



ARTICLE



<https://doi.org/10.1057/s41599-024-03319-4>

OPEN

# The reduction of Netspeak in Mandarin computer-mediated communication: a least effort motivation at the utterance level

Yong Zhou<sup>1✉</sup> & Yicheng Wu<sup>2</sup>

This article addresses the phenomenon of Netspeak reductions with special reference to Mandarin computer-mediated communication. A tentative classification of Chinese Netspeak reductions is first provided, namely, two-, three-, or four-character reductions occurring at both the lexical and the syntactic levels and other atypical reductions. It is then proposed that Zipf's Principle of Least Effort, which usually works at the lexical level, can work as well at the utterance level and can account for the increasing number of Netspeak reductions in daily communication, that is, the frequency of using Netspeak reduction determines its vitality and distribution.

<sup>1</sup> College of International Studies, Yangzhou University, Yangzhou, China. <sup>2</sup> School of International Studies, Zhejiang University, Hangzhou, China.  
✉email: [zhouyongjoe@yzu.edu.cn](mailto:zhouyongjoe@yzu.edu.cn)

## Introduction

The explosion of Internet information has overwhelmingly shaped our lives and brought a profound impact on our language, which accordingly helps to form a new electronic medium of communication, so-called “Cyberspeak” or “Netspeak,” a term coined by Crystal (Crystal, 2001, 2008; see also Chao, 2012 for a detailed discussion) following the Orwellian introduction of “Newspeak” and “Oldspeak” in his famous novel 1984. As a novel type of language form displaying Internet-unique features, Netspeak is described as “succinct,” “functional enough,” “ubiquitous,” and is indicative of an imminent language revolution (Crystal, 2001, p. 19).

The Internet has been indispensable in our communicative life, and contributes to the formation of the so-called Internet-mediated living of modern society (Xie and Yus, 2018, 2021). In China, by June 2022, the number of Internet users had arrived at 1051 million. Thanks to the development of Weibo “microblog” and other chatting vehicles like QQ and WeChat in China, online communications have become increasingly efficient, and it seems tempting for people to type characters as few as possible. One of the striking features of Netspeak is the omission of unnecessary words for communication, which seems to help Mandarin speakers become better language users from the perspective of social pragmatics (Mey, 1994). Unfortunately, Netspeak reduction has been little studied, albeit it figures prominently in daily communication. This present study attempts to provide a pilot account for Netspeak reductions in Mandarin Chinese, focusing on their classification, motivation, and the trend of their development. It is then claimed that Zipf’s Principle of Least Effort can adequately account for how Netspeak reductions emerge and work in daily communication.

The article is organized as follows: “Literature review” section reviews the relevant studies on “computer-mediated communication” (henceforth CMC) and “Netspeak.” “Methodology and sampling” section introduces the method and sampling, i.e., a classification of Chinese Netspeak reductions. In section “The Least Effort motivation for reductions,” it is claimed that Zipf’s Principle of Least Effort can adequately account for how Netspeak reductions emerge and work in daily communication. Section “Conclusion” concludes the article with the implications of the study discussed.

## Literature review

CMC is defined as any human communication that occurs through two or more electronic devices (Bayer et al., 2020), which has attracted linguists’ attention since the 1980s. While the term has traditionally referred to those modes of communication that occur via computer-mediated formats (e.g., instant messaging, emails, chat rooms, online forums and social network services), it has also been applied to other forms of text-based interaction, such as text messaging (Thurlow et al., 2004; Zsido et al., 2020). Relevant studies on CMC mainly focus on the social effects of different computer-supported communication technologies and Internet-based social networking supported by social software, i.e., politeness or impoliteness in the CMC (see Herring, 1996), language choice and code switching (e.g., Paolillo, 1996) and variation in usage—especially of typography and orthography (e.g., Herring and Zelenkauskaitė, 2009; Ling and Baron, 2007; Burke and Kraut, 2016 inter alia). Except for a few studies of language-oriented CMC research in Japanese, Chinese and other languages (e.g., Werry, 1996; Sugimoto and Levin, 2000; Pan, 2002; Nishimura, 2003; Gao, 2007; Xie et al., 2021; among others), the vast majority of others, however, has been conducted mainly regarding about the English language.

Since the Internet is an electronic, global, and interactive medium, the kind of language used on the Internet should not place limits on imagination, or therefore on new words with new meanings and new forms, the so-called principle of limited novelty (Hudson, 2000, p. 241), i.e., new meanings are preferred in old forms, and new forms are preferred in old meanings. With the wide application of CMC technology in modern society, Netspeak, as a new electronic medium of communication, is becoming an inevitable trend and consequently has appeared in many CMC environments like Blog, Twitter and online chatting. However, from a linguistic point of view, CMC tends to include all forms of communication, such as music, photographs, line-drawings, videos as well as languages utilized by CMC. A general term as it is, Netspeak should still be regarded as a subcategory of CMC, just as Crystal has described in his book (Crystal, 2001, p. 242):

The Internet is going to record this linguistic diversity more fully and accurately than was ever possible before. What is truly remarkable is that so many people have learned so quickly to adapt their language to meet the demands of the new situations, and to exploit the potential of the new medium so creatively to form new areas of expression.

The emergence of Netspeak will definitely enrich the range of communicative options available to us, which is also where the significance of Netspeak study lies. Reductions of language have been considered to be characteristic of communication in various modes of CMC and Netspeak (see Crystal, 2001, 2008; Danet and Herring, 2007, p. 27; Anis, 2007; Jonathan, 2015; Yus, 2022 inter alia). In this literature, two general types of process have been investigated: syntactic reduction, on the one hand, and morphological or orthographical reduction, on the other. Syntactic processes include the deletion of subjects and modal auxiliaries (Murray, 2000) and ellipsis (Lee, 2002). Our focus here, though, is on morphological or orthographical reduction (for code-switching in Arabic CMC, see Boudjemaa, 2022). As has been mentioned above, many types of morphological or orthographical reduction occur in CMC or Netspeak discourse, for example, initialisms such as *lol* for “laughing out loud”, vowel deletion as in *ovr* for English “over” or *tjrs* for French “toujours” (“always”), contractions as in *don’t* for English “do not”, and letter/number homophones such as *b* for English “be” and *2* for English “to” or “too” (see Crystal, 2001, 2008, 2011; Anis, 2007; Bieswanger, 2007; Don, 2014; Yus, 2022, among others).

