Lexical Transfer from Chinese to English in Terms of Collocation for Chinese University Students

Yanxin Chang

Suzhou New Oriental School, Suzhou Jiangsu

Email: changyanxin3@xdf.cn

Received: Nov. 3rd, 2019; accepted: Nov. 19th, 2019; published: Nov. 27th, 2019

Abstract

Collocation plays an important role in language learning for language learners. However, difficulties are usually found in using collocations correctly or like a native speaker for language learners, which further obstructs the learners in their language skills, such as writing, reading, listening and speaking. One of the reasons for the collocation errors is the lexical transfer from L1 to L2, which means the language learners transfer the elements of their native language onto that of the learned language in the collocations. In this research, 52 English writings from Chinese university students are collected to investigate the number of collocation errors and errors for each collocation type in a quantitative method. Qualitatively, collocation errors which are possibly judged to be due to the lexical transfer are discussed by examples. Furthermore, as a corpus-based study, the process of picking up and checking collocation errors are under the assistance of online corpora. The main finding points out that there might be a high error frequency for Chinese learners to use collocation, and lexical transfer accounts for the most reason. The research findings benefit both language teachers and language learners.

Keywords

Lexical Transfer, Collocation, Language Learning, Corpus

中国大学生写作中母语对英语词汇搭配的词汇 迁移现象

常言心

苏州新东方学校,江苏 苏州 Email: changyanxin3@xdf.cn

收稿日期: 2019年11月3日; 录用日期: 2019年11月19日; 发布日期: 2019年11月27日

文章引用: 常言心. 中国大学生写作中母语对英语词汇搭配的词汇迁移现象[J]. 国外英语考试教学与研究, 2019, 1(3): 121-129. DOI: 10.12677/oetpr.2019.13015

摘要

词汇搭配在语言学习中起着重要的作用。然而对于语言学习者来说很难正确使用或像母语者那样使用词汇搭配,这进一步阻碍了学习者听说读写等语言技能的提高。产生错误的其中一个原因是词汇迁移,这意味着语言学习者将母语的元素迁移到所学第二语言中的元素。本研究以52篇中国大学生英语作文为研究对象,首先采用定量的方法,分析了词汇搭配错误的频率和词汇搭配错误的种类。另外,定性地分析了由于词汇迁移所带来的词汇搭配错误。整个过程中,词汇搭配错误的提取和检查全部基于在线语料库的帮助下进行。中国学习者使用搭配错误的频率较高,词汇迁移是最主要的原因。研究结果表明,中国大学生词汇搭配错误频率较高,且大多受母语的词汇迁移所影响,为语言教学提供了新思路。

关键词

词汇迁移, 词汇搭配, 语言学习, 语料库

Copyright © 2019 by author(s) and Hans Publishers Inc.

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

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



Open Access

1. 简介

词汇对于学习语言来说是一项基本知识,在语言学习过程中占有主导地位。因此,词汇的搭配也被认为是非母语者在语言学习过程中的一个重要而又困难的问题。在 2010 年,Arnon 和 Snider 的研究表明 [1],语言学习者在使用词汇搭配时,是逐字考虑进行选择的,而不是把搭配中的所有元素作为一个整体来考虑。基于这种情况下,语言学习者的写作中经常会发现词汇搭配错误,从而影响着他们的写作能力。

当研究语言学习者词汇搭配错误的原因时,部分研究者提出了由于第一语言影响所学外语的词汇迁移的概念。Gas 提出[2],语言学习者经常会把母语的元素迁移到外语的使用。特别是对于母语为汉语的学习者而言,汉语词汇搭配中的词素顺序与英语词汇搭配的差异性,以及汉语多义词的翻译困难,增加了因母语词汇迁移而造成英语词汇搭配错误的可能性。因此,本研究旨在调查中国大学写作中英语词汇搭配错误的证据,并对词汇迁移对英语词汇搭配可能造成的影响进行调查。同时,会对不同类型的英语词汇搭配错误进行划分和比较。

2. 研究问题

本研究将解决一下三个研究问题,分别为:

- 1) 中国大学生英语作文中常见的词汇搭配错误类型有哪些?
- 2) 哪些词汇搭配错误可能与词汇迁移有关?
- 3) 对中国学生的英语词汇搭配学习和教学有什么启示?

