

Loughborough University
Institutional Repository

*The use and prescription of
epicene pronouns: a
corpus-based approach to
generic he and singular they
in British English*

This item was submitted to Loughborough University's Institutional Repository by the/an author.

Additional Information:

- A Doctoral Thesis. Submitted in partial fulfillment of the requirements for the award of Doctor of Philosophy of Loughborough University.

Metadata Record: <https://dspace.lboro.ac.uk/2134/9118>

Publisher: © by Laura Louise Paterson

Please cite the published version.

This item was submitted to Loughborough's Institutional Repository (<https://dspace.lboro.ac.uk/>) by the author and is made available under the following Creative Commons Licence conditions.



CC creative commons
COMMONS DEED

Attribution-NonCommercial-NoDerivs 2.5

You are free:

- to copy, distribute, display, and perform the work

Under the following conditions:

 **Attribution.** You must attribute the work in the manner specified by the author or licensor.

 **Noncommercial.** You may not use this work for commercial purposes.

 **No Derivative Works.** You may not alter, transform, or build upon this work.

- For any reuse or distribution, you must make clear to others the license terms of this work.
- Any of these conditions can be waived if you get permission from the copyright holder.

Your fair use and other rights are in no way affected by the above.

This is a human-readable summary of the [Legal Code \(the full license\)](#).

[Disclaimer](#) 

For the full text of this licence, please go to:
<http://creativecommons.org/licenses/by-nc-nd/2.5/>

Thesis Access Form

Copy No.....Location.....

Author.....

Title.....

Status of access OPEN / RESTRICTED / CONFIDENTIAL

Moratorium Period:.....years, ending...../.....200.....

Conditions of access approved by (CAPITALS):.....

Supervisor (Signature).....

School of.....

Author's Declaration: *I agree the following conditions:*

Open access work shall be made available (in the University and externally) and reproduced as necessary at the discretion of the University Librarian or Dean of School. It may also be digitised by the British Library and made freely available on the Internet to registered users of the EThOS service subject to the EThOS supply agreements.

The statement itself shall apply to ALL copies including electronic copies:

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

Restricted/confidential work: All access and any photocopying shall be strictly subject to written permission from the University Dean of School and any external sponsor, if any.

Author's signature.....Date.....

users declaration: for signature during any Moratorium period (Not Open work): <i>I undertake to uphold the above conditions:</i>			
Date	Name (CAPITALS)	Signature	Address

The use and prescription of epicene pronouns: A corpus-based
approach to generic *he* and singular *they* in British English

by

Laura Louise Paterson

Doctoral Thesis

Submitted in partial fulfilment of the requirements
for the award of
Doctor of Philosophy of Loughborough University

10th October 2011

© by Laura Louise Paterson (2011)

CERTIFICATE OF ORIGINALITY

This is to certify that I am responsible for the work submitted in this thesis, that the original work is my own except as specified in acknowledgments or in footnotes, and that neither the thesis nor the original work contained therein has been submitted to this or any other institution for a degree.

..... (Signed)

..... (Date)

ACKNOWLEDGEMENTS

This research was funded by a three year studentship from the Department of English and Drama at Loughborough University. My thanks go to Dr Chris Christie and Dr Arianna Maiorani for providing expert guidance and thorough feedback, and for being wonderful friends. I would also like to thank both the academic staff and the support staff in the Department of English and Drama. In addition, I wish to express my gratitude to Dr Paul Baker and Dr Andrew Hardie for allowing me access to the BE06 corpus held at Lancaster University for the purpose of this research. To Geordie Mackintosh, Thomas Alexander Paterson, Katy Lever and Katie Normanton-Bell I appreciate the constant support (and the reading of drafts), and to Eileen and Thomas Paterson I am extremely grateful for the opportunities your support has afforded me.

ABSTRACT

In English the personal pronouns are morphologically marked for grammatical number, whilst the third-person singular pronouns are also obligatorily marked for gender. As a result, the use of any singular animate antecedent coindexed with a third-person pronoun forces a choice between *he* and *she*, whether or not the biological sex of the intended referent is known. This forced choice of gender, and the corresponding lack of a gender-neutral third-person singular pronoun where gender is not formally marked, is the primary focus of this thesis. I compare and contrast the use of the two main candidates for epicene status, singular *they* and generic *he*, which are found consistently opposed in the wider literature.

Using corpus-based methods I analyse current epicene usage in written British English, and investigate which epicene pronouns are given to language-acquiring children in their L1 input. I also consider current prescriptions on epicene usage in grammar texts published post-2000 and investigate whether there is any evidence that language-external factors impact upon epicene choice. The synthesis of my findings with the wider literature on epicene pronouns leads me to the conclusion that, despite the restrictions imposed on the written pronoun paradigm evident in grammatical prescriptivism, singular *they* is the epicene pronoun of British English.

