

When number is not neutral: Translating Chinese bare nouns into English

Ludovica Lena - Xiamen University

ludolena@xmu.edu.cn

Yue Liu* - Xiamen University

liuyue@xmu.edu.cn

Rebut /Received: 11-01-24

Acceptat / Accepted: 09-04-24

Resum. *Quan el número no és neutre: Traduir noms escuets xinesos a l'anglès.* Aquest estudi explora els noms escuets xinesos, centrant-se específicament en la seva característica de neutralitat numèrica. Com a cas d'estudi, investiguem el nom escuet més utilitzat per expressar la referència humana general, *rén* 'persona, persones'. Basant-nos en una anàlisi d'un corpus paral·lel xinès-anglès, la recerca examina les propietats a nivell de frase i a nivell de discurs que influeixen en la interpretació de significats singulars vs. plurals inclosos en el nom escuet font quan es tradueix a un idioma com l'anglès, que marca gramaticalment el número. L'anàlisi revela la relació multifactorial entre la marca de número en anglès i l'estructura semàntica (en termes de partició i genericitat), així com les propietats pragmàtiques (relacionades amb l'estat del discurs) del nom escuet font. Es demostra que certs contextos exclouen inequívocament una interpretació singular o plural, especialment en afirmacions genèric-partitives i en oracions que introdueixen línies de discurs directe.

Paraules clau: noms escuets, xinès, número, traducció.

Abstract. *When number is not neutral: Translating Chinese bare nouns into English.* This study explores Chinese bare nouns, with a specific focus on their feature of number neutrality. As a case study, we investigate the bare noun most frequently used to convey general human reference, i.e., *rén* 'person, people'. Building upon a parallel corpus composed of Mo Yan's novels and their aligned English translations, the research investigates the sentence-level and discourse-level properties influencing the interpretation of singular vs. plural meanings included within the number-neutral source bare noun when translated into a target language such as English, which

grammatically marks number. The analysis uncovers the multifactorial relationship between number marking in English and the semantic structure (in terms of partitivity and genericity) as well as the pragmatic properties (related to the discourse status) of the source bare noun. It is shown that certain contexts unambiguously preclude a singular or plural interpretation, particularly in generic-partitive statements and sentences introducing direct speech lines.

Keywords: bare noun, Chinese, number, translation.

1. Introduction

This corpus-based investigation is concerned with Standard Chinese (hereafter Chinese) bare nouns, that is, nouns that appear without any prenominal determiner. Bare nouns in Chinese, as in other languages, are highly multifunctional, since they are undetermined with respect not only to number, but also to definiteness, specificity and genericity (Cheng & Sybesma, 1999; Hole, 2012; LaPolla, 1995; Lena, 2023b; Li, 1997; Rullmann & You, 2006). Specifically, as pointed out by Rullmann & You (2006), Chinese bare nouns convey a general number, referring to both singular and plural referents.

This article builds upon a parallel corpus of Chinese-to-English aligned translations to observe how Chinese bare nouns are translated into a target language that grammatically marks number such as English, and specifically in which contexts the singular vs. plural meaning included in the source bare noun is activated. The bare noun *ren* ‘person, people’ is taken as our case study.

Our working hypothesis is that number marking in the target language is influenced by the semantic-pragmatic properties of the source bare noun, of which it is an indirect representation. Since prior research has pointed out that sentence position is relevant for the interpretation of bare nouns (see Sect. 2), the bare noun *ren* is examined in different syntactic positions, in the light of its semantic and pragmatic properties. It will be shown that, although one-to-one correspondences between a given property and the aligned form in English cannot be expected in most cases, tendencies emerge when the semantic properties of genericity (Kind vs. Entity), partitivity (Whole vs. Part), and the discourse status (Discourse-new vs. Discourse-old) of the source bare noun are considered.

The rest of the article is structured as follows. Sect. 2 illustrates how (in)definiteness oppositions are realized in an article-less language such as Chinese, with a focus on the properties of bare nouns. After presenting the parallel corpus and the methodology adopted in this study (Sect. 3), the article turns to the presentation of the data (Sect. 4), discussing the distribution and types of bare *ren* in the source subcorpus (4.1), the overall number marking in the aligned subcorpus (4.2), and finally linking number marking to the semantic-pragmatic properties of the source bare noun (4.3). A discussion is

offered in Section 5. Finally, Section 6 concludes the article and sums up its main contributions.

2. Background: Bare nouns in Chinese

As an article-less language, Chinese lacks the grammatical category of (in)definiteness, and relies on different strategies to mark (in)definiteness oppositions (Chen, 2004). The following examples are extracted from the same novels used to build our corpus (described in Section 3) to briefly illustrate the distinction between bare and non-bare nouns in Chinese. As this contribution is concerned with the bare noun *ren* ‘person, people’, the sentences below will focus on noun phrases (NPs) with *ren* as the nominal head.

