A Corpus-Based Examination of the Translation of the Suffix –ism into Chinese

Lily Lim

Abstract The suffix –ism is one of the most productive suffixes in English, which presents complexities and challenges in English-to-Chinese translation. This study examines corpus tools and resources that translators can have access to, aiming at revealing the potential values of corpus resources that translators can draw on to better deal with the suffix in their practice. Our investigation is focused on interrogating a parallel corpus named "Education", which contains 450,000 pairs of English and Chinese sentences. Compared with traditional tools such as monolingual and English-Chinese (reverse) dictionaries, the parallel corpus exhibits remarkable values. It is easy to access at Sketch Engine, enabling systematic queries on decent coverage of –ism words. The English-Chinese concordance lines present a wealth of resources with which translators can tease out the senses conveyed by the –ism words, retrieve rich translation candidates in meaningful contexts that previous translators have worked with. The technical skills required appeared to be very attainable to translators with average IT competence. The merits of corpus resources are discussed in the context of translator training.

Keywords Suffix –ism · Translation equivalent · Lexicology · Parallel corpus

1   Introduction

This study on –ism words was first intrigued by the present investigator's experience of teaching translators/interpreters at the tertiary level. When translating a passage entitled "Sukarno: A Fallen Hero's Legacy" that contains sentence (1) into Chinese (Loveard 1994), almost every student rendered 'nationalism' into 民族主义 *mínzú zhǔyì* 'national/ethical group's ideology/ism', a noun phrase in which 民族 *mínzú* 'nation, ethnical group' modifies 主义 *zhǔyì* 'ideology or ism'. The suffix

L. Lim (✉)

School of Languages and Translation, Macao Polytechnic Institute, Rua de Luis Gonzaga Gomes, Macau
e-mail: llim@ipm.edu.mo

–ism is therefore rendered into 主义 *zhǔyì*, an easy-to-find corresponding item in Chinese. However, this is a rather simplistic solution.

(1) Across the 5,000 km breadth of Indonesia, one name is synonymous with Indonesian **nationalism**: Sukarno, nation-builder and architect of independence (boldfaced fonts added for emphasis).

More experienced translators would consider other expressions that better express the meaning of "nationalism" in (1) because the collocation 印度尼西亚民族主义 *yìnní mínzú zhǔyì* 'Indonesian nationalism' sounds neutral or even moderately negative to Chinese ears, suggesting an ideology that entertains a narrower perspective that prioritises national interests. In fact, in Cheng's (郑宝璇 2004) textbook titled *Translation for Media*, a sample translation is provided for the passage from which (1) was extracted, in which "nationalism" is translated into 民族精神 *mínzú jīngshén* 'the spirit or essence of the nation' (ibid: 69), rather than 民族主义 *mínzú zhǔyì*. This translation sounds positive and even heroic in Chinese, in line with the sentiment of awe and admiration on Sukarno expressed in the original text.

As illustrated by example (1), more experienced translators tend to have more varied ways to render –ism words into Chinese. There is therefore a real need for student translators to expand their lexical repertoire so that they can convey the denotations and connotations entailed by the –ism words in context more aptly and effectively. In this study, we examine some of the most accessible tools and resources that translators can use and trail through the discovery processes to reveal the potential of the resources in solving translation problems surrounding –ism words. We aim at casting light on the following research questions:

(a) To what extent do traditional tools—e.g., English-Chinese dictionaries and reverse dictionaries—enable translators to understand the senses of –ism words and to obtain useful translation candidates in Chinese for rendering the words?
(b) Compared with the traditional tools, what advantages do monolingual and parallel corpora exhibit in translators' understanding of and dealing with –ism words?
(c) What level of technical skills—e.g., advanced, sophisticated, or basic—are required so that translators can carry out self-directed corpus-based studies to attain reasonable outcomes in terms of understanding the –ism words and retrieving translation candidates?

2 A Brief Literature Survey

Research over the last three decades has attested the value of using monolingual, parallel and comparative corpora in lexical studies and lexicography (e.g., Fillmore 1992; Kovář et al. 2016; Kubicka 2019; Lefever and Hoste 2014; Li 2017; Li et al.

2020b; Li and Wang 李龙兴 & 王宪 2021; Sinclair 2001; Teubert 2001; Wang and Chu-Ren 2017; Wang 2018; Wang 2021 in this volume) and also in translation practice and translation studies (e.g., Baker 1999; Chen et al. 2020; Johansson 2007; Li et al. 2020a; Mauranen 2004; Wang et al. 2020; Xiao 2010; Zanettin 2014: Chaps. 2 and 5). Recent studies on translator and interpreter training underscore the importance of fostering (trainee) translators' competence in information and communication technology (ICT) (e.g., Doval and Nieto 2019: 3–4; Laviosa and Falco 2021 in this volume; Wang and Lim 2017). A stream of the studies stresses on the notion of ICT *literacy*, which is argued to be an integrated part in the translator/interpreter's training curriculum (e.g., Laviosa & Falco 2021: Sect. 3; Lim 2020: 152), referring to the competence such as working with computer-assisted translation (CAT) tools, corpus-based tools and resources, text analysis tools, and translation memory systems. It is widely acknowledged that translators should not only consult conventional resources such as dictionaries and glossaries but also make effective use of corpus-based tools and resources in their practice.

In addition, corpora have been used in research on the translation of affixations across languages (e.g., Defrancq and Rawoens 2016; Lefer 2012; Lefer and Grabar 2015; Quah 1999). Although we were unable to find systematic research on the translation of the suffix –ism across languages, Lim's (2019) corpus-based investigation on "terror*ism*" casts light in this respect. Lim (ibid) found that "terrorism" in English and 恐怖主义 *kǒngbù zhǔyì* in Chinese differ from each other at least in two ways. First, "terrorism" is markedly more frequently used than 恐怖主义 based on corpus evidence. Second, 恐怖主义 tends to be complemented by an NP (noun phrase) much more than "terrorism" does. More precisely, 恐怖主义 occurs in complex NPs—e.g., 恐怖主义活动 *kǒngbù zhǔyì huódòng* 'terrorist act/activity' much more frequently than "terrorism" does, since "terrorism" often entails the meaning of terrorist activity in context. This points to the observation that "terrorism" spans a wider range of semantic meanings than 恐怖主义 does and therefore tends to stand alone, while, by contrast, 恐怖主义 needs to rely on an additional word such as a following NP to express more specific meanings. This further leads to the working hypothesis that the suffix –ism tends to convey wider semantic meanings than 主义 does although it needs to be studied and verified with (corpus) evidence. Given the fact that English-Chinese bilingual corpora have been steadily developed and some are easily accessible via the Internet, we can now examine some major resources in this chapter that translators can use to better deal with –ism words.

3 Conventional Tools for Understanding –ism and Identifying Its Chinese Translations

In this section, we examine a range of conventional tools and resources that translators may access for reaching a better understanding of the senses of –ism words and for retrieving translation candidates for –ism words. We examine the

usefulness of monolingual and bilingual dictionaries and also monolingual corpora of both English and Chinese. From the results of this section, we will move on to explore in Sect. 4 the special contributions that a large parallel corpus can make in the hands of inquisitive translators with regard to their needs.

3.1 The Senses of –ism in Monolingual Dictionary

In terms of dictionaries, we examine a commonly used monolingual dictionary of English in this section and will explore an English-to-Chinese reverse dictionary and also large online dictionaries with English-Chinese sentence pairs as bilingual examples in the next Sect. (3.2). Of all the conventional resources, English dictionaries arguably present the most essential and traditional tools for translators to look up the definition and explanations of the suffix –ism regarding its main senses. For example, the Merriam Webster Dictionary (https://www.merriam-webster.com/dictionary/-ism) provides the definition of the noun suffix –ism in four sense groups and gives examples for each sense:

Senses	Examples	Chinese annotations[1]
1a: act: practice: process	// critic*ism* // plagiar*ism*	做法 *zuòfǎ* 'practice', 行为 *xíngwéi* 'act'
b: manner of action or behavior characteristic of a (specified) person or thing	// animal*ism*	做的方式 *zuò de fāngshì* 'manner of doing / acting'
c: prejudice or discrimination on the basis of a (specified) attribute	// rac*ism* // sex*ism*	歧视行为 *qíshì xíngwéi* 'discriminatory act'
2a: state: condition: property	// barbarian*ism*	状态 *zhuàngtài* 'state', 情况 *qíngkuàng* 'situation', 属性 *shǔxìng* 'property'
b: abnormal state or condition resulting from excess of a (specified) thing or marked by resemblance to (such) a person or thing	// alcohol*ism* // giant*ism*	病症 *bìngzhèng* 'disease, syndrome' …型 *xíng* '… shape, appearance'
3a: doctrine: theory: religion	// Buddh*ism*	学说 *xuéshuō* 'doctrine', 理论*lǐlùn* 'theory', 信仰 *xìnyǎng* 'religion, faith'
b: adherence to a system or a class of principles	// stoic*ism*	保持系统/原则 *bǎochí xìtǒng/yuánzé* 'adhere to (a) system/principle/s'
4: characteristic or peculiar feature or trait	// colloquial*ism*	语言特征等 *yǔyán tèzhēng děng* 'language features, etc.'