Keywords: Epicene pronouns, singular *they*, generic *he*, corpus linguistics, pronoun acquisition, traditional grammatical prescriptivism, second-wave feminism, non-sexist language, language reform.

Word count: 79, 768 (excluding tables and figures)

TABLE OF CONTENTS

INTRODUCTION	1
CHAPTER ONE: EXPLORING EPICENE PRONOUNS	6
1.1: CHARACTERISTICS OF THE PERSONAL PRONOUNS	7
1.1.1: THE PARADIGM AS A CLOSED-CLASS	8
1.1.2: PREVIOUS CHANGES: TRENDS THROUGHOUT HISTORY	12
1.1.3: ACQUISITION AND PRONOMINAL CHANGE	21
1.2: THE PSYCHOLOGICAL REALITY OF EPICENE PRONOUNS	25
1.2.1: THE SEMANTIC SALIENCE OF GENDER MARKING	26
1.2.2: WHAT EPICENES MEAN	29
1.3: DIFFERENCES BETWEEN GENERIC HE AND SINGULAR THEY	40
1.3.1: AGREEMENT WITH DIFFERENT TYPES OF ANTECEDENT	41
1.3.2: NOTIONAL NUMBER AND OTHER EXPLANATIONS	48
1.3.3: HOMONYMY THEORY	53
CONCLUSIONS	58
CHAPTER TWO: SOCIAL MOVEMENTS AND EPICENES	62
2.1: CONTEXTUALISING EPICENE PRESCRIPTIONS	63
2.1.1: INITIATING FORMAL PRESCRIPTIVE PRACTICE	64
2.1.2: TRACING EPICENE HISTORY	70
2.2: THE EFFECTS OF SEXUAL POLITICS	78
2.2.1: FEMINISM AND LANGUAGE	79
2.2.2: FEMINIST LANGUAGE REFORM AND EPICENES	84
2.3: EPICENES IN EDUCATION	90
2.3.1: "CORRECT" LANGUAGE IN EDUCATIONAL HANDBOOKS	91
2.3.2: MODERN APPROACHES TO EPICENES	96
CONCLUSIONS	103
CHAPTER 3: INVESTIGATING EPICENES	107
3.1: CORPUS LINGUISTICS: A PROBLEMATIC METHODOLOGY?	108
3.1.1: CORPUS LINGUISTICS AND GRAMMATICAL FEATURES	109
3.1.2: INTRODUCING THE CORPORA	113
3.2: WHAT THE RESEARCH QUESTIONS COVER	119
3.2.1: ANALYSING EPICENE USAGE AND PRESCRIPTIONS	120
3.2.2: INVESTIGATING LANGUAGE-EXTERNAL INFLUENCES	123
3.2.3: EPICENES AND PARADIGM ACQUISITION	125
CONCLUSIONS	127
CHAPTER 4: EPICENES IN THE 21ST CENTURY	130
4.1: STANDARD ENGLISH CORPORA	131
4.1.1: OCCURRENCES OF SINGULAR <i>THEY</i>	131
4.1.2: OCCURRENCES OF GENERIC <i>HE</i>	137

4.2: COMPARING <i>THEY</i> WITH <i>HE</i> IN BE06	143
4.2.1: QUANTITATIVE ANALYSIS	143
4.2.2: QUALITATIVE ANTECEDENT ANALYSIS	149
4.2.3: GENDER STEREOTYPED ANTECEDENTS	156
CONCLUSIONS	164
CHAPTER 5: STANDARD ENGLISH AND PRESCRIPTIONS	166
5.1: ANALYSING PRESCRIPTIONS	167
5.1.1: ENDORSING EPICENE CANDIDATES	168
5.1.2: ANTECEDENT ANALYSIS	174
5.1.3: DEALING WITH STEREOTYPES	179
5.2: SOCIAL INTERFERENCE?	182
5.2.1: CORPUS COMPARATIVE STATISTICAL KEYWORDS	182
5.2.2: KEYWORDS IN THE CORPUS	186
CONCLUSIONS	194
CHAPTER 6: ACQUIRING EPICENES	197
6.1 THE CHILDES CORPUS	198
6.1.1: SELECTING THE DATASET	198
6.1.2: SAMPLING THE DATASET	202
6.2: PRONOUNS PRESENT IN ACQUISITION	206
6.2.1: SINGULAR AND PLURAL <i>YOU</i>	208
6.2.2: SINGULAR AND PLURAL <i>THEY</i>	209
6.2.3: MASCULINE AND GENERIC <i>HE</i>	211
6.3: ANTECEDENTS IN ACQUISITION	214
6.3.1: SYNTACTIC CLASSIFICATION OF ANTECEDENTS	215
6.3.2: PROBLEMS WITH PERSONIFICATION	217
CONCLUSIONS	220
CHAPTER 7: FURTHERING EPICENE RESEARCH	223
7.1: THE ENGLISH EPICENE	226
7.2: EPICENES AND LANGUAGE-EXTERNAL FORCES	232
7.3: THE ACQUISITION OF THE EPICENE PRONOUN	238
7.4: EXPLAINING EPICENE REFERENCE: THE HOMONYMY OF <i>THEY</i>	243
CONCLUSIONS	247
CONCLUSION	250
APPENDIX A: THE GRAMMAR CORPUS	254
REFERENCES	256

