languages

MDPI

Article

Language Change and Morphological Processes in Contemporary Chinese: The Case of 健康码 (Health QR Code)

Yun Xiao [1,*] and Yi Ren [2]

[1] Department of Modern Language, Bryant University, Smithfield, RI 02917, USA
[2] Department of Asian Languages and Cultures, University of California, Los Angeles, CA 90095, USA
* Correspondence: yxiao@bryant.edu

Abstract: The Chinese language has witnessed remarkable changes in the past several decades, marked by a rapid rise of new words, frequent innovation of pseudo-affixes, and a notable increase in word length. By analyzing the creation, spread, and expansion of the new word 健康码 *Jiankang-ma* "health QR code", this study sheds light on how language change takes place and how nonce formation is brought into being. Following the explosion of COVID-19 in China, 健康码 *Jiankang-ma* "health QR code" was created and promoted by the local and central governments and subsequently generated a large XX-码 XX-*ma* "XX-code" word family through various morphological processes, such as abbreviation, clipping, derivation, and analogy, where -码 *-ma* "-code" has acquired some new meaning distinct from its original form as a bound root. Linguistically, 健康码 *Jiankang-ma* "health QR code" is a three-morpheme word in a 2 + 1 length pattern, which phonologically consists of a single super foot that makes the expression catchy and appealing. The highly productive AB-C internal structure makes -码 *-ma* "-code", although not yet a fully grammaticalized affix, a strong pseudo-suffix that has high morphological productivity with a fixed suffix positioning. Given the high frequency of the lexical 码 *ma* "code" in contemporary Chinese language use, the pseudo-suffix -码 *-ma* "-code" may, however, not completely replace it in the long run, as principles of grammaticalization theories predict. Instead, it is likely that both the lexical 码 *ma* "code" and the pseudo-suffix -码 *-ma* "-code" will coexist side by side along split pathways.

Keywords: language change; grammaticalization; 健康码 *Jiankang-ma*; pseudo-suffix; word length pattern; word family

Citation: Xiao, Yun, and Yi Ren. 2023. Language Change and Morphological Processes in Contemporary Chinese: The Case of 健康码 (Health QR Code). *Languages* 8: 123. https://doi.org/10.3390/languages8020123

Academic Editors: Jidong Chen and Hongyin Tao

Received: 31 March 2022
Revised: 30 March 2023
Accepted: 20 April 2023
Published: 4 May 2023

1. Introduction

Language change concerns how linguistic structures are innovated, come into being, evolve, and spread. The process opens a window on language users' mental processes and human cognition from a historical perspective, but "language change is not a phenomenon of the distant past" but current and ongoing (Bybee 2015, p. 1). As new concepts and new events enter a culture, language users find ways to interpret new data and create new words or new expressions through linguistic internal or external resources such as compounding, derivation, and borrowing (Matthews 1991; Bybee 2015). When a new word or new expression is created and spreads, language change takes place.

Studies in Chinese linguistics reveal that compounding is prevalent and is the most productive process of word formation in the Modern Chinese language (Li and Thompson 1981; Norman 1988; Dong 2004, among others). Drawing data from a corpus of 100 million characters entitled "Core Wordlist of Contemporary Chinese for Information Processing", Dong (2004) concurs that the majority of Chinese words are compounds and that compounding is the leading word formation process in Chinese, which typically combines independent words. Furthermore, Packard (2000) maintains that, in addition to compounding, there are several other word formation devices in Chinese, such as abbreviation, borrowing, shifting, and numerical formulae. Recent data further show that Chinese nonce

formation is constrained by word length patterns and shaped through various morphological processes such as abbreviation, derivation, analogy, blending, clipping, ellipsis, paraphrasing, and semantic reduction or extension (Xiao 2015, 2016, 2019).

Since China thrived in the decades-long economic reform and international integration, it has made many social developments and progress, which brought about remarkable changes in the Chinese language, characterized by a rapid increase in new words, a notable expansion of word length, frequent innovations of derivational morphemes that created a large number of new pseudo-prefixes/suffixes (Jiang 2009; Zeng and Wang 2011). Lexemes (i.e., content or function words) become affixoid or affixes, which, in turn, generate large word families through derivation and analogy, such as 被-XX *bei*-XX "passive voice-XX", 大-XX *da*-XX "big-XX", XX-族 XX-*zu* "XX-group/class" (Xiao 2015, 2016). Out of the new words/expressions, many were created by Chinese Internet Language (CIL) users (Xing and Wang 2008; W. Yang 2011; X. Yang 2012; Gao 2012; Xiao 2015, among others). The CIL was officially branded as 平民化 *pingminhua*, "grassroots, ordinary", 低门槛 *dimenkan*, "low-standard, vulgar" (Language Situation in China, 2012:205) and hence, discouraged in the official language use.

The recent new word 健康码 *Jiankang-ma* "health QR code", was initially created by Chinese language users to cope with the coronavirus outbreak in China. It was, however, immediately adopted by the central government and demanded as the "Pass" for commuters and travelers. Out of the Chinese media new words in 2020, 健康码 *Jiankang-ma* "health QR code" was ranked as top-five and swiftly generated a large word family of XX-码 XX-*ma* "XX-code". With 健康码 *Jiankang-ma* "health QR code" as an example, this study explores how Chinese language users create new words with various morphological devices to interpret the new data and promote the new products. Drawing on theoretical perspectives in grammaticalization, lexicalization, and language change (Hopper and Traugott 2003; Briton and Traugott 2006; Bybee 2015, among others), this study takes on a synchronic analysis and intends to answer questions such as: How is a nonce form, such as 健康码 *Jiankang-ma* "health QR code" being innovated, gaining acceptance, spreading, and deriving a large word family? What morphological processes contribute to its development and productivity? Data for analysis are drawn from current Chinese media reports, official documents, or directives, including recent data from *Language Situation in China* 《中国语言生活状况报告》 (GYWGW n.d.).

