

基于语料库的华为年报翻译风格研究

毛 誉

上海海事大学外国语学院，上海

收稿日期：2024年2月25日；录用日期：2024年3月28日；发布日期：2024年4月8日

摘要

企业年报是一个企业对过去某个阶段内企业经营状况的总结和展示，对于吸引外部投资，增强市场信心，树立良好形象有着重要意义。过往有关翻译风格研究的关注点常常集中在文学作品和艺术作品，本文关注企业年报翻译过程中体现出来的翻译风格，以类比语料库的方式对中国企业年报英文翻译和美国本土企业的年报，利用WordSmith, CLAWS等语料分析和处理工具，在平均句长，可读性，词汇密度，类型符比以及名词化现象等方面将中美企业年报翻译风格进行对比分析。研究结果表明，华为企业年报英文翻译风格与美国本土企业接近，翻译水平较高。本研究有助于提高中国本土企业年报的英文翻译质量，更好地向世界展现中国企业的服务水平和企业面貌，从而推动中国企业的国际化进程。

关键词

语料库，对比，翻译风格，企业年报

A Corpus-Based Analysis of Translation Style of HUAWEI Annual Report

Yu Mao

College of Foreign Languages, Shanghai Maritime University, Shanghai

Received: Feb. 25th, 2024; accepted: Mar. 28th, 2024; published: Apr. 8th, 2024

Abstract

The annual report of a company is a summary and demonstration of the operation during a period of time, which is of great significance in attracting external investment, enhancing market confidence and establishing a good company image. While past studies on translation style often focus on literary and artistic works, this paper focuses on the translation style of annual reports. By comparing the English translations of Chinese company's annual reports with the one of American local enterprise and with the help of corpus, Word Smith, CLAWS and other corpus analyzing tools,

this paper analyzes Chinese and American annual reports and compares corpora from the lexical and sentence levels. The results found that the English translation of HUAWEI's annual report is close to that of the U.S. local enterprises in terms of translation style. This study helps to improve the quality of English translation of the annual reports of Chinese local enterprises, to better present the development level and corporate outlook of Chinese enterprises to the world, and thus to promote the internationalisation of Chinese enterprises.

Keywords

Corpus, Translation Style, Comparison, Annual Report

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

企业年报是上市公司与外部利益相关者进行沟通的重要文件。致股东的信位于年报的开头部分，以文字形式叙述了企业上一年度并展望下一年度的财务表现及运营情况[1]。企业年报作为一种特殊体裁，受到了来自学界的诸多关注。有研究表明，企业年报和致股东的信已成为企业树立良好形象的工具，管理层通过语言表达等方式影响读者对企业的感知和评价。华为作为中国走向世界市场上中的龙头企业，其年报的内容以及质量收到诸多关注，作为一家产品面向世界的企业，华为的年报不仅有中文版本，还有对应的英文版本。英文版本的年报有助于增强华为对外的宣传能力，吸引来自国外的投资者和关注。华为年报的英文翻译版本在走向世界的过程中有着重要的桥梁意义。华为年报英文版本的翻译质量的高低以及整体的翻译风格直接影响到英语受众读者的观感，地道准确的英文翻译有助于树立华为良好的企业形象和对外形象[2]。进行推动该企业的发展和进步。通过将中国本土企业的英文翻译年报和美国本土企业的年报进行对比，有助于了解中国企业机构的翻译风格和翻译能力，研究结果有助于中国本土企业改进英文年报或者英文材料的翻译水平，为中国企业走向世界提供参考和借鉴。

2. 相关文献综述

邵寒，王立非[3]在语料库的基础上对比研究中美企业财务语篇可读性分析。此外企业年报的研究多聚焦于企业年报中的语篇功能和态度资源[4]。很少有研究从年报的翻译风格和翻译效果出发探究中国企业年报的质量和水平[5]。本文将翻译风格的研究扩大到非文学作品范围，非文学作品在翻译中风格的研究还有待深入，非文学翻译受译者主观能动性影响较少，因此适合采用语料库以及定量与定性的研究方式。Baker [6]认为基于语料库的翻译研究方法可以提供客观有效的数据支持，从而帮助研究者们更好地识别译者风格，探索译者的社会文化处境。通过对语料库分析中美年报翻译风格差异，有助于确立华为的机构身份，厘清华为意图呈现出来的企业形象。