LIST OF TABLES

TABLE 1.1: THE OLD ENGLISH PERSONAL PRONOUN PARADIGM (500-1100)	13
TABLE 1.2: THE MIDDLE ENGLISH PERSONAL PRONOUN PARADIGM (1100-1500)	14
TABLE 1.3: THE EARLY MODERN ENGLISH PERSONAL PRONOUN PARADIGM (1500-1800)	16
TABLE 1.4: THE CURRENT STANDARD ENGLISH PERSONAL PRONOUN PARADIGM (1800-)	19
TABLE 4.1: DISTRIBUTION OF TOKENS OF SINGULAR <i>THEY</i> IN PRONOUNS_PRESS	132
TABLE 4.2: DISTRIBUTION OF TOKENS OF SINGULAR <i>THEY</i> IN PRONOUNS_GENERAL_PROSE	132
TABLE 4.3: DISTRIBUTION OF TOKENS OF SINGULAR <i>THEY</i> IN BOTH SUB-CORPORA	133
TABLE 4.4: DISTRIBUTION OF ANTECEDENTS OF SINGULAR <i>THEY</i> AFTER THE CONDENSATION OF DUPLICATES	134
TABLE 4.5: PERCENTAGE DISTRIBUTION OF ANTECEDENTS OF SINGULAR <i>THEY</i> BEFORE THE CONDENSATION OF DUPLICATES	135
TABLE 4.6: PERCENTAGE DISTRIBUTION OF ANTECEDENTS OF SINGULAR <i>THEY</i> AFTER THE CONDENSATION OF DUPLICATES	135
TABLE 4.7: DISTRIBUTION OF TOKENS OF GENERIC <i>HE</i> IN PRONOUNS_PRESS	137
TABLE 4.8: DISTRIBUTION OF TOKENS OF GENERIC <i>HE</i> IN PRONOUNS_GENERAL_PROSE	138
TABLE 4.9: DISTRIBUTION OF TOKENS OF GENERIC <i>HE</i> IN BOTH SUB-CORPORA	140
TABLE 4.10: DISTRIBUTION OF ANTECEDENTS OF GENERIC <i>HE</i> AFTER THE CONDENSATION OF DUPLICATES	140
TABLE 4.11: PERCENTAGE DISTRIBUTION OF ANTECEDENTS OF GENERIC <i>HE</i> BEFORE THE CONDENSATION OF DUPLICATES	141
TABLE 4.12: PERCENTAGE DISTRIBUTION OF ANTECEDENTS OF GENERIC <i>HE</i> AFTER THE CONDENSATION OF DUPLICATES	142
TABLE 4.13: DISTRIBUTION OF TOKENS OF SINGULAR <i>THEY</i> AND GENERIC <i>HE</i> IN PRONOUNS_PRESS AND PRONOUNS_GENERAL_PROSE	144
TABLE 4.14: DISTRIBUTION OF TOKENS OF SINGULAR <i>THEY</i> AND GENERIC <i>HE</i> IN PRONOUNS_PRESS AND PRONOUNS_GENERAL_PROSE AFTER CONDENSATION OF ANTECEDENTS	146
TABLE 4.15: ANTECEDENTS OF GENERIC <i>HE</i>	149
TABLE 4.16: ANTECEDENTS OF SINGULAR <i>THEY</i>	152
TABLE 4.17: STEREOTYPES WITH GENERIC <i>HE</i>	159
TABLE 4.18: STEREOTYPES WITH SINGULAR <i>THEY</i>	161
TABLE 5.1: SINGULAR <i>THEY</i> AND GENERIC <i>HE</i> IN THE GRAMMAR CORPUS	171
TABLE 5.2: ALTERNATIVES TO EPICENES	173
TABLE 5.3: PRONOUN CHOICE AND ANTECEDENT CHOICE IN THE GRAMMAR CORPUS EXAMPLES	175