Various types of reductions have just been one of the most salient features of Netspeak. Based on Thurlow and Brown’s (2003) study, however, even in a heavily abbreviated language like English, the reductions make up less than 20% of the content of the messages in their data. Additionally, the type of reductions also varies. Crystal (2001, p. 84) claims that “acronyms are common in Netspeak.” He also lists over 100 shortenings used in Netspeak conversations, which consists almost exclusively of initialisms. However, based on a cross-linguistic survey, Bieswanger (2007, 2013) conducts quantitative analyses, showing that initialisms are rarely used. Baron (2008) has found only eight initialisms in 191 text messages written by undergraduates at the University of Michigan, or 0.04 initialisms per message. Apart from the cross-linguistic differences shown by reductions, Bieswanger (2007) maintains that the overall number of shortenings used in the German data was even lower than that in the English corpus. Specifically, in the German corpus, there are 0.86 shortenings in each message on average, whereas the average number of shortenings per message is 5.57 in English, whose frequency is thus six times that of analyzed

**Table 1** The samples of three- and four-character reductions.

Three-character reductions	Four-character reductions
<i>ǎi cuò qióng</i> "a short, ugly and poor guy"; <i>bái fū měi</i> "a fair-skinned, rich and beautiful lady"; <i>bái gǔ jīng</i> "female office elites"; <i>chéng huì wán</i> "You urban folks are genuinely born to have fun"; <i>gāo fù shuài</i> "a tall, rich and handsome guy"; <i>huó jiǔ jiàn</i> "The longevities see more things"; <i>lěng wú quē</i> "a cold-hearted guy without belief"; <i>rán bìng luǎn</i> "But it is of no damn use"; <i>suī xiàn gǒu</i> "Though I remain single now"; <i>yǔ shī zǎo</i> "Chinese teacher died early"; <i>xiǎo què xìng</i> "little dose of genuine happiness"	<i>bú míng jué lì</i> "I don't quite get it, but I think you are terrific"; <i>bú yuē ér tóng</i> "He or she becomes a gay (lesbian) after such a long time without dating"; <i>lèi jué bú ài</i> "Too tired to love"; <i>nán mò nǚ lèi</i> "males would stop talking and females would shed tears when seeing this"; <i>rén jiān bú chāi</i> "Some lies are better not exposed, as life is already so hard."; <i>shí dòng rán jù</i> "I was scared though I was deeply moved but still rejected him"; <i>shuō nà jué yú</i> "I feel myself an unwelcome outsider, while the other people are having a good time talking and frolicking"; <i>wú zhī shǎo nǚ</i> "an unaffiliated intellectual female official"; <i>xī dà pǔ bēn</i> "The news is so exhilarating that everyone is celebrating and spreading it to the rest of the world"; <i>xì sī kǒng jí</i> "If you think it over, you would feel extremely horrible"

German text messages. Herring and Zelenkauskaitė (2009) contend that abbreviations also demonstrate gender variation. For instance, in mobile phone text messages (SMS) posted to a public Italian interactive television (iTV) program, female participants used more instances of abbreviations (776 cases) than male participants did (611 cases), as Table 2 illustrates in their study. Lotherington and Xu (2004) have investigated the reductions in English and Chinese, revealing that homophones were popularly used in Chinese. Xiao (2015) further claims that the rise of Chinese Internet Language (CIL) is determined by globalization and the language use of grassroots people, i.e., the social changes "from below." Arndt-Lappe (2018) explores the variability of shortening processes across languages, such as Italian, German, and English, proving that clipping is a unified phenomenon that could not be predicted. While Hilpert et al. (2021) maintain that English clipping followed predictable tendencies and conformed to a probabilistic, multifactorial model, which could be explained functionally in terms of cognitive, discourse-pragmatic, and phonological factors. The findings above clearly demonstrate the general characteristics of the Indo-European language Netspeak.

To sum up, existing studies have revealed considerable language-related variations concerning the overall frequency and the types of reductions used in Netspeak. But nearly all findings are based on the data of Indo-European languages. It is a different picture in isolating languages like Chinese. Due to its unique writing system, there is a substantial propensity for lexical reductions and syntactic reductions. Lexical reduction here is similar to initialism, and usually contains the first character of each phrase.

### Methodology and sampling

**Methodology.** This present study attempts to provide a pilot account for Netspeak reductions in Mandarin Chinese, with a particular focus on their classification, motivation, and the trend of their development. The following two research questions will be addressed:

- What is the classification and distribution of Netspeak reductions in Mandarin CMC?
- How are Netspeak reductions in Mandarin CMC motivated?

To address the two questions, two study variables, viz., the distribution of Netspeak reductions in Mandarin CMC and the frequency of using Netspeak reductions, will be taken into consideration by analyzing the collected data from the Internet. Specifically, this study attempts to examine the classification, distribution and motivation of Mandarin Netspeak reductions under a hybrid approach, i.e., qualitative and quantitative.

In terms of qualitative analysis, the data are first carefully selected from BCC corpus (<http://bcc.bcu.edu.cn>), a 15-billion-word Chinese corpus set up by Beijing Language and Culture University, which encompasses data from newspapers, literature, technology as well as microblogs like *Sina Weibo* (Xun et al., 2016). Yet, due to the limitation of the update process, some of the data are directly taken from *Sina Weibo* search. Then, a tentative classification of Chinese Netspeak reductions is provided, namely, two-, three-, or four-character reductions occurring at the lexical and the syntactic levels and other atypical reductions. Due to the disyllabification of vocabulary, a remarkable feature of Mandarin Chinese, it would be impossible to carry out an exhaustive analysis of two-character reductions, whereas the number of three- and four-character reductions is somewhat limited and thus could be thoroughly investigated. Accordingly, we will examine the distribution of Netspeak reduction in Mandarin CMC with a particular focus on three- and four-character reductions composed of 11 three-character reductions and 10 four-character reductions, respectively, illustrated and exemplified in Table 1.