The [numeral + classifier] sequence has both referential and non-referential uses (Chen, 2004; LaPolla, 1995). Accordingly, the NP *yī-ge ren* [one-CL person] ‘a person’ denotes a specific individual in (1), while in (2) each occurrence points to a generic individual (through this contribution, the relevant referring expression – if available – in the original English sentence is marked in italics):

- (1) Ěr=biān túrán xiǎng-qǐ-le **yī-ge rén** huānkǔ
 ear=side suddenly sound-INC-PFV¹ one-CL person cheerful
 de hǎnjiào shēng: (...).
 SUB shout sound
 Suddenly my ears filled with the happy shouts of *a man* somewhere: (...).
 (Life and Death are Wearing Me Out)
- (2) **Yī-ge rén** hèn lìng **yī-ge rén** jìng néng
 one-CL person hate another one-CL person actually can
 dádào rúcǐ qiángliè de chéngdù, zhè wúyí
 reach such intense SUB degree this undoubtedly
 shì yī zhǒng měi, yī zhǒng duìyú quán
 be one kind beauty one kind to all
 rénlèi de wěidà gòngxiàn.
 humanity SUB great contribution
 If *one person's* hatred for *another* could reach such proportions, it was an
 unquestioned form of beauty, a magnificent contribution to humanity.
 (The Republic of Wine)

Nouns modified by demonstrative determiners, i.e. the proximal demonstrative *zhe* ‘this’ and the distal demonstrative *na* ‘that’, are used to encode identifiable referents,

discourse-old or inferable, (LaPolla, 1995)¹. Example (3) features two ‘definite’ NPs with *ren* as their nominal head, marked with the proximal (*na-ge ren* ‘that person’) and distal (*zhe-ge ren* ‘this person’) demonstrative determiner, respectively.

- (3) Fùqīn qiánbian **nà-ge rén** kēngkēng-de kēsou-qilai,
 father in.front that-CL person ONOM-ADV cough-INC
zhè-ge rén de kēsou shēng fēicháng shúxī.
 this-CL person SUB cough sound very be.familiar.with
The man in front of Father coughed loudly. It was a familiar cough (Lit.
 ‘The/this person’s cough was familiar’), (...). (Red Sorghum Clan)

Besides nouns that are marked either by a quantifying expression or by a demonstrative determiner, Chinese speakers can make use of bare nouns as referring expressions, as mentioned above. Bare nouns are an interesting site of study because of their formally unmarked status and their highly multifunctional nature. Bare nouns in Chinese are not only number-neutral, they can also denote discourse-old, discourse-new and generic referents. They resemble English bare plurals in that they can have a generic or specific interpretation, but in addition, they can also convey definite reference (Bremmers et al., 2021; Cheng & Sybesma, 1999; Dayal & Jiang, 2022; Jenks 2018; Li, 1997).

From the point of view of propositional semantics, it has been acknowledged in the literature that Chinese bare nouns have both Kind-level and Entity-level interpretations (Li, 2013; Li & Thompson, 1981; Rullman & You, 2006; Yang, 1998). Previous studies also pointed out a semantic correlation between the interpretation of bare nouns and the type of predicates (Li, 1997; Kuo, 2008). That is, as subjects of individual-level predicates (i.e. predicates denoting characterizing properties), bare nouns have a generic interpretation, while as subjects of stage-level predicates (i.e. predicates denoting transient properties), they are interpreted as Entity-referring (Carlson, 1977; Kratzer, 1988; Krifka et al., 1995).

In turn, the Entity-level meaning can be both definite and indefinite, and pragmatics play a role in the interpretation (Kuo, 2008). From the point of view of the discourse structure, bare nouns as referring expressions present a high degree of ambiguity, given that they can designate both unidentifiable and identifiable referents (Sun, 1988; LaPolla, 1995)². Generally speaking, bare nouns in preverbal position tend to receive

1. The glosses used through the article are as follows. ADV: adverbial particle, CL: classifier, COLL: collective marker; CRS: Current Relevant State aspect marker, DUR: durative aspect marker, INC: inchoative aspect marker; NEG: negation, ONOM: onomatopoeia; PASS: passive marker, PFV: perfective aspect marker, PROG: progressive aspect marker, SFP: sentence final particle, SG: singular, SUB: subordinative particle, VEN: venitive marker.

2. On accessibility and identifiability (Prince 1981; Ariel 1990; Gundel et al., 1993; Chafe 1994; Lambrecht 1994).

a definite (i.e., generic or discourse-old) interpretation, while they are interpreted as indefinite in postverbal position (Cheng & Sybesma, 1999; Hole 2012; Lena, 2020a).

On these theoretical grounds, the current work argues that number marking can be understood in the light of the following interacting properties of the source bare noun: GENERICITY (Kind vs. Entity), PARTITIVITY (Whole vs. Part), and DISCOURSE STATUS (Discourse-old vs. Discourse-new). Let us begin by presenting the corpus and methodology adopted, before turning to the discussion of the data.

3. Corpus and Method

The parallel corpus compiled for this study consists of six Chinese novels written by the Nobel literature prize laureate Mo Yan (i.e. *The Garlic Ballads*, *Sandalwood Death*, *The Republic of Wine*, *Big Breasts & Wide Hips*, *Life and Death are Wearing Me Out*, *The Red Sorghum Clan*) and their aligned translations into English (all translated by Howard Goldblatt). The corpus was constituted via the *English-Chinese Parallel Concordancer* made available by the Department of Linguistics and Modern Language Studies at the Education University of Hong Kong (<https://corpus.eduhk.hk/paraconc>).