TABLE 5.4: SYNTACTIC ANTECEDENT CLASSIFICATION FOR GENERIC <i>HE</i> AND SINGULAR <i>THEY</i> IN THE GRAMMAR CORPUS EXAMPLES	176
TABLE 5.5: ANTECEDENTS OF GENERIC <i>HE</i> AND SINGULAR <i>THEY</i> IN THE GRAMMAR CORPUS EXAMPLES	178
TABLE 5.6: STEREOTYPES WITH GENERIC <i>HE</i>	180
TABLE 5.7: STEREOTYPES WITH SINGULAR <i>THEY</i>	180
TABLE 5.8: OPEN-CLASS CCSK	184
TABLE 5.9: CLOSED-CLASS CCSK	185
TABLE 5.10: KEYWORDS IN THE GRAMMAR CORPUS	187
TABLE 5.11: GROUPED KEYWORDS IN THE GRAMMAR CORPUS	187
TABLE 5.12: KEY TERMS IN THE BE06 SUB-CORPORA	188
TABLE 5.13: CONCORDANCE LINES FOR KEYWORDS IN THE GRAMMAR CORPUS	190
TABLE 5.14: TOP TEN COLLOCATES OF <i>THEY</i> AND <i>HE</i>	192
TABLE 6.1: FREQUENCY COUNTS FOR TESTED PRONOUNS IN BECKY MANCHESTER DATA	203
TABLE 6.2: FREQUENCY COUNTS FOR ALL CHILDREN AND CAREGIVERS	205
TABLE 6.3: DISTRIBUTION OF PRONOMINAL TOKENS ACROSS CHILDREN AND CAREGIVERS	206
TABLE 6.4: DISTRIBUTION OF SINGULAR AND PLURAL <i>YOU</i>	208
TABLE 6.5: DISTRIBUTION OF SINGULAR AND PLURAL <i>THEY</i>	210
TABLE 6.6: DISTRIBUTION OF MASCULINE AND GENERIC <i>HE</i>	212
TABLE 6.7: SYNTACTIC DISTRIBUTION OF ANTECEDENTS IN THE CHILDES CORPUS	215
TABLE 6.8: SYNTACTIC DISTRIBUTION OF ANTECEDENTS IN CHILDES CORPUS AFTER CONDENSATION	216
TABLE 6.9: ANTECEDENTS OF SINGULAR <i>THEY</i> IN THE CHILDES CORPUS	217
TABLE 6.10: ANTECEDENTS OF GENERIC <i>HE</i> IN THE CHILDES CORPUS	218

LIST OF FIGURES

FIGURE 1.1: NOTIONAL NUMBER	50
FIGURE 1.2: NOTIONAL NUMBER AND SYNTAX	52
FIGURE 1.3: THE PHI-FEATURE SPECIFICATIONS OF PLURAL AND SINGULAR <i>THEY</i>	54
FIGURE 1.4: THE PHI-FEATURE SPECIFICATIONS OF MASCULINE AND GENERIC <i>HE</i>	55
FIGURE 2.1: HARVARD CRIMSON	88
FIGURE 2.2: GUIDELINES FOR THE AVOIDANCE OF GENERIC <i>HE</i>	97
FIGURE 3.1: RESEARCH QUESTION ONE	120
FIGURE 3.2: RESEARCH QUESTION TWO	121
FIGURE 3.3: RESEARCH QUESTION THREE	123
FIGURE 3.4: RESEARCH QUESTION FOUR	126
FIGURE 4.1: PERCENTAGE COMPARISON OF DISTRIBUTION OF ANTECEDENTS OF SINGULAR <i>THEY</i>	133
FIGURE 4.2: PERCENTAGE COMPARISON OF ANTECEDENT DISTRIBUTION BEFORE AND AFTER CONDENSATION	135
FIGURE 4.3: PERCENTAGE COMPARISON OF DISTRIBUTION OF ANTECEDENTS OF GENERIC <i>HE</i>	139
FIGURE 4.4: EXAMPLE OF LEGAL TEXT	141
FIGURE 4.5: PERCENTAGE COMPARISON OF ANTECEDENT DISTRIBUTION BEFORE AND AFTER CONDENSATION	142
FIGURE 4.6: PERCENTAGE COMPARISON OF ANTECEDENT DISTRIBUTION FOR SINGULAR <i>THEY</i> AND GENERIC <i>HE</i> ACROSS BOTH SUB-CORPORA	145
FIGURE 4.7: EPICENE PRONOUNS USED WITH DIFFERENT SYNTACTIC FORMS	146
FIGURE 4.8: PERCENTAGE COMPARISON OF ANTECEDENT DISTRIBUTION FOR SINGULAR <i>THEY</i> AND GENERIC <i>HE</i> ACROSS BOTH SUB-CORPORA AFTER CONDENSATION PROCESS	147
FIGURE 4.9: EPICENE PRONOUNS USED WITH DIFFERENT SYNTACTIC FORMS AFTER THE CONDENSATION OF ANTECEDENTS	148
FIGURE 4.10: SEMANTIC CLASSIFICATION OF ANTECEDENTS OF GENERIC <i>HE</i>	150
FIGURE 4.11: SEMANTIC CLASSIFICATION OF ANTECEDENTS OF SINGULAR <i>THEY</i>	155
FIGURE 5.1: GRAPHOLOGICAL SALIENCE OF DISCUSSIONS OF EPICENE PRONOUNS (TEXT 8)	169
FIGURE 5.2: DISTRIBUTION OF GENERIC <i>HE</i> AND SINGULAR <i>THEY</i> ACROSS ANTECEDENT TYPES	177