Note [1]Chinese annotations (with pinyin and gloss in English) are added by the investigator for distinguishing sense groups and suggesting potential translation candidates

Of the four sense groups, only the sense group 3 (on doctrine and principles) is closely related to 主义 in Chinese, while the other three sense groups (1, 2 and 4) are rarely so, e.g., "criticism" of sense 1a, "colloquialism" of sense 4. There are exceptional cases under the latter sense groups though, which mainly have to do with paraphrasing translation, e.g., "sexism", a word of sense 1c, may be paraphrased into 大男子主义 *dà nánzǐ zhǔyì* 'male chauvinism' in certain contexts, rather than the more commonly used translation 性别歧视 *xìngbié qíshì* 'sex(-based) discrimination'. The –ism words of sense group 3—i.e., on a doctrine, theory, or religion and on the adherence to a system or a set of principles—tend to be rendered into 主义 in Chinese. But still, not all the words under the sense group 3 are translatable into 主义 in Chinese, e.g., Budd*hism* is rendered as 佛教 rather than 佛主义. From the observations above, it is clear that not all –ism words are translatable into 主义 in Chinese, and, more precisely, we can (and translators in general should) reach the hypothesis that the semantic range of –ism is broader than that of the corresponding suffix 主义 in Chinese. The hypothesis needs to be tested with more bilingual language evidence, while the observations clearly point to the need for identifying translation candidates in Chinese for –ism apart from 主义.

3.2 Translation Candidates for –ism Words in English-Chinese Dictionaries

Reverse dictionaries facilitate the access of –ism words, given the fact that –ism is a suffix and all the –ism words appear in a continuous sequence in reverse dictionaries. We consulted *A Reverse English-Chinese Dictionary* by Sun (孙梅 1993), in which –ism words appear in eight consecutive pages, from "Lamaism" (ibid: 361) to "Nazism" (ibid: 368). Most of the ism words are annotated by only one or two Chinese terms, which explain the –ism word and may also serve as translation equivalents, while a few words are provided with three (e.g., "anarchism", "nationalism") or even four (e.g., "criticism") Chinese correspondents. The number of the Chinese correspondent terms for each –ism word is therefore not large, which is in fact far lower than the repertoire retrievable in the parallel corpus we will investigate (see Sect. 4). However, from the Chinese expressions corresponding to the –ism words over the eight pages in Sun (ibid), observant translators can still find a considerable range of Chinese expressions ending with recurring characters such as 教 *jiào* 'religion, sect', 会 *huì* 'school, association, society', 学 *xué* 'study, school', 法 *fǎ* 'method', 论 *lùn* 'discourse, doctrine', 者 *zhě* 'people, scholar', 风气 *fēngqì* 'trend', 运动 *yùndòng* 'movement', 行为 *xíngwéi* 'behaviour', and 精神 *jīngshén* 'spirit', 性 *xìng* 'character', apart from 主义. These expressions are informative to the translators about the diversified translation candidates for –ism words.

In addition to reverse dictionaries and monolingual dictionaries, large English-Chinese dictionaries are undoubtedly one of the most used bilingual

resources by translators. Lu Gusun's (陆谷孙 2007) 英汉大词典 (*The English-Chinese Dictionary*) is one of the most reliable English-Chinese dictionaries, which is often considered as a must-have tool by professional translators. It exhibits very conscientious work in terms of the coverage of the senses of each entry, using Chinese to finely and precisely explain the senses, which are further illustrated by examples in English with Chinese translations. Translators can look up all the information for potential translation candidates they can employ, or at least for inspiration to sort out useful translations. However, translators cannot expect that they can find the precise items immediately useful for the text they are translating, since the dictionary is more about the English language in terms of the senses of the entries (cf. Kubicka 2019: 84), rather than being designed for providing translation solutions. In other words, it is not a dictionary for English-Chinese translators.

Moreover, online English-Chinese dictionaries tend to be large, updated with free access, which come with bilingual examples of phrases and sentences that often exceed printed dictionaries in number. However, the online dictionaries do not allow translators to search all the –ism words together as they can do with reverse dictionaries, and translators can only search one –ism word at a time. Large online English-Chinese dictionaries provide examples of English-Chinese sentence pairs. However, the number and variety of the examples may still not be large enough for translators to harvest a decent range of translation candidates, although they may be good enough for language learners in general. For example, in YouDao dictionary, one of the largest online dictionaries developed in China (http://dict.youdao.com/), we retrieved only 27 English-Chinese bilingual sentence pairs for "nationalism". Of the 27 tokens of Chinese terms that correspond to "nationalism", only three different types emerged. We therefore consulted large online dictionaries with bilingual sentence pairs (e.g., YouDao dictionary) and also online bilingual corpus portals—e.g., CCL Chinese-English bilingual corpus of Peking University (http://ccl.pku.edu.cn:8080/ccl_corpus/index_bi.jsp), and BCC bilingual corpus of the BLCU (Beijing Language and Culture University) Corpus Center (although it is not accessible from 2021) (http://bcc.blcu.edu.cn/lang/bi)—to search the most frequently occurring –ism words (cf. Sect. 3.3) one at a time. We were able to retrieve various Chinese expressions other than 主义—e.g., 精神 *jingshen* 'spirit', 行为 *xingwei* 'behaviour', 制度 *zhidu* 'system', 运动 *yundong* 'movement', 性 *xing* 'character' – to render –ism words such as "professionalism", "feminism", "vandalism", and "terrorism". However, searching one –ism word at a time in these resources is laborious, and the translation equivalents identified tend to be low in both tokens and types compared to those contained in the parallel corpus "Education" (see Sect. 4), which also enables all –ism words to be retrieved by one search in the Sketch Engine platform (cf. Fig. 1).
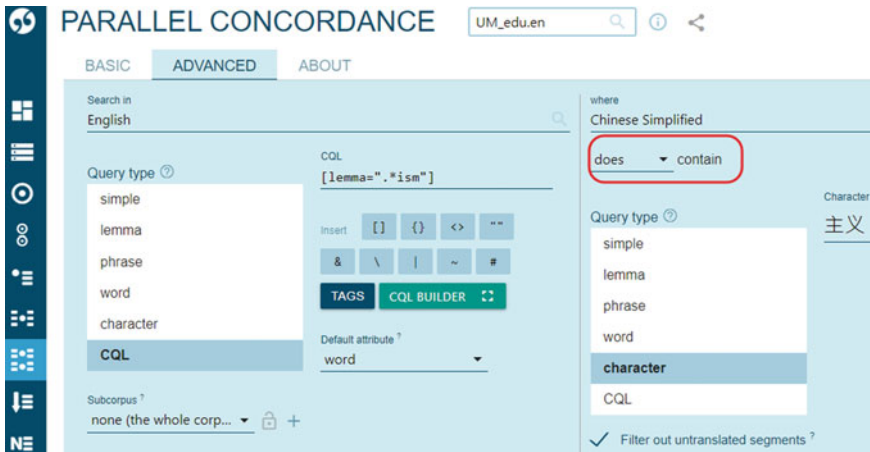
Fig. 1 The "containing" query that searches both the –ism words in English and the characters 主义 in Chinese

3.3 Monolingual Corpora of English and Chinese: Comparing –ism and -主义 Words

Accessing monolingual English corpora enables translators to gain basic information about the frequency of occurrence of the –ism words taken together and also to identify the most frequently used ones. We used the British National Corpus (BNC) for the purpose, in view of its coverage of a variety of genres of both written and spoken texts. The results indicate that the suffix –ism is highly productive in English. The tokens of –ism words amount to 56,588 in total in BNC, which translates into a rate of 503.7 per million tokens (PMT). This frequency of occurrence is closest to that of the lemmas of "company" (57,118 tokens) and "course" (56,776), which are ranked the 21st and 22nd most commonly used nouns in BNC. Therefore, statistically, translators have to deal with various –ism words in their day-to-day practices. The top 30 most frequently used –ism words are listed in Table 1, according to the Wordlist resulting from the search of lemma ending with "ism" in BNC, a preloaded corpus accessible at Sketch Engine (SkE). In view of the much varied senses entailed by the large array of –ism words (cf. Sect. 3.1), translators should work more efficiently and effectively by gaining a better understanding of –ism words in terms of their senses and potential translation candidates.