TABLE 1. THE CORPUS

Novels	English title	Tokens (words)	Sentences
<i>Tiāntáng suàntái zhī gē</i>	The Garlic Ballads	126.075 (101.939)	6.571
<i>Tǎnxiāng xíng</i>	Sandalwood Death	195.849 (163.225)	8.954
<i>Jiǔ guó</i>	The Republic of Wine	137.288 (113.921)	6.491
<i>Fēng rǔ fēi tún</i>	Big Breasts and Wide Hips	330.308 (271.730)	17.542
<i>Shēngsǐ píáo</i>	Life and Death are Wearing Me Out	276.532 (229.790)	11.430
<i>Hóng gāoliang jiāzú</i>	The Red Sorghum Clan	161.140 (134.257)	7.730
Total		1.227.192 (1.014.862)	74.243

The data extraction was done in two steps. Firstly, a search of the character 人 [rén] was performed within the corpus, which gave 7895 raw results. A randomized selection of 200 hits was then coded by the author of this work with the objective of identifying instances of bare *ren*, as distinguished by both *ren* preceded by a prenominal modifier and instances of lexicalized items including *ren* as a bounded morpheme (table 2).

This procedure allowed us to identify the contexts of occurrence of *rén*, i.e. as a subject (*ren_S*), an object (*ren_O*), as a pivot (*ren_P*), as the agent in passive constructions

(*ren_A*). “Pivot” is a term used in the literature on existential constructions (Bentley et al., 2015; McNally, 2011; Sarda & Lena, 2023). In this study it denotes the bare noun *ren* appearing in existential-presentational constructions, namely *you ren* [exist person] ‘there’s someone, there are people’ (Lena, 2023b). The label “agent” (*ren_A*) is used for commodity and denotes a specific syntactic position whereby *ren* is introduced by the passive marker *bei*.

Next, a random sample of each *ren*-type (*ren_S*, *ren_O*, *ren_P*, *ren_A*, *ren_CAUS*) was extracted from the corpus (n = 50; N = 250; cf. table 3).

The English aligned sentences were then coded according to the type of referring expression aligned with *ren*. Assuming that “a high quality translation creates a grammatically correct mirror image of the original meaning in the target language” (Tellings et al., 2022), the parallel corpus methodology aims at identifying some of the interacting factors that contribute to activating, in the English translation, the singular or plural meaning included in the general number of the source bare noun *ren* ‘person, people’.

As a study based on a limited corpus featuring novels from a single author, this work serves as exploratory research that lays the groundwork for future studies utilizing larger and more diverse datasets.

4. Results

4.1. The source subcorpus

Table 2 below shows the general distribution of the morpheme [ren] in the randomized sample of 200 occurrences extracted from the corpus (the extraction method was presented in Sect. 3). The occurrence of bare *ren* is rather infrequent (14 %) with respect to its appearance with prenominal modifiers (41.5 %) and as a lexicalized item (36 %).

TABLE 2. EXTRACTION OF THE MORPHEME [REN] FROM THE CORPUS

Category	Forms	Examples	Count	%
Pre-nominal modifiers + N	RC + <i>ren</i>	<i>huo-zhe de ren</i> [live-DUR SUB person]	27	13.5 %
	ADJ + <i>ren</i>	<i>hao ren</i> [good person]	24	12 %
	QUANT + <i>ren</i>	<i>yi ren</i> [one person]	21	10.5 %
	DEM + <i>ren</i>	<i>na ren</i> [that person]	11	5.5 %
	Subtotal		83	41.5 %

Bare noun	<i>ren_S</i>	<i>ren</i> V [person V]	2	1 %
	<i>ren_O</i>	V <i>ren</i> [V person]	7	3.5 %
	<i>ren_P</i>	<i>you ren</i> [exist person]	10	5 %
	<i>ren_A</i>	<i>bei ren</i> V [PASS person V]	3	1.5 %
	<i>ren_CAUS</i>	<i>rang ren</i> V [make person V]	5	2.5 %
	<i>ren_DITR</i>	<i>gei ren</i> O [give person O]	1	0.5 %
	Subtotal		28	14 %
<i>Ren-men</i>	<i>ren-men_S</i>	<i>ren-men</i> V [person-COLL V]	6	3 %
Lexicalized	<i>x-ren</i>	<i>fanren</i> ‘criminal’	44	22 %
	<i>ren-x</i>	<i>renjia</i> ‘the others’	28	14 %
	Subtotal		72	36 %
Other		<i>ge ren</i> ‘personally’	11	5.5 %
Total			200	100%

Figure 1 below illustrates the distribution of *ren* as a bare noun with respect to the other macro-categories.