INTRODUCTION

The current standard English British personal pronouns have been relatively fixed for the last two hundred years. The pronouns are marked for number across the first, second¹, and third-persons, whilst the third-person singular forms are also obligatorily marked for gender. As a result, the use of any singular animate antecedent coindexed with a third-person pronoun forces a choice between *he* and *she*, whether or not the biological sex of the intended referent is known. This forced choice of gender, and the corresponding lack of a gender-neutral third-person singular pronoun, is the primary focus of this thesis².

The absence of a formally-endorsed genderless third-person singular pronoun can be theorised, using Hartmann and James' (1998:84) definition of a lexical gap, as an "absence of a word to express a particular meaning". According to Cruse (2006:93) such gaps are hypothesised when "a language might be expected to have a word and express a particular idea, but no such word exists". Furthermore, Cruse (2006:93) argues that a "lexical gap has to be internally motivated: typically, it results from a nearly-consistent structural pattern in the language which in exceptional cases is not followed". In terms of pronouns the exceptional case in question is the overt marking of gender in the third-person singular, which is anomalous to the rest of the paradigm. Therefore, in line with the arguments of Weidmann (1984), it is possible to theorise a gap in the pronoun paradigm for an epicene (gender-neutral) third-person singular form that would correspond to the lack of gender marking across the rest of the paradigm.

However, subscription to the notion of a lexical gap does not underpin all research on gender-neutral pronouns. Newman (1992:228) cites Corbett's (1991) study which suggested that epicene (gender-neutral) pronouns "are very rare - if they exist at all" and argues that there may not be a gap in the

1 I will discuss the syntactic number marking of the second-person *you* in comparison with the syntactic number marking of singular and plural *they* throughout this thesis, although I accept that there is no morphological difference between the morphological forms of singular and plural *you*.

2 Holmes (2000:142) argues that languages that include "a distinction between female and male, animate and inanimate, in the choice of third person pronoun" serve to make gender "a relevant semantic feature", as through its rarity, any form of morphological gender marking is significant. Indeed, gender marking on the third-person is rare cross-linguistically (see Bhat 2004, in section 1.2.1), and thus it is not a language universal (Greenberg 1963) needed for effective communication.

pronoun paradigm, dismissing the concept that “a gap must be filled or avoided” as a “misguided notion” (Newman, 1992:469). I address this claim further in chapter one.

Following Newman (1998:357) I use the term *epicene* for pronouns that are “coreferent with a singular antecedent”, which in turn refers to “a referent of unknown or indeterminate sex”. I thus classify an epicene pronoun as a pronoun which is not overtly marked for gender and which is coindexed with an animate singular antecedent. Working with this definition, I concentrate on two candidates that have historically enjoyed high status in discussions of epicene pronouns (henceforth the *epicene debate*). I contrast generic *he*³ - where the masculine singular pronoun is taken to include the feminine if an antecedent is of unknown sex - with singular *they* - where the formally plural pronoun is used with singular antecedents.

- (1) a) [A musician]_i should take care of [*his*]_i instrument⁴.
b) [A musician]_i should take care of [*their*]_i instrument.

The two candidates are differentiated by the fact that singular *they* involves the issue of number concord and generic *he* involves an extension of gender concord between a pronoun and its antecedent. However, the sentences in (1) illustrate that each of these epicene candidates can be coindexed with a singular indefinite antecedent, although research has shown that disputes over the interpretation of the two pronouns means that a gender-neutral reading, and/or a singular reading may not occur.