3. 文献综述

3.1. 词汇迁移

词汇搭配学习是语言学习的重要组成部分。然而,由于词汇的复杂性,许多二语学习者不能完全掌握二语词汇,他们在二语口语和书面产出中经常出现母语迁移错误[3]。在 L2 词汇学习的一些初步研究

中,发现了词汇层面上母语对二语习得的强烈影响。正如 Yip 所说[4],从教学的角度来看,词汇迁移错误是非常具有启示性的,因为它们反映了母语知识会被迁移且母语迁移会如何影响二语习得。在词汇搭配方面,人们普遍认为它很难学习,尤其是对于非母语人士来说[5]。其中一种解释是,语言学习者必须了解整个短语在源语言中的意义以及目标语言中的典型用法,而不是简单地逐字翻译。对于语言学习的初学者来说,他们的母语是他们获取学习外语的唯一语言信息,因此他们必须经历一个认知过程,将他们所知道的知识转化为他们正在学习的知识[6]。研究发现,语言学习者在母语和二语之间建立了对等的关系,并在不同语言的两种搭配中迁移了他们认为共同的元素[7]。另外研究者们还指出了词汇搭配迁移的其他原因,如不同语言中语义韵律的差异,这在 Dorna 和 Emele 的研究中得到了支持[8]。为了解释语言学习者,尤其是外语学习者,在词汇或配的学习和教学中遇到的困难,原因之一可以归结为母语向二语的迁移。因此,为了提高词汇或交际教学水平,研究者们进行了大量有关母语迁移的研究。

3.2. 母语迁移

如果语言学习者被要求用外语进行交流,他们通常会表现出一种倾向,即把第一语言中的成分无意识地"迁移"到第二语言(即目标语言)的言语模式上。尽管母语迁移现象已被广泛接受,但不同领域的学者却从不同的角度来探讨这一问题,即使在今天,"迁移"一词本身的定义和恰当性也一直备受争议。例如,Kellerman 建议,这个术语只用于指从一种语言到另一种语言的特征整合过程,它和另一些术语如"干扰"和"跨语言影响"表述的是同一现象[9]。Grosjean 指出,术语"迁移"应该用来指母语对二语的永久影响,而"干扰"则用来描述二语中的语言生产中不时出现母语特征的情况[10]。大多数学者在心理语言学领域使用"干扰"一词,而二语习得领域的语言学家更喜欢使用"迁移"或"跨语言影响",因为他们认为前者在本质上包含负面含义[11]。一个被广泛引用的关于迁移的定义是Odlin在他的著作《语言迁移》中所述[12]:

迁移是由于目标语言和任何其他语言之间的相似性和差异而产生的影响,这些语言以前(也许是不完全)习得过。 具体来说,本研究中的迁移指的是学习者母语(汉语)对二语(英语)词汇学习的影响。

虽然在一般的迁移现象中存在着很大程度的一致性,但不同的学者对迁移的各个方面给予了具体的重视。例如,McWhinney 指出,在词汇、概念化、语用学、语音和音韵学等层面上,迁移比其他层次更容易发现,综合性和生产性的过程都表现出迁移的证据[13]。当谈到迁移发生的阶段时,Pienemann 提出了一种发展观点,即学习者在获得足够的二语知识之前不会从母语迁移元素[14]。换言之,中级和高级学习者比初级学习者更容易进行母语迁移,后者仍处于学习第二语言的初级阶段,缺乏第二语言的熟练程度。相比之下,Lefebvre、White 和 Jourdan 认为,一般情况下,母语在二语学习的早期阶段起着至关重要的作用,而在后期阶段,如果缺乏积极的证据,它仍然会影响二语习得[15]。此外,研究发现,二语学习者或双语者之间的迁移方向不仅限于从一语到二语或从二语到一语[16],多语种之间的情况更为复杂,因为它们不仅可能将一语元素迁移到二语,还可能从一语迁移到三语,从二语迁移到三语[17]。因此,母语极有可能被迁移到目的语,从而影响学生的词汇搭配学习和语言学习。