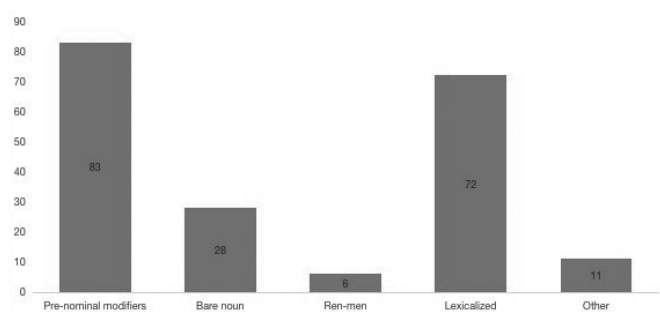


FIGURE 1. RELATIVE FREQUENCY OF *REN* AS A BARE NOUN

This exploratory search further allowed us to identify the contexts of occurrence of *ren* as a bare noun. As previously discussed, *ren* can appear as a subject (*ren_S*) – although infrequently, as an object (*ren_O*), including the separately-calculated object position after a causative verb (*ren_CAUS*) and the pivot position in existentials (*ren_P*). It can also

be an agent introduced by the passive marker (*ren_A*) or, marginally, it can be a recipient in ditransitive constructions (*ren_DITR*).

On these grounds, we then extracted 50 random examples for each type of bare *ren* based on its position and function in the sentence, with the exception of *ren* appearing in ditransitive constructions (*ren_DITR*), due to an insufficient number of relevant occurrences, as seen in table 3. The total 250 Chinese sentences were then analysed by considering the referring expressions aligned with *ren* in the corresponding English sentences, as discussed in the next subsection.

TABLE 3. THE BARE-REN CORPUS

Types	Count	%
<i>Ren_S</i>	50	20 %
<i>Ren_O</i>	50	20 %
<i>Ren_A</i>	50	20 %
<i>Ren_P</i>	50	20 %
<i>Ren_CAUS</i>	50	20 %
Subotal	250	100 %

4.2. Number marking in the aligned subcorpus

Let us now turn to the discussion of the aligned data. To begin with, table 4 below illustrates the distribution of the number marking in the English aligned referring expressions according to the types of *ren*. Importantly, the categories “singular” and “plural” co-exist with the category “irrelevant”. This encompasses all the cases in which number is not expressed because a relevant referring expression is not provided in the English sentence.

TABLE 4. NUMBER MARKING IN THE ALIGNED SENTENCES

	<i>Ren_S</i>	<i>Ren_O</i>	<i>Ren_A</i>	<i>Ren_P</i>	<i>Ren_CAUS</i>	Total
Plural	41	13	11	15	13	93 (37.2 %)
Singular	9	15	6	26	16	72 (28.8 %)
Irrelevant	0	22	33	9	21	85 (34 %)
Subotal	50	50	50	50	50	250 (100 %)

When all sentences are considered together, the mean values do not show a preference toward one particular category (the plural category being only slightly more represented

at 37.2 %). However, when each type of *ren* is considered separately, interesting tendencies emerge. For the sake of clarity, the data are represented in figure 2 below. First, the relevant human referent is always expressed in the aligned translation when the source bare noun *ren* has the function of a subject (*ren_S*). Secondly, the aligned referring expression strongly tend to have plural number. By contrast, when *ren* is an existential pivot (*ren_P*), the aligned referring expression has more often singular number. The categories *ren_O* and *ren_CAUS* show a similar distributive behaviour, with similar proportions of singular and plural marking, and a relatively more frequent portion of sentences in which number is unmarked. Finally, when *ren* is the agent in passive constructions (*ren_A*), it strongly tends to remain unexpressed in the aligned sentence.

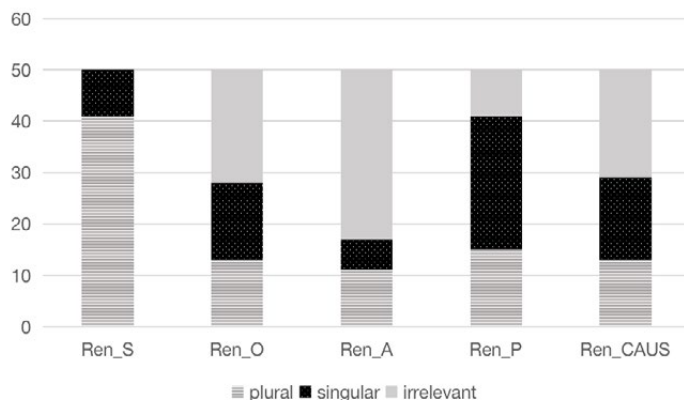


FIGURE 2. NUMBER MARKING IN THE ALIGNED SENTENCES

4.3. Linking number marking to the semantic-pragmatic properties of the source bare noun

4.3.1. General observations

Let us begin the discussion by considering the types and frequency of the aligned referring expressions in the English sentences, as seen in Table 4. A great deal of variation can be observed, in which potentially Kind-referring expressions such as *people* coexist with potentially Entity-referring expressions, in turn likely to be Discourse-new (e.g. *someone*, $a + N_{sg}$) or Discourse-old (e.g. personal pronouns). Interestingly, lexical categories such as *humans* or *every human* (and also the bare singular *man*) unambiguously point to the class of human beings in their entirety.