Laitinen (2007:100) argues that “both grammatical and sociolinguistic factors” are points of contention in the epicene debate. Therefore, research focusing narrowly on just one area of the epicene debate, such as syntactic agreement, will exclude too much valuable and relevant data, such as sociolinguistic perspectives on language prescriptions⁵. To this end, in chapter one I focus on experimental, quantifiable, and empirical data,

3 The term generic *he* is used throughout epicene literature, and as such I have adopted the conventional terminology. However, I must stress that the use of *generic* is to be interpreted specifically as referring to gender inclusivity. Thus, it is used as a specialised term, rather than in its standard usage as defined in the OED as “characteristic of or relating to a class or type of objects, phenomena, etc.; applicable to a large group or class, or any member of it; not specific, general”.

4 Subscript indices denote syntactic coreference.

5 In addition, the epicene debate is not isolated from wider discussions in linguistics, sociology and beyond. For example, Bennett-Kastor (1996) looks at children’s use of generic pronouns, thus framing the debate in terms of language acquisition (section 1.1.3).

concerned with *language-internal* factors affecting epicene choice⁶, such as syntactic agreement, pronoun acquisition, and the characteristics associated with closed-classes. In chapter two I discuss the impact of *language-external* factors on the epicene debate, including non-sexist language campaigns and prescriptions on language use. However, the two areas are not mutually exclusive, and therefore a full analysis of third-person epicene pronouns in written British English must consider both of these areas.

In chapter one I contextualise the development of epicene pronouns within the wider history of the whole personal pronoun paradigm. I survey the literature on previous pronoun changes, from Old English through to the present day, and argue that the resistance to change that is characteristic of closed-classes like the pronouns may explain why such changes have been infrequent and slow. Yet at the same time, based on a summary of experimental studies and empirical evidence, I challenge the argument that the closed-class pronouns have low semantic value by noting what appears to be a default masculine value of generic *he*. I also consider the argument that closed-classes become fixed in a person's mental grammar during first language acquisition, noting that there is very little literature on the acquisition of epicene pronouns, an area of research I address in chapter six. In the final part of chapter one I discuss different hypotheses that attempt to account for the agreement between the pronoun *they* and singular antecedents, referring to phi-features and Harley's (2008) proposition of pronominal (meta)syncretisms. I also propose an adaptation of Whitley's (1978) Homonymy Theory which attempts to explain how *they* can agree with singular antecedents.

In chapter two I focus on the prescription of particular epicene pronouns from the eighteenth century onwards. Structuring the chapter chronologically, I document the development of traditional grammatical prescriptivism, before moving on to the non-sexist language reforms born out of second-wave feminism. I show that the traditional promotion of generic *he* as the English epicene is now in conflict with the argument that a masculine generic excludes women as possible referents. I conclude the chapter with an evaluation of more modern studies, looking at prescriptions in grammar handbooks from the 1970s onwards. The results of such studies

⁶ I use the labels *language-internal* and *language-external* following Labov (2001).

illustrate the continuation of the epicene debate up until the twenty-first century. I also include a short discussion of the impact of epicene prescriptions in education, noting that, despite the promotion of generic *he* in educational texts and grammar handbooks, it is not necessarily used within the classroom.

At the end of each of the first two chapters I set out a total of four hypotheses addressing areas of the epicene debate which cannot be fully explained by current scholarship. My hypotheses focus on current epicene usage, current epicene prescriptions, the impact of language-external forces on epicene choice, and epicene acquisition. Based on these hypotheses I pose a series of research questions, which I address throughout the rest of the thesis. In response to the research questions, my primary aim in chapter three is to justify my choice to use corpus linguistic methods as my primary tools of analysis. I introduce the three corpora that I use in chapters four through six, before taking each of my research questions in turn and documenting how I will tackle each part. I also note the limits of this current research, insofar as it is not within the bounds of this study to directly test the validity of Homonymy Theory. Nevertheless, the data I provide when addressing my research questions allows me to expand on my adaptation of Homonymy Theory in chapter seven.

In chapter four I use two sub-corpora of the BE06 corpus (based at Lancaster University) to document epicene usage in written British standard English post-2000. My analysis indicates that singular *they* is the overwhelming choice of epicene pronoun in British English, although generic *he* does still occur in the corpus. In chapter five my analysis of the epicene prescriptions in the grammar corpus (the structure of which is detailed in chapter three) indicates that, contrary to previous research, generic *he* has fallen out of favour and is no longer formally endorsed in the majority of grammars. My research also shows that, although it may receive limited favourable treatment, singular *they* is still not formally endorsed in the majority of grammars despite its consistent usage.