4. 数据收集与语料库软件介绍

4.1. 学习者语料库

本研究为试图找出英语词汇搭配模式,并找出可能由于母语向二语负迁移而导致的词汇搭配错误,在 SketchEngine 建立了一个学习者语料库(图 1),包括至少 52 篇由中国大学生撰写的学术英语论文,其中大部分学生就读于英语相关的专业。此外,为了保证数据的充分性,要求每篇论文的字书均在 2000 字以上,总字数为 146,591,确保可以找到大量的词汇搭配以供分析。

Counts		General info		Lexicon sizes		Tags legend		Lempos suffixes	
Tokens	184,821	Language	English	word	14,356	adjective	J.*	adjective	-j
Words	146,591	Encoding	UTF-8	tag	62	adverb	RB.?	adverb	-a
Sentences	7,634	Compiled	08/31/2017 10:30:44	lempos	10,293	conjunction	CC	conjunction	-c
Documents	52	Tagset	Description	lemma	9,482	determiner	DT	noun	-n
		Word sketch grammar	Definition	lempos_lc @	9,479	noun	N.*	numeral	-m
				lemma_lc @	8,449	noun singular	NN	preposition	-i
				lc @	12,518	noun plural	NNS	pronoun	-d
						numeral	CD	verb	-v
						narticle	DD		
Structures s 0	and attri	butes							
file 52		*							

Figure 1. A summary of the learner corpus built in the SketchEngine 图 1. SketchEngine 中构建的学习者语料库摘要

4.2. 参考语料库

g O

本研究首先通过笔者手动筛选高亮学习这语料库中的任何可能的词汇搭配错误,随后辅助Grammarly 把学习者语料库进行再一次筛选,两次筛选尽可能标记出所有可能的词汇搭配错误。然后,所标记的词汇搭配到底正确与否,需要使用参考语料库进行进一步的检验。参考语料库为 concordance English v.8.3 (http://www.lextutor.ca)。concordance English 拥有大量的英语语言样本语料库,因此可以通过它来识别英语中地道的词汇搭配,如为了检验"make efforts"这个搭配是否正确,我们可以通过concordance English v.8.3 对词汇搭配正确与否检验(见图 2)。如图 4 所示,可以将关键字"make"设置为一个家族词,以便可以搜索其所有不同形式的家族词(例如"making""made"和"makes"),以获得搜索结果。此外,将关联词"effort"设置为在任意一方 3 个词内搜索与关键字相伴的,从而扩展搜索区域,以获得更准确的检验。结果根据图 3 中的搜索结果,可以发现尽管它在不同的上下文中以不同的形式出现,词汇搭配"make efforts"在语料库中的使用频率很高(百万分之 25),因此这种词汇搭配对于学习者来说是正确的,是地道的或标准的英语。

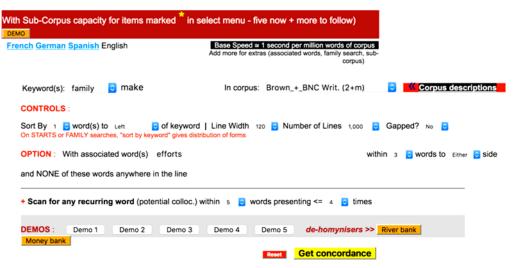


Figure 2. Making concordance in The Corpus Concordance English v.8.3
图 2. 在语料库 Concordance English v.8.3 中检验词汇搭配正确性, "make efforts"为例

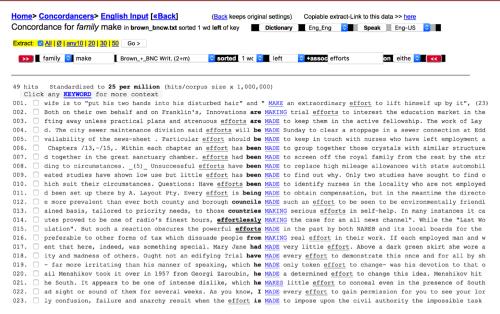


Figure 3. Concordance results on 'make efforts' in The Corpus Concordance English v.8.3 图 3. 在语料库 Concordance English v.8.3 中 "make efforts" 的检验结果