To compare with the results of –ism words in BNC, a Chinese monolingual corpus to query主义expressions was accessed. We used the Chinese GigaWord 2 Corpus (Mainland, simplified characters), created in 2005, preloaded to SkE and tagged with the Chinese GigaWord tagset which covers newswire texts. Our Character search of主义 under the Concordance tab returned 146,816 tokens, at a rate of 587.0 per million tokens (PMT). The 30 most frequently used Chinese terms

Table 1 BNC Wordlist: lemma ending with "ism"

Rank	Lemma	Freq	PMT	Rank	Lemma	Freq	PMT
1	criticism	5806	51.68	16	scepticism	637	5.67
2	mechanism	4921	43.80	17	journalism	594	5.29
3	**capitalism**	1884	16.77	18	liberalism	504	4.49
4	organism	1793	15.96	19	professionalism	497	4.42
5	**socialism**	1644	14.63	20	metabolism	493	4.39
6	tourism	1464	13.03	21	conservatism	477	4.25
7	**racism**	1093	9.73	22	baptism	476	4.24
8	**nationalism**	1024	9.11	23	**imperialism**	420	3.74
9	**communism**	949	8.45	24	unionism	393	3.50
10	**realism**	888	7.90	25	modernism	390	3.47
11	optimism	840	7.48	26	pluralism	384	3.42
12	**Marxism**	732	6.52	27	**individualism**	380	3.38
13	feminism	693ara>	6.17	28	symbolism	378	3.36
14	**terrorism**	690	6.14	29	vandalism	338	3.01
15	fascism	654	5.82	30	Catholicism	328	2.92

Abbreviation PMT = per million tokens

are presented in Table 2, based on both the keyword in context (KWIC) list and the list of the first word on the left of 主义. The -主义 expressions in Chinese occur at a relatively higher frequency than that (503.7 PMT) of –ism words in BNC. However, we should note that the term 社会主义 *shèhuì zhǔyì* 'socialism' accounts for 54% of all the tokens of 主义 expressions in the Chinese GigaWord 2 Corpus (cf. Table 2), which reflects the tendency in lexical choices in mainland China. There are some other expressions in this Chinese corpus—e.g., 马克思主义 *mǎkèsī zhǔyì* 'Marxism' (ranked 4), 共产主义 *gòngchǎn zhǔyì* 'communism' (ranked 6), and 资本主义 *zīběn zhǔyì* 'capitalism' (ranked 9)—that occur at markedly higher frequencies than their English correspondents do in BNC. We paid particular attention to the –ism words in Table 1 that have translation equivalents in Table 2, and identified ten and highlighted them with bold-faced fonts in Tables 1 and 2. The ten –ism words tend to be rendered into -主义 expressions in Chinese, according to the intuition of proficient Chinese-English bilingual speakers. By contrast, some other words—e.g., "criticism", 'tourism', "metabolism"—are the most improbable words to be translated into主义 expressions in Chinese. However, this intuition needs to be tested with the evidence from parallel corpora (cf. Sect. 4.2).

Table 2 The most frequently used -主义 expressions in the Chinese GigaWord 2 Corpus

Rank	Chinese terms	*Pinyi*	Gloss in English	Freq	PMT
1	社会主义	*shèhuì zhǔyì*	**socialism**	78,629	314.36
2	爱国主义	*àiguó zhǔyì*	patriotism	9524	38.08
3	恐怖主义	*kǒngbù zhǔyì*	**terrorism**	8749	34.98
4	马克思主义	*mǎkèsī zhǔyì*	**Marxism**	8032	32.11
5	人道主义	*réndào zhǔyì*	humanitarian	6432	25.72
6	共产主义	*gòngchǎn zhǔyì*	**communism**	3040	12.15
7	保护主义	*bǎohù zhǔyì*	protectionism	2455	9.82
8	霸权主义	*bàquán zhǔyì*	hegemonism	2241	8.96
9	资本主义	*zīběn zhǔyì*	**capitalism**	2056	8.22
10	帝国主义	*dìguó zhǔyì*	**imperialism**	1904	7.61
11	民主主义	*mínzhǔ zhǔyì*	democracy	1802	7.20
12	马列主义	*mǎliè zhǔyì*	Marxism-Leninism	1792	7.16
13	种族主义	*zhǒngzú zhǔyì*	**racism**	1635	6.54
14	形式主义	*xíngshì zhǔyì*	formalism	1454	5.81
15	军国主义	*jūnguó zhǔyì*	militarism	1355	5.42
16	集体主义	*jítǐ zhǔyì*	collectivism	1053	4.21
17	殖民主义	*zhímín zhǔyì*	colonialism	1002	4.01
18	官僚主义	*guānliáo zhǔyì*	bureaucracy	946	3.78
19	分裂主义	*fēnliè zhǔyì*	separatism	898	3.59
20	马克思列宁主义	*mǎkèsī lièníng zhǔyì*	Marxism-Leninism	870	3.48
21	唯物主义	*wéiwù zhǔyì*	materialism	865	3.46
22	极端主义	*jíduān zhǔyì*	extremism	809	3.23
23	民族主义	*mínzú zhǔyì*	**nationalism**	637	2.55
24	拜金主义	*bàijīn zhǔyì*	money-worship-ism	545	2.18
25	英雄主义	*yīngxióng zhǔyì*	heroism	506	2.02
26	分离主义	*fēnlí zhǔyì*	separatism	382	1.53
27	个人主义	*gèrén zhǔyì*	**individualism**	296	1.18
28	享乐主义	*xiǎnglè zhǔyì*	hedonism	295	1.18
29	现实主义	*xiànshí zhǔyì*	**realism**	285	1.14
30	国际主义	*guójì zhǔyì*	internationalism	253	1.01

Abbreviation PMT = per million tokens

4 Querying a Parallel Corpus: The UM-Corpus at SkE

We selected an English-Chinese parallel corpus named "Education" and uploaded it to SkE for our queries and analysis. The "Education" corpus is one of the largest components of the UM-Corpus, a multi-domain and balanced parallel corpus constructed at the University of Macau (UM) (see Tian et al. 2014). The "Education" corpus consists of texts "acquired from online teaching materials, such as language teaching resources, and dictionaries, which can be served as language

education" (ibid: 1840). We selected this corpus owing to its reasonable language and translation quality, decent size, and relevance to language. The "Education" corpus is downloaded from http://nlp2ct.cis.umac.mo/um-corpus/index.html, the website in which a 2.2 million sentence-pair version of the UM corpus is released to the community for research purposes. In addition, the "Education" corpus is one of the largest subcorpora of the released version, which consists of 450,000 pairs of sentences in English and Chinese with 8,401,095 tokens in English and 13,749,570 tokens in Chinese (ibid: 1840, Table 4). The direction of translation is not specified in the released version though. Therefore, the translation candidates for –ism words we intended to identify in the "Education" corpus cannot be specified in terms of the direction of translation. That is, the words and expressions identified in Chinese can either be the translations of –ism words from English into Chinese—i.e., "translation equivalents" in Mikhailov's (2021) terminology in this volume—or the lexical items in Chinese that are rendered into –ism words in English—i.e., "translation stimuli" (ibid).

Most of the sentence pairs in the "Education" corpus sound like English-to-Chinese rather than Chinese-to-English translation, as evidenced by their subject matter, style and wording. The investigator googled various sentence pairs on the web and confirmed that in most of the pairs the English sentences were the original, taken from books, articles, Wikipedia entries, U.S. presidents' speeches, and transcripts of lectures given at Yale University. A small portion of the sentence pairs is Chinese-to-English translation, e.g., the bilingual texts extracted from the Selected Works of Mao Tse-Tung, e.g., his speeches delivered on 3 May 1937 and on 7 April 1944. Therefore, the translation candidates identified in the "Education" corpus for rendering –ism words into Chinese predominantly consist of the instances of "translation equivalents" and also a small number of "translation stimuli", according to Mikhailov's terminology (ibid). In this section, the "translation" of –ism words is taken in a broad sense.

In the following sub-sections, we explore the values of the "Education" corpus in terms of the coverage of –ism words, the richness of translation candidates available, and the potential for translators to devise queries in order to answer their questions or tease out some patterns or regularities on the translation of –ism words. We also consider the technical competence and skills needed in the translators to conduct inquisitive learning with the corpus.

4.1 The Coverage of –ism Words

We retrieved all the instances of –ism words in the "Education" corpus under the Concordance tab using the following query devised in corpus query language (CQL):

[lemma = " *ism"]

which returned 6,724 tokens of –ism words (i.e., 716 PMT), excluding an erroneous instance that was manually identified, in which the sentence boundary is mistaken

by the system. Based on the Frequency List of the keyword in context (KWIC) for the concordance lines, we manually detected the –ism words with capitalised initials and merged them to lower-case ones. This leads to a total of 615 types of the lemmas of –ism words in English.

Compared with BNC (96,134,547 words in total), which returns 56,694 tokens and 1,543 types of –ism words by the same CQL query indicated above, the English subcorpora (8,168,479 words) of the "Education" corpus exhibits a considerable coverage of the commonly-used –ism words in both types (n = 615) and tokens (n = 6,724). The Chinese sentences aligned with the English ones constitute the wealth of resources from which translation candidates can be harvested.