2. Chinese Language Change with the Rise of New Words and New Pseudo-Affixes

2.1. Di-Syllabicity Tendency and Word-Length Patterns in Modern Chinese

The traditional view was that Old Chinese was an isolating and monosyllabic language, which lacked inflection or derivation, and each word constituted one syllable/morpheme only (Chao 1968; Li and Thompson 1981; Norman 1988, among others) ("syllable" and "morpheme" to be used interchangeably thereafter). However, Feng (1998) notes that the Old Chinese syllable structure started to change in Han Dynasty (206 BC-220AD), during which two-syllable unit became preferred, and he suggests that such a change was caused by the loss of bimoraic feet and the replacement of disyllabic feet in Middle Chinese. Due to such historical phonological change, di-syllabicity has become the main rhythmic tendency in Modern Chinese (Lü 1963, 2004). As disyllabic words increased, so did the multi-syllable words. Lü notes that, in the *List of 3000 Common Words in Putonghua* (Standard Mandarin), 2263 of them are di- or multi-syllabic, out of which disyllabic words are dominating (1963 and 2004).

Drawing on data from the 1.3 million-character Lancaster Corpus of Mandarin Chinese (LCMC, McEnery and Xiao 2004), Duanmu (2012) notes that there are four preferred length patterns in the two-word expressions, such as 2 + 2, 2 + 1, 1 + 2, and 1 + 1 (the number standing for the number of syllables). Out of them, 1 + 1 [N N] two-syllable combinations are the majority (55.8% of total token counts and 63% of total type counts). Within the three-morpheme combinations, the 2 + 1 pattern has a higher frequency than the 1+2 counterpart (16.2% and 0.8%, respectively), and 1 + 2 [N N] is disfavored (Duanmu 2012). Such

findings support Lü's position (1963 and 2004) that, in Modern Chinese, 1+1 two-syllable combinations are the majority, and 2 + 1 combinations have a much higher frequency than those of 1 + 2.

Duanmu (2012) further proposes that Chinese word combinations are more motivated by phonological constraints than by syntactic or semantic requirements, which lends support to Feng's notions of "prosodic syntax" (Feng 2013) and "prosodic morphology" (Feng 2017). According to Feng (2013), phonology prohibits certain word orders and length patterns in Chinese word formation, in that two-syllable combinations are prosodic words with the smallest and standard metrical foot, while three-syllable combinations constitute a super foot comprised a two-syllable (big) foot and a one-syllable (small) foot, out of which the small foot must be non-lexical and bound. In addition, the 2+1 [N N] three-syllable combinations are derivative and productive in word formation, while the 1+2 [N N] pattern is disfavored.

2.2. Emergence of Pseudo-Affixes in Chinese

As an isolating language, Chinese is found to have few established affixes (Lü 1979, 2004; Li and Thompson 1981). Lü reports that, in Chinese grammar, there are only five prefixes (阿- *ah-* "affectionate prefix", 第- *di-* "ordinal prefix", 初- *chu-* "start-", 老- *lao-* "old-", 小- *xiao-* "little"), seven suffixes (-子 *-zi* "nominal suffix", -儿 *-er* "affectionate suffix", -头 *-tou* "-head", -巴 *-ba* "nominal suffix", -者 *-zhe* "agentive suffix", -们 *-men* "plural suffix", -然 *-ran* "adverbial suffix"), and two infixes (得 *de* "potential infix", 不*bu* "negative potential infix") (Lü 1979, 2004). Accordingly, Chinese users resort to "pseudo-prefixes/suffixes" to create a nonce formation for new products or new events, for which he identified 18 pseudo-prefixes, such as 可- *ke-* "able-", 好- *hao-* "good-", 难- *nan-* "difficult-", 准- *zhun-* "prospective-", 类- *lei-* "quasi-", 亚- *ya-* "sub-", 次- *ci-* "second", 超- *chao-* "ultra-", 单- *dan-* "single-", 非- *fei-* "negative prefix", etc., and 23 pseudo-suffixes such as -员 *-yuan* "-member", -家 *-jia* "-expert", -人 *-ren* "-person", -界 *-jie* "-field", -品 *-pin* "-type", -种 *-zhong* "-kind", -学 *-xue* "-field of study", -性 *-xing* "-nature", -化 *-hua* "-formation", etc. According to Lü, these "pseudos" are lexical words that play a grammatical function role in the new contexts but retain some of their original semantic meanings without "being completely grammaticalized" (Lü 1979, 2004).

Recent Chinese language studies have identified a large number of pseudo-prefixes and suffixes and found them to be created through downgrading or grammaticalization from content or functional words, which further generate large numbers of "new word families" through analogy (A. Yang 2005; Y. Yang 2007; Sun 2007; Cao 2010; Zeng and Wang 2011; Xiao 2015, 2016, 2019, among others). Drawing on data from the database of《现代汉语语法信息词典》*Contemporary Chinese Grammar Dictionary*, which contains 80,000 words, Zeng and Wang (2011) identified 34 pseudo-prefixes and 46 pseudo-suffixes that met three criteria, such as the degree of morphological productivity, fixation of positioning, and semantic bleaching. The results reveal a higher frequency of recent "psudos" than those in Lü's account shown above. In addition, the researchers believed that the criteria used for the study were able to separate quasi-affixes from content morphemes and resolve the fuzzy borderline issues between derivation and compounding in Chinese. Likewise, using data from Modern Chinese literature readings, A. Yang (2005) found a large quantity of newly-derived three-morpheme compound nouns that consist of an internal structure of either 1+2 pseudo-prefix + root or 2+1 root + pseudo-suffix (the number standing for the number of morphemes), with the latter having a much higher frequency than the former. Specifically, out of the 319 derived NPs in the researcher's corpus, 53 are in the pattern of 1+2 pseudo-prefix + root (16.61%) and 266 in 2+1 root + pseudo-suffix (83.39%) (2005). Such results support the aforementioned finding that pseudo-affixes are largely increasing in Modern Chinese, and 2+1 three-morpheme compound nouns are the more productive formation pattern.