另一个使用的检验程序是 SkerchEnigne,这是前面提到的创建学习者语料库的数据库。然而,在这里,英国国家语料库(BNC)被用作参考语料库,以确定某些英语搭配在中国学习者写作中的有效性和正确性。在 SketchEngine 中,有两种方法可以用来检查词汇搭配的正确性。仍以词汇搭配"make efforts"为例,如图 4 显示,可通过在"make"和"efforts"之间加星号,来搜索这个词汇搭配中间存在其他词汇的情况。另一方面,SketchEngine 还可以设置为搜索"make"和"efforts"之间的任意数量的单词。将主要引导词"make"两边的单词距离设置为 5,如图 5 所示,搜索结果可以显示出更多与搭配"make efforts"相关的结果(例如,"efforts should be made"和"make an ever greater efforts"),这有助于进一步研究词汇搭配在参考语料库中的使用频率,判断其是否正确。

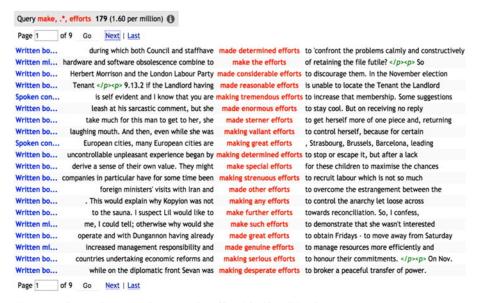


Figure 4. Concordance results on 'make efforts' in SketchEngine 1 图 4. "make efforts" 在 SketchEngine 语料库中的搜索结果 1

DOI: 10.12677/oetpr.2019.13015

mple query:	make	Make Concordance			
	Query types Context Text types ②				
Query type	Query type simple lemma phrase word character CQL				
Lemma:		PoS: unspecified 😊			
Phrase:					
Word form:		PoS: unspecified 💿 🗆 match case			
Character:					
Character: CQL:		Default attribute: c = word (lowercase)			
	Tagset summary CQL builder	Default attribute: c = word (lowercase)			
CQL:	Tagset summary CQL builder	Default attribute: Ic = word (lowercase)			
		Default attribute: Ic = word (lowercase)			
CQL:	ter				
CQL:	Tagset summary CQL builder	Default attribute: Ic = word (lowercase)			

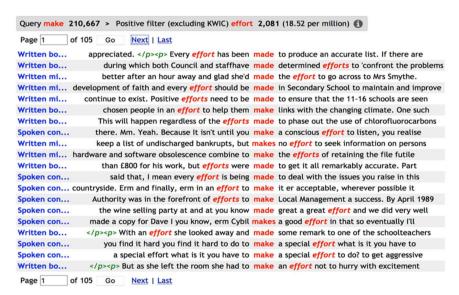


Figure 5. Concordance results on "make" and lemma "efforts" in SketchEngine 2 图 5. "make efforts" 在 SketchEngine 语料库中的搜索结果 2

通过这两个参考语料库,对学习者语料库中所有可能出现的词汇搭配错误进行了检查,并对搭配错误进行了识别。词汇搭配错误使用的原因可能是由于掌握不足、母语迁移或其他原因,因此研究者将进一步探讨那些被认为可能是母语迁移组引起的词汇搭配错误,并给出它们的母语等价物和判断的原因。此外,考虑到共有 6 种词汇搭配类型,"错误库"中的所有词汇搭配错误都将标记其类型,并在 Excel 电子表格中单独分组。在识别出母语迁移引起的词汇搭配错误后,将其输入 Excel 电子表格作为一个"库",并将其分为不同的类型,运用这种定性研究方法可以解决第一个研究问题。随后,采用定量研究方法进行计数和计算。如前所述,所有可能由母语迁移引起的词汇搭配错误都被标记并归类为不同类型,因此不同类型的词汇搭配错误将被分别计算。通过比较不同词汇搭配错误类型的数量,可以回答第二个研究问题,即中国学习者通常会犯哪些类型的搭配错误。此外,还将计算每种类型的错误频率,并提供如下公式:

错误频率 = 错误数量 ×1000

通过比较,得出中国学习者最常误用的词汇搭配类型,对中国学习者的词汇搭配教学具有一定的指导意义。

5. 调查结果与分析

表 1 是对所收集的论文数量、总词数、可能的词汇搭配错误以及被认为是由于词汇迁移所导致的词汇搭配错误作出的简要总结。通过参考语料库(语料库 Concordance English v.8.3 和 SketchEngine)检查所有潜在的词汇搭配错误后,学习者语料库中共有 102 个搭配很可能是错误,如表 1 所示。其中有 48 个词汇搭配被认为是由于词汇迁移所导致的,部分示例可见图 6,占所有潜在词汇搭配错误的 47%。用公式计算词汇迁移错误的错误频率,其结果为 0.327。