4.2   Differentiating Three Types of –ism Words

Once the "Education" corpus was uploaded to SkE as a Multilingual Corpus, it was split into two corpora—an English one (8,168,479 words) and a Chinese one (8,020,107 words) with simplified Chinese characters—while the two are interlinked at the sentence level. The two work together as an English-Chinese parallel corpus that allows simultaneous searches of both corpora. The combined searches are crucial for translators to query the –ism words and the –主义 expressions in the English-Chinese paired sentences. The searches lead to quantitative results, which are far more informative than those provided by conventional resources (cf. Sect. 3).

The queries lead to an immediate finding that there are far more instances of –ism words that are *not* translated into words containing the Chinese characters 主义 *zhǔyì* 'doctrine, dogma' than those that are rendered into –主义 words. The investigator reached the finding from the quantitative results returned by the "containing" query and the "not-containing" query (detailed below), while translators can easily devise similar queries with the built-in query features at SkE. The "containing" query simultaneously searches the –ism words in the English corpus and the characters 主义 in the Chinese corpus (see Fig. 1), while the two corpora are aligned at the sentence level. It therefore retrieves the sentence pairs (n = 1,833) in which an –ism word occurs in the English sentence, while a -主义 word occurs in the aligned Chinese sentence. This involves "noises" in which the -主义 word in Chinese is the translation of a word other than the –ism word in the English sentence, for example, an –ist word (e.g., "socialist") rather than an –ism word in the sentence is translated into 主义. The parallel concordance lines were manually checked through and a low number of instances of the noise were found. Exclusion of the noises will further reduce the number of 1,833 sentence pairs returned by the "containing" query, while the main finding of the trends is valid and even strengthened. We use the term "*zhuyi*-words" to refer to those –ism words in English that are either exclusively or predominantly translated into 主义 in Chinese, e.g., "socialism", "capitalism", and "Marxism".

The "not-containing" query returned a much larger number of sentence pairs (n = 4,892) than the "containing" query (n = 1,833) did. The former query can be considered a simple modification of the latter in which the search of the Chinese characters 主义 is switched from "does contain" (like in Fig. 1) to "does *not* contain", gathering all the instances of the –ism words with no 主义 in the corresponding Chinese sentences. The result indicates that there are at least 2.67 times as many –ism words that do not correspond to 主义 in Chinese as those that do in our data. We term the former type of –ism words "non-*zhuyi*" words to indicate their tendency of not being rendered into any word of 主义. Given the predominance of the "non-*zhuyi*" words over the "*zhuyi*" ones, translators need to give particular attention to the former group. Apart from the typical "*zhuyi*" and "non-*zhuyi*" words, there are –ism words that yield notable results in both the "containing" and "not-containing" queries, which we termed "partly *zhuyi*-words", e.g., "terrorism", "realism", and "activism".

By tabulating the frequencies of occurrence of each –ism word in both the "containing" and the "not-containing" queries, its tendency emerges regarding whether or not it tends to be rendered into (or, at least, corresponds to) –主义 words in Chinese. The –ism words in English can therefore be classified into three types— i.e., the "*zhuyi*", "non-*zhuyi*", and "partly-*zhuyi*" words. Classifying the –ism words into the three types entails practical value to translators, who can thereby deal with the three differently.

The first type—i.e., the *zhuyi*-words—is the easiest to render into Chinese, and competent translators should be able to produce a list of these words rather easily with SkE. Once the parallel concordance is generated by the "containing query" (cf. Fig. 1), one can yield the list of *zhuyi*-words using the Frequency of keyword in context (KWIC) feature for the English corpus (on the left in the parallel-concordance display), and, similarly, producing the list of –主义 words in the Chinese corpus (on the right). The English and Chinese lists match each other closely in the most frequently occurring items, e.g., "socialism" in English to 社会主义 *shèhuì zhǔyì* 'socialism' in Chinese, "capitalism" to 资本主义 *zīběn zhǔyì* 'capitalism', "terrorism" to 恐怖主义 *kǒngbù zhǔyì* 'terrorism', "imperialism" to 帝国主义 *dìguó zhǔyì I* 'imperialism', "Marxism" 马克思主义 to *mǎkèsī zhǔyì* 'Marxism', and "anarchism" to 无政府主义 *wú zhèngfǔ zhǔyì* 'anarchism'. Once the "noises" that contain mismatches between –ism and –主义 words were cleaned, we found that these words are almost exclusively rendered into –主义 words in Chinese. Translators can handle them straightaway without spending their time considering alternative translations.

Similarly, we can retrieve the list of non-*zhuyi* words using the concordance lines in English generated by the "not-containing" query. Top on the list by descending frequency order are the words including "mechanism", "criticism", "organism", "tourism", "metabolism", "Buddhism", "autism", and "baptism". These words are rarely rendered into –主义 in our data. This appears to be useful information to translators, although native speakers of Chinese should intuitively know this. Most of the words have nothing to do with a doctrine, a system of beliefs or an ideology, and therefore are understandably non-*zhuyi* words. In addition,

some words—e.g., "Taoism" and "Confucianism"—that definitely pertain to some belief systems or doctrines are ranked high on the non-*zhuyi* list as well. To sort out the ways and regularities for rendering the –ism words, parallel corpus presents primary resources that translators can turn to for information and inspiration.

In terms of the third type—the partly *zhuyi*-words—their meaning spans from the sense of *zhuyi* (doctrine, dogma, ideology) to other senses, such as the practice of a certain kind or an institution or a system established under the guidance of a certain dogma or ideology. Their most frequently occurring correspondent items in Chinese serve as important indicators of the range of the senses expressed by these –ism words.

4.3 Translation Candidates in Chinese: The Repertoire and Retrieval Methods

Since *zhuyi*-words are the easiest to translate into Chinese (cf. Sect. 4.2), we focus on the non-*zhuyi* and partly-*zhuyi* words and their potential translations in this section. We found the parallel corpus particularly useful for providing the translation candidates of individual –ism words as well as for differentiating major types of –ism words that tend to be rendered into some commonly used Chinese expressions other than 主义 (cf. Sect. 4.4).

The parallel concordance lines retrieved by the "not-containing" query display the English and Chinese sentences in juxtaposition, which present a quick reference for diversified translations of each –ism word. The English sentences can be alphabetically sorted by the –ism words they contain, which are the KWIC in the English concordance lines on the left, giving a display of all the sentences that contain each "non-*zhuyi*" –ism word (see Fig. 2), beginning with "absenteeism", "academism", and "activism" onward. Browsing the Chinese correspondents (on the right) to the concordances of "absenteeism" (on the left) immediately reveals the markedly frequent occurrences of both 旷工 *kuànggōng* 'absent from work' and 旷课 *kuàngkè* 'absent from school/class'. Translators can use the "Find on Page" or the "Find in This Page" feature in the commonly used web browsers (e.g., Firefox, Google Chrome, Microsoft Edge) to highlight *all* the instances of the character旷 *kuàng* on the page, as illustrated in Fig. 2. In so doing, they can quickly scan through the Chinese expressions with highlighted instances of 旷 and then draw their attention to the lines in which 旷 is absent to discover other translation candidates for "absenteeism". For instance, 缺勤率 *quēqín lǜ* 'absence/absenteeism rate' that occurs in the second last line of the Chinese concordances (underlined by the investigator) in Fig. 2 is a case in point. In the same way, translators can search the individual –ism words they need to study for their interest or for their translation tasks and look into the Chinese concordances for a wealth of useful translation candidates.
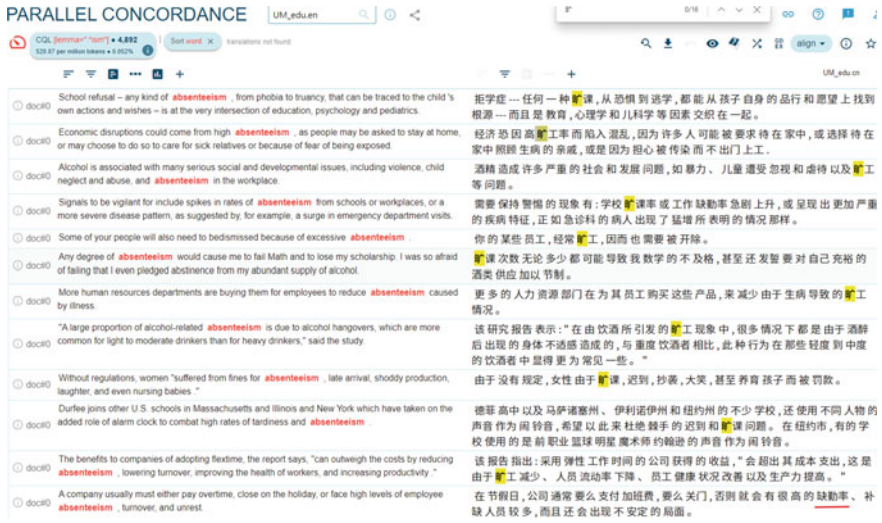
Fig. 2 Parallel concordance of "non-*zhuyi*" –ism words with highlighted characters in the Chinese concordances

Those frequently occurring –ism words lead to a large number of concordance lines, e.g., "mechanism" has 1,321 lines and "criticism" 334. Browsing through the concordance lines takes a considerable amount of time and efforts, although this allows translators to thoroughly identify all the translation candidates in the corpus. We were able to manually collect a large repertoire of translation candidates for "mechanism", of which the frequently occurring ones are presented in Table 3. Meanwhile, there are also other much less frequently used ones such as 病机 *bìngjī*, 'pathogenesis, 心理 *xīnlǐ* 'psychology', 方法 *fāngfǎ* 'method, means' and so on, which are not included in Table 3 though.