3. 研究设计

3.1. 研究问题

本文拟回答以下问题：1) 华为企业年报网站英译文是否具有明显的风格特征，以及具有哪些语言特

征？2) 同为英语文本，以母语为代表的微软企业年报和与中国机构译者译出的中国企业年报翻译文本在语言层面上有何异同？3) 通过语料对比，华为企业年报网站英译文是否存在能够改进的地方？

3.2. 数据来源

本文以中国企业华为 2018~2022 年间的英文年报网站英译文为研究对象，自建小型语料库，并以美国微软 2022 年企业年报作为参照语料库。语料文本均选自各网站中的年报系列内容，确保了信息种类的一致性。通过下载历年中美企业年报并进行人工筛查，剔除网站页面的图片，序号等非相关信息。所选文本数据采用同一规则的数据处理法，尽量保留译文原始数据。最终中国华为年报语料库由五个小型文本组成，分别是华为 2018、2019、2020、2021 和 2022 年英文年报。美国企业的语料库由微软公司 2022 年的企业年报组成。由于年报由众多不同的章节和汇报内容，为了确保语料组成的内容的一致性，主要选取年报中的致股东信以及关于可持续发展等介绍性内容较多的文字部分，舍弃牵扯到报表或图形等部分。通过这样的方式，最终形成了华为年报和微软年报的小型语料库。选取篇数众多的华为年报有助于减少年份对于华为年报英文翻译风格一致性的影响。

3.3. 研究工具及方法

在对所建平行语料库进行分词赋码时，本研究采用工具为英国兰卡斯特大学开发的 CLAWS 分词工具。在所建语料库基础上，通过语料库检索分析软件 Word Smith 4.0 和 AntConc 3.5.8 分别从词汇、句子和语篇三个层面对两个语料库中的众多文本进行定量和定性分析，并对相关差异的成因进行探讨。

4. 研究结果与讨论

4.1. 类符 - 形符比

在语料库翻译学中，类符/形符比(Type/Token Ratio, TTR)反映文本的词汇丰富度[7]。形符(Token)指的是语料库中的总词数，而类符(Type)指的是不重复出现的形符数。一般而言，TTR 越高，词汇丰富度越高，反之越低。然而，当文本篇幅较长时，通常使用标准类符/形符比(Standardized Type/Token Ratio, STTR)来衡量词汇丰富度。STTR 指的是文本中每千词的类符和形符比例，可以排除文章总字数对结果造成的影响，使统计结果更加准确，结果见表 1。

Table 1. Type/token ratio

表 1. 类符 - 形符比

	2018 HUAWEI	2019 HUAWEI	2020 HUAWEI	2021 HUAWEI	2022 HUAWEI	2022 Microsoft
形符	5178	5483	5344	5213	5448	5459
类符	1277	1275	1336	1231	1265	1426
类符/形符比	25	24	26	24	24	27
标准类符/形符比	39.6	41.7	40.7	43.18	41.46	44.36

在标准类符/形符比值方面，华为连续五年年报英译本的标准类符/型符比的比值均低于微软 2022 年年报，其中以 2022 年的微软年报为最高。这表明华为年报的译本词汇使用丰富度不如微软，总的来说，华为连续五年年报英译本与微软 2022 年年报语料库的类符/形符比值整体相差不大，比较具有原创英语的特征。

4.2. 平均词长

平均词长是指语料文本中全部单词的平均长度，它反映着词汇运用的复杂度[5]。一般来说，平均词长越长，代表长词的使用越频繁，因此文本越复杂。通过 WordSmith 4.0 的 wordlist 功能对比可以比较语篇的平均词长对比。见表 2。

Table 2. Average word length

表 2. 平均词长

2018 HUAWEI	2019 HUAWEI	2020 HUAWEI	2021 HUAWEI	2022 HUAWEI	2022 Microsoft
6	6	5	5	6	5

根据 WordSmith 4.0 的研究结果表明，2022 年微软年报的平均词长为 5，2020 年和 2021 年华为英文年报的平均词长为 5。在 2018 年、2019 年和 2022 年，华为英文年报的平均词长为 6 个单位。总体来看，华为 5 年间，平均词长与微软年报单词的平均长度相似，因此可以说双方在词汇运用的复杂度方面十分接近。