Table 1. A summary of the main research findings

表 1. 主要研究发现概括

研究结	
总篇数	52
总字数	146,591
潜在的词汇搭配错误	102
由词汇迁移导致的词汇搭配错误	48

Expression as found	Chinese equivale	Chinese equivalent in pinyi	ral translation of Chinese equ	Correct English	Type ▼
dogs shout	狗大叫	gou da jiao	dogs bark	dogs bark	N + V
carry out	带出去	dai chu qu	take out	take out	V + Adv
come out (an idea)	想出 (一个想法)	xiang chu (yi ge xiang fa)	come up with (an idea)	come up with (an idea)	V + Adv
humour sense	幽默感	you mo gan	a sense of humour	a sense of humour	N + N
expesive price	昂贵的价格	ang gui de jia ge	high price	high price	Adj + N
old customers	老顾客	lao gu ke	regular customers	regular customers	Adj + N
learn knowledge	学到知识	xue dao zhi shi	acquire knowledge	acquire knowledge	V + N
raise a requst	提出要求	ti chu yao qiu	make a request	make a request	V + N
satisfy the requirements	满足要求	man zu yao qiu	meet/fulfill one's requirements	fulfill the requirements	V + N
quicken development	加速发展	jia su fa zhan	accelerate the development	accelerate development	V + N

Figure 6. Extracted examples on collocation potentially to be due to lexical transfer 图 6. 由于词汇迁移导致的错误搭配示例

在 Excel 电子表格中对所有词汇搭配错误进行分组和标注的过程中,分别计算了每种错误类型的数量和错误频率,如下面图 7 和图 8 所示。此外,还计算了每种错误词汇搭配类型的比例,如图 9 所示,以便进行比较,检查哪种类型可能被中国学习者误用。在 48 个搭配迁移错误中,动词 + 名词类型的错误出现频率最高,其中 25 个词汇搭配错误是由于汉语到英语的负迁移,占所有词汇搭配迁移错误的一半以上(53%)接下来,形容词 + 名词类型占据了 29%。但是,副词 + 形容词类型在学习者语料库中没有得到数据。

6. 结论

综上所述,本研究首先证实了词汇搭配学习对语言学习者的意义。考虑到学习者语料库中的所发现的词汇搭配错误,这意味着语言学习者在撰写英语论文时极易出现词汇搭配错误。此外,研究表明几乎一半的词汇搭配错误是由于从母语到第二语言的词汇迁移,因此母语在语言学习或词汇搭配学习中的作用应受到更多的重视。此外,研究结果还强调,动词相关搭配最常被误用。

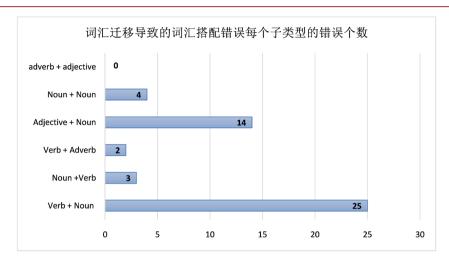


Figure 7. The number of collocation transfer errors for each subtype 图 7. 词汇迁移导致的词汇搭配错误每个子类型的错误个数

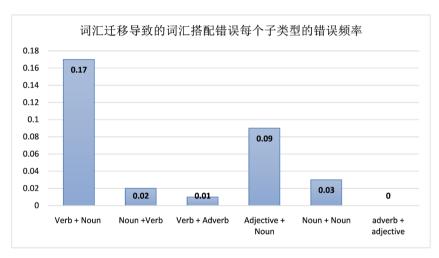


Figure 8. The error frequency of collocation transfer errors for each subtype 图 8. 词汇迁移导致的词汇搭配错误每个子类型的错误频率

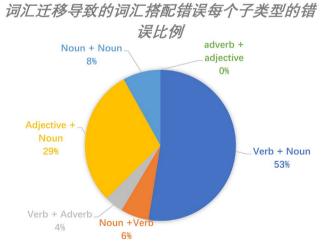


Figure 9. The ratio of collocation transfer errors for each subtype 图 9. 词汇迁移导致的词汇搭配错误每个子类型的错误比例