To save the labour of manually picking through the concordance lines, translators can create a subcorpus of the Chinese concordances of the specific –ism word/s they want to focus on. Using the subcorpus they can generate the lists of keywords and key terms in which the frequently used Chinese item/s corresponding to the –ism word under investigation should emerge. We take "mechanism" as an example again, since it is by far the most frequently used –ism word in our sample, with 1,321 tokens in total and taking both singular and plural forms. It corresponds to a wide range of terms in Chinese, as previously discussed. One can perform parallel-text concordancing by searching the lemma of "mechanism", using the English corpus as the primary corpus and the Chinese one as the secondary. From the query results, one can create a subcorpus of the Chinese concordance lines displayed on the right hand of the parallel-concordance interface and switch to the Chinese corpus (from using the English corpus) so as to have access to the subcorpus of "mechanism". The subcorpus can then be processed under the keywords tab to extract keywords (termed "single-words" in SkE) and key terms (termed

Table 3 The list of Chinese translation correspondents of "mechanism"

Chinese terms with *Pinyin*	Gloss in English
机制 *jīzhì*	'mechanism'
机理 *jīlǐ*	'principle, (reaction) mechanism"
作用 *zuòyòng*	'effect, function'
机构 *jīgòu*,	'device, mechanism'
机械 *jīxiè*	'mechanical (device)'
结构 *jiégòu*	'structure'
装置 *zhuāngzhì*	'device'
原理 *yuánlǐ*	'principle, theory'
系统 *xìtǒng*	'system"
机器 *jīqì*	'machine'

"multi-words terms" in SkE), using Chinese Web (enTenTen17, simplified characters) as the reference corpus, which comes as the default. The top translation candidates for "mechanism" we manually identified above also appear in the keyword list—e.g., 机制 (n = 103), 机理 (n = 41), 作用 (n = 28), 机构 (n = 20), 装置 (n = 9) (cf. Table 3). The key-term list further indicates that some of the keywords tend to join each other to form collocations, e.g., 作用机制 *zuòyòng jīzhì* 'mechanism of action', 作用机理 *zuòyòng jīlǐ* 'mechanism or principle of action', 机械装置 *jīxiè zhuāngzhì* 'mechanical device'. The results not only attest to the significance of the translation candidates but also provide useful collocational patterns of the items. To reiterate, creating and using the subcorpus enables translators to harvest the frequently occurring Chinese translation candidates rather efficiently and effectively at SkE.

Once the list of the translation candidates has been produced, inquisitive translators would then need to sort out the (typical) situations in which the major translation candidates tend to be used. In the example of "mechanism", the list of the Chinese translation candidates contains items relating to a variety of technical fields and domains, of which translators may not have specific knowledge. The translators, however, can always draw on the subcorpus that consists of the English and Chinese concordances as described above to better understand the translation candidates in terms of their context of use. With this in mind, we now examine five frequently occurring translation candidates 机械, 机理, 作用, 机构 (cf. Table 3) and also 病机, which is a truncated form.

The term 机械 *jīxiè* 'mechanical, machinery' emerges in our data with close relevance to mechanical engineering. Turning on "Find on page" in the web browser to highlight all the instances of 机械 in the parallel concordances of the subcorpus on "mechanism", translators can quickly locate the typical co-texts in which the word "mechanism" occurs in English while the term 机械 is used in Chinese. The examples include "good mechanism performance" in English, which is aligned to 优良的机械物理性能 *yōuliáng de jīxiè wùlǐ xìngnéng* 'excellent

mechanical (and) physical properties/performance' in Chinese, and "[a] mechanism (for launching aircraft)" in English corresponding to (发射航空器的)一种机械装置 '(fāshè hángkōngqì de) yī zhǒng jīxiè zhuāngzhì' 'a mechanical device (for launching an aircraft)' in Chinese. From the contexts of use, we can observe that 机械, when used as a translation correspondent for "mechanism", denotes mechanical devices, machines, or appliances. Similarly, by "finding" the term 机理 on the concordance page, one can soon realise that it tends to refer to the scientific or theoretical understanding, basis, or principles that explain certain phenomena or observations in natural sciences as well as in pathology. The parallel sentences show that 机理 can refer to the mechanism for a gravitational action in physics, that for crops breeding in biology, that for the ageing of materials in chemistry, and that for a virus infection or brain's reactions to drugs in pathology. In addition, searching the translation candidate 作用 on the concordance page shows that it tends to form some frequently occurring collocations such as 作用机制 or 作用机理 (cf. Table 3). Such collocations also appear in the multi-term list at SkE, which refer to the reason for which, or the underlying processes by which, a (pharmaceutical or pathological) effect or outcome takes place. These Chinese collocations closely correspond to the English expression "mechanism of action" in terms of word-for-word meaning—i.e., "action" for 作用 and "mechanism" for 机理 or 机制 (cf. Table 3). However, the phrase "mechanism of action" occurs infrequently in our English data, while the single word "mechanism" tends to be commonly used in context to refer to the action of drugs. For example, "mechanism" in the English concordance line matches a longer expression (药物) 作用机制 (yàowù) zuòyòng jīzhì '(drug's) action/effect mechanism' in the Chinese line. There are also situations in which 作用 is used alone to translate "mechanism", e.g., "the anti-inflammatory mechanisms (may be related to…)" in English is rendered into 抗炎作用 (可能与…有关) kàng yán zuòyòng (kěnéng yǔ…yǒuguān) 'anti-inflammatory effect (may be related to…)" in Chinese. This shows that the collocation 作用机制 in Chinese is simplified into 作用, while the meaning of 机制 can be still derivable from the context. Moreover, translators can easily note from the concordance lines that, when 机构 is used to render "mechanism", it refers to a *mechanical* device, tool, or structure rather than an institution or (administrative, educational) establishment. Furthermore, our concordance data on "mechanism" contains both (a) the compound expressions such as 发病机制 fābìng jīzhì, 致病机理 zhì bìng jīlǐ, and 疾病机制 jíbìng jīzhì, which all have the meaning of 'pathogenesis' and (b) the shortened form 病机 bìng jī. Sharp-eyed translators should quickly figure out that the latter is a truncated form of any the three compounds.

From a wide range of translation candidates for "mechanism", we can observe that each of the expressions in Chinese tends to point to a certain sense of "mechanism" in English. Sorting out different translation candidates turns out to be a process of laying out various senses that the word "mechanism" entails. This learning process allows translators to understand "mechanism" in terms of its fine-grained shadings of meaning by its translation correspondents.

4.4 From Chinese Translations to the Senses of –ism Words

The parallel corpus not only allows translators to conduct self-directed learning of the translation of –ism words from English into Chinese (cf. Sect. 4.3), and it can also be used for queries in the other direction—i.e., from Chinese to English—translators can investigate the commonly used Chinese expressions and seek out the (type of) –ism words that tend to be rendered into these expressions. For example, 论 *lùn* 'doctrine, discourse, theory' is a very notable character on the keyword list of the subcorpus of the Chinese concordances corresponding to –ism words in English, which appears in various expressions in Chinese—e.g., 论 in 不可知论 *bùkězhī lùn* 'agnosticism', 无神论 *wúshén lùn* 'atheism', 建构论 *jiàngòu lùn* 'constructivism'. Translators may further sort out the (types of) –ism words that tend to be rendered into the …论 expressions in Chinese. An exhaustive search can be conducted by a query similar to the "containing" query in Fig. 1, in which 论 replaces 主义 to be the character contained in the Chinese corpus. The query returns 749 parallel concordance lines in the "Education" corpus, while the English concordances present the resources in which the –ism words corresponding to 论 are to be retrieved. From the frequency list of KWIC lemmas of the English concordances, translators can spot and then confirm with examples in the concordances the –ism words that correspond to …论 terms (see Table 4).