Subsequent to that, to justify the reliability of the collected data, this study conducts semi-structured interviews with 20 L1 Chinese speakers, 10 males and 10 females, who are recruited from the English Department at a Chinese university. Interviews consist of questions concerning the acceptability of the collected data. Specifically, the participants would be asked whether the reductions are acceptable when seeing them, partially acceptable when placed in the context or unacceptable. Only those reductions which are accepted by 90% of participants will be chosen. Some old-fashioned or ambiguous reductions would be excluded. For example, *yǔ shī zǎo*, which is contracted from the sentence *Yǔwén lǎoshī sǐ de zǎo* "Chinese teacher died early," satirically means someone has a poor command of his/her native language, and *bái gǔ jīng*, which is shortened from *báilǐng gǔgàn jīngyīng* "female office elites," originally means "white-boned devil" in the classic Chinese novel *Journey to the West*. Finally, Zipf's Principle of Least Effort is applied to offer a reasonable explanation of Netspeak reductions in Mandarin Chinese. In the meantime, some predictions on the tendency of Netspeak reductions in Mandarin Chinese will be tentatively provided.

In terms of quantitative analysis, the frequency of some Mandarin Netspeak reductions in the corpus of BCC and *Sina Weibo* will be initially checked, proving the prevalence and vitality of some types of reductions. In addition, some index data like *Sougou* (the most popular Chinese input software available from <http://zhishu.sougou.com>) and *Baidu* (the number one search engine in China available from <http://index.baidu>) will be employed to find the frequencies of words and

phrases for the purpose of predicting the distribution and the diachronic change of the Mandarin Netspeak reductions (see also La, 2019), which will provide the empirical evidence for the qualitative discussion.

To sum up, a hybrid method, qualitative and quantitative, is adopted to examine the developing tendency of the selected Mandarin Netspeak reductions retrieved from the Internet, so as to offer a holistic picture of Netspeak reductions in Mandarin CMC.

**Sampling.** The developments of the Internet and the limitation of character space (no more than 160 characters) for *Weibo* have motivated the evolution of reduction forms emerging in virtual worlds. In order to speed up communication and display creativity by playing with language, reductions in Chinese Netspeak discourse are springing up like mushrooms. Before addressing the classification of Netspeak reductions in Chinese, we first take a look at corresponding reductions in English, which usually represent words, phrases and sentences (Crystal, 2008).

(1) a. “LIKE” if you agree with me...!!!! then (Pls. copy and paste and spread this message) # To President Osama?

b. IMO his future is at another position, that’s all. I do not say he is not a good player, just not a great one.

c. As consequences, you need to kill it from Activity Monitor (or Terminal) to shut down the app. # HTH.

As indicated by the above examples, which are adopted from COCA (<https://www.english-corpora.org/coca/>), *Pls* in (1)a is shortened for “please”, *IMO* is the abbreviated form of the phrase “in my opinion”, whereas *HTH* in (1)c is contracted from the clause “Hope this helps”. Due to its unique orthographical convention and writing system, Chinese reductions in Netspeak can be classified into two-, three-, four-character reductions and other atypical reductions as indicated by the examples below<sup>1</sup>. First, consider the two- and three-character reductions below:

(2) a. Wǎngluò hóng rén → Wǎng hóng  
Internet red person Internet red

“Internet celebrity”

b. *Tài xiàng le, zhè děng měitú shuǐpíng bú dāng wǎng hóng tài kě xī le!*

“The photo looks exactly like the original. What a pity if you don’t become an online star because you are so good at beautifying the pictures.”

(3) a. Gāoduān dàqì shàng dàngcì → Gāo dà shàng  
high-end magnificent classy high big up

“high-end and groovy”

b. *Zhè chē búshì tǔháojīn, bú gòu gāo dà shàng.*

“This car is not luxury-gold, which is not in a high-end atmospheric grade.”

(4) Chéng lǐ rén zhēn huì wán. → Chéng huì wán.  
a. city inside people true can play city can play

“You urban folks are genuinely born to have fun.”

b. *Chéng huì wán, yǐhòu bùyòng huāqián mǎi wánjù le.*

“You urban folks are genuinely born to have fun. And you won’t spend money on the toys in the future.”

As indicated by example (2), 70% of lexicons in Modern Chinese are disyllabic (see Huang and Liao, 2007), which reflects

the tendency of word formation in modern Chinese. Nevertheless, most new words or so-called network catchwords are trisyllabic, because it is efficient to create new words out of existing ones by blending a morpheme with a bisyllabic word, which could take place both in lexical and sentential levels as shown in (3) and (4) respectively. Lexical reduction as in (3) is similar to initialism in the way it only retains the initial character of each expression, which is usually used to describe something high-end and groovy, and has occupied the second place of Top Ten Netspeaks of China in 2013. While *chéng huì wán* in (4) which literally means “You urban folks are genuinely born to have fun” is selected as one of the Top Ten Netspeaks of the Year 2015 released by the State Language Commission of China. As regards the idiom-like four-character reductions, they could also occur at both the lexical and the sentential levels, as shown below:

(5) a. Xǐ wén lè jiàn; dà kuài rén xīn;  
happy see joyful see big delight people heart  
pǔ tiān tóng qīng; bēn zǒu xiāng gào  
universal sky together celebrate run walk mutual tell  
→ Xǐ dà pǔ bēn  
happy big universal run  
“The news is so exhilarating that everyone is celebrating and spreading it to the rest of the world.”

b. *Jīntiān bèi nánshén diǎnmíng le, xǐ dà pǔ bēn.*

“It was so exhilarating that the male god called my name today.”

(6) a. Zīxī xiǎng xiǎng jué dé kǒngbù zhì jí  
careful think think feel GEN horrify until extreme  
→ Xī sī kǒng jí  
careful think horrible extreme

“If you think it over, you would feel extremely horrible.”

b. *Yě zhēnde yǒu rén wéi le fángzi jiá líhūn. xī sī kǒng jí.*

“It happens that some people want to get divorced for the house! You would feel extremely horrible when thinking it over.”

(7) Wǒ Xǐhuān nǐ, nǐ zhīdào ma?  
I like you you know  
→ Wǒ xuān nǐ, nǐ zào ma?  
I declare you you make PRT

“Do you know that I love you?”