因此,研究表明了学习词汇搭配对于语言学习者的重要性。可以推断,如果语言学习者在校对时更多地关注词汇搭配,那么在学习者语料库中发现的一些词汇搭配错误是可以避免的。总体来说,母语为汉语的语言学习者应该注意多义词,因为它是词汇搭配中错误出现的主要原因。可以利用在线参考语料库在使用词汇搭配时进行检查,而不是直接逐字将词汇搭配从汉语翻译成英语。从语言教师的角度出发,建议语言教师在教学前特别设计"以词汇搭配为中心"的活动,使学习者热身,为后续的学习做好准备。这些研究结果都反映了词汇迁移作用对语言学习者的重要性。教师只有了解学习者产生词汇搭配错误的原因,才能给出解决方法。在语言教学之前,教学大纲和教材可能需要根据学习者的母语和文化进行调整。

参考文献

- [1] Arnon, I. and Snider, N. (2010) More than Words: Frequency Effects for Multi-Word Phrases. *Journal of Memory and Language*, **62**, 67-82. https://doi.org/10.1016/j.jml.2009.09.005
- [2] Gas, S. (1979) Language Transfer and Universal Grammatical Relations. Language Learning, 29, 327-344. https://doi.org/10.1111/j.1467-1770.1979.tb01073.x
- [3] Ellis, N.C. (2008) The Periphery and the Heart of Language. In: *Phraseology: An Interdisciplinary Perspective*, John Benjamins Publishing, Amsterdam, 1-13. https://doi.org/10.1075/z.138.02ell
- [4] Yip, V. (1995) Interlanguage and Learnability: From Chinese to English. Volume 11, John Benjamins Publishing, Amsterdam. https://doi.org/10.1075/lald.11
- [5] Leed, R.L. and Nakhimovsky, A.D. (1979) Lexical Functions and Language Learning. *Slavic and East European Journal*, 23, 104-113. https://doi.org/10.2307/307804
- [6] Llach, M.P.A. (2010) An Overview of Variables Affecting Lexical Transfer in Writing: A Review Study. *International Journal of Linguistics*, **2**, 1-17. https://doi.org/10.5296/ijl.v2i1.445
- [7] Ringbom, H. (2006) The Importance of Different Types of Similarity in Transfer Studies. In: Arabski, J., Ed., Cross-Linguistic Influences in the Second Language Lexicon, Multilingual Matters, Clevedon, 36-45. https://doi.org/10.21832/9781853598579-007
- [8] Dorna, M. and Emele, M.C. (1996) Semantic-Based Transfer. *Proceedings of the 16th conference on Computational Linguistics*, 1, 316-321. https://doi.org/10.3115/992628.992684
- [9] Kellerman, E. (1987) Aspects of Transferability in Second Language Acquisition.
- [10] Grosjean, F. (2001) The Bilingual's Language Modes. One Mind, Two Language.
- [11] Jarvis, S. and Pavlenko, A. (2008) Crosslinguistic Influence in Language and Cognition. Routledge, London. https://doi.org/10.4324/9780203935927
- [12] Odlin, T. (1989) Language Transfer: Cross-Linguistic Influence in Language Learning. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9781139524537
- [13] MacWhinney, B. (2008) How Mental Models Encode Embodied Linguistic Perspectives. Department of Psychology, 172.
- [14] Pienemann, M. (2005) Cross-Linguistic Aspects of Processability Theory. Volume 30, John Benjamins Publishing, Amsterdam. https://doi.org/10.1075/sibil.30
- [15] Lefebvre, C., White, L. and Jourdan, C. (2006) L2 Acquisition and Creole Genesis: Dialogues. Volume 42, John Benjamins Publishing, Amsterdam. https://doi.org/10.1075/lald.42
- [16] Grosjean, F. and Py, B. (1991) La restructuration d'une première langue: L'intégration de variantes de contact dans la compétence de migrants bilingues. *La linguistique*, **27**, 35-60.
- [17] Cenoz, J., Hufeisen, B. and Jessner, U. (2001) Cross-Linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives. Volume 31, Multilingual Matters, De Gruyter. https://doi.org/10.21832/9781853595509