Table 4 The –ism words that correspond to …论 terms in Chinese

The –ism words in English	The corresponding Chinese terms and *pinyin*	The gloss of Chinese terms
aestheticism	美学理论 *měixué lǐlùn*	'aesthetic theory or doctrine'
agnosticism	不可知论 *bù kězhī lùn*	'agnostic doctrine'
animism	泛灵论 *fàn líng lùn*	'animist, or pan-spirit, doctrine'
anthropomorphism	神人同性论 *shén rén tóngxìng lùn*, 拟人论 *nǐrén lùn*	'anthropophuistic doctrine', 'personification doctrine'
atheism	无神论 *wúshén lùn*	'atheist doctrine'
bloodline-ism	血统论 *xuètǒng lùn*	'blood-line doctrine, pedigree'
cognosciblism	可知论 *kězhī lùn*	'gnostic doctrine'
constructionism	建构论 *jiàngòu lùn*	'constructivist doctrine'
dogmatism	教条论 *jiàotiáo lùn*	'dogma(tist) doctrine'
electromagnetism	电磁理论 *diàncí lǐlùn*	'electromagnetic theory'
esotericism	隐微理论 *yǐn wēi lǐlùn*	'esoteric or implicit theory'
functionalism	德国功能派翻译理论 *déguó gōngnéng pài fānyì lǐlùn*	'German functionalist translation theory'
Marxism	马克思主义理论 *mǎkèsī zhǔyì lǐlùn*	'Marxist theory'
nihilism	国家主权虚无理论 *guójiā zhǔquán xūwú lǐlùn*	'national sovereignty nihilist theory'
syllogism	三段论 *sānduàn lùn*	'three-statements theory'

The –ism words in Table 4 strongly suggest that the …论 expressions correspond to those –ism words that denote schools of thoughts, scholarly or ideological positions and theories. The sense expressed is very close to that of 主义, and indeed several –ism words in Table 4 correspond to both 论 and 主义. For example, "aestheticism" can be also rendered into 唯美主义 *wéiměi zhǔyì*, "dogmatism" into 教条主义 *jiàotiáo zhǔyì*, "nihilism" into 虚无主义 *xūwú zhǔyì*, and "functionalism" into 功能主义 *gōngnéng zhǔyì*. These words can be considered "partly–*zhuyi*" words, for which both 主义 and 论 are major translation candidates. In addition, the word "Marxism" can be rendered into 马克思 主义理论 in some contexts (cf. Table 4), in which 主义 and 论 co-occur to form a Chinese word cluster. Translators can certainly hold that "Marxism" is a *zhuyi*-word, given the fact that its Chinese correspondents almost always contain 主义 (cf. Sect. 4.2), though truncated forms can be used too. Moreover, from the corpus evidence, translators can understand that the Chinese correspondents of "Marxism" can be expanded from 马克思主义 to 马克思主义 理论 (cf. Table 4) and 马克思主义思想 *mǎkèsī zhǔyì sīxiǎng* 'Marxist thoughts or ideology'. Furthermore, unlike the *zhuyi*- and partly-*zhuyi* words, there are words in Table 4 that are rendered into …论 but *not* …主义 expressions in our data, e.g., "agnosticism", "electromagnetism", and "syllogism", which are clearly non-*zhuyi* words.

Using the same methods as described for querying 论 above, translators can investigate other frequently occurring Chinese expressions and their corresponding –ism words in English. Some main findings of our searches are presented in Table 5, although, given the space of this chapter, we are unable to enlist all the notable recurring Chinese expressions that emerged in our data. From the Chinese expressions, we can observe that the –ism words entail at least three major senses— i.e., (a) action, (b) doctrine, and (c) characteristics or particularities of language, disease, and so on—which we can now illustrate with examples.

The first sense pertains to action/s or behaviour/s of some kinds. For example, 无意识行为 *wúyìshí xíngwéi* 'non-conscious act' corresponds to "automatism", which denotes actions performed unconsciously. The actions can be performed collectively in the form of a campaign or a social movement as well—e.g. 反恐运动 *fǎnkǒng yùndòng* 'antiterrorist campaign' that corresponds to "counterterrorism" in English. In addition, the action(s) can be denoted by the actors or performers engaged in the action, and, for this reason, we can put this particular sense to the first sense of –ism words. For example, the word "activism" in (2) is rendered into 活动家 *huódòng jiā* 'activists', while "futurism" in (3) is rendered into 未来学家 *wèiláixué jiā* 'futurology scholar/s', which sounds appropriate and idiomatic in context. In fact, translators can discover that the "Education" corpus contains rich examples in which the Chinese expressions with …家 *jiā* 'master, scholar' or …者 *zhě* 'practitioner, participant' correspond to –ism words in English. For example, 儒家 *rú jiā* 'Confucius followers or doctrine' corresponds to "Confucianism", referring to both the followers and the school of Confucianism. Similarly, 不可知论者 *bù kězhīlùn zhě* 'not-knowable-doctrine believer' corresponds to "agnosticism" in context.

Table 5 Chinese expressions and the corresponding –ism words in major sense (groups)

Chinese expressions, *pinyin* and gloss	Chinese Examples, *pinyin* and gloss in English	The corresponding –ism words
(a) Action		
…行为 *xíngwéi* 'act, action, behaviour'	无意识行为 *wúyìshí xíngwéi* 'non-conscious act'	automatism
	破坏 (公共财物) 行为 *pòhuài (gōnggòng cáiwù) xíngwéi* 'damaging (public assets) acts or behaviours'	vandalism
	反犹太行为 *fǎn yóutài xíngwéi* 'anti-Semitic act'	anti-Semitism
	恐怖主义行为 'terrorism act'	terrorism
	利他行为 *lìtā xíngwéi* 'benefitting others act'	altruism
(a1) Actor or practitioner		
…家 *jiā* 'a master, expert, or scholar of a doctrine, or the school of thoughts promoted by the practitioner'	活动家 *huódòng jiā* 'activist'	activism
	儒家 *rú jiā* 'Confucius followers or the school'	Confucianism
	道家 *dào jiā* 'Daoist person or doctrine'	Daoism
	佛家 *fó jiā* 'Buddhist person or doctrine'	Buddhism
	法家 'legalist person or doctrine'	Legalism
	批评家 *pīpíng jiā* 'criticising person, critics'	criticism
	未来学家 *wèiláixué jiā* 'futurology scholar'	futurism
(b) Doctrine, school of thoughts; mentality		
…学 *xué* 'study, doctrine'	孔子学说 *kǒngzǐ xuéshuō*, 儒学 *rúxué* 'Confucius doctrine or studies'	Confucianism
	汉学 *hànxué* 'Han or Chinese studies'	sinologism
	行为学派 *xíngwéi xuépài* 'the behaviourist doctrine or school'	behaviourism
	(生物)电磁学 *(shēngwù) diàncí xué* '(bio-) electromagnetic studies'	(bio)electromagnetism
	东方学 *dōngfāng xué* 'Oriental studies'	Orientalism

Table 5 (continued)

Chinese expressions, *pinyin* and gloss	Chinese Examples, *pinyin* and gloss in English	The corresponding – ism words
…精神 *jīngshén* 'spirit, mentality, or principle'	体育精神 *tǐyù jīngshén* 'athletic spirit'	athleticism
	乐观精神 *lèguān jīngshén* 'optimistic spirit'	optimism
	爱国精神 *àiguó jīngshén* 'patriotic spirit'	patriotism
	人道主义精神 *réndào zhǔyì jīngshén* 'humanitarian spirit'	humanitarianism
	敬业精神 *jìngyè jīngshén* 'respect-profession spirit'	professionalism
…思想 *sīxiǎng* 'thoughts, ideas, ideology'	儒家思想 *rújiā sīxiǎng*, 孔子思想 *kǒngzǐ sīxiǎng* 'Confucius thoughts'	Confucianism
	宪政思想 *xiànzhèng sīxiǎng* 'constitutional thoughts'	constitutionalism
	君权思想 *jūnquán sīxiǎng* 'kingship thoughts, ideology'	royalism
(b1) Systems or institutions established under certain doctrines or ideologies		
…制 *zhì* 'system, institution, or establishment'	宪政体制 *xiànzhèng tǐzhì* 'constitutional system'	constitutionalism
	工会制度 'union system'	unionism
	集权制度 *jíquán zhìdù* 'centralised-power system'	authoritarianism
	联邦制度 *liánbāng zhìdù* 'federal system'	federalism
	封建制度 *fēngjiàn zhìdù* 'feudal system'	feudalism
(c) Characteristics of language, pathology…		
…语 *yǔ* 'language, word/s, discourse'	双语 *shuāngyǔ* 'bilingual'	bilingualism
	警语 *jǐng yǔ* 'warning or pithy words'	aphorism
…症 *zhèng* 'disease'	自闭症 *zì bì zhèng* 'self-closed disease'	autism
	酒精中毒症 *jiǔjīng zhòngdú zhèng* 'alcohol poisoning disease'	alcoholism
	梦游症 *mèngyóu zhèng* 'dream-travel disease'	somnambulism

(2)
English   FAIR's Media Activism Resources
Chinese   众多媒体活动家资源
                                           Zhòngduō méitǐ huódòng jiā zīyuán
                                           'Zhongduo media activist resources'

(3)
English   Futurism rejected all traditions
Chinese   未来学家拒绝一切传统
                                           Wèilái xué jiā jùjué yīqiè chuántǒng
                                           'Futurology scholars rejected all traditions''

The second sense denotes certain doctrines, thoughts or ideologies. The expressions with 学 *xué* 'studies, doctrine' are commonly used, e.g., 汉学 *hànxué* 'Han or Chinese studies' corresponds to "Sinologism" in English. Also broadly under the second sense, we can find expressions denoting people's personality traits, mentalities, convictions and the principles they subscribe to, e.g., 爱国精神 *àiguó jīngshén* 'patriotic spirit, mind' in (4) corresponds to "patriotism" in English. In addition, the systems or institutions established under certain doctrines or ideologies can be put under the second sense. For example, in (5), "constitutionalism" in English is rendered into 宪政制度 *xiànzhèng zhìdù* 'constitutional system' in Chinese. More examples can be found in Table 5.