Example (5a) is the reduction of several phrases, which denotes that “The news is so exhilarating that everyone is celebrating and spreading it to the rest of the world”. In the sentential reduction of (6a), *xiǎng* “think” is substituted by its synonym *sī*, simply because *xiǎng* is seldom modified by *xī* “careful”, which indicates that the Netspeak reductions tend to observe the collocation rules of Mandarin Chinese. Except for lexical and syntactic reductions discussed above, reductions sometimes occur at the phonological level. There are two reductions in (7). One is *xuān* contracted from *xǐhuān* “like”, and the other is *zào* contracted from *zhīdào* “know”. The reductions come into being, simply owing to the fast speed of utterance and the assimilation of the two syllables<sup>2</sup>.

So far, we have discussed all the classifications of reductions in the Chinese Netspeak discourse. A question that naturally arises is, what motivates such reductions? In what follows, we will address this issue.

### The Least Effort motivation for reductions

As proposed in the previous section, the reductions are initially motivated by character space limitations and typing speed. The question is, are there any other factors that motivate Mandarin



Netspeak reductions? In this section, we will first discuss such possible alternatives, and then explain why a least-effort account might be superior to others. Then, a reasonable explanation of Netspeak reductions in Mandarin Chinese is presented in terms of the *Principle of Least Effort*. Finally, some predictions on the tendency of Netspeak reductions in Mandarin Chinese are tentatively made, that is, the vitality of Netspeak reductions should be substantially determined by the frequency of their use.

In general, reductions are used deliberately in Netspeak discourse, which, to a great extent, reflects technological and social motivations. Obviously, reductions could economize on typing effort and speed up typing, particularly in synchronous Netspeak discourse (Nishimura, 2003). Nevertheless, it should be noted that the linguistic system of languages also determines reductions. For instance, we have to use *sī* “think” instead of *xiǎng* in *xì sī kǒng jí* “If you think it over, you would feel extremely terrified”, due to Mandarin idiomatic collocation of *sī* “think” with *xì* “careful.” This means technological innovation or typing efforts could only be regarded as an immediate or external factor governing the formation of Netspeak reductions. On the other hand, reductions display creativity and “indicate familiarity and intimacy between users” (Lee, 2002, p. 10). A case in point is the so-called new translanguaging script (Li and Zhu, 2019), such as *rè skr rén* (*rè sǐ gè rén*) “terribly hot” in which *skr* is homophonic with the Chinese words *sǐ gè* due to the similarity of pronunciation. It should be noted that such reductions are also a distinctive feature of Mandarin asynchronous discourse like blogs or email exchanges in Chinese. Though these forms are novel, they are rather difficult to follow if there is a lack of transparent context because the manipulation of linguistic forms is essentially playful and “the motivations behind it may be quite personal” (Li and Zhu, 2019). Creativity, therefore, is only one of the characteristics, and it is not the primary motivation behind Netspeak reductions in Mandarin Chinese.

The reductions could also be explicated within pragmatic theories like the Gricean Cooperative Principle and Sperber and Wilson’s relevance theory. Bieswanger (2013) claims that the use of abbreviations in CMC is in parallel with the Grice (1975) maxims of manner, the demand for brevity and quantity, whereas Yus (2011) examines the CMC contexts from the perspective of Relevance Theory. It is worth mentioning that the two theories both take the conveyed messages as their focus of study, and the only task for the interlocutors is to figure out the implicatures from the literal meanings, which is sometimes not the case for the Netspeak reductions in Mandarin Chinese. In reductions like *chéng huì wán* “You urban folks are genuinely born to have fun” and *rán bìng luǎn* “But it is of no damn use,” the crucial information *rén* “folks” and *méi yǒu* “no” remain missing. While in reductions like *biǎo jiàngzi* “Do not behave like this,” only articulation effort works as we could not determine the implied meaning out of their literal ones. More importantly, both Gricean Cooperative Principle and Sperber and Wilson’s relevance theory mainly deal with how to infer suggested meanings from the viewpoint of the addressee instead of the addresser, who initially creates novel utterances like the reductions. As Xiang (2017) has noted, “Every effective utterance is the result of the interlocutor’s efforts to strive after the optimal equilibrium of the cost and utility in their verbal communication.” The survival and flourishing of Netspeak reductions in Mandarin Chinese should be guaranteed by achieving a compromise between addresser and addressee at the cost of least effort.

From what has been discussed above, typing speed and creativity, which are the manifestations of language economy, could only partially account for the emergence of Netspeak reductions

in Mandarin Chinese. Albeit some participants (5 out of 20) in semi-structured interviews admitted that the typing speed and creativity, to some extent, had attracted them, and it was really cool to use these reductions in online communication. A safe and plausible solution is to work out the common ground shared by all alternative accounts, that is, the idea of minimizing the vocabulary guided by the least effort. Up to now, it is safe to contend that Netspeak reductions in Mandarin Chinese should be motivated and accounted for by the Zipfian Principle of Least Effort.

The Harvard linguist Zipf (1935, 1949) first noticed that the same regularities in all languages on use frequencies of words follow a universal property of the human mind. He postulates a principle that would explain the equilibrium between uniformity and diversity in the use of words. He explicates the principle, which is entitled *Principle of Least Effort* as follows: “the primary principle that governs our entire individual and collective behavior of all sorts, including the behavior of our language ...” (Zipf, 1949, p. viii). According to Zipf, the distribution of word use was due to the tendency to communicate efficiently with least effort.

In the realm of linguistic behavior, “a force pressing in the direction of a single word or sound that would express all possible meanings, thereby sparing the speaker the mental effort that is involved in selecting the appropriate linguistic forms for the meanings she wants to convey” (Zipf, 1949: 20), which would be contradictory between speaker’s and hearer’s economies of effort. Zipf further distinguished such efforts, one being oriented toward minimal linguistic articulation, and the other being geared toward maximal explicitness. Only by achieving the balance of the cognitive effort made by the speaker and hearer could the *Principle of Least Effort* be fully realized.