2.3. Rise of Three-Morpheme Words in Contemporary Chinese Language Use

Contrary to the long-held belief that two-morpheme words are preferred in Modern Chinese, recent data show that three-morpheme words have become the preferred model for media neologisms in recent years. *Language Situation in China*《中国语言生活状况报告》(GYWGW n.d.) reports that there were 5733 media new words/expressions created in the past 10 years (from 2005 to 2015) and that, out of those created in 2009–2011, despite other forms that share a relatively small proportion, three-morpheme words were the majority (51.01% in 2009, 53% in 2010, and 51.77% in 2011), which had a much higher frequency than the two-morpheme counterparts (18.69%, 17%, and 15.51%, respectively). As a follow-up, Hou and Zou (2015)[1] further studied the media new words in 2012–2014. The results show that, out of the new words created in these three years, three-morpheme words remained to be the majority (44.45% for 2012, 45.45% for 2013, and 41.75% for 2014) and had a much higher frequency than the two-morpheme counterparts (13.16%, 23.97%, and 28.54%, respectively).

In summary, three-morpheme words are the majority and consistently have a higher frequency than their two-morpheme counterparts for six consecutive years. In addition, data from *Language Situation in China*《中国语言生活状况报告》(GYWGW n.d.) show that many of the new usages that emerged in 2011 are formed by derivation. For instance, of the 592 new words created in 2011, the top ones mostly consisted of three morphemes formed with root word + pseudo-suffixes, such as:

(1) XX-族 (*-zu* "-group/class", 5.06% of the total), e.g., 上班-族 (*shangban-zu* "working class")
(2) X-哥 (*-ge* "-brother", 5.06% of the total), e.g., 手表-哥 (*shoubiao-ge* "watch-brother")
(3) XX-门 (*-men* "-scandal", 3.37% of the total), e.g., 嫖娼-门 (*piaochang-men* "prostitution-scandal")
(4) XX-帝 (*di* "-king", 0.84% of the total), e.g., 高考-帝 (*gaokao-di* "college entrance examination-king", i.e., a student who repeated the college entrance exam numerous times.)

It is noted that the above top-rated three-morpheme words are all compound nouns, formed in the pattern of 2 + 1, in which the two-syllable elements are all lexical morphemes comprising either an NP or a VP, such as 手表 (NP) *shoubiao* "watch", 高考 (NP) *gaokao* "college entrance exam", 上班 (VP) *shangban* "to go to work". According to Feng's notion of "Prosodic Syntax" (Feng 1995, 2013), each of these three-morpheme combinations consists of a "super foot," i.e., a "big foot" and a "small foot". In his discussion on the concept of "Prosodic Morphology", Feng (1995, 2017) maintains that the 2 + 1 super foot in Mandarin Chinese is subject to the Trisyllabic Foot Constraint (TFC), by which "super foot formation is allowed only if standard foot formation leaves stray syllables" (p. 169). According to the TFC, the standard foot consists of two-syllable combinations formed with fully-stressed content morphemes, while the stray syllable is non-lexical, bound ("unable to be independent"), and phonologically destressed or neutralized (p. 170). Accordingly, such 2 + 1 three-morpheme NPs can be better indicated as AB-C, i.e., (NP/VP-affix), instead of (N N), as proposed by Feng (2013), in which AB can be an NP or a VP (root word), while C is a destressed/neutralized grammatical affix.

By drawing on the data from *Language Situation in China* 《中国语言生活状况报告》(GYWGW n.d.) and analyzing the list of 2015 media new words for their features in word-length patterns, word-internal structures, and their formation processes, Xiao (2019) further lends support to the aforementioned findings. She notes that the average word length of the 446 words in the list is 3.34, out of which 176 are three-morpheme (39.46% of total), 161 are four-morpheme (36.10%), and 75 are two-morpheme (16.82%). In addition, the three-morpheme words in the list consist of three types of word patterns: 2 + 1 (n = 121, 68.75%); 1 + 2 (n = 51, 18.98%); and 1 + 1 + 1 (n = 4, 2.27%), in which the majority are in the 2 + 1 (AB-C) pattern, forming compound nouns, such as 获得-感 (VP-N) *huode-gan* "sense of gaining", 创客-群 (VP-N) *chuangke-qun* "maker and inventor group", 高考-帝 (NP-N) *gaokao-di* "college entrance examination-king", i.e., a student who repeated the college entrance exam for numerous times. An insertion test shows that the internal structure of

AB-C is no longer accessible, namely, adding a genitive particle 的 *de* would cause ungrammaticality, e.g., 获得感 *huode-gan* "sense of gaining" →* 获得 *de* 感 *huode de gan*, 高考帝 *gaokao-di* "college entrance examination-king" →* 高考 *de* 帝 *gaokao de di*. By the Lexical Integrity Hypothesis (Huang 1984; Duanmu 1998, p. 137), "no phrase-level rule may affect a proper subpart of a word". As such, the AB-C word pattern forms lexical compound nouns instead of phrases.