(4)
English   We've seen patriotism slide into jingoism
Chinese   我们看到爱国精神不知不觉地陷入侵略主义
            Wǒmen kàn dào àiguó jīngshén bùzhī bùjué de xiànrù qīnlüè zhǔyì
            'We see patriotic spirit unnoticeably falling into invad-ism'

(5)
English   Impeachment system is one of the important parts of western
            constitutionalism
Chinese   弹劾制度是西方宪政制度中的一项重要制度 。
            Tánhé zhìdù shì xīfāng xiànzhèng zhìdù zhōng de yī xiàng zhòngyào zhìdù
            'Impeachment system is an important system in the western
            constitutional system''

Finally, the third sense of –ism words denotes (the characteristics, style, or features of) language varieties, diseases, and pathological conditions and so on. The typical examples are "colloquialism", "autism", and "alcoholism", which are usually rendered into Chinese expressions with 语 *yǔ* 'language, discourse', 言 *yán* 'words, language' and 症 *zhèng* 'disease, syndrome of certain disease'. The examples include 口语 *kǒuyǔ* 'spoken language (style)' for "colloquialism", 自闭症 *zì bì zhèng* 'self-closed disease' for "autism", and 酒精中毒症 *jiǔjīng zhòngdú zhèng* 'alcohol poisoning disease' for "alcoholism".

From the Chinese lexical correspondents of –ism words, we can identify three major sense groups of –ism words. The "Education" corpus therefore exhibits much

potential for translators to tease out the senses of –ism words and harvest useful translation candidates. The skills and competence for conducting corpus-based queries as demonstrated in Sect. 4 are not technically advanced or highly sophisticated. As a result, we believe that self-directed inquisitive translators are able to reach discoveries in their interested areas insomuch as they are guided by their questions and are ready to make efforts for seeking out their answers in the corpus.

5 Discussion

We can now discuss the results of our study in relation to the research questions outlined in Sect. 1.

5.1 Traditional Tools Versus Parallel Corpora for Translators

From our examination of the conventional tools for translators to deal with –ism words, monolingual English dictionaries are valuable in laying out the major senses of the suffix, while reverse English-Chinese dictionaries are particularly useful for providing key translation candidates for a collection of –ism words presented altogether. These tools and resources assist translators to gain an overview of –ism words in terms of their senses and allow them to understand that a variety of Chinese translations tends to correspond to different –ism words. However, large parallel corpora have unique merits for translators to conduct (exhaustive) searches that yield quantitative results. The quantitative results lead to the emergence of the patterns and distributions of –ism words and their Chinese translations, while the parallel concordances allow –ism words and their translations to be systematically gathered, analysed, and examined in textual contexts. Large parallel corpora therefore extend a rather solid basis for teasing out the senses of –ism words, classifying the words into major types in relation to their Chinese translations, and harvesting translation candidates. More importantly, inquisitive translators can devise their own queries, in particular, with corpus query language (CQL) at the SkE interface and sort out their interested issues. The values of parallel corpus in these respects, especially the flexibility provided to translators for in-depth investigations, clearly exceed those that traditional resources can offer.

In terms of the denotations entailed by –ism words and their connotations invoked in context, dictionaries tend to give definitions with emphasis placed on the former, while large parallel corpora tend to capture the latter by both the –ism words used in various contexts and their Chinese correspondents. Translators certainly need to gain a sound understanding of the denotative meanings, but still, the connotative meanings brought to light by the Chinese correspondents of –ism

words are of practical use to them as well. The translation candidates strongly suggest that there are different alternatives that previous translators have utilised in the meaningful contexts in which they have worked. The alternatives bring to life authentic translation situations, to which translators can relate and be more resourceful, imaginative, and inspired.

5.2   Parallel Corpus for Translators: Drawbacks and Merits

There have been scholarly debates in the literature on the drawbacks and merits of parallel corpora for lexicography and for linguistic studies, e.g., an early debate arose between Teubert (1996) and Mauranen (2002), both leading researchers of the field. Parallel corpora tend to be much more modest in size compared to both monolingual corpora and comparable corpora, with a narrower scope of genres and text types available. Automatic alignment between the source and the target texts tends to involve a scope of inaccuracy, while manual checking of the aligned texts, especially at sentence level, is laborious and time-consuming. In addition, the issue of the direction of translation has long been raised, while a recent discussion can be found in Mikhailov (2021 in this volume). The direction of translation indeed should be specified at the stage of the construction of the corpus so that the translation equivalents can be specified accordingly, and both the frequency of occurrence and the distributions of the array of translation equivalents can make real sense.

Having said this, however, parallel corpora with unspecified direction of translation like the "UM-Education corpus" we explored in this study can still be valuable to translators. Most translators use corpora not for carrying out translation studies as translation scholars do—e.g., systematically describing translation equivalents or translation stimuli. Neither are they aiming at conducting contrastive linguistics analysis for generalisable results as attempted by linguists. Their primary need is to find their way to translation solutions, in particular, gathering useful translation candidates for their translation assignments. The parallel corpora that provide them with words used in the contexts similar to the ones they need to deal with can be potentially useful resources, with which they can select the most useful translation candidates. If the parallel corpora come with rich meta-data on the direction of translation and information on the institutional or the technical context in which the text is produced, that will certainly be context-rich information that translators can make sense of and derive insights from.

5.3   Technological Skills for Translators

Using monolingual corpora to generate wordlists with SkE (cf. Sect. 3.3) only requires basic skills. Querying parallel corpus at the interface of Parallel

Concordance is very attainable by translators with average computer skills. SkE-based access to the aligned bilingual texts is rather user-friendly, and the setting for various built-in query methods (esp. CQL) gives translators flexibility to devise and carry out queries that answer their questions surrounding the translation of –ism words. Self-motivated translators can acquire the skills and benefit a great deal by switching back and forth between different interfaces and resources at SkE for answering the questions they want to pursue. For example, by shifting between parallel concordance and the subcorpora that can be created from the parallel concordance lines, translators can produce keyword and key-term lists to spot recurrent Chinese expressions that correspond to the –ism words they seek to investigate (cf. Sect. 4.3). Translators' competence in making effective use of corpus tools and resources give them clear advantages in their professional practices, compared with those who rely only on traditional resources such as dictionaries and glossaries.

6   Conclusion

This study examines conventional and more recent corpus tools and resources for translators to gain practical knowledge of –ism words in English and retrieve their translation candidates in Chinese. We discovered that the traditional tools such as monolingual and bilingual dictionaries, especially reverse English-Chinese dictionaries, present translators the major senses of the suffix –ism and diversified translations of –ism words. In addition, large monolingual corpora of both English and Chinese, when accessed by translators with basic corpus skills, provide the lists of both –ism words and –主义 words in English and Chinese respectively, sortable by frequency to enable comparing and contrasting between the two. Moreover, large-scale parallel corpora exhibited unique value in terms of providing a large repertoire of translation candidates for a decent coverage of –ism words in a variety of contexts. The parallel concordance lines present a wealth of resources with which inquisitive translators can devise their queries to reach revealing results. The findings would allow translators to tease out their practical classification of different types of –ism words, appreciate the different senses denoted by the –ism words, and identify their typical translations into Chinese in certain situations or cotexts. We would argue that the ability to utilise corpus tools and resources, in particular, large-scale or subject-specific parallel corpora, constitutes an important and integral part of professional translators' competence in the IT age, and this competence needs to be fostered in the translator training curriculum. Subsequent studies can draw closer attention to (trainee) translators' actual use of the corpus-based tools and the ways they develop their skills and competence in this respect.

References

Baker, M. 1999. The role of corpora in investigating the linguistic behaviour of professional translators. *International Journal of Corpus Linguistics* 4: 281–298. https://doi.org/10.1075/ijcl.4.2.05bak.