Martinet (1962) also iterated the *Principle of Least Effort* and described a primary mechanism of language change as coming from the interaction of two factors: “first, the requirements of communication, the need for the speaker to convey his message, and second, the principle of least effort, which makes him restrict his output of energy, both mental and physical, to the minimum compatible with achieving his ends” (Martinet, 1962, p. 139). In addition, the *Principle of Least Effort* has also been applied to account for human articulation. Levinson (2000, p. 27–30) calls it an “articulatory bottleneck,” which refers to the slow transmission rate of human speech compared with other aspects of speech production and comprehension like prearticulation or mental preparation processes. In other words, human beings should minimize articulation and reduce the risk of creating disturbances caused by too much information. That is the very story of why Netspeak reduction happens.

Mandarin Chinese, which is often labeled as a pragmatically-oriented, paratactic language, is characterized by hidden complexity (Bisang, 2014, 2015), in the sense that what is morpho-syntactically represented often does not match what is semantically expressed, a sort of form-meaning mismatch that needs to be resolved by resorting to pragmatic inferencing, precisely pragmatic enrichment, which is partially responsible for the creation of Netspeak reductions. It also suggests that any smooth communication could be guaranteed by achieving a balance between the language system itself and language use, which implies that Least Effort also works at the utterance level. A vivid case in point is Netspeak reduction in Mandarin Chinese, which observes the *Principle of Least Effort*, simply because it is energy-consuming to maintain a large vocabulary. In Netspeak, netizens could then use reduced forms to convey the message in an economical way, and audiences could obtain the intended meanings with relative ease in the online communicative context, which facilitates the widespread of such reductions on the Internet and

even in daily conversation. In what follows, we will examine the procedure on how the Netspeak reductions in Mandarin Chinese evolve and get motivated.

**Least Effort motivation of Netspeak reductions in Mandarin Chinese.** It has been observed that a concept is a bundle of semantic features or “genes of meaning”, in Zipf’s words. If a particular bundle occurs frequently enough in a specific community, they will be assigned a phonological representation, a word. If it occurs infrequently, no specific word would be available. When this concept emerges, therefore, people will have to describe it using a string of words or a phrase. As language evolves, a relationship would be established between the length of words and the frequency of occurrences of the phenomena they describe, that is, the frequency will minimize the expected length of a word or sentence (see Rousseau, 2002).

The Netspeak reductions in Mandarin Chinese should supposedly have the same working mechanism. The more frequently an expression is used on the Internet, the more likely a reduced form comes into being. In addition, the word limit of Weibo and some other online chatting vehicles drive the development of Netspeak reductions in Mandarin Chinese to a great extent. Here, we cite some examples from Sina Weibo search.

(8)	Aoliào	shàng	jìngrán	kěyǐ	zuò	huà?!	chéng	huì	wán!!!
	Oreo	on	even	may	draw	picture	city	can	play
	"You can even draw on Oreo cookies?! You urban folks are genuinely born to have fun!!!!"								
(9)	Zhōngyú	dào	shēnzhèn	běi	le,	rán	bìng	luǎn	
	finally	arrive	Shenzhen	north	PRT	but	furthermore	ovum	
	wǒ	zuò	fǎn	le,	dièr	cì	le.		
	I	seat	opposite	PRT	second	CLF	PRT		
	"Finally, I arrived at the north of Shenzhen, but it was of no damn use. I took the wrong direction for the second time."								
(10)	Bān	wán	jiā	yǐ	lèi	jué	bú	ài	le,
	move	finish	home	already	tired	feel	no	love	PRT
	suìbiàn	shài	shài	wǒ	de	xīn	shuìyī	...	
	casually	show	show	my	GEN	new	pajamas		
	"I'm too tired to love after moving to a new place. I just show my new pajamas for relaxation..."								

Following the *principle of least effort*, the examples above are increasingly popular in online communication, whose use frequency, the occurrences of *chéng huìwán*, *rán bìng luǎn* and *lèi jué bú ài* in Sina Weibo on 21 February 2020, arrive at 45, 92 and 60 times, respectively. This persistent popularity of such reductions means both addresser and addressee can grasp the exact meaning of *chéng huìwán*, *rán bìng luǎn* and *lèi jué bú ài* easily, thus achieving a better compromise between the efforts of both sides, which is also in line with the speaker and auditor’s economy by Zipf (1949). As has been reasonably pointed out in Carston (2005), however, there is a split here between the speaker’s and hearer’s effort in the communication process. The unification of two sides is only possible if the effort at issue is that of the hearer rather than that of the speaker, that is, the hearer should take less effort to process, which will be supported by corpora data shown in Table 2<sup>3</sup>.

In Table 2, it is not surprising that some three-character reductions such as *gāo fù shuài* “a tall, rich and handsome guy” and *bái fù měi* “a fair-skinned, rich and beautiful lady”, together with some four-character reductions such as *rén jiān bú chāi* “Some lies are better not exposed, as life is already so hard”, *lèi jué bú ài* “Too tired to love” and *bú míng jué lì* “I don’t quite get it, but I think you are really terrific,” occupy the leading position in the corpora. This is possible because people could quickly grasp the connotative meanings of these reductions. As we have noticed in the semi-structured interview, most of the three-character reductions could be identified by all participants simultaneously, suggesting that the hearer or receiver’s effort matters most during

Table 2 The frequency of three- and four-character reductions taken from the corpus of BCC.			
Types of reductions	Cases of reductions	Numbers of tokens	Ranks
Three-character reductions	<i>gāo dà shàng</i>	907	7
	<i>lěng wú quē</i>	31	16
	<i>bái fù měi</i>	7851	2
	<i>gāo fù shuài</i>	9742	1
	<i>ǎi cuó qióng</i>	235	11
	<i>xiǎo què xìng</i>	499	9
	<i>qǐng yǐn bēi</i>	226	12
	<i>yǔ shī zǎo</i>	140	13
Four-character reductions	<i>xǐ dà pǔ bēn</i>	1413	6
	<i>lèi jué bú ài</i>	3181	4
	<i>bú míng jué lì</i>	1420	5
	<i>xì sī kǒng jí</i>	507	8
	<i>rén jiān bú chāi</i>	3563	3
	<i>shí dòng rán jù</i>	421	10
	<i>shuō nà jué yú</i>	75	15
	<i>nán mò nǚ lèi</i>	127	14

the interpretive process of Netspeak reductions. Among the top five reductions in the table above, it is worth noting that three-character reductions like *gāo fù shuài* “a fair-skinned, rich and beautiful lady” and *bái fù měi* “a tall, rich and handsome guy,” the cases of the lexical reduction came into being and became popular in the year of 2008, whereas the left three four-character reductions were coined in the year of 2013. If such external factors were excluded, the utilization of four-character reductions, on average, seems to be more frequent and productive than that of three-character reductions, which is determined not only by the prevalence of four-character idioms in Mandarin Chinese, but also by the hearer-friendliness of such four-character reductions, i.e., the hearers have no trouble understanding the intended meanings.

In a word, the Netspeak reductions in Mandarin Chinese are generally in accord with the *Principle of Least Effort*. In the meantime, it should be noted that the speaker and hearer, as different individuals, are equal participants in the online communicative process. The comprehension of Netspeak reductions heavily relies on the most accessible and salient knowledge in online communicative contexts. This means the balance between the speaker’s and the hearer’s efforts must be achieved. If the hearer’s effort overweighs that of the speaker’s, the Netspeak communication would possibly collapse, and such reductions will be eliminated for the time being. In relevance-theoretic terms, in all instances, the extra processing effort required of the hearer is offset by particular cognitive effects that the less effort-demanding variant would not have had (Sperber and Wilson, 1995). Only those Netspeak reductions in Mandarin Chinese observing such least effort motivation could survive and flourish.

**The tendency of using Netspeak reductions in Mandarin Chinese.** As mentioned in the previous section, Netspeak reductions are highly popular in Sina Weibo. Nine of the 28 specific samples of Netspeak reductions have become the hot words of Weibo data (available from <http://data.weibo.com>). In order to present an accurate description of the frequency of using Netspeak reductions, we pick out seven of them as the subject of our study based on the classification in previous sections, six of which are chosen from the top ten reductions in Table 2. In contrast, *jiàng zǐ* “behaving like this” representing the phonological reduction is listed as the reference item. The data utilized in the following part

are accessed from the *Sougou* index and *Baidu* index, which will plausibly manifest the latest tendency for Netspeak reductions in Mandarin Chinese.

First, let us examine the data from the *Sougou* index, which lasts from 23 January 2022 to 22 February 2022. Table 3 shows the variations of sample reductions within 31 days delivered by the *Sougou* index.

In Table 3, the frequency of reductions is evident in the span of 30 days. Lexical reductions such as *gāo fù shuài* “a tall, rich and handsome guy” and *xiǎo què xìng* “little dose of genuine happiness,” whose numbers reach 5801 and 4990 times, rank the highest in the average index, while another *xǐ dà pǔ běn* “The news is so exhilarating that everyone is celebrating and spreading it to the rest of the world,” by contrast, occupies the fifth position in the table with an average index of 262 times, which possibly means it costs more processing efforts in typing four-character lexical reductions. Then the phonological reduction, *jiàngzǐ* “behave like this”, reaches 16892 times and ranks No. 3 in the list. Finally, syntactic reductions represented by *rén jiān bú chāi* “Some lies are better not exposed, as life is already so hard”, *rán bìng luǎn* “But it is of no damn use” and *lèi jué bú ài* “Too tired to love” have only been typed 298, 210, and 170 times, respectively. The highest single-time frequency is in line with the average one, which generally depicts the frequency order of the sample Netspeak reductions, that is, lexical reductions > phonological reductions > syntactic reductions. In the long run, we could also obtain a similar result as given in Table 3, based on *Baidu* index data spanning from year 2011 to 2023.

In Table 4, we can clearly see the change in the search index over the past nine years. The frequencies are as follows: As for the highest index, three-character lexical reduction *xiǎo què xìng* “little dose of genuine happiness” occupies the first place with 39,536 times per day, and three-character syntactic reduction *rán bìng luǎn* “But it is of no damn use” the second place with 33,625 times, then four-character lexical reduction *rén jiān bú chāi* “Some lies are better not exposed, as life is already so hard” with 15,334 times followed by two lexical reductions, *xǐ dà pǔ běn* “The news is so exhilarating that everyone is celebrating and spreading it to the rest of the world”

and *gāo fù shuài* “a tall, rich and handsome guy”, which reach 10,152 and 10,303 times. After that, another four-character syntactic reduction like *lèi jué bú ài* “Too tired to love” with 2394 times follows, and finally the phonological reduction *jiàngzǐ* “behave like this” is searched only 1824 times. In terms of average index, Table 4 is in line with Table 3 with a minor change in the position of phonological reduction following the order of lexical reductions > syntactic reductions > phonological reductions. Simply because the obtained data in Table 3 is based on input software, and it is more convenient for people to type *jiàngzǐ*, a two-syllable string, instead of *zhè yàngzǐ*, a three-syllable string. While the data collected in Table 4 are based on *Baidu*, a search software, it partially reflects the heat of some reductions and the netizens’ curiosity about its meanings. This accounts for why some newly and vaguely coined reductions like *rán bìng luǎn* and *rén jiān bú chāi* are searched so frequently. It should also be pointed out that Netspeak reductions bear the imprint of Internet-spawned neologisms, which are generally deemed cool, superficial, and short-lived. That is why all the reductions in Table 4 are on a declining trend after reaching their peak. Taking the average index in Table 4 into consideration, we can safely conclude that the variations of lexical reductions together with syntactic reductions are more stable than that of phonological reductions, for the simple reason that what is reserved in the process of lexical or syntactic reductions contains more information under the guidance of Least Effort Principle. In other words, it is difficult for phonological reductions to achieve brevity without the loss of useful information, which clearly shows the tendency of Netspeak reductions in Mandarin Chinese. That is, lexical reductions and syntactic reductions spread widely and persistently.

Conclusion

In this article, we have addressed some issues concerning the Netspeak reductions in Mandarin Chinese. First, we have provided a classification of Chinese Netspeak reductions, viz., two-character, three-character or four-character reductions occurring at both the lexical and the syntactic levels and other types of phonological reduction. It has then been pointed out that the vitality of Netspeak reductions is substantially determined by the frequency of their use. The frequency of using Netspeak reductions is, in turn, mainly determined by the mechanism behind the Netspeak reduction, i.e., the two competing motivations: economy and explicitness. It is noteworthy that other factors like coolness and creativity could also contribute to the diffusion and frequency of the Netspeak reductions<sup>4</sup>.

In this pilot study, we venture to offer a holistic picture of Netspeak reductions in Mandarin CMC with special reference to its distribution and developing tendency. Admittedly, there may be room for improving the proposed analysis. Due to the lack of online Netspeak corpora, the material analyzed is somewhat limited, and the trend of Mandarin Netspeak

Table 3 The use frequency of sample Netspeak reductions based on the <i>Sougou</i> index.		
Reductions	The average index	The highest index/date
<i>xiǎo què xìng</i>	4990	11,347/06.02
<i>rán bìng luǎn</i>	170	242/16.02
<i>gāo fù shuài</i>	5801	7843/13.02
<i>xǐ dà pǔ běn</i>	262	426/01.02
<i>lèi jué bú ài</i>	210	338/21.02
<i>rén jiān bú chāi</i>	298	492/19.02
<i>jiàng zǐ</i>	392	522/07.02

Table 4 The use frequency of sample netspeak reductions based on the <i>Baidu</i> index.			
Reductions	Lowest index/date	Highest index/date	Average index
<i>xiǎo què xìng</i>	137/30.01.2012	39,536/31.12.2017	1030
<i>rán bìng luǎn</i>	8/29.03.2015	34,794/29.06.2015	1138
<i>gāo fù shuài</i>	17/04.04.2011	10,303/09.04.2012	1269
<i>xǐ dà pǔ běn</i>	8/24.09.2012	10,512/21.12.2013	1224
<i>lèi jué bú ài</i>	8/10.09.2012	2394/05.04.2014	717
<i>rén jiān bú chāi</i>	408/12.04.2013	15,334/16.09.2013	1659
<i>jiàng zǐ</i>	261/23.12.2017	1824/16.06.2012	465



reductions needs to be further observed, which is what we will do in the future study.

## Data availability

All data analyzed are contained in this article.

Received: 2 February 2024; Accepted: 10 June 2024;

Published online: 18 June 2024

## Notes

- 1 The Netspeak reductions addressed in this article are first given a word-for-word gloss and then an English translation of their intended meanings. And the symbol → should be construed as “be contracted into”. Abbreviations used in the gloss are as follows: CLF, classifier; COP, copular; GEN, genitive; NEG, negative; PRT, particle.
- 2 In some local varieties of *Putonghua* (Standard Chinese), such as Nanping in Fujian Province, the syllable *yàng* is usually linked to the syllable of the preceding morpheme *zhè*, which results in the reduction *jiàng* and which is possibly the origin of such phonological reductions (see Liu, 2012).
- 3 It should be noted that the corpus of BCC only contains *Weibo* data for the year 2013, so in this table we have to exclude some reductions that did not emerge at the time, viz., *rán bìng luǎn* “But it is of no damn use”, *chéng huì wán* “You urban folks are genuinely born to have fun”, *huó jiǔ jiàn* “The longevities see more things” and *sù xiàn gǒu* “Though I remain single now” as well as some ambiguous ones like *bái gǔ jīng* “female office elites”, *wú zhī shǎo nǚ* “an unaffiliated intellectual female official” and *bú yuē ér tóng* “He or she becomes a gay (lesbian) after such a long time without dating”.
- 4 As an anonymous reviewer correctly points out, economy influences the original creation of the Netspeak reduction, whereas their diffusion and frequency could be related to other factors like coolness.