Having documented written epicene usage and the content of modern prescriptions, my final corpus, which I use in chapter six, is a set of child language transcripts sourced from the CHILDES database (MacWhinney 2000). I analyse the pronoun usage of a set of four children still within the age

boundaries of pronoun acquisition, focusing also on the pronouns they receive as input from their primary caregivers. My results in this chapter show that, although there are tokens of generic *he* in the data, its use in the corpus is problematic. When such tokens are eliminated it becomes clear that the children in the corpus do not receive generic *he* as input, although their caregivers do use singular *they*. I also use the CHILDES corpus to investigate parallels between the use of second and third-person pronouns in the acquisition data, showing that tokens of singular and plural *they* occur in relative distribution to occurrences of singular and plural *you*.

Finally, in chapter seven I contextualise the results from my corpus analyses in chapters four through six within the wider literature of the epicene debate I evaluated in the first two chapters. I show how my corpus analyses impact upon the epicene debate by providing up-to-date documentation of the current usage and prescription of epicene pronouns. My results indicate that the current epicene of choice in written British standard English is singular *they*, which, incidentally, is now received more favourably in grammar books than previous research has suggested. Furthermore, my work on the CHILDES corpus indicates that children receive tokens of singular *they* as L1 input, thus arguably making it possible to acquire the form as a member of a person's internal pronoun paradigm. Finally, my results can also support the argument that Homonymy Theory can account for coreference between singular *they* and formally singular antecedents. However, as is evident from the discussion of the theory in chapter one, this concept is still in its infancy.

CHAPTER ONE: EXPLORING EPICENE PRONOUNS

1.1: CHARACTERISTICS OF THE PERSONAL PRONOUNS	7
1.1.1: THE PARADIGM AS A CLOSED-CLASS	8
1.1.2: PREVIOUS CHANGES: TRENDS THROUGHOUT HISTORY	12
1.1.3: ACQUISITION AND PRONOMINAL CHANGE	21
1.2: THE PSYCHOLOGICAL REALITY OF EPICENE PRONOUNS	25
1.2.1: THE SEMANTIC SALIENCE OF GENDER MARKING	26
1.2.2: WHAT EPICENES MEAN	29
1.3: DIFFERENCES BETWEEN GENERIC HE AND SINGULAR THEY	40
1.3.1: AGREEMENT WITH DIFFERENT TYPES OF ANTECEDENT	41
1.3.2: NOTIONAL NUMBER AND OTHER EXPLANATIONS	48
1.3.3: HOMONYMY THEORY	53
CONCLUSIONS	58

In this first chapter I survey the existing literature on the language-internal elements of the epicene debate, which is founded primarily upon empirical research and experimental data, in order to identify the key topics and points of contention within the wider scholarship on epicene pronouns. My goal in the first section of this chapter is to contextualise epicene pronouns within the whole personal pronoun paradigm, acknowledging that it is widely accepted within the wider literature that the pronouns are a closed-class. Furthermore, I consider whether epicene pronouns possess the characteristics associated with closed-classes, focusing specifically on the arguments that closed-classes are resistant to change and unlikely to admit new members. In relation to these characteristics I survey the literature on the acquisition of pronouns and the related debates about when a closed-class becomes fixed in a person's mental grammar.

In the second section of the chapter, I note that, despite their closed-class status, previous studies have indicated that the personal pronouns are atypical in their characteristics. Although they display some characteristics that are common to closed-classes, such as their resistance to change, scholarship indicates that the pronouns do not necessarily possess all characteristics associated with closed-classes. For example, a review of the

literature indicates that closed-classes are accepted to be low in semantic value, but there is a large body of work in support of the position that the gender marking on third-person pronouns is salient to comprehension. My review of previous studies on epicene production and comprehension in section 1.2.2 indicates a high level of consensus in the literature in support of the position that generic *he* has a default masculine interpretation, and therefore carries its own semantic value.

Following on from my review of these studies, I focus more closely on the two prominent epicene candidates, singular *they* and generic *he*, which are found repeatedly in the literature. By focusing on scholarship that illustrates how the two candidates (dis)agree with different types of antecedent I highlight the syntactic differences between the two primary contenders for epicene status. Based upon my acknowledgement of the key issues in the literature on epicene agreement I then evaluate the current explanations for the apparent acceptability of coreference between singular *they* and formally singular antecedents, focusing specifically on the fact that no current theory appears to be able to explain coreference between singular *they* and definite singular NPs. I then attempt to address this problematic area of epicene agreement by evaluating an existing theory of personal and impersonal pronouns, as proposed by Whitley (1978), and argue that it is possible to adapt this existing research to account for the apparent suitability of singular *they* as the English epicene that is evident in existing scholarship.