4.3. 平均句长

平均句长是指语料文本中全部句子的平均长度，它与句子运用的复杂程度紧密相关。通常情况下，语料平均句长越大，句式越复杂，因而可读性越低，对读者来说，阅读难度越高。巴特勒[8]曾根据句子长度将句子分为三类：短句、中句和长句，其中短句一般包含 1~9 个单词，中句包含 10~24 个单词，而长句则包含 25 个单词以上，平均句长见表 3。

Table 3. Average sentence length

表 3. 平均句长

	2018 HUAWEI	2019 HUAWEI	2020 HUAWEI	2021 HUAWEI	2022 HUAWEI	2022 Microsoft
句数	250	246	242	250	236	219
平均句长	20	22	21	20	22	24

表 3 统计数据显示，五个华为年报语篇的平均句长均少于微软 2022 年年报。微软 2022 年年报的平均句长为 24，这说明整体来说，美国微软企业年报整体上句子平均长度更长，根据巴特勒的定义，不论是华为语料库还是微软的语料库，其句子长度均处于中句的范围内，因此两者在平均句长方面相差不大。

4.4. 可读性分析

使用到 Laurence Anthony 开发的用于检测网页易读性的软件 Ant Word Profiler (以下简称 AWP)。AWP 自带三个对比文本语料词表(level lists)，前两个词表为英语常用的 2000 个词族(General Service List)，每个词表 1000 个词族，第三个词表为学术用语词表(Academic Word List)，包含 570 个词族。这三个语料词表的难度依次递增。通过 AWP 软件可以将研究语料导入，与上述三个词表进行词族覆盖，得到待研究文本中各级词汇所占比例。企业年报的体裁专业性较强，但选词主要仍为现代常用词汇。常用词的比例越高，文本越容易阅读。因此，通过语料库软件 Ant Word Profiler 统计年报中各级词汇中所占比例，可以反映出年报的易读程度，数据结果见表 4。

Table 4. Readability analysis**表 4. 可读性分析**

	一级词汇占比	二级词汇占比	三级词汇占比	总占比
2018 HUAWEI	65.2	7.8	14.4	87.4
2019 HUAWEI	64.6	5.7	14.8	85.1
2020 HUAWEI	67.6	6.0	13.2	86.8
2021 HUAWEI	69.3	5.8	11.8	86.9
2022 HUAWEI	62.9	5.7	14.9	83.5
2022 Microsoft	70.2	5.5	12.2	87.9

可以看到，微软 2022 的年报在一级词汇中所占比例较高，在二级词汇中占比较少。这表明微软的报告中所使用的常用词族更多。但是可以看到六份年报总占比相差不大，因此中美双方在可读性方面相差不大。

4.5. 名词化分析

名词化是指某个词类转换为名词的语法过程，一般通过添加词缀或词性转化等方式实现。王晋军[9]通过对五种不同类型文本的名词化现象，得出结论：名词化现象出现的频率与文本正式程度成正比。由于添加词缀构成名词的实例很多，具有典型性且易于统计[5]，因此本文利用 WordSmith 软件中的 Concord 功能，选取“-ment”“-tion”“-ity”“-ness”“-sion”五个常见的名词化后缀形式进行检索分析。检索结果见表 5。

Table 5. Nominalisation analysis**表 5. 名词化分析**

后缀	2018 HUAWEI	2019 HUAWEI	2020 HUAWEI	2021 HUAWEI	2022 HUAWEI	2022 Microsoft
-tions(s)	98	144	91	107	175	80
-ment(s)	84	54	47	46	62	35
-ity(ities)	71	43	71	44	56	90
-ness	59	23	26	21	15	19
-sion(s)	7	11	12	9	12	16
总计	319	275	247	227	320	240
占总词数百分比	6.3	5.1	4.7	4.5	6	4.5

可以看到，除了 2020 年的华为英语年报以上名词占总词数的百分比与微软 2022 年年报持平以外，其余年份的华为年报英语译文中具有名词化性质的名词占总词数的百分比均大于微软 2022 年年报。由于名词化现象的出现与文本的正式程度有极大关联，因此可以得出这样的结论即中国企业年报在用词上更倾向于使用词语的名词形式，而美国企业更倾向于使用动词等更为精简的表达方式，这与中美双方英文的行文习惯密不可分。