## References

- Anis J (2007) Neography: unconventional spelling in French SMS text messages. In: Danet B, Herring SC (eds) *The multilingual Internet: language, culture, and communication online*. Oxford University Press, New York, p 87–115
- Arndt-Lappe S (2018) Expanding the lexicon by truncation: variability, recoverability, and productivity. In: Arndt-Lappe S, Braun A, Moulin C, Winter-Froemel E (eds) *Expanding the lexicon: linguistic innovation, morphological productivity, and ludicity*. Mouton de Gruyter, Berlin/Boston, p 141–170
- Baron NS (2008) *Always on: language in an online and mobile world*. Oxford University Press, New York
- Bayer JB, Triep P, Ellison NB (2020) Social media elements, ecologies and effects. *Annu Rev Psychol* 71(1):471–497
- Bisang W (2014) Overt and hidden complexity—two types of complexity and their implications. *Pozn Stud Contemp Linguist* 50(2):127–143
- Bisang W (2015) Problems with primary vs. secondary grammaticalization: the case of East and mainland Southeast Asian languages. *Lang Sci* 47:132–147
- Bieswanger M (2007) 2 abbrevi8 or not 2 abbrevi8: a contrastive analysis of different shortening strategies in English and German text messages. *Texas Linguistic Forum* 50. Retrieved from <http://studentorgs.utexas.edu/salsa/proceedings/2006/Bieswanger.pdf> on 5 March, 2017
- Bieswanger M (2013) Micro-linguistic structural features of computer-mediated communication. In: Herring SC, Stein D, Virtanen T (eds) *Pragmatics of computer-mediated communication*. Mouton de Gruyter, Berlin/Boston, p 463–485
- Boudjemaa H (2022) Deconstructing the linguistic features of hybrid text-based online social network communication among Algerian Facebook users: a case study. *Top Linguist* 23(1):62–71
- Burke M, Kraut RE (2016) The relationship between Facebook use and well-being depends on communication type and tie strength. *J Comput-Mediat Commun* 21(4):265–281
- Carston R (2005) Relevance theory, Grice and the neo-Griceans: a response to Laurence Horn’s current issues in neo-Gricean pragmatics. *Intercult Pragmat* 2(3):303–319
- Chao J (2012) *Wangluo Yuyan Chuanbo Daolun* [A multi-dimensional study on the netspeak communication]. Tsinghua University Press, Beijing
- Crystal D (2001) *Language and the Internet*. Cambridge University Press, Cambridge
- Crystal D (2008) *Txting: the Gr8 Db8*. Oxford University Press, Oxford
- Crystal D (2011) *Internet linguistics*. Routledge, London
- Danet B, Herring SC (2007) Introduction. In: Danet B, Herring SC (eds) *The multilingual Internet: language, culture, and communication online*. Oxford University Press, New York, p 1–39
- Don J (2014) *Morphological theory and the morphology of English*. Edinburgh University Press, Edinburgh
- Gao LW (2007) *Chinese Internet language: a study of identity constructions*. Lincom Europa, Munich
- Grice P (1975) Logic and conversation. In: Cole P, Morgan LJ (eds) *Syntax and semantics*. Academic Press, New York, p 41–58
- Herring SC (1996) Two variants of an electronic message schema. In: Herring SC (ed) *Computer-mediated communication: linguistic, social and cross-cultural perspectives*. John Benjamins, Amsterdam/Philadelphia, p 81–108
- Herring SC, Zelenkauskaitė A (2009) Symbolic capital in a virtual heterosexual market: abbreviation and insertion in Italian iTV SMS. *Writ Commun* 26(1):5–31
- Hilpert M, David CS, Jennifer R (2021) A multivariate approach to English clip-pings. *Glossa* 6(1):1–30
- Huang BR, Liao XD (2007) *Xiandai Hanyu* [Modern Chinese]. Higher Education Press, Beijing
- Hudson G (2000) *Essential introductory linguistics*. Blackwell, Oxford
- Jonathan RW (2015) Processes and variations in language economisation. *Ampersand* 2:72–82
- La AS (2019) *The instant generation: Chinese internet language and identity trends*. Unpublished Master thesis, University of California, Santa Barbara
- Lee C (2002) Literacy practices in computer-mediated communication in Hong Kong. *Read Matrix* 2(2):1–25
- Levinson SC (2000) Presumptive meanings: the theory of generalized conversational implicatures. The MIT Press, Cambridge
- Li W, Zhu H (2019) Transcribing: playful subversion with Chinese characters. *Int J Multiling* 16(2):145–161
- Ling R, Baron NS (2007) Text messaging and IM: linguistic comparison of American college data. *J Lang Soc Psychol* 26(3):291–298
- Liu YJ (2012) Renzhi shiyu xia wangluo xinci de chanshi [On network neologisms from the perspective of cognition]. *J Changsha Univ* 26:94–96
- Lotherington H, Xu YJ (2004) How to chat in English and Chinese: emerging digital language conventions. *ReCALL* 16(2):308–329
- Martinet A (1962) *A functional view of language*. Clarendon Press, Oxford
- Mey JL (1994) How to do good things with words: a social pragmatics for survival. *Pragmatics* 4(2):236–263
- Murray DE (2000) Protean communication: the language of computer-mediated communication. *Tesol Quart* 34(3):397–421
- Nishimura Y (2003). Linguistic innovations and interactional features of casual online communication in Japanese. *J Comput-Mediat Commun*. <https://doi.org/10.1111/j.1083-6101.2003.tb00356.x>
- Pan M (2002) *Wangluo yuyan zhi chutan yanjiu* [The exploratory on Internet language—the approach of Chinese]. Conference on Information and Communications. Shih Hsin University, Taipei, p 67–76
- Paolillo JC (1996) Language choice on soc.culture.punjab. *Electron J Commun* 6:EJ550464
- Rousseau R (2002) George Kingsley Zipf: life, idea, his law and informetrics. *Glottometrics* 3:11–18
- Sperber D, Wilson D (1995) *Relevance: communication and cognition*. Blackwell, Oxford
- Sugimoto T, Levin AJ (2000) Multiple literacies and multimedia: a comparison of Japanese and American uses of the Internet. In: Gail H, Selfe CL (eds) *Global literacies and the World-Wide Web*. Routledge, London, p 133–153
- Thurlow C, Brown A (2003) Generation txt? The sociolinguistics of young people’s text-messaging. *Discourse Anal Online*. Retrieved from <http://extra.shu.ac.uk/daol/articles/v1/n1/a3/thurlow2002003-t.html>
- Thurlow C, Lengel L, Tomic A (2004) *Computer-mediated communication: social interaction online*. Sage, London
- Xiang MY (2017) Toward a Neo-economy Principle in pragmatics. *J Pragmat* 107:31–45
- Xiao Y (2015) The rise of Chinese internet language: product and agent of globalization from below. *Glob Chin* 1(1):111–137
- Xie CQ, Yus F (2018) Introducing Internet pragmatics. *Internet Pragmat* 1(1):1–12
- Xie CQ, Yus F (2021) Digitally-mediated communication. In: Haugh M, Kádár DZ, Terkourafi M (eds) *The Cambridge handbook of sociopragmatics*. Cambridge University Press, Cambridge, p 454–474
- Xie CQ, Yus F, Haberland H (2021) Approaches to Internet pragmatics: theory and practice. John Benjamins, Amsterdam
- Xun ED, Rao GQ, Xiao XY, Zang JJ (2016) Da shuju beijing xia BCC yuliao ku de yanzhi [The construction of the BCC corpus in the age of big data]. *Corpus Linguist* 3:93–109
- Werry CC (1996) Linguistic and international features of Internet relay chat. In: Herring SC (ed) *Computer-mediated communication: linguistic, social and cross-cultural perspectives*. John Benjamins, Amsterdam/Philadelphia, p 47–64



- Yus F (2011) *Cyberpragmatics: Internet-mediated communication in context*. John Benjamins, Amsterdam
- Yus F (2022) *Smartphone communication: interactions in the App ecosystem*. Routledge, Abingdon
- Zipf GK (1935) *The Psycho-biology of language: an introduction to dynamic philology*. MIT Press, Cambridge
- Zipf GK (1949) *Human behavior and the Principle of Least Effort*. Addison-Wesley Press, Cambridge
- Zsido AN, Arato N, Lang A, Labadi B, Stecina D, Bandi SA (2020) The connection and background mechanisms of social fears and problematic social networking site use: a structural equation modeling analysis. *Psychiat Res* 292:113323

## Acknowledgements

This work is supported by the funding from Social Science Foundation of Jiangsu Province of China (21YYB015).

## Author contributions

Material preparation, data collection, and data analysis were performed by YZ. The first draft of this manuscript was written by YZ and YW proofread and commented on previous versions of this manuscript. All authors read and approved the final manuscript.

## Competing interests

The authors declare no competing interests.

## Ethical approval

This article does not contain any studies with human participants performed by any of the authors.

## Informed consent

This article does not contain any studies with human participants performed by any of the authors.

## Additional information

**Correspondence** and requests for materials should be addressed to Yong Zhou.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024