As shown above, in the three-morpheme AB-C combinations, AB can be an NP or a VP, while C is an affixoid that can be downgraded from a noun or an adjective, such as 哥 (N) *ge* "brother" in 手表-哥 *shoubiao-ge* "watch-brother, a person who owns many watches", and 难 (adj.) *nan* "difficult" in 入学-难 *ruxue-nan* "hard/difficult to find a school for kids". Under the present COVID-19 threat in China, language users have created many AB-Adj. new words to describe their tough challenges, daily hardships, and sufferings with XX-难 XX-*nan* "difficult to-XX", such as:

(1) 看病-难 (VP-Adj.) *kanbing-nan* "hard/difficult to get a doctor";
(2) 扫码-难(VP-Adj.) *saoma-nan* "hard/difficult to scan the code";
(3) 买菜-难(VP-Adj.) *maicai-nan* "hard/difficult to get groceries";
(4) 吃菜-难(VP-Adj.) *chicai-nan* "hard/difficult to get vegetables";
(5) 吃饭-难(VP-Adj.) *chifan-nan* "hard/difficult to get food";
(6) 出门-难(VP-Adj.) *chumen-nan* "hard/difficult to get out of the house";
(7) 乘车-难(VP-Adj.) *chengce-nan* "hard/difficult to commute around";
(8) 回家-难(VP-Adj.) *huijia-nan* "hard/difficult to go home".

As a matter of fact, these words can be easily and precisely expressed in more elaborate phrases or sentences. For instance, 看病-难 *kanbing-nan* "hard to see a doctor" can be more explicitly expressed as 看病很难 *kanbing hennan* "very hard to find a doctor", or 这些日子看病很难，找不到医生 *zhe xie rizi kan bing hen nan, zhao bu dao yisheng* "It is difficult to see a doctor because it is hard to find one". Why do language users choose the terse AB-C format instead? From a historical perspective, Zhou (2014, p. 258) holds that the Chinese language evolved to two- and four-morpheme combinations in Old Chinese to represent the elitist upper culture (雅文化 *yawenhua*). On the other hand, three-morpheme combinations did not rise until the Tang Dynasty (i.e., Medieval Chinese) when the vulgar common culture of towns (市井文化 *shijing wenhua*) started to thrive. In other words, while the two- and four-morpheme words are formal and standard in the Chinese language, the three-morpheme expressions are informal and idiomatic. Phonologically, the three-morpheme AB-C combination carries a single super foot that makes the expression more catchy and colloquial than the others. This may be one of the factors that many day-to-day Chinese terms are usually three-morpheme, such as 身份证 *shenfen-zheng* "ID card", 通行证*tongxing-zheng* "Transportation Card", 银行卡 *yinhang-ka* "credit card", 手机号 *shouji-hao* "cell phone number", 住房卡 *zhufang-ka* "hotel room card", including the most recent market products, such as 支付宝 *Zhifu-bao* "Alipay", 一维码 *yiwei-ma* "one-dimensional code", 二维码 *erwei-ma* "QR code", 扫描器 *saomiao-qi* "scanner", 读取器 *duqu-qi* "reader", etc.

As shown above, the 2 + 1 (AB-C) word pattern has been a popular and productive NP maker in Chinese language use. Now we turn to the newly created AB-C combination 健康码 *Jiankang-ma* "health QR code" and explore how it was created, adopted, spread, and swept through entire China overnight.

3. The Creation, Spread, and Thriving of 健康码 *Jiankang-ma*

3.1. The Explosion of COVID-19 and the Establishment of the National HEALTH QR Code System

The Coronavirus disease 2019 (COVID-19), an acute contagious disease, was first identified in Wuhan, China, in December 2019, which caused serious illness and was transmitted swiftly. Following its outbreak, the disease traveled nationwide, pushing the death toll to thousands. To control its spread, Chinese citizens were required to get tested, vaccinated, or quarantined, for which there needed to be a mechanism to collect, install, and

utilize the data and a new expression(s) to interpret the data to communicate with the administrators and common people. As an experiment, the local government of Hangzhou, the home City of Alibaba Group, developed a health App (digital health certificate) with the model of the Alipay wallet App 支付宝 *Zhifubao*. The Alipay is a barcode/QR Code payment solution that enables merchants to collect money by scanning the barcode/QR code shown on a customer's Alipay wallet App. In addition, the code is auto-refreshed every minute in the App. Alipay was established in Hangzhou, China, in 2004 and has spread nationwide ever since. To help common people understand the new health App and know how to use it, the health App was originally presented as an affiliate to Alipay and named 支付宝健康码 *Zhifubao Jiankang-ma* "Alipay Health Code".