4.6. 词汇密度

词汇密度是译文风格分析的重要指标之一，它代表的是实义词在整个译文文本中所占的比重。词汇密度占比越高，表明译文文本中传达的信息量越大，同时词汇密度的高低也是判读文本难易度的指标之一。词汇密度反映文本的词汇丰富度及信息含量[10]。高词汇密度表明译文中实词比重大，信息承载量大，企业年报中词汇密度高代表着对原文翻译的忠实以及原文年报的内容专业程度较高。词汇密度较高可能代表着对原文含义的准确表达。英语中的实词通常包括主动词、名词、形容词、副词和数词等。通过 CLAWS 软件对五篇华为英语年报和一篇微软年报进行分词，在把经过分词处理的几篇语料导入 AntConc，通过 AntConc 的高级检索功能计算出语篇中实词所占比例，结果见表 6。

Table 6. Lexical density

表 6. 词汇密度

	2018 HUAWEI	2019 HUAWEI	2020 HUAWEI	2021 HUAWEI	2022 HUAWEI	2022 Microsoft
词汇密度	48%	44%	42%	40%	45%	43%

可以看到，微软 2022 年的年报中的词汇密度比 2018, 2019 和 2022 年的华为年报的词汇密度低。与此同时，微软高于 2020 和 2021 年华为英语年报中的词汇密度。整体来看，华为 5 年英文年报的词汇密度均超过 50%，且中美企业双方词汇密度相差范围不大，这表明华为企业英文年报中的英语文本传达的信息量较大，比较忠实于原文商业年报的写作风格。

5. 结语

通过建立语料库的方式，将华为历年的年报英文译文与美国本土企业微软的年报进行型符比，类符比，平均词长，平均句长，词汇密度，名词化以及可读性方面的比较。可以看到华为英文年报的翻译风格十分接近美国本土企业，英文年报翻译水平和质量较高。由于年报体裁的特殊性，对应译文的翻译风格受原文的影响很大。译者较少有自由发挥的空间。根据语料库工具 WordsSmith 4.0 以及 AntConc 工具的统计结果可以发现，在平均词长、平均句长、可读性、名词化和词汇密度等方面中国华为企业与美国微软企业相差不大，没有显著的区分。在标准类型符比方面微软略高于华为，这表明微软在年报中的用语更加丰富多样。在名词化程度方面，华为年报的名词化程度更高，文体更加正式。这与双方在年报中体现的语言风格有关，微软年报中的用语更加贴近生活，更加亲切自然。总体来说，华为年报中的英文译本在翻译风格方面与美国本土化企业相差不大，整体翻译质量较高。

参考文献

- [1] 钱嘉慧, 王立非. 英文企业年报语篇中态度评价的语料库考察及分析[J]. 商务外语研究, 2015(1): 1-7.
- [2] 付海燕. 中美博物馆机构身份的话语建构——基于语料库的批评话语分析研究[J]. 天津外国语大学学报, 2016, 23(6): 35-39+79.
- [3] 部寒, 王立非. 基于语料库的中美企业财务语篇可读性对比分析[J]. 解放军外国语学院学报, 2021, 44(1): 71-78+128
- [4] 王立非, 部寒. DICTION 文本分析工具及其在情感分析中的应用[J]. 外国语文, 2018, 34(1): 67-75.
- [5] 胡东磊. 基于类比语料库的中国国家博物馆网站文本译者风格研究[J]. 英语广场, 2023(14): 11-15.
- [6] Baker, P. (2004) Querying Keywords: Questions of Difference, Frequency and Sense in Keywords Analysis. *Journal of English Linguistics*, 32, 346-359. <https://doi.org/10.1177/0075424204269894>
- [7] 吕鹏飞, 陈道胜. 基于语料库的《论语》英译本翻译风格比较研究——以辜鸿铭和亚瑟·威利两译本为例[J]. 上

- 海翻译, 2021(3): 61-65.
- [8] Butler, C. (1985) Statistics in Linguistics. Basil Blackwell, Oxford.
- [9] 王晋军. 名词化在语篇类型中的体现[J]. 外语学刊, 2003(2): 74-78.
- [10] 祝朝伟, 李润丰. 基于语料库的庞德中国典籍英译译者风格探析[J]. 外语教学, 2023, 44(4): 75-82.