Chen, X., V.X. Wang, and C.R. Huang. 2020. Sketching the English translations of Kumārajīva's The Diamond Sutra: A comparison of individual translators and translation teams. In *Proceedings of the 34th pacific asia conference on language, information and computation*, ed. L.M. Nguyen, C.M. Luong, and S. Song, 30–41.

Defrancq, B., and G. Rawoens. 2016. Assessing morphologically motivated transfer in parallel corpora. *Target-International Journal of Translation Studies* 28 (3): 372–398. https://doi.org/10.1075/target.28.3.02def.

Doval, I., and M.T.S. Nieto. 2019. Parallel corpora in focus: An account of current achievements and challenges. In *Parallel corpora for contrastive and translation studies: New resources and applications*, ed. I. Doval and M.T.S. Nieto, vol. 90, 1–15. John Benjamins.

Fillmore, C.J. 1992. Corpus Linguistics or Computer-aided armchair linguistics. In *Directions in corpus linguistics: Proceedings of Nobel Symposium*, ed. Jan Svartvik, vol. 82, 35–60. Stockholm: Walter de Gruyter.

Johansson, S. 2007. *Seeing through multilingual corpora: On the use of corpora in contrastive studies*. John Benjamins.

Kovář, V., V. Baisa, and M. Jakubíček. 2016. Sketch engine for bilingual lexicography. *International Journal of Lexicography*, 29 (3): 339–352. https://doi.org/10.1093/ijl/ecw029

Kubicka, E. 2019. So-called dictionary equivalents confronted with parallel corpora (and the consequences for bilingual lexicography). *Glottodidactica. An International Journal of Applied Linguistics*, 46 (2): 75–89.

Laviosa, S., and G. Falco. 2021. Using corpora in translation pedagogy. In *New perspectives on corpus translation studies*, ed. V.X. Wang, L. Lim, and D. Li, 3–27. https://doi.org/10.1007/978-981-16-4918-9_1.

Lefer, M.A. 2012. Word-formation in translated language: The impact of language-pair specific features and genre variation.*Across Languages and Cultures* 13 (2): 145–172. https://doi.org/10.1556/Acr.13.2012.2.2.

Lefer, M.A., and N. Grabar. 2015. Super-creative and over-bureaucratic : A cross-genre corpus-based study on the use and translation of evaluative prefixation in TED talks and EU parliamentary debates. *Across Languages and Cultures* 16 (2): 187–206. https://doi.org/10.1556/084.2015.16.2.3.

Lefever, Els, and Véronique Hoste. 2014. Parallel corpora make sense: Bypassing the knowledge acquisition bottleneck for Word Sense Disambiguation. *International Journal of Corpus Linguistics* 19 (3): 333–367.

Li, D. 2017. Translator style a corpus-assisted approach. In *Corpus methodologies explained an empirical approach to translation studies*, ed. M. Ji, M. Oakes, D. Li, and L. Hareide, 103–136. Routledge.

Li, L., S. Dong, and V.X. Wang. 2020a. Gaige and reform: A Chinese-English comparative keywords study. In *From minimal contrast to meaning construct: Corpus-based, near synonym driven approaches to Chinese lexical semantics*, ed. Q. Su, and W. Zhan, 321–332. Springer/Peking University Press.

Li, L., C.R. Huang, and V.X. Wang. 2020b. Lexical variations and human behavior changes: A corpus-assisted Investigation of GAMBLING and GAMING in the Past Centuries. *SAGE Open* 10 (3): 1–14. https://doi.org/10.1177/2158244020951272

Lim, L. 2019. Are terrorism and *kongbu zhuyi* translation equivalents? A corpus-based investigation of meaning, structure and alternative translations. In *Proceedings of the 33rd pacific asia conference on language, information and computation*, ed. R. Otoguro, M. Komachi, and T. Ohkuma, 516–523.

Lim, L. 2020. Interpreting training in China: Past, present and future, In *Key issues in translation studies in China: Reflections and new insights*, ed. L. Lim, and D. Li, 143–160. https://doi.org/10.1007/978-981-15-5865-8_7

Loveard, K. 1994. Sukarno: A fallen hero's legacy. *Asiaweek*, 36.

Mauranen, A. 2002. Will 'translationese' ruin a contrastive study? *Languages in Contrast* 2: 161–185.

Mauranen, A. 2004. Contrasting languages and varieties with translational corpora. *Languages in Contrast* 5 (1): 73–93.

Mikhailov, M. 2021. Mind the source data! Translation equivalents and translation stimuli from parallel corpora. In *New perspectives on corpus translation studies*, ed. V.X. Wang, L. Lim, and D. Li, 259–279. Springer. https://doi.org/10.1007/978-981-16-4918-9_10

Quah, C.K. 1999. Issues in the translation of English affixes into Malay. *Meta* 44 (4): 604–616. https://doi.org/10.7202/003881ar.

Sinclair, J. 2001. Data-derived multilingual lexicons. *International Journal of Corpus Linguistics*, 6 (Special Issue): 79–94.

Teubert, W. 1996. Comparable or parallel corpora? *International Journal of Lexicography* 9 (3): 238–264. https://doi.org/10.1093/ijl/9.3.238.

Teubert, W. 2001. Corpus linguistics and lexicography. *International Journal of Corpus Linguistics*, 6 (SI): 125–153. https://doi.org/10.1075/ijcl.6.si.11teu

Tian, L., D.F. Wong, L.S. Chao, P. Quaresma, F. Oliveira, and L. Yi. 2014. UM-Corpus: A large English-Chinese parallel corpus for statistical machine translation. In *Proceedings of the ninth international conference on language resources and evaluation (LREC'14)*, ed. N. Calzolari, K. Choukri, T. Declerck, H. Loftsson, B. Maegaard, J. Mariani, A. Moreno, J. Odijk, and S. Piperidis, 1837–1842. European Language Resources Association (ELRA).

Wang, Shan, and Huang Chu-Ren. 2017. Word sketch lexicography: New perspectives on lexicographic studies of Chinese near synonyms. *Lingua Sinica* 3 (1): 1–22.

Wang, V.X. 2018. Sketching a Chinese writer's vocabulary profile in English: The case of Ha Jin. In *Proceedings of the 32nd pacific asia conference on language, information and computation*, ed. S. Politzer-Ahles, Y.Y. Hsu, C.R. Huang, and Y. Yao, 722–728.

Wang, V.X. 2021. Examining prefix de- with its translations in Chinese: Making sense with parallel corpora. In *New perspectives on corpus translation studies*, ed. L. Lim, D. Li, and V.X. Wang, 299–318. Springer. https://doi.org/10.1007/978-981-16-4918-9_12.

Wang, V.X., and L. Lim. 2017. How do translators use web resources? Evidence from the performance of English-Chinese translators. In *Human issues in translation technology*, ed. D. Kelly, 63–79. Routledge.

Wang, V.X., X. Chen, S. Quan, and C.R. Huang. 2020. A parallel corpus-driven approach to bilingual oenology term banks: How culture differences influence wine tasting terms. In *Proceedings of the 34th pacific asia conference on language, information and computation*, ed. L.M. Nguyen, C.M. Luong, and S. Song, 318–328.

Xiao, R. 2010. How different is translated Chinese from native Chinese? A corpus-based study of translation universals. *International Journal of Corpus Linguistics* 15 (1): 5–35.

Zanettin, F. 2014. *Translation-driven corpora: Corpus resources for descriptive and applied translation studies*. Routledge.

References in Chinese

陆谷孙. 2007. 英汉大词典 (2nd ed.). 上海译文出版社.

李龙兴, & 王宪. 2021. 应急语言服务视角下的新冠肺炎医学英语专题术语表开发 (The development of COVID-19 word list from the perspective of emergency language services). 中国科技术语 (*China Terminology*), *23*(2), 32–41.

孙梅. 1993. 英汉倒排词典 (*A reverse English-Chinese dictionary*). 广西教育出版社.
荀恩东, 饶高琦, 肖晓悦, & 臧娇娇. 2016. 大数据背景下 BCC 语料库的研制. 语料库语言学 (*Corpus Linguistics*), *3*(1), 93–118.
郑宝璇. 2004. 传媒翻译. 香港城市大学出版社.

Lily Lim holds a PhD in Applied Linguistics (University of Queensland), a Master's Degree in Software Engineering (University of Macau), Certificate of Training Techniques (Escolas da Armada, Portugal), and Certificate of Chinese-Portuguese Conference Interpreting (Comissão Europeia). She has been both a practising interpreter and trainer for conference interpreters for nearly two decades. She is currently an Associate Professor and Coordinator of the Chinese-English Translation Program at the School of Languages and Translation, Macao Polytechnic Institute. Her recent research covers computer-assisted interpreter and translator training, and corpus-based language studies. She has published papers in *ReCALL*, *Babel*, and *The Interpreter and Translator Trainer*; book chapters with Rodopi, Springer, Cambridge Scholars Publishing, and Routledge; and a monograph with Bookman.