Soon after that, similar mechanisms exploded, which were created one after another by various local governments, such as 深圳防疫健康码 *Shenzhen Fangyi Jiankang-ma* "Shenzhen Epidemic Prevention Health Code", 天津健康码 *Tianjin Jiankang-m* "Tianjin Health Code", 北京健康宝 *Beijing Jiangkang-bao* "Beijing Health Code", 粤康码 *Yuekang-ma* "Guangdong Health Code", etc. To unify the various local codes, the Chinese central government issued a directive of 互联网＋医疗健康 *hulianwang + yiliao jiangkan* "Internet + Medical Health" data system in December 2020, derived from the government strategic plan 互联网+ 行动 *hulianwang + xingdong*" Internet + action" launched in 2015 (Xiao 2019). Accordingly, a national Epidemic Prevention Health QR Code System (国家防疫健康信息二维代码系统 *Guojia Fangyi Jiankang Xinxi Erwei Daima Xitong*) was created, which collected data from individuals that included their COVID-19 testing/vaccine/quarantine status as well as personal information, residence, medical records, travel history, and recent contacts. However, the name of this new system was too technical for common people and had too many syllables for daily communication. Thus, morphological devices were employed to reshape it. First, 11 out of the 14 morphemes were clipped, such as ~~国家~~ ~~防疫~~ ***健康*** 信息 ~~二维~~ ~~代~~***码*** ~~系统~~ ~~*Guojia Fangyi*~~ ***Jiankang*** ~~*Xinxi Erwei Daima*~~ ~~*Xitong*~~ "National Epidemic Prevention Health QR Code System", at the expense of heavy semantic reduction. Secondly, to make the expression catchy and colloquial, the retained three morphemes were selected to fit the 2 + 1 (AB-C) pattern, i.e., 健康-码 *Jiankang-ma* "health-code". Thus, 健康-码 *Jiankang-ma* "health-code" officially came into being, which overrides the (2 + 2) four-morpheme counterpart 健康代码 *Jiankang Daima* "Health QR Code", as it is supposed to be from the literal translation. By comparison, the Health QR Code is presented with a four-morpheme expression 场所代码 *Changsuo Daima* "Location QR Code" in Taiwan, where Mandarin Chinese is the official language as well. Although they both indicate the same system, 健康码 *Jiankang-ma* and 健康代码 *Jiankang Daima* differ phonologically and morphologically. Phonologically, the former is comprised a single super foot, while the latter is made up of two standard metrical feet.

3.2. *健康码 Jiankang-ma "Health QR Code" and Its Word Family XX-码 XX-ma "XX-code"*

According to the "Contemporary Chinese Dictionary" (Dictionary Editorial Office of Institute of Linguistics and Chinese Academy of Social Sciences 2012), the character 码 *ma* serves as different morphemes when used as a noun and a verb. 码 *ma* may be listed as a verb, as in 码放 *mafang* "to stack", a noun, such as 码 *ma* "mark, code", or a measure word, such as 码 *ma* "yard". A survey of the database of *Language Situation in China* 《中国语言生活状况报告》 (GYWGW n.d.) reveals that 码 *ma* has a stable and strong presence in the Chinese language use from 2017 to 2019. In the total characters of annual media use (媒体用字总表 *Meiti Yongzi Zongbiao* "list of total characters in the media") (GYWGW n.d.), 码 is ranked top #904 in 2016, top #873 in 2017, and top #897 in 2018, respectively. On the other hand, 健康码 *Jiankang-ma* had no occurrence in the list of media new words from 2016 to 2018. However, 健康码 *Jiankang-ma* "health QR code" jumped to the top #5 in the media new words of 2020 (National Language Resource Monitoring and Research Center 2020). Using a Google search, the researcher found 1.02 billion tokens of 健康码 *Jiankang-ma* "health QR code" in 0.44 s (13 January 2022), which shows that the frequency token of this new expression boomed from zero to over 1 billion in only one year.

Ever since its launch, 健康码 *Jiankang-ma* "health QR code" has been highlighted in all public documents, newspapers, TV news, social media, and government notices, including the Chinese Embassy's directives for travelers flying to and out of China. Nationwide, it has become a daily necessity for citizens, who need to scan it before entering any public spaces, such as buses, subways, stores, schools, restaurants, etc. A typical example was a sick old lady who did not have a smartphone to install the 健康码 *Jiankang-ma* "health QR code" App and was rejected by the city buses one after the other. In the end, she had to walk miles to get home. With such a regulating and dominating power, 健康码 *Jiankang-ma* "health QR code" has reached every sphere of Chinese citizens' lives and generated a large number of new words associated with the health QR code system and beyond, as shown in below Table 1:

Table 1. Sample New words associated with COVID-19 and beyond.

Word Length Patterns	**Internal Structures**	**New Words Associated with COVID-19**
2 + 1	VP-*ma*	验证码 *Yanzheng-ma* "verification code" 登机码 *Dengji-ma* "boarding code" 畅行码 *Changxing-ma* "free travel code" 散装码 *Sanzhuang-ma* "scattered (non-national) code"
2 + 1	NP-*ma*	健康码 *Jiankang-ma* "health QR code" 核酸码 *Hesuan-ma* "COVID-19 PCR code" 行程码 *Xingcheng-ma* "itinerary code" 海关码 *Haiguan-ma* "customs code" 全国码 *Quanguo-ma* "national code" 地方码 *Difang-ma* "local code" 公司码 *Gongsi-ma* "company code" 商场码 *Shangchang-ma* "shopping mall code" 社区码 *Shequ-ma* "community code" 企业码 *Qiye-ma* "enterprise code"
1 + 1	Adj.-*ma*	绿码 *Lü-ma* "low-risk health code" 黄码 *Huang-ma* "moderate-risk health code" 红码 *Hong-ma* "high-risk health code"
1 + 1	V *ma*	扫码 *sao ma* "to scan code" 解码 *jie ma* "to decode"
2 + 2	N-*ma* + XX/XX + N-*ma*	双码同验 *Shuang-ma tongyan* "double decoding" 一地多码 *Yidi duo-ma* "multiple codes in one place" 一码通行 *Yi-ma tongxing* "to pass with one code" 多码融合 *Duo-ma ronghe* "multi-code merging"
Word length patterns	**Internal structures**	**Sample New words with - 码 unassociated with COVID-19**
2 + 1	VP-*ma*	支付码 *Zhifu-ma* "payment code" 防火码 *Fanghuo-ma* "fire prevention code" 收款码 *Shoukuan-ma* "receipt code"
	NP-*ma*	信用码 *Xinyong-ma* "credit code"