TABLE 4. FORM OF THE ALIGNED REFERRING EXPRESSIONS

Aligned referring expressions (REs)	Total
people	36
personal pronoun	27
<i>someone (somebody)</i>	24
humans / every human	12
<i>a</i> + N _{SG}	12
anyone (anybody, whoever)	8
everyone (everybody)	7
<i>some</i> (+ N _{PL})	7
the people	6
<i>no one / nobody / no</i> N _{SG}	4
others	4
<i>one of</i> + N _{PL}	3
<i>the</i> + N _{PL}	3
<i>another</i> (+N _{SG})	2
man	2
<i>the</i> + N _{SG}	2
Other REs (n=1)	8
Unexpressed	83
Total	250

In what follows, each type of *ren* is analysed individually, since interpretative differences – likely to influence number marking in the aligned sentences – are expected between the different positions. In particular, recall from Sect. 2 that previous studies pointed out a correlation with (in)definiteness: bare nouns tend to receive a definite reading in preverbal position and are interpreted as indefinite in postverbal position.

4.3.2. The subject position

Preverbal *ren* (*ren*_S) is often found in generic statements where it refers to the kind of “human beings”. In (4), the English sentence features a noun modified by the indefinite article (i.e. *a man*) – although a bare plural (e.g. *men* or *people*) would be a perfectly acceptable option (and often is, cf. 4.2). By contrast, in (5) *ren* is Entity-referring, as it points to some particular individual. Additionally, this entity is evoked in discourse, hence the alignment with the English pronoun *him* is made possible, signaling an

element in high activation state (Gundel et al., 1993). In other words, its referent is Entity-referring and Discourse-old, identifiable by the interlocutor.

- (4) Kind-referring (Whole).

Rén guò liú míng, yàn guò liú shēng.
 person pass leave name wild.goose pass leave sound
A man leaves behind his good name, a wild goose leaves behind its call.
 (Big Breasts and Wide Hips)

- (5) Entity-referring (Discourse-old)

Rén fǎnzhèng sǐ-le, yīqiè dōu yào kǎolǜ
 person anyway die-PFV everything all must consider
 huó-zhe de rén.
 live-DUR SUB person
 Shuōchuān le, jiù shì qián!
 disclose CRS then be money
 It's too late to do anything for *him* now (Lit. 'The person has died anyway'),
 so it's time to think about yourselves – in other words, money.
 (The Garlic Ballads)

In its Entity interpretation, number marking in the aligned sentence is unpredictable due to the context-dependent retrieval. However, *ren_S* is prevalently Kind-referring in the corpus, and that motivates its frequent alignment with plural referring expressions in English (cf. Sect. 4.2). In both cases, *ren_S* is treated as a topical element and cannot remain unexpressed in the English translation.

4.3.3. The object position

The occurrence of *ren_O* was examined after the verbs *da* 'hit', *ma* 'insult', *jian* 'see', *sha* 'kill', i.e., the most frequent verbs observed in the exploratory corpus search (Sect. 4.1). In this context, *ren_O* is often Entity-referring and nonspecific. In such non-specific uses, it is either unexpressed (e.g. (6)), either aligned with *someone* or *anyone* (e.g. (7)).

- (6) Entity-referring, nonspecific.

“Bù xǔ mà rén!”
 NEG allow insult person
 “No cursing allowed,” (...). (The Garlic Ballads)

- (7) Entity-referring, nonspecific.

Wǒ shuō Zhào Jiǎ, shā rén shí nǐ shì

1SG say Zhao Jia kill person time 2SG be
 pà háishì bú pà?
 be.afraid or NEG be.afraid
 I say, Zhao Jia, when you kill *someone*, are you afraid? (Sandalwood Death).

However, if *ren_O* appears in contrastive constructions where human beings are contrasted with other species, the Kind-interpretation can be activated:

(8) Kind-referring (Whole).

Jiàn rén dǎ rén, jiàn gǒu dǎ gǒu.
 see person hit person see dog hit dog
Every human who got in their way tasted the whips, as did the dogs. (Sandalwood Death)

(9) Kind-referring (Whole).

Wǒ zhīdào tāmen huì kāi qiāng shā rén, dàn
 1SG know 3PL can open fire kill person but
 tāmen bú huì kāi qiāng shā lǘ.
 3PL NEG can open fire kill donkey
 I knew they wouldn't hesitate to shoot *a person*, but would never shoot a donkey. (Life and Death Are Wearing Me Out).

Despite the object position, *ren_O* can then receive a Kind (definite) interpretation, although less frequently than *ren_S* in our corpus.

As the object of the causative verb *ràng* 'let, make', *ren_CAUS* does not behave much differently than *ren_O*. The only difference being that when *ren_CAUS* is Entity-referring and specific, it can denote discourse-old entities:

(10) Entity-referring (Discourse-old)

...dāihuì ràng rén sòng-lai.
 in.a.while make person deliver-VEN
 I'll have *one of my men* bring it over. (Big Breasts & Wide Hips).

4.3.4. The agent position

As previously mentioned, *ren_A* stands out insofar as it is most often unexpressed in the aligned sentence (Sect 4.2), as seen in (11). In a few cases, however, the Kind interpretation can be activated by the context, as in (12), and the aligned English sentence must express the relevant human referent.

(11) Entity-referring, non-specific.

“Shì-bù-shì bèi rén yān le?”

be-NEG-be PASS person castrate PFV
 Not castrated \emptyset , are you? (The Republic of Wine).

- (12) Entity-referring, non-specific.