1.1: CHARACTERISTICS OF THE PERSONAL PRONOUNS

In this first section I focus on the personal pronoun paradigm as a whole in order to contextualise the epicene debate within wider current literature on pronouns. I begin in section 1.1.1 with a consideration of the closed-class status of pronouns, which goes almost unanimously unchallenged in previous scholarship, focusing on the properties generally associated with closed-classes, such as their resistance to change and their lack of independent semantic value (see section 1.2 for more detail). In

section 1.1.2 I review the scholarship on the development of the pronoun paradigm from Old English to the present day to illustrate that, despite their apparent closed-class status, pronoun changes have occurred throughout history, and I highlight patterns in what types of changes have been identified by other scholars. In section 1.1.3 I use the identification of these patterns, combined with scholarship on language acquisition to propose that the acquisition of singular *they* is entirely possible, and furthermore suggest that theories of closed-class acquisition may hold the key to explaining why certain changes in the pronouns have been adopted into the standard paradigm and others have not.

1.1.1: THE PARADIGM AS A CLOSED-CLASS

It is widely established within linguistics that languages are made up of open and closed-classes (Schultz 1975; Shapiro and Jensen 1986; Cutler 1993; Chafetz 1994; Segalowitz and Lane 2000; Rizzi 2004). Herron and Bates (1997) review the different, but generally overlapping, definitions that attempt to account for these class divisions in language and argue that “closed-class words are used primarily to express grammatical and semantic relations, are generally very high in frequency and low in semantic content” (1997:217). Significantly for this research, personal pronouns are classified in the wider literature as members of a closed-class paradigm, although Geurts (2000:728) notes that “there is no sharp divide between content words and function words”. Similarly, Rizzi (2004:438-439) suggests that there are “some intermediate cases” where forms can display both open- and closed-class characteristics. My discussion of the literature on the semantic contribution of epicene pronouns in section 1.2 suggests that there is evidence that Rizzi’s argument can be applied to third-person singular pronouns, but in the present section I am concerned with the conceptualisation of the whole personal pronoun paradigm more generally.

Importantly, there is evidence that the concept of open- and closed-classes is more than just linguistic theory, as results from brain imaging studies (see Neville, Mills and Lawson 1992; ter Keurs, Brown, Hagoort, and Stegeman 1999), support the argument “that distinct neural structures are

involved in the access and use of elements of the two classes” (Rizzi 2004:440). Empirical data supports Rizzi’s (2004:440) assertion that “the open and closed lexicons are segregated in the brain” (2004:440)⁷. Therefore, there is evidence to support the argument that pronouns, as a closed-class, are processed differently to open-class words, which in turn suggests that the closed-class nature of pronouns is a key issue within the epicene debate. Yet Segalowitz and Lane (2000:337) argue that “showing that the brain must respect this distinction” between open- and closed-classes is, at best difficult, and at worst, impossible, given the current available technology. However, three key studies, discussed below, provide evidence indicating that there is a processing difference between how the brain treats open- and closed-classes.

Based on the data of Bradley and Garrett (1983) who showed that the left hemisphere, which contains areas linked to language (i.e. Broca and Wernicke’s areas), appeared to be sensitive to the difference between open- and closed-class items, Shapiro and Jensen (1986:321) investigated whether the left hemisphere “selectively supports a distinct closed class recognition device” (1986:321). They presented forty-eight ‘nonwords’⁸ to sixteen participants, testing one visual field at a time based on the assumption that data presented to the right eye is initially processed by the left hemisphere and vice versa. Shapiro and Jensen’s results showed that differences in reaction times for identifying open and closed-class words were “more robust for presentations to the RVF [right visual field]” (1986:323). In other words, the left hemisphere of the brain was more sensitive to the open/closed-class distinction⁹, a finding which Shapiro and Jensen linked to its “role in processing grammatical structure” (1986:324). Their results support the general hypothesis that the brain processes open- and closed-class words differently¹⁰.

7 See Pulvermüller (1999) for an extensive discussion and hypotheses of how language is stored in the brain.

8 These are words created from English roots but which do not occur in the language. In Shapiro and Jensen’s study the nonwords had either an open-class word or a closed-class word as their stem, as in *lostner* and *mostner* respectively (1986:321).

9 This was also true for the 48 real English words which acted as their experimental controls, indicating that the left hemisphere reacted more quickly to language than the right hemisphere.

10 Segalowitz and Lane (2000:378) review the literature on where different classes are stored in the brain, concentrating on the results of ERP experiments, which show that open-class words cause “a large bilateral negative peak