As shown in the table above, there are dozens of new words associated with the COVID-19 threat and beyond, out of which XX-码 XX-*ma* "XX-code" in a 2 + 1 (AB-C) pattern is the majority (n = 18). In addition, there are five in the 1 + 1 word pattern, four in the 2 + 2 word pattern, but none in the 1 + 2 word pattern. To distinguish the degrees of COVID risks, 健康码 *Jiankang-ma* "health QR code" is classified into 绿码 *Lü-ma* "green health code" (low-risk health code), 黄码 *Huang-ma* "yellow health code" (moderate-risk health code), and 红码 *Hong-ma* "red health code" (high-risk health code), which are all abbreviated from multiple-morpheme phrases, e.g., ~~标明~~ ***绿*** ~~颜色的低风险健康~~ ***码*** ~~*biao ming*~~ ***lü*** ~~*yanse de di fengxian jiankang*~~ ***ma*** "health code marked with green color", in which a good number of morphemes are clipped. With the morphological treatment, the sur-

vived combinations contain one standard prosodic foot each, forming a 1 + 1 compound noun. Moreover, 码 *ma* "code" also forms two-syllable verb phrases related to COVID management, such as 扫码 *sao ma* "to scan code", 解码 *jie ma* "to decode", and a few 2 + 2 four-morpheme combinations such as 双码同验 *Shuang-ma tongyan* "double decoding", 一地多码 *Yidi duo-ma* "multiple codes in one place", 一码通行 *Yi-ma tongxing* "to pass with one code", and 多码融合 *Duo-ma ronghe* "multi-code merging", etc. Furthermore, the XX-*ma* family even extends to events or products not related to the pandemic, such as 防火码 *Fanghuo-ma* "fire prevention code".

Morphologically, in the XX-码 XX-*ma* "XX-code" family, -码 *-ma* "-code" is very productive as it can be attached or bound to various constructions, such as an NP or a VP, and changes their word categories to NP. Furthermore, it is noteworthy that the positioning of -码 *-ma* "-code" is also highly fixed as a suffix. Semantically, -码 *-ma* "-code" in 健康码 *Jiankang-ma* "health QR code" is acquiring some new meaning that is restricted to the specific COVID-19 context, while the lexical 码 *ma* "code" as in 代码 *daima* "code" is used more broadly without such a restriction. Given such abilities, 码 *-ma* "-code" can be classified as a pseudo-suffix that generates a large word family through abbreviation and analogy (Zeng and Wang 2011).

3.3. The Transitory Pseudo-Suffix -码 -ma "-code" and the Lexical 码ma "Code"

As discussed above, -码 *-ma* functions as a morphologically bound marker in the new AB-C context. While retaining some of the meaning of the original form 码 *ma* (i.e., mark, code), it can be considered as a transitory pseudo-suffix at the present time due to two factors: (1) it becomes a morphologically bound pseudo-suffix in the new linguistic context; and (2) it has gained some new meaning as the abbreviation of the 健康码 *Jiankang-ma* "health QR code" in the new context and co-exists with the original form 码 *ma*.

By the "functional–typological" approach (Briton and Traugott 2006), language change involves complete replacement over time (i.e., A → B), but changes always involve variation: older forms and newer forms co-exist side by side in the formulation of A → A ~ B → B (p. 6). In the present study, -码 *-ma* "-code" was created as a variant of the lexical 码 *ma* by the Chinese language users in response to the urgent need for clear, concise, and consistent communication in the COVID-19 emergency. While the current usages do not present a strong grammaticalization tendency of -码 *-ma* "-code" into a derivational suffix, ample language data give evidence that this form, highly productive in contexts associated with COVID-19 and the health QR code system, is becoming a pseudo-suffix that acquires new semantics and undergoes fixation of positioning.

If the size of the -码 *-ma* "-code" word family continues to grow with an increasing frequency of use, it is plausible to predict that this form may even continue to evolve through the transition through a "pseudo-suffix" to an "established" suffix eventually.

However, given its stable and strong presence as a content word in the Chinese language use, as shown in the database of *Language Situation in China*《中国语言生活状况报告》(GYWGW n.d.), 码 *ma* may continue to be a lexical word leading a split pathway with its variant pseudo-suffix -码 *-ma*. Cross-linguistic data show that such a development is likely. For instance, in English, the lexical word "back" has quite a few variant forms side by side each other, such as a noun in "my *back* hurts", an adverb in "go *back* home", an adjective in "go through the *back* door", and a verb in "*back* the car out of the garage". As such, chances are that form A may not be completely replaced by form B; instead, it is likely that they will stand side by side in the long run.

4. Discussion and Conclusions

By examining the creation and spread of the nonce formation 健康码 *Jiankang-ma* "health QR code", the study sheds light on the directions of new word creation, which can be bottom up or top down. Unlike the new words created by the Chinese Internet language users in the past decade, which were branded as grassroots and shunned away by the government and mainstream media (Zhuo and Hsieh 2019), the nonce formation

健康码 *Jiankang-ma* "health QR code" was created by the local governments to cope with the explosion of COVID-19 and was adopted by the central government as a dominate measure for the common people. As such, it obtained fast-track access to the mainstream speech community and gained the momentum of spread and expansion. Accordingly, it swept the nation far and wide and generated a large word family XX 码 *XX-ma* associated with the pandemic and beyond.