Xiànzài tā shì yī-piàn fèixū le, rén chuàngzào
 now 3SG be one-CL ruins CRS person build
 de, yòu bèi rén cuīhuǐ.
 SUB again PASS person destroy
 Now the village lay in ruins; man had created it, and *man* had destroyed it.
 (Red Sorghum Clan).

4.3.5. The pivot position

Finally, when appearing after the existential verb *you* ‘there be, exist’, *ren*_P can be Entity-referring (as in (13)) or Kind-referring (as in (14))³:

- (13) Entity-referring (Discourse-new).

“Nà shì shéi?” Hēi‘àn=li yǒu rén hǎn.
 that be who darkness=in exist person scream
 “Who’s there?” came a voice out of the darkness. (Lit. ‘...someone shouted
 in the darkness’). (The Garlic Ballads).

- (14) Kind-referring (Part).

...yǒu rén shuō nà xiē huāwén shì tiānshàng
 exist person say that few pattern be heaven
 de kēdǒu wén.
 SUB tadpole script

A design had been scorched into her back, which *some people* said was the script of heavenly tadpoles. (Red Sorghum Clan)

Hence, *ren*_P can denote both Entity-referring and Kind-referring referents, with however some peculiarities. Kind-referring *ren*_P belong to the generic-partitive subtype, that is, in the overwhelming majority of cases, they refer to a part of the generic set of “people”. This contrasts with sentences previously seen in which Kind-referring *ren* (often *ren*_S, but also possible in the other positions except for *ren*_P) refer to the overall generic set of “people”. The data show that when the source bare noun is positive with respect to the two parameters of partitivity and genericity, the plural meaning of the aligned referring expression is systematically activated in the English sentence (e.g. in (14)).

3. See Lena (2024a, 2024b) for a detailed analysis of the bare noun *ren* in the pivotal position.

One aspect complicates the picture emerged so far. The partitive interpretation is not limited to generic sentences. Entity-referring *ren_P* can also denote discourse entities that stand in Set-Member relation with a Discourse-old referential group, as in (15):

(15) Entity-referring (Discourse-anchored).

Zhàlán=li	chòuqìxūntiān,	yǒu	rén	zài	dǎhūlu,
fence=in	stink.to.high.heaven	exist	person	PROG	snore
yǒu	rén	wǎng	zhàlán=biān	jiǎo=shang	
exist	person	towards	fence=side	corner=on	
nà-ge	tiěpí	shuītǒng=li	sāniào, (...).		
that-CL iron	pail=in	urinate			

Some of the men snored loudly; *others* got up to piss in a tin pail, (...).
(Red Sorghum Clan).

Both (13) and (15) include entity-referring participants, which are Discourse-new in the former case, and Discourse-anchored in the latter (Prince, 1981). This also means that, although *ren_P* can pick up a member of a discourse-old group, it cannot directly denote a discourse-old referent, hence the “definite” (discourse-old) interpretation is not available. As in the case of discourse-old *ren_S*, nonetheless, the interpretation of discourse-anchored entities is also context-dependent, and the choice of the aligned referring expression in English, including the singular vs. plural marking, must take the previous context into account.

In the case of Discourse-new referents, number marking is dependent on the propositional semantics at the sentence level. An illustrative example is provided in (16), where the sentence-initial locative expression *daolu liang bian* ‘at the two sides of the road’ implies that at least one person is located on each side of the road. The English translation consequently uses the plural referring expression *crowds of gawkers*:

(16) Entity referring (Discourse-new)

Dàolù	liǎng	biān	shǐzhōng	yǒu	rén	guānkàn (...).
road	two	side	all.along	exist	person	watch

...but always there were *crowds of gawkers*, (...). (The Garlic Ballads)
Lit. ‘There were people watching on both sides of the road.’

In this respect, one context systematically activates the singular meaning in the aligned English sentence, that is, when the sentence introduces a direct speech line, as in (17–18):

(17) Entity referring (Discourse-new)

Xià	yǔ	le,	yǒu	rén	rǎngjiào.
fall	rain	CRS	exist	person	scream

“It’s raining,” *someone* shouted, (...). (Big Breasts & Wide Hips).

(18) Entity referring (Discourse-new)

Ān tīngdào yǒu rén shuō: “Dūtǒng dàrén,
 1SG hear exist person say Commander Great
 zhè shì zhīxiàn dàrén de sīzhái!”
 this be Magistrate Great SUB private.residence
 “Commander,” *someone* said, “these are the Magistrate’s private quarters.”
 (Sandalwood Death).

Direct speech not only unambiguously activates Entity-referring reading of *ren*_P, it also systematically identifies the relevant entity as a single participant.

5. Discussion

In different syntactic positions, the bare noun *ren* can receive Kind or Entity readings. The Kind reading can be further distinguished in UNIVERSAL (when it denotes the Whole) and PARTITIVE (when it denotes the Part). Universal kind readings can be found in all syntactic positions except for *ren*_P, but are particularly represented in the subject position (*ren*_S). The relevant referents are typically encoded by plural referring expressions in the English sentence (e.g., *people*, *humans*), although singular forms with generic interpretation are not excluded (e.g., *a man*, *a person*).

In its Entity-referring interpretation, *ren* can be Discourse-new, Discourse-old (or Discourse-anchored, as far as *ren*_P is concerned). Number marking for Discourse-old *ren* is dependent on the previous context, and can result in forms typically high in activation scales (e.g., personal pronouns). In the case of Discourse-new *ren*, by contrast, the number interpretation is dependent on the propositional meaning of the sentence.

One context, in particular, systematically activates the singular meaning of the aligned referring expression, namely, when *ren*_P introduces a direct speech line. Despite the number-neutrality property of the source bare noun *ren*, the choice of an aligned referring expression with plural number in these contexts would often result in an infelicitous outcome. The opposite situation is observed with Kind-referring pivot *ren*. When *ren* is a pivot in existential constructions (*ren*_P), it can only receive a generic-partitive reading, as far as the Kind interpretation is concerned. By definition, the generic-partitive interpretation refers to more than one individual, and the corresponding aligned referring expression must be plural (*some people*, *certain individuals*). In this scenario, selecting a singular referring expression in English would similarly result in an inappropriate choice.

In most cases, however, English has to choose one option (e.g., a referring expression with singular reference) while the other option (e.g., a referring expression with plural reference) could also be accepted, as the relevant opposition are semantic-pragmatic in nature, of which the singular vs. plural morphological marking is only an opaque representation. As a language morphologically marking number, the number-neutral

property of bare nouns in Chinese is generally lost in translation. Different strategies, however, allow to keep the number indeterminacy. Recall in this respect the high portion of sentences in which *ren* was *not* expressed in the English sentence (Sect. 4.2). One such strategy is the use of an agent-less passive construction, in which the relevant entity is inferred from the meaning of the predicate, as previously seen. Another strategy is found in sentences such as (19), involving the thematization of the location, while the human referent is only inferred from the meaning conveyed by the predicate:

- (19) Wéiqiáng=shang hěn píngjìng, sìhū méi yǒu **rén**.
 enclosing.wall=on very quiet seemingly NEG exist person
 They pressed forward, drawing closer and closer to the wall, which appeared **unmanned**. (Sandalwood Death)⁴

In sentences such as (19), the essential meaning is the human presence (or absence), rather than the specific number of individuals involved. This concept is further exemplified in English sentences like (20), where the term *manned* synthetically captures this idea without overloading the sentence with momentarily irrelevant information such as the number of astronauts involved.

- (20) Elon Musk's SpaceX COO says **manned** missions will reach Mars by 2030.⁵

6. Conclusion

Building upon parallel corpus methodology, this research delved into the number-neutral nature of Chinese bare nouns, with specific attention on the translation of the bare noun *ren* 'person, people' into English.

It revealed that the distinction between singular and plural markings in the English translations indirectly reflects the semantic and pragmatic properties of the source bare noun, which in turn are dependent from its position in the sentence.

Certain contexts unequivocally rule out the singular or plural interpretation respectively. When the sentence introduces a direct speech line, the English translation invariably opts for singular referring expressions. Conversely, the Kind-referring reading of the generic-partitive subtype activates the use of plural referring expressions in English (e.g., *some people*, *certain individuals*, etc.). This stands in contrast to the generic-universal interpretation of *ren*, which can be conveyed in translation through the generic

4. The backgrounding of the human entity is often found in negative you ren constructions, as in (19) (see Lena, 2024c, for a full discussion).

5. Retrieved from: <https://www.euronews.com/next/2022/05/09/elon-musk-s-spacex-coo-says-manned-missions-will-reach-mars-by-2030-nasa-says-otherwise>.

usage of the indefinite article (e.g., *a man*), although the use of the bare plural *people* is dominant.

Finally, it has been pointed out that the referent of *ren* is often left unexpressed in the aligned subcorpus, resulting in several strategies allowing the number-neutral properties of the source bare noun to be kept.

In conclusion, this work has served as a preliminary investigation on the translation strategies of Chinese bare nouns. However, it is important to highlight that this case study – based on works by one author and translated by a single translator – requires corroboration from larger and more diversified datasets to substantiate the hypotheses it presents.

References

- Ariel, Mira (1990). *Accessing Noun-Phrase Antecedents*. London & New York: Routledge.
- Bremmers, David, Liu, Jianan, van der Klis, Martijn & Le Bruyn, Bert (2021). Translation mining: Definiteness across languages (A reply to Jenks 2018). *Linguistic Inquiry*, 1-18.
- Bentley, Delia, Ciconte, Francesco Maria & Cruschina, Silvio (2015). *Existentials and Locatives in Romance Dialects of Italy*. Oxford University Press.
- Carlson, Gregory N. (1977). *Reference to kinds in English*. PhD dissertation, University of Massachusetts at Amherst.
- Chafe, Wallace L. (1994). *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. The University of Chicago Press.
- Chen, Ping (2004). Identifiability and definiteness in Chinese. *Linguistics* 42(6), 1129-1184.
- Cheng, Lisa Lai-Shen & Sybesma, Rint (1999). Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry*, 30(4), 509-542.
- Dayal, Veneeta & Jiang, Li Julie. (2022). The puzzle of anaphoric bare nouns in Mandarin: A counterpoint to *Index!*. *Linguistic inquiry*, 54(1), 147-167.
- Gundel, Jeanette K., Hedberg, Nancy & Zacharski, Ron (1993). Cognitive status and the form of referring expressions in discourse. *Language* 69 (2), 274-307.
- Hole, Daniel (2012). The information structure of Chinese. In Manfred Krifka & Renate Musan (Eds.), *The expression of information structure*, 45-70. De Gruyter.
- Jenks, Peter (2018). Articulated definiteness without articles. *Linguistic Inquiry* 49(3), 501-536.
- Kratzer, Angelika (1988). Stage-level and individual-level predicates. In Manfred Krifka (ed.), *Genericity in natural language*, 247-84. University of Tubingen.
- Krifka, Manfred, Pelletier, Francis J., Carlson Gregory N., ter Meulen, Alice, Link, Godehard & Chierchia, Gennaro. (1995). Genericity: An introduction. In Gregory N. Carlson & Francis J. Pelletier (eds.), *The generic book*, 1-124. The University of Chicago Press.

