, two, several): these nouns can refer to single entities as well (examples (8a)-(8d)). In this 'member' sense, the X-族 *zú* noun does not denote a collective whole, but rather a class/category, indicating the nature of the individuals (see the discussion above on English nouns of collective origin). This is confirmed by examples like (10a): sentences like 我是上班族/背包族/爱车族 *wǒ shì shàngbānzú/bèibāozú/àichēzú* 'I am an office worker/a backpacker/a car lover' express belonging to a category (member-class relation), rather than being part of a group (part-whole relation).

The meaning shift from collective to individual underwent by 族 -zú can thus be described as follows: group > members of a category/class > individual (a member of the category/class).

## 4.2 The Evolution of 党 dǎng

In § 3.2 we have seen that the meaning of 党 dǎng as the right-hand constituent of complex words indicating groups of people with common characteristics or behaviour probably originates from the meaning ‘clique’, though the meaning of ‘political party’ contributed to the development of this new sense as well (Chen, Zhu 2010). As we have shown, X-党 dǎng words can indicate a variety of referents: people with a particular behaviour or habit in common; people addicted to something or who love something; people with some characteristics in common. We argue that all these meanings can be subsumed under the meaning ‘category/group of people with common characteristics or behaviour’. Given this generalisation of meaning, and the variety of referents it can designate, we conclude that 党 dǎng underwent a grammaticalisation process and should be then considered as a suffix. However, as we have pointed out in § 3.2, some X-党 dǎng words retain the negative nuance of the original meaning ‘clique’, arguably reflecting an earlier stage in the semantic evolution of this formative.

Furthermore, we argue that 党 dǎng, much like 族 zú, also underwent a semantic shift from collective to individual. As a matter of fact, we pointed out in § 3.4, that 党 dǎng can be used to refer to ‘members’ rather than to a collective whole (see examples (6a)-(6b) and (7b)), and to single entities as well (see examples (9a)-(9b)). We can conclude that, like 族 zú, it refers to a class/category, indicating the nature of the individuals, as emerges from examples like the one in (10b). See also the following example:

14. 因为我是游戏党,所以当初买这部手机时最先看中的就是它的性能配置。<sup>40</sup>  
yīnwèi wǒ shì yóuxì-dǎng suǒyǐ dāngchū mǎi zhè  
because 1SG be game-DANG so at.first buy this  
bù shǒujī shí zuìxiān kànzhōng de jùshì  
CLF mobile.phone time very.first take.a.fancy.to SP be  
tā de xìngnéng pèizhì  
3SG.N SP function configuration  
‘Since I am an online videogame player, when I bought this mobile phone, what I first considered was its performance settings’.

<sup>40</sup> <https://c.m.163.com/news/a/FNA6N0CJ05318V7C.html>.

The meaning shift from collective to individual underwent by 党 *dǎng* is similar to the one underwent by 族 *zú*: group > members of a category/class > individual (a member of the category/class).

Therefore, given the meaning generalisation and semantic shift underwent by 党 *dǎng*, we conclude that it can be included among affixes.

### 4.3 The Evolution of 客 *kè*

As highlighted by Wu (2010), Basciano (2017), and Arcodia and Basciano (2018), the pattern X-客 *kè* already existed in previous stages of the language. The basic meaning of 客 *kè*, as we have seen (see § 3.3), is 'guest, visitor'; however, if we look at its meaning in Classical Chinese, we also find 'person specialised in a certain activity', 'person engaged in a particular pursuit' (see 古汉语大词典 *Gu Hanyu da cidian* 'Great Dictionary of Classical Chinese', 1999), as it is evident e.g. in words like 侠客 *xiá-kè* 'chivalrous-KE, knight errant', 掮客 *qián-kè* 'serve.as.broker-KE, broker', 剑客 *jiàn-kè* 'sword-KE, swordsman'. Thus, it can be argued that the use of 客 *kè* as the right-hand constituent of complex words indicating 'a person doing a certain activity', or 'a person with certain characteristics', developed from this meaning.

Wu (2010) argues that the meaning 'guest, visitor' is the oldest one, which is attested since the pre-Qin period (before 221 BC); it then underwent extension of meaning, and its scope widened, beginning to indicate not only home visitors, but also travellers, people travelling or residing away from home, and even emissaries and invaders or aggressors. Later, 客 *kè*, while preserving the original meaning of 'guest', also developed other meanings: for example, Wu observes that for 水客 *shuǐ-kè* 'water-KE' the meaning 'boatman' emerged in the Wei-Jin period (220-420). Then, it underwent further extension of meaning: in the Tang period (618-907), for example, the word 瘦客 *shòu-kè* 'thin-KE, emaciated' emerged. Therefore, this morpheme underwent gradual generalisation of meaning, departing from its original meaning and starting to indicate 'a person involved in some activity' (e.g. 刺客 *cì-kè* 'assassinate-KE, assassin', 说客 *shuō-kè* 'speak-KE, persuasive talker') or a 'person with certain characteristics' (e.g. 醉客 *zuì-kè* 'drunk-KE, drunkard').

Thus, apparently the influence of English and netspeak gave an impulse to the development of an already existing pattern, rather than leading to the creation of a new one. Arcodia and Basciano (2018, 248) even speculate that the choice of 客 *kè* as a phonetic adaptation of the second syllable of English *hacker*, among many other morphemes which are commonly used in Modern Chinese for phonetic adaptations in loanwords (e.g. 克 *kè* 'overcome', 科 *kē* 'department' etc.), could have been motivated also by the meaning which 客 *kè* already had in word formation.

At present, 客 *kè* as the right-hand constituent of complex words can form nouns indicating different types of persons doing any kind of activity (not only on the web) or having certain characteristics. According to Arcodia and Basciano (2018), the general word-formation schema for these words is 'person related to X' ('person doing X' or 'person characterised by X'). Given the gradual extension of meaning underwent by this item, we consider 客 *kè* as an affix.

However, Arcodia and Basciano (2018) point out that those neologisms where the whole X-客 *kè* word is a phonetic adaptation of an English word not indicating a person, as e.g. 切客 *qiē-kè* 'cut-KE, fan of location-based services who regularly checks in to keep friends and relatives posted on her/his whereabouts', do not fit well this schema. As we have seen in § 3.3, the whole word is a phonetic adaptation of English *check-in*, but it indicates a person involved in an activity connected to the semantic of the phonetic adaptation as a whole ('person doing X-客 *kè*', rather than 'person doing X'). The role of 客 *kè*, thus, is not only phonetic: as we mentioned earlier, it contributes the meaning of 'person' as well. Therefore, Arcodia and Basciano (2018) consider these words as a special case of the X-客 *kè* construction.

As suggested by an anonymous reviewer, an alternative explanation could be that there are two routes of generalisation of 客 *kè*: one is the native route; the other one is the loan route, possibly resulting from the introduction of 黑客 *hēikè* 'hacker'. The native route ('person doing X' or 'person characterised by X') may be argued to have developed from a gradual extension of the meaning 'person specialised in a certain activity' and contributes to form words as e.g. 排客 *pái-kè* 'line. up-KE, a person paid to stand in a queue for others', 必胜客 *bì-shèng-kè* 'certainly-remain-KE, person doomed to remain single'. The loan route ('person doing X-客 *kè*'), instead, can be argued to have developed from the 黑客 *hēi-kè* 'hacker' model. As we have seen, while 黑客 *hēi-kè* 'hacker' is a phonetic adaptation of an English word indicating a kind of person (see also e.g. 极客 *jí-kè* 'extremely-KE, geek'), in many cases X-客 *kè* is not a phonetic adaptation of a word indicating a person; the meaning 'person' is rather conveyed by 客 *kè*, which is not only a phonetic component. We may hypothesise that 客 *kè*, originally part of a loanword, over time developed as an affix, whose meaning ('a person doing a certain activity') is somehow connected to the one of the loanword it was part of (i.e. *hacker*, a person engaged in a particular kind of activity). The development from the meaning 'hacker' could also explain the quite high number of X-客 *kè* words indicating persons doing activities on the web, or anyway using computers or new technologies (see § 3.3): a *hacker* is someone who does a particular activity online, i.e. someone who uses computers to get access to data in somebody else's computer or phone system without permission.

Even in this scenario, though, it cannot be excluded that the choice of 客 *kè* for the phonetic adaptation and its development into an agen-

tive suffix in these words have been influenced by the meaning this item already had in word formation, as mentioned above, and that the influence of English simply gave a new impulse to its development: thus, 黑客 *hēi-kè* 'hacker' and other words indicating different kinds of hackers may have had a role in reinforcing the word-formation schema at issue, given their basic agentive meaning, rather than being the source of it (Arcodia, Basciano 2018). Needless to say, the issue requires further investigation.

#### 4.4 A Comparison of the Three Word-Formation Patterns

In the previous sections, we described the evolution of 族 *-zú*, 党 *-dǎng* and 客 *-kè*, and we argued for their affixal status, since they underwent a gradual generalisation of meaning and can now be used to indicate a wider variety of referents. In addition, we have shown that 族 *-zú* and 党 *-dǎng* also underwent a semantic shift from collective to individual and can be currently used to refer to single individuals as well.

The development of these three affixes also shows the different mechanisms at work in grammaticalisation processes and the interplay between native patterns and foreign models. As for 族 *zú*, its affixal use was apparently imported from Japanese, a source language for many neologisms, as well as for new word-formation patterns, especially in the period between the end of 19th and the beginning of the 20th century (Masini 1993). As for 客 *-kè*, we pointed out that English had a key role in its development, as suggested also by the high proportion of phonetic adaptations of English words among X-客 *kè* neologisms. At the same time, though, this word-formation pattern was already present in Chinese and developed through a grammaticalisation process inner to the language; thus, it may be argued that English favoured the development of an existing pattern, rather than creating a new one.

The grammaticalisation paths followed by 族 *-zú* and 党 *-dǎng* are very similar, and actually the two affixes are very close in meaning; they may be found attached to the same bases without apparent changes in meaning (see Chen, Zhu 2010, and § 4.2). However, we pointed out that 党 *-dǎng* appeared later as a suffix, and conveys a more modern flavour; in addition, in some words it still retains the negative nuance of the original meaning 'clique' (see Chen, Zhu 2010; § 4.2). This suffix is not as established as 族 *-zú*, and apparently its use is typical of user-generated content (i.e. created by the users of an online system). This is quite clear if we compare the number of types (i.e. the number of different words created by a word-formation process) found in the dictionary of neologisms (XCY) with those found in the corpus (LWC) for X-族 *zú* and X-党 *dǎng* words:



**Table 2** X-族 *zú* and X-党 *dǎng* neologisms in XCY and LWC<sup>41</sup>

	XCY	LWC
X-族 <i>zú</i> words	215	142
X-党 <i>dǎng</i> words	3	184

As may be seen in table 2, we find an abundance of X-族 *zú* words in the dictionary of neologisms (XCY), which is a hint of the fact that this affix has been consistently and continuously used over the last thirty years, and its use is widespread in society. This word formation pattern is now established in the Chinese lexicon, and many X-族 *zú* words have been ‘institutionalised’, i.e. after having been widely employed for a reasonable amount of time, they have started to be accepted and recognised by language users as items of their regular vocabulary (see Bauer 1983; Fernández-Domínguez 2010). In contrast, only 3 X-党 *dǎng* words are listed in the XCY, which is in line with the relatively young age of this suffix: this word-formation pattern is not as established as X-族 *zú*, and most of these neologisms are not ‘institutionalised’. Coinages may be produced and used for some time, and then disappear: these words are known by the speakers who coined them, and perhaps to the speaking community around, but remain unnoticed for most language users (Hohenhaus 2005; Fernández-Domínguez 2010). Blocking could avoid the institutionalisation among speakers of part of X-党 *dǎng* words: it is possible that some X-党 *dǎng* words appear in the language, are used for a short period of time, and then disappear in favour of the previously existing X-族 *zú* words which are already widely used in the community (e.g. 上班族 *shàngbāndǎng* vs 上班族 *shàngbānzú* ‘office workers’; see Fernández-Domínguez, Díaz-Negrillo and Štekauer 2007). Blocking, indeed, does not avoid the coinage of words, but rather their institutionalisation, i.e. their wide usage in the community (Bauer [1988] 2003, 80-1).

However, if we look at the XCY and LWC columns in table 2, we can see that, despite the much greater number of distinct X-族 *zú* words in the XCY, in the LWC there are actually more X-党 *dǎng* words than X-族 *zú* words. This suggests that 党 *-dǎng* as a suffix is particularly frequent in user-generated texts. The preference for X-党 *dǎng* words can be stylistically motivated, since it represents a more fashionable pattern (see § 3.2). According to Plag (2006b, 550), productivity is also influenced by fashion, regardless of any need to name things (social factors or pragmatic needs can motivate new word creation; see

<sup>41</sup> We excluded from the count words in which 党 *dǎng* and 族 *zú* bear their original meaning, as e.g. 政党 *zhèngdǎng* ‘political party’ or 藏族 *zàngzú* ‘Tibetan ethnic group’.

Dal, Namer 2016). We will go back to this issue in the next section.

As for X-客 *kè*, we showed that it followed a peculiar grammaticalisation path, in which a native pattern interacted with a foreign model. The agentive meaning it acquired ('person doing a certain activity' or 'person with certain characteristics') is close to that conveyed by 党 *-dǎng* and 族 *-zú* (compare 刷书客 *shuā-shū-kè* 'scan-book-KE' and 刷书族 *shuā-shū-zú* 'scan-book-ZU', both referring to someone who scans with a mini-scanner the content from the books in a bookstore or a library); in addition, we remarked that the two suffixes may combine in the same word (see § 3.4). Nevertheless, we have stressed the fact that, in our corpus, many X-客 *kè* words indicate persons involved in online activities, or anyway activities connected to technology, and that there is a high proportion of loanwords among them, differently from X-党 *dǎng* and X-族 *zú* words, highlighting the role of English and of the word 黑客 *hēikè* 'hacker' in the development of this pattern (§ 4.3). What about the diffusion of this word-formation pattern? Judging from the number of X-客 *kè* types found in the XCY and in the LWC [tab. 3], this pattern is not particularly established and widespread in the language, neither it is particularly common in netspeak, if compared to X-族 *zú* and X-党 *dǎng* [tab. 2].

**Table 3** X-客 *kè* words in XCY and LWC<sup>42</sup>

	XCY	LWC
X-客 <i>kè</i> words	21	56

The figures in table 3 shows that the number of X-客 *kè* 'institutionalised' words is not high: the words listed in the XCY are more than those listed for X-党 *dǎng*, which is quite expected, since X-党 *dǎng* is the newest word-formation pattern among those considered; however, they are very few if compared to X-族 *zú* words. In addition, the number of types of X-客 *kè* found in the LWC is quite low compared to the other suffixes at issue, suggesting that the number of new words coined by means of this process is relatively limited: as pointed out by Fernández-Domínguez,

accepting the assumption that corpora are reliable reflections of language (Bauer 2001: 47; Plag 2003: 52), V [type frequency] should be a good indicator of the number of words coined by a pro-

<sup>42</sup> We excluded those words in which 客 *kè* bears the meaning of 'guest' or 'client', as e.g. 顾客 *gùkè* 'customer', and compounds in which the right hand constituent is a X-客 *kè* word, as 心理黑客 *xīnlǐ-hēikè* 'psychology-hacker, a person who helps others solve psychological issues'. Also, we decided to exclude all words indicating different kinds of 'hackers', for the reasons explained in fn. 20.

cess, so that the higher the figure of types, the more units a process has formed. (2010, 198)

We will return to this issue in the next section.

All in all, the morphological processes involving the three suffixes at issue are all productive, in the sense that they are ‘available’, i.e. they can be used in the present stage of the language to build new words (Bauer 2001, 205-11). But to what extent is their availability exploited in language use, i.e. to what extent are they ‘profitable’ (Bauer 2001, 205-11)? In the next section, we will compare their productivity by assessing their ‘profitability’ in the LWC: while availability is a qualitative notion (a process is either available or not), profitability is a quantitative notion because it deals with how many lexemes an available process coins, thus one process may be more profitable than another (Fernández-Domínguez 2010; for an overview on qualitative and quantitative approaches to productivity, see Dal, Namer 2016).

As pointed out by Plag (2006a, 124), “it is well known that certain affixes are more commonly found in certain types of texts than in others”: given the characteristics of the three affixes illustrated here, LWC is best suited to assess their profitability, since it is quite recent and is made up of user-generated content. As the LWC collects all the posts by Weibo users within a certain period of time, it reflects how words are actually, spontaneously and creatively used, and consequently the vitality of the three suffixes. The use of corpora rather than dictionaries as a source of data is motivated by the fact that in a corpus we may find productively formed derivatives which are not listed in dictionaries, and thus “corpus-based descriptions of productivity reflect how words are actually used” (Nishimoto 2003, 51).

## 5 A Comparison of the Productivity of 族 -zú, 党 -dǎng and 客 -kè

Several methods have been proposed in the literature to measure the profitability of a given process (for an overview, see Plag 2006a, 2006b). The same affix may score differently for different measures, thus yielding different productivity rankings, depending on the method used (for a summary, see Plag 2006b, 544-6). This is because each measure “highlights a special aspect of productivity” (Plag 2006a, 123).

As we have already shown in the previous section, if we look at type frequency, widely used as a productivity measure in the literature (see Fernández-Domínguez 2013), then 党 -dǎng is the most productive suffix, while 客 -kè is the least productive one.

**Table 4** X-族 *zú*, X-党 *dǎng* and X-客 *kè*: type frequency in the LWC

	<b>Types</b>
X-党 <i>dǎng</i> words	184
X-族 <i>zú</i> words	142
X-客 <i>kè</i> words	56

However, as observed by Fernández-Domínguez (2013), this measure may tell us something about the degree of generalisation (the degree to which a process has spread its derivatives in language) of a process but does not say anything about its availability, ignoring the synchronic status of word-formation processes: it focuses on the attestation of lexemes. This measure describes past productivity, i.e. the productivity of a process up to the present, and it is independent of its actual use (see also Dal, Namer 2016).

Other approaches to productivity look at this notion from a probabilistic-statistical perspective and focus on the likeliness of a given pattern to coin new words in the future (see the overview in Fernández-Domínguez 2013). Here we adopt Baayen's hapax-based index of productivity (P-index),<sup>43</sup> which is based on the number of *hapax legomena* (Baayen 1992): if an affix is very productive, we expect to find many *hapax legomena* containing that affix in a large text corpus, since it is typically among hapaxes that we find the higher proportion of neologisms (Renouf, Baayen 1996). Therefore, the crucial assumption behind this method is that the number of hapaxes of a given morphological category correlates with the number of neologisms of that category. In this sense, the number of hapaxes can be seen as an indicator of productivity.

Baayen's P-index is obtained by dividing the number of *hapax legomena* with a given affix ( $n_1$ ) by the number of tokens containing that affix ( $N$ ) in the corpus considered:

$$15. P = n_1 / N$$

If all of the words found in a text sample are hapaxes, the P-index will be 1 (maximal productivity), while many high frequency words increase the value of  $N$ , leading to a low productivity index.<sup>44</sup> Thus, high token frequency is connected with a high degree of lexicalisa-

---

<sup>43</sup> Baayen's models have undergone a number of modifications over the years, but in all of them *hapaxes* occupy a central position (for an overview, see Fernández-Domínguez 2013; Dal, Namer 2016).

<sup>44</sup> Several shortcomings of this hapax-based measure of productivity have been pointed out (see e.g. Bauer 2001; Fernández-Domínguez 2013; Dal, Namer 2016). Generally speaking, larger corpora lead to increased accuracy in calculating the P-index.

tion (storage in the lexicon) and low productivity, while low token frequency is connected with a low degree of lexicalisation and high productivity: as observed by Plag (2006a, 123), the presence of a large number of low-frequency words keeps the rule alive, since they force speakers to segment the derivatives, strengthening the existence of the affix. *Hapax legomena* are often unfamiliar words, but they are understandable for the hearer or reader if the process which created them is still ‘active’.

Table 5 shows the P-index of 族 -*zú*, 党 -*dǎng* and 客 -*kè* in the LWC.<sup>45</sup>

**Table 5** P-index of X-族 *zú*, X-党 *dǎng* and X-客 *kè* in the LWC

	<b>Tokens (N)</b>	<b>Hapax legomena (n1)</b>	<b>P-index</b>
X-党 <i>dǎng</i> words	342	137	0,400
X-族 <i>zú</i> words	1335	80	0.059
X-客 <i>kè</i> words	469	23	0,049

As we can see from the figures in table 5, the P-index of 党 -*dǎng* ranks the highest, while that of 客 -*kè* ranks the lowest, in line with the productivity ranking obtained by calculating type frequency [tab. 4]. 党 -*dǎng* has the highest number of hapaxes but the lowest number of tokens, meaning that among X-党 *dǎng* words there are not many high frequency words, leading to a very high P-index: this means that this pattern has a high potential to be used for the coinage of new forms, if needed. In contrast, 族 -*zú* displays a number of tokens significantly higher than that of the other two suffixes, meaning that many X-族 *zú* words are quite frequently used, leading to a large number of tokens and, consequently, an overall decrease of the P-index. As for 客 -*kè*, it is characterised by a low number of hapaxes (the lowest among the three suffixes) but a relatively high number of tokens (higher than 党 -*dǎng*), meaning that some of these words are frequently used; this leads to a low P-index.

These data confirm what already emerged from the discussion in the previous sections, i.e. that the X-族 *zú* pattern is quite established, and that many X-族 *zú* words are widespread in the language

<sup>45</sup> We must remark that the Leiden Weibo corpus has one major problem, namely that many messages are simply reposted from other users, and thus there are many cases of duplicated messages. This leads to an increase in the number of tokens; we thus manually removed the repeated messages, in order to get a more reliable picture.

Furthermore, we had to exclude the word 博客 *bókè* from the count. The overall number of tokens in the LWC is 3,006, but it includes both the meaning of ‘blog’ and that of ‘blogger’. Since it was not feasible to separate manually ‘blogger’ from ‘blog’, given the high number of tokens, we decided to exclude it. However, at a cursory look, we noticed that the meaning of ‘blog’ is predominant.

and have become ‘institutionalised’. Also, the high productivity displayed by 党 *-dǎng* is in line with the young age of this pattern and with its current popularity among netizens. The X-族 *zú* pattern was widely used for a certain period of time, producing many words which eventually became accepted as part of the common language and ‘institutionalised’ (as confirmed by the high number of types in the XCY; § 4.4), but it has apparently been superseded by the newly popular X-党 *dǎng* pattern, confirming what observed by Chen and Zhu (2010) on the two patterns. Its P-index predicts a high potential to build new words in future, much higher than that of X-族 *zú*.

As for the X-客 *kè* pattern, it is not as established as X-族 *zú*, but, at the same time, it displays limited productivity. The reasons of its low productivity should be investigated in depth: what are the factors restricting its productivity? Since many affixal elements indicating a type of person are currently found in Chinese, especially in user-generated texts, pragmatic factors, sociological factors, and blocking phenomena should be probably taken into account in order to get a clearer picture.

## 6 Conclusions

The influence of foreign languages and netspeak in the past few years not only led to the creation of a large number of neologisms, but also to the development of new word-formation patterns in Chinese, with the creation of many derivational affixes. Some of these items may be widely used at a given time but are then superseded after a while by a newer word-formation pattern. In this paper, we examined three suffixes emerged in the last thirty years, i.e. 族 *-zú*, 党 *-dǎng* and 客 *-kè*, all forming nouns referring to persons. After describing the three word-formation patterns, we focused on the evolution of the three formatives, characterised by meaning generalisation, arguing that at present they can all be considered as suffixes, based on their fixed position in complex words (to the right) and on the meaning generalisation observed.

The suffixes 族 *-zú* and 党 *-dǎng* both form nouns indicating a variety of referents, which we argued can all be subsumed under the general meaning ‘people with common characteristics or behaviour’. These two suffixes can also be attached to the same base without any change in meaning. In addition, we also remarked that both of them underwent a meaning shift from collective to individual, and thus they can be used to refer to single entities as well. However, from the point of view of meaning, the two suffixes do not exactly overlap: differently from X-族 *zú* words, some X-党 *dǎng* words retain the negative nuance of the original meaning ‘clique’, indicating a series of illegal activities. In addition, we pointed out that X-党 *dǎng* is

currently a more popular and fashionable pattern, thus possessing a more modern flavour.

Through the analysis of productivity based on the data of the LWC, we also showed that, while both word-formation patterns are 'available' to form new words at the present stage of the language, the degree of profitability of the X-党 *dǎng* pattern is much higher, meaning that it has a high potential to build new words: the X-族 *zú* pattern was widely used for a certain period of time, producing many words which eventually became accepted as part of the common language, but it has been apparently superseded by the newly popular X-党 *dǎng* pattern.

As for 客 *kè*, a number of words containing this suffix emerged starting from the 2000s in user-generated content. We argued that the influence of English and netspeak gave impulse to an already existent, though limitedly productive, word-formation pattern. From the analysis of the data in our corpus, the X-客 *kè* word-formation pattern is not particularly established and widespread in the language, neither it is particularly frequent in user-generated content: in the LWC it ranks the lowest for type frequency, number of hapaxes, and P-index, while ranks higher than 党 *-dǎng* in terms of token frequency, meaning that some X-客 *kè* words are frequently used. Even though this pattern is available for the creation of neologisms, its potential to create new words is quite limited.

As we mentioned, besides those investigated in this study, at present we can find a number of emerging suffixes indicating people in Chinese. One may wonder why so many different affixes are needed to create words referring to persons. A broader investigation comparing the properties and usage differences of different suffixes would be welcome. Since the creation of neologisms is not always meant to satisfy naming needs, it would be worth investigating the role of social factors, pragmatic needs, as well as language trends, in the development of these suffixes.

## Bibliography

- Arcodia, G.F. (2011). "A Construction Morphology Account of Derivation in Mandarin Chinese". *Morphology*, 21, 89-130. <https://doi.org/10.1007/s11525-010-9173-2>.
- Arcodia, G.F.; Basciano, B. (2012). "On the Productivity of the Chinese Affixes -兒 *-r*, -化 *-huà* and -頭 *-tóu*". *Taiwan Journal of Linguistics*, 10(2), 89-117. [http://dx.doi.org/10.6519/TJL.2012.10\(2\).3](http://dx.doi.org/10.6519/TJL.2012.10(2).3).
- Arcodia, G.F.; Basciano, B. (2018). "The Construction Morphology Analysis of Chinese Word Formation". Booij, G. (ed.), *The Construction of Words. Advances in Construction Morphology*. Berlin: Springer, 219-53. [https://doi.org/10.1007/978-3-319-74394-3\\_9](https://doi.org/10.1007/978-3-319-74394-3_9).

- Audring, J.; Booij, G. (2016). "Cooperation and Coercion". *Linguistics*, 54(4), 617-37. <https://doi.org/10.1515/Ling-2016-0012>.
- Baayen, R.H. (1992). "Quantitative Aspects of Morphological Productivity". Booij, G.; Van Marle, J. (eds), *Yearbook of Morphology 1991*. Dordrecht; London: Kluwer Academic Publishers, 109-49. [https://doi.org/10.1007/978-94-011-2516-1\\_8](https://doi.org/10.1007/978-94-011-2516-1_8).
- Bareato, S. (2017). *La derivazione in cinese. Uno studio su corpora dei formanti 族 zú e 党 dǎng* [MA dissertation]. Venice: Ca' Foscari University of Venice.
- Basciano, B. (2017). "黑客 hēikè, 白客 báikè, 红客 hóngkè. Hacker e altri 'ospiti' tra i neologismi del cinese moderno". Bulfoni, C. et al. (eds), 文心 *Wenxin. L'essenza della scrittura. Contributi in onore di Alessandra Cristina Lavagnino*. Milano: FrancoAngeli, 384-95.
- Bauer, L. (1983). *English Word-Formation*. Cambridge: Cambridge University Press.
- Bauer, L. (1998). "Is there a Class of Neoclassical Compounds, and if so Is It Productive?". *Linguistics*, 36(3), 403-22. <https://doi.org/10.1515/Ling.1998.36.3.403>.
- Bauer, L. (2001). *Morphological Productivity*. Cambridge: Cambridge University Press.
- Bauer, L. [1988] (2003). *Introducing Linguistic Morphology*. 2nd ed. Edinburgh: Edinburgh University Press.
- Bauer, L. (2005). "The Borderline between Derivation and Compounding". Dressler, W. et al. (eds), *Morphology and Its Demarcations*. Amsterdam; Philadelphia: John Benjamins, 97-108. <https://doi.org/10.1075/cilt.264.07bau>.
- Bauer, L. (2006). "Compound". Brown, K. (ed.), *Encyclopedia of Language and Linguistics*, vol. 2. Oxford: Elsevier, 719-26.
- Bisang, W. (1996). "Areal Typology and Grammaticalization. Processes of Grammaticalization Based on Nouns and Verbs in East and Mainland South East Asian Languages". *Studies in Language*, 20(3), 519-97. <https://doi.org/10.1075/sL.20.3.03bis>.
- Booij, G. (2005). "Compounding and Derivation. Evidence for Construction Morphology". Dressler, W. et al. (eds), *Morphology and Its Demarcations*. Amsterdam; Philadelphia: John Benjamins, 109-32. <https://doi.org/10.1075/cilt.264.08boo>.
- Booij, G. (2010). *Construction Morphology*. Oxford: Oxford University Press.
- Cao D. 曹大为 (2007). "Zu' de leicizhuihua shiyong fenxi" "族" 的类词缀化使用分析 (Analysis of the Pseudo-Affixal Use of 族 zú). *Shandong shehui kexue*, 5, 150-2.
- Cao T. 曹铁根; Mo W. 莫伟勇 (2012). "Wangluo xinciyu 'X-kong' yuyi jixi" 网络新词语 "X控" 语义解析 (Analysis of the Meaning of X-控 kòng Neologisms on the Web). *Hunan keji daxue xuebao*, 15(1), 122-4.
- Ceccagno, A.; Basciano, B. (2007). "Compound Headedness in Chinese. An Analysis of Neologisms". *Morphology*, 17, 207-31. <https://doi.org/10.1007/s11525-008-9119-0>.
- Chen C. 陈昌来; Zhu Y. 朱艳霞 (2010). "Shuo liuxingyu 'X-dang' — yi jian lun zhi ren yusu de leicizhuihua" 说流行语 "X党" — 兼论指人语素的类词缀化 (On the Buzzword 'X-党 dǎng'. A Discussion on the Affixal Status of Morphemes Referring to Persons). *Dangdai xiucixue*, 159, 64-70.



- Cheng, L.L.-S.; Sybesma, R. (1999). "Bare and Not-so-bare Nouns and the Structure of NP". *Linguistic Inquiry*, 30(4), 509-42. <https://doi.org/10.1162/002438999554192>.
- Cheung, C.C.-H. (2016). *Parts of Speech in Mandarin. The State of the Art*. Singapore: Springer.
- Croft, W. (1994). "Semantic Universals in Classifier Systems". *Word*, 45(2), 145-71. <https://doi.org/10.1080/00437956.1994.11435922>.
- Dal, G.; Namer, F. (2016). "Productivity". Hippiusley, A.; Stump, G. (eds), *The Cambridge Handbook of Morphology*. Cambridge: Cambridge University Press, 70-89. <https://doi.org/10.1017/9781139814720.004>.
- Dong X. 董秀芳 (2004). *Hanyu de ciku yu cifa* 汉语的词库与词法 (Chinese Lexicon and Morphology). Beijing: Peking University Press.
- Fabb, N. (1998). "Compounding". Spencer, A.; Zwicky, A.M. (eds), *Handbook of Morphology*. Oxford: Blackwell, 66-83. [https://www.blackwell-publishing.com/content/BPL/Images/Content\\_Store/WWW\\_Content/9780631226949/01Prelim.pdf](https://www.blackwell-publishing.com/content/BPL/Images/Content_Store/WWW_Content/9780631226949/01Prelim.pdf).
- Fernández-Domínguez, J. (2010). "Productivity vs. Lexicalisation. Frequency-Based Hypotheses on Word-Formation". *Poznan Studies in Contemporary Linguistics*, 46(2), 193-219. <https://doi.org/10.2478/v10010-010-0010-x>.
- Fernández-Domínguez, J. (2013). "Morphological Productivity Measurement. Exploring Qualitative versus Quantitative Approaches". *English Studies*, 94(4), 422-47. <https://doi.org/10.1080/0013838x.2013.780823>.
- Fernández-Domínguez, J.; Díaz-Negrillo, A.; Štekauer, P. (2007). "How Is Low Morphological Productivity Measured?". *Atlantis*, 29, 29-54. <https://www.jstor.org/stable/41055264>.
- Gardelle, L. (2019). *Semantic Plurality. English Collective Nouns and Other Ways of Denoting Pluralities of Entities*. Amsterdam; Philadelphia: John Benjamins.
- Haspelmath, M. (2002). *Understanding Morphology*. Oxford: Oxford University Press.
- Heine, B.; Claudi, U.; Hünnemeyer, F. (1991). *Grammaticalization. A Conceptual Framework*. Chicago: University of Chicago Press.
- Hohenhaus, P. (2005). "Lexicalization and Institutionalization". Štekauer, P.; Lieber, R. (eds), *Handbook of Word-Formation*. Dordrecht: Springer, 353-73. [https://doi.org/10.1007/1-4020-3596-9\\_15](https://doi.org/10.1007/1-4020-3596-9_15).
- Iljic, R. (1994). "Quantification in Mandarin Chinese. Two Markers of Plurality". *Linguistics*, 32(1), 91-116. <https://doi.org/10.1515/ling.1994.32.1.91>.
- Katamba, F. (1993). *Morphology*. New York: Martin's Press.
- Li J. 李杰 (2013). "'Zu' leici zhong cizhuhua qingxiang xianxiang" "族"类词中词缀化倾向现象 (The Development of 族 *zú* Words into Derivates). *Xi'an hangkong xueyuan xuebao*, 2(31), 21-3.
- Li, C.; Thompson, S.A. (1981). *Mandarin Chinese. A Functional Reference Grammar*. Berkeley: University of California Press.
- Lu Y. 鲁瑛 (2010). "'xx-zu' cilei de yuyanxue yanjiu" "xx族"词类的语言学研究 (Linguistic Research on "xx-zu" Neologisms). *Waiguo yuwen*, 2(26), 71-5.
- Lü S. 吕叔湘 (1941). *Zhongguo wenfa yaolue* 中国文法要略 (An Outline of Chinese Grammar). Shanghai: The Commercial Press.
- Ma M. 马渺沙 (2016). "Cong ci de xingcheng ji qi jiegou de jiaodu qianxi wangluo liuxing leicizhui de xianxiang — yi 'X-zu', 'X-kong' wei lie" 从词的形成及其结构的角度的浅析网络流行类词缀的现象——以“X族”“X控”为例 (An

- Analysis of the Phenomenon of Popular Affixoids on the Web from the Perspective of the Form and Structure of Words. The Case of X-族 *zú* and X-控 *kòng*). *Meijie yu wenhua yanjiu*, 12, 99-101.
- Ma Q. 马庆株 (1995). "Xiandai Hanyu cizhui de xingzhi, fanwei he fenlei" 现代汉语词缀的性质, 范围和分类 (Nature, Domain and Classification of Affixes in Modern Chinese). *Zhongguo yuyan xuebao*, 6, 101-37.
- Masini, F. (1993), *The Formation of Modern Chinese Lexicon and Its Evolution Towards a National Language. The Period from 1840 to 1898*. Berkeley: University of California.
- Naumann, B.; Vogel, P.M. (2000). "Derivation". Booij, G.; Lehmann, C.; Mugdan, J. (eds), *Morphologie-Morphology*. Berlin; New York: Mouton de Gruyter, 929-43.
- Nishimoto E. (2003). "Measuring and Comparing the Productivity of Mandarin Chinese Suffixes". *Computational Linguistics and Chinese Language Processing*, 8(1), 49-76.
- Packard, J. (2000). *The Morphology of Chinese*. Cambridge: Cambridge University Press, 71-3.
- Pan W. 潘文国; Ye B. 叶步青; Han Y. 韩洋 (2004). *Hanyu de goucifa yanjiu* 汉语的构词法研究 (The Research on Word Formation in Chinese). Shanghai: Huadong Shifan Daxue Chubanshe.
- Peirsman, Y.; Geeraerts, D. (2006). "Metonymy as a Prototypical Category". *Cognitive Linguistics*, 17(3), 269-316. <https://doi.org/10.1515/cog.2006.007>.
- Peyraube, A. (1998). "On the History of Classifiers in Archaic and Medieval Chinese". T'sou, Benjamin (ed.), *Studia Linguistica Serica*. Hong Kong: City University of Hong Kong, 39-68.
- Plag, I. (2003). *Word-Formation in English*. Cambridge: Cambridge University Press.
- Plag, I. (2006a). "Productivity". Brown, K. (ed.), *Encyclopedia of Language and Linguistics*, vol. 10. 2nd ed. Oxford: Elsevier, 121-8. <https://doi.org/10.1016/B0-08-044854-2/00125-5>.
- Plag, I. (2006b). "Productivity". Aarts, B.; McMahon, A. (eds), *The Handbook of English Linguistics*. Oxford: Blackwell Publishing, 537-56.
- Renouf, A.; Baayen, R.H. (1996). "Aviating among the Hapax Legomena. Morphological Grammaticalisation in Current British Newspaper English". Renouf, A. (ed.), *Explorations in Corpus Linguistics*. Amsterdam; Atlanta, GA: Rodopi, 181-9.
- Shanghai Daily* (2010). *Chinese Buzzwords. With English Explanations*. 2nd ed. Shanghai: Shanghai Press.
- Shen G. 沈光浩 (2015). *Hanyu paishengshi xinciyu yanjiu* 汉语派生式新词语研究 (Research on Derived Neologisms in Chinese). Beijing: Zhongguo Shehui Kexue Chubanshe.
- Sproat, R.; Shih, C. (1996). "A Corpus-Based Analysis of Mandarin Nominal Root Compound". *Journal of East Asian Linguistics*, 5, 49-71. <https://doi.org/10.1007/bf00129805>.
- Steffen Chung, K. (2006). *Mandarin Compound Verbs*. Taipei: Crane.
- Sun Y. 孙艳 (2000). "Xiandai Hanyu cizhui wenti tantao" 现代汉语词缀问题探讨 (A Study on Some Problems Related to Modern Chinese Affixes). *Hebei Shifan Daxue xuebao*, 23, 55-8.

- Sybesma, R. (2017). "Classifiers, Nominal". Sybesma, R. et al. (eds), *Encyclopedia of Chinese Language and Linguistics*, vol. 1. Leiden: Brill, 620-7. [http://dx.doi.org/10.1163/2210-7363\\_ecLL\\_COM\\_00000091](http://dx.doi.org/10.1163/2210-7363_ecLL_COM_00000091).
- Wang L. 王力 [1980] (2002). *Hanyu shi gao* 汉语史稿 (History of Chinese). Beijing: Zhonghua Shuju.
- Wu L. 吴琅琅 (2010). "Xiao xi Wenzhou fangyan 'ke' zu ci — cong gu Hanyu jiaodu zhengming 'ke' houzhui cunzai keneng" 小析温州方言“客”族词——从古汉语角度证明“客”后缀存在可能 (A Brief Analysis of the 客 kè Family of Words in Wenzhou Dialect. Proving the Existence of the 客 -kè Suffix from the Perspective of Old Chinese). *Yuyan wenxue yanjiu*, 3, 27-8.
- Xiao Y. 肖遥遥 (2009). "'Zu' lei Hanyu xinci yufahua qianxi" "族"类汉语新词语法化浅析 (Analysis of the Grammaticalization of 族 zú Neologisms in Chinese). *Zhongzhou Daxue xuebao*, 2(26), 75-7.
- Yang H. 杨海明; Chen Q. 陈倩仪 (2012). "'-zu' ci yu xinwen yuyan de shidai gan" "族"词与新闻语言的时代感 (~族 zú Words and the Sense of Time in the Language of the News). *Xinwen aihaozhe*, 92-3.
- Yip P.-C. (2000). *The Chinese Lexicon. A Comprehensive Survey*. London; New York: Routledge.
- Zhang Y. 张谊生; Xu X. 许歆媛 (2008). "Qian xi 'X ke' ci zu — cihuihua he yufahua de guanxi xin tan" 浅析“X客”词族——词汇化和语法化的关系新探 (A Preliminary Analysis of the “X-客 kè” Family of Words. A New Look at the Relation between Lexicalization and Grammaticalization). *Yuyan wenzi yingyong*, 4, 77-82.
- Zhao A. 赵爱萍 (2009). "Shimao 'zuci' jiqi shehui, xinli toudi" 时髦“族词”及其社会、心理透视 (The Fashionable 族 zú Words. A Sociological and Psychological Perspective). *Qiqihar Shifan Gaodeng Zhuanke Xuexiao xuebao*, 4, 35-6.

## Dictionaries

- Gu Hanyu da cidian* 古汉语大词典 (Great Dictionary of Classical Chinese) (1999). Shanghai: Shanghai Cishu Chubanshe.
- The Contemporary Chinese Dictionary* (Chinese-English edition) (2002). Beijing: Foreign Language Teaching and Research Press.
- Xiandai Hanyu cidian* 现代汉语词典 (The Contemporary Chinese Dictionary) (2005). Beijing: The Commercial Press.
- Xin shiji xinciyu da cidian* 新世纪新词语大词典 (New Century Comprehensive Dictionary of Neologisms) (2015). Shanghai: Shanghai Cishu Chubanshe.
- Xinhua xinciyu cidian* 新华新词语词典 (The Xinhua Dictionary of New Words) (2003). Beijing: The Commercial Press.