Linguistically, 健康码 *Jiankang-ma* "health QR code" is in a 2 + 1 word length pattern and has an AB-C internal structure. As the data show, out of the many new words associated with COVID-19, words in the 2 + 1 (AB-C) pattern are the majority, while 1 + 1 and 2 + 2 are the minority. Additionally, there is not a single occurrence of the 1 + 2 pattern. Such findings support the previous studies (Duanmu 2012; Feng 2013, 2017) that the 2 + 1 form is a highly productive word-forming pattern in Chinese grammar, while 1 + 2 (N N) is less favored.

The study also shows that, in the creation of 健康码 *Jiankang-ma* "health QR code" and its word family, a number of morphological devices were employed, such as abbreviation, clipping, and analogy. Moreover, the formation process was mostly constrained by word length pattern and phonological requirements, as Duanmu suggested (Duanmu 2012). To obtain the word patterns that met the phonological requirements, unwanted morphemes were clipped at the expense of semantic reduction. In the process of generating a large word family associated with COVID-19, -码 *-ma* "-code" has also undergone analogical changes and stands for the specific health QR system in the compound words of the XX-码 XX-*ma* "XX-code" family.

According to the three criteria of Chinese pseudo-affixes offered by Zeng and Wang (2011), including morphological productivity, fixation of positioning, and semantic bleaching, it has been observed that the bound root 码 *ma* "code" has developed a new variant, the pseudo-suffix -码 *-ma* "-code" in the creation of 健康码 *Jiankang-ma* "health QR code", which coexists with its original form side by side yet along split pathways.

By examining the semantic, morphological, and grammatical attributes of the pseudo-suffix 码 *-ma* "-code", this study provides an interesting case of a Chinese neologism that shows strong productivity, similar to other pseudo-affix neologisms (e.g., -哥 *-ge* "-brother" and -姐 *-jie* "-sister" in Depner (2020), XX-难 XX-*nan*, "difficult to-XX"). Differing from the political coinages developed top–down or Internet lexical innovations spread bottom–up, we have witnessed a new pattern of creating neologisms in the rise and spread of 健康码 *Jiankang-ma* "health QR code" and the XX-码 *XX-ma* "XX-code" word family, that is, to designate clear and concise labels for the urgent measures taken in the governmental response to public crisis during drastic social changes.

Finally, the development of the pseudo-suffix -码 *-ma* "-code" can be taken as an instance of Emergent Grammar (Hopper 1988, 1998), in which new language patterns surface from discourse as a result of frequent uses. Whether XX-码 XX-*ma* "XX-code" becomes even more wide-spread or dies out in the future, this case study sheds light on the transitory nature of grammar and meaning and, thus, supports a dynamic and emergent view of language use.

Author Contributions: Conceptualization, Y.X. and Y.R.; methodology, Y.X.; data analysis, Y.X.; writing- original draft preparation, Y.X.; writing-review and editing, Y.X. and Y.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are openly available in *Language Situation in China* 2012–2019, at ISBN 9787100095914; 9787100116206; 9787100122610; 9787100139793; 9787100161220; 9787100175067, reference number [15].

Conflicts of Interest: The authors declare no conflict of interest.

Note

[1] Min Hou and Yu Zou are both co-editors/authors of *Language Situation in China* 《中国语言生活状况报告》.

References

Briton, Laurel J., and Elizabeth Closs Traugott. 2006. *Lexicalization and Language Change*. Cambridge: Cambridge University Press.

Bybee, Joan. 2015. *Language Change*. Cambridge: Cambridge University Press.

Cao, Wei 曹炜. 2010. *Modern Chinese Etymology*《现代汉语词汇研究》. Guangzhou: Jinan University Press 暨南大学出版社.

Chao, Yuen Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley and Los Angeles: University of California Press.

Depner, Shelley Ching-yu. 2020. We are family through morphology: The affixoids-ge and-jie in Mandarin Chinese. *Lingua* 241: 102847. [CrossRef]

Dictionary Editorial Office of Institute of Linguistics, and Chinese Academy of Social Sciences. 2012. *Contemporary Chinese Dictionary*, 6th ed. 《现代汉语词典 (第六版) 》. Beijing: The Commercial Press (商务印书馆).

Dong, Xiufang 董秀芳. 2004. *Chinese Lexicon and Morphology*《汉语的词库与词法》. Beijing: Beijing University Press (北京大学出版社).

Duanmu, San. 1998. Wordhood in Chinese. In *New Approaches to Chinese Word Formation: Morphology, Phonology and the Lexicon in Modern and Ancient Chinese*. Edited by Jerome L. Packard. New York: Mouton de Gruyter, pp. 135–96.

Duanmu, San. 2012. Word-length preferences in Chinese: A corpus study. *East Asian Linguistics* 21: 89–114. [CrossRef]

Feng, Shengli. 1995. *Prosodic Structure and Prosodically Constrained Syntax in Chinese*. Ann Arbor: UMI Dissertation Services.

Feng, Shengli. 1998. Prosodic structure and compound words in Classical Chinese. In *New Approaches to Chinese Word Formation: Morphology, Phonology and the Lexicon in Modern and Ancient Chinese*. Edited by Jerome L. Packard. New York: Mouton de Gruyter, pp. 197–260.

Feng, Shengli. 2013. *Chinese Prosodic Syntax*《汉语韵律句法学》. Beijing: The Commercial Press (商务印书馆).

Feng, Shengli. 2017. *Prosodic Morphology in Mandarin Chinese*. London: Routledge.

Gao, Liwei. 2012. Synchronic variation or diachronic change: A sociolinguistic study of Chinese Internet language. In *Chinese under Globalization: Emerging Trends in Language Use in China*. Edited by Jin Liu and Hongyin Tao. Hackensack: World Scientific Publishing Company, pp. 7–28.

Guojia Yuyan Wenzi Gongzuo Weiyuanhui 国家语言文字工作委员会. n.d. *Language Situation in China: 2005–2019* 《中国语言生活状况报告》. Beijing: Commercial Press (商务印书馆).

Hopper, Paul J. 1988. Emergent Grammar and the a priori Grammar Postulate. In *Linguistics in Context*. Edited by Deborah Tannen. Norwood: Ablex, pp. 117–34.

Hopper, Paul J. 1998. Emergent Grammar. In *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure*. Edited by Michael Tomasello. Mahwah: Lawrence Erlbaum, pp. 155–75.

Hopper, Paul J., and Elizabeth Closs Traugott. 2003. *Grammaticalization*, 2nd ed. Cambridge: Cambridge University Press.

Hou, Min 侯敏, and Yu Zou 邹煜. 2015. *Chinese New Words in 2014* 《2014 汉语新词语》. Beijing: Commercial Press (商务印书馆).

Huang, C. T. James. 1984. Phrase structure, lexical integrity, and Chinese compounds. *Journal of the Chinese Language Teachers Association* 19: 53–78.

Jiang, Zongxu 蒋宗许. 2009. *Study on Chinese Affixes*《汉语词缀研究》. Chengdu: Sichuan Publishing Corp. (四川出版社集团巴蜀书社).

Li, Charles N., and Sandra A. Thompson. 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press.

Lü, Shuxiang 吕叔湘. 1963. (Chinese Language), Volume 1. 现代汉语单双音节问题初探，《中国语文》第一期. In《吕叔湘文集：汉语语法论文集》第二卷. Beijing: Commercial Press (商务印书馆), pp. 415–44.

Lü, Shuxiang 吕叔湘. 1979. 汉语语法分析问题. In《吕叔湘文集：汉语语法论文集》第二卷 *Volume 2*. Beijing: Commercial Press (商务印书馆), pp. 483–571.

Lü, Shuxiang 吕叔湘. 2004.《吕叔湘文集：汉语语法论文集》第二卷 *Volume 2*. Beijing: Commercial Press (商务印书馆).

Matthews, Peter H. 1991. *Morphology*, 2nd ed. Cambridge: Cambridge University Press.

McEnery, Anthony, and Zhonghua Xiao. 2004. Swearing in Modern British English: The Case of Fuck in the BNC. *Language and Literature* 13: 235–68. [CrossRef]

National Language Resource Monitoring and Research Center. 2020. *Chinese Stocktaking: Top 10 New Media Words of 2020 in China*; The Website of Ministry of Education of People's Republic of China, December 16. Available online: http://www.moe.gov.cn/jyb_xwfb/gzdt_gzdt/s5987/202012/t20201216_505872.html (accessed on 1 February 2023).

Norman, Jerry. 1988. *Chinese*. Cambridge: Cambridge University Press.

Packard, Jerome L. 2000. *The Morphology of Chinese: A Linguistic and Cognitive Approach*. Cambridge: Cambridge University Press.

Sun, Xin'ai. 2007. (Thoughts on the development Modern Chinese prefixes and suffixes) 对现代汉语词缀问题的几点思考. *Journal of Yanshan University (Philosophy and Social Science Edition)* 8: 80–84.

Xiao, Yun. 2015. The rise of Chinese Internet Language: Product and agent of globalization from below. *Journal of Global Chinese* 1: 111–37. [CrossRef]

Xiao, Yun. 2016. The rise of catch-word "被 *bei*-XX:" Grammaticalization and re-analysis. *Chinese as a Second Language* 《汉语教学研究》 51: 79–97.

Xiao, Yun. 2019. New Words in Contemporary Chinese Language Use: Linguistic Features and Formation Processes. In *Current Studies in Chinese Language and Discourse*. Edited by Yun Xiao and Linda Tsung. Amsterdam: John Benjamins Publishing Company, pp. 5–25.

Xing, Fuyi 邢福义, and Guosheng Wang 汪国胜. 2008. *Modern Chinese Grammar and Rhetoric* 《现代汉语语法修辞》. Beijing: The Higher Education Press (高等教育出版社).

Yang, Aijiao 杨爱蛟. 2005. *Modern Chinese 3-Morpheme Words* 《近代汉语三音词研究》. Wuhan: Wuhan University Press (武汉大学出版社).

Yang, Wenquan 杨文全. 2011. *The Development and Change of Contemporary Chinese Words* 《当代汉语词汇衍生变异研究》. Chengdu: Sichuan Publishing Company (四川出版集团).

Yang, Xiaoping 杨小平. 2012. *New Words and Expressions in Contemporary Chinese*《当代汉语新词新语研究》. Beijing: The Chinese Academy of Social Sciences Press (中国社会科学出版社).

Yang, Yifei. 2007. Properties, functions and development of pseudo prefixes in the new Chinese 3-morepheme usages (新兴三音节准前缀词的性质，功能与发展). *Linguistics Science* 6: 18–57.

Zeng, Liying, and Pingli Wang. 2011. Use a lexical database for identifying quasi-affixes in contemporary Chinese. *International Journal of Knowledge and Language Processing* 2: 53–61.

Zhou, Jian 周荐. 2014. *On Chinese Lexical Structures*《汉语词汇结构论》. Beijing: The People's Education Press (人民教育出版社).

Zhuo, Jing-Schmidt, and Shu-Kai Hsieh. 2019. Chinese neologism. In *The Routledge Handbook of Chinese Applied Linguistics*. Edited by Chu-Ren Huang, Zhuo Jing-Schmidt and Barbara Meisterernst. London: Routledge, pp. 514–34.