- Kuo, Jenny Yi-chun (2008). A pragmatic approach to the interpretations of Mandarin bare nouns. *Journal of Pragmatics* 40(6). 1082-1102.
- Lambrecht, Knud (1994). *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge University Press.
- LaPolla, Randy J. (1995). Pragmatic relations and word order in Chinese. In Pamela A. Dowling & Michael Noonan (eds.), *Word order in discourse* (pp. 297-329). John Benjamins.
- Lena, Ludovica (2017). Les énoncés « présentatifs » chez les apprenants sinophones de français L2. *SHS Web of Conferences*, 38. EDP Sciences.
- Lena, Ludovica (2020a). Presentational sentences: A study on the information structure of Path verbs in spoken discourse. In Dongyan Chen & Daniel Bell (eds.), *Explorations of Chinese theoretical and applied linguistics*, 45-81. Cambridge Scholars.
- Lena, Ludovica (2020b). Referent introducing strategies in advanced L2 usage: A bi-directional study on French learners of Chinese and Chinese learners of French. In Jonathon Ryan & Peter Crosthwaite (eds.), *Referring in a second language: Studies on reference to person in a multilingual world*, 164-183. Routledge.
- Lena, Ludovica (2022). An unbridgeable gap? The treatment of definiteness restrictions in French and Chinese presentational constructions. *Discours*, 31.
- Lena, Ludovica (2023a). The (non-)acquisition of the Chinese Definiteness Effect: A usage-based account. In Howard Hao-Jan Chen, Keiko Mochizuki & Hongyin Tao (eds), *Learner corpora: Construction and explorations in Chinese and related languages*. Springer.
- Lena, Ludovica (2023b). Partition and existence: The case of *you ren* 'there's someone, there are people' in Chinese. In L. Sarda and L. Lena (eds.), *Existential constructions across languages: Forms, meanings and functions*, 245-282. John Benjamins.
- Lena, Ludovica (2024a). Encoding indefinite human reference without indefinite pronouns: the case of Chinese presentationals. *Folia Linguistica*, 58(3), 659-698.
- Lena, Ludovica (2024b). Looking for someone: The encoding of indefinite human reference in Chinese/English aligned translation. *Lingua*, 308, 1-24.
- Lena, Ludovica (2024c). Negative existential and presentationals in Chinese: A parallel corpus investigation. *Languages in contrast* 24(2), 226-247.
- Li, Charles. N. & Thompson, Sandra. A. (1981). *Mandarin Chinese: A functional reference grammar*. University of California Press.
- Li, Jie (1997). Predicate type and the semantics of bare nouns in Chinese. In Liejiong Xu (ed.), *The referential properties of Chinese noun phrases*, 99-115. Ecole des hautes études en Sciences sociales, Centre de recherches linguistiques sur l'Asie orientale.
- Li, Xuping (2013). *Numeral classifiers in Chinese: The syntax-semantics interface*. De Gruyter.
- McNally, Louise (2011). Existential Sentences. In Klaus von Heusinger, Claudia Maienborn & Paul Portner (eds.), *Semantics: An international handbook of natural language meaning*, 1829-48. De Gruyter.

- Prince, Ellen F. (1981). Toward a Taxonomy of Given-New Information. In Peter Cole (ed.), *Radical Pragmatics*, 223-256. Academic Press.
- Rullmann, Hotze & You, Aili (2006). General number and the semantics and pragmatics of indefinite bare nouns in Mandarin Chinese. In Klaus von Heusinger & Ken Turner (eds.), *Where Semantics Meets Pragmatics*, 175-196. Brill.
- Sarda, Laure & Lena, Ludovica (2023). Existential constructions: In search of a definition. In Laure Sarda & Ludovica Lena (Eds), *Existential constructions across languages: Forms, meanings and functions*, 1-32. John Benjamins.
- Sun, Chaofen (1988). The discourse function of numeral classifiers in Mandarin Chinese. *Journal of Chinese Linguistics* 16(2), 298-323.
- Tellings, Jos, Fuchs, Martín, van der Klis, Martijn, Le Bruyn, Bert & de Swart, Henriëtte. (2022). Perfect variations in dialogue: A parallel corpus approach. *Proceedings of SALT* 32.
- Yang, Ronglan (1998). Chinese bare nouns as kind-denoting terms. *RuLing Papers*, 1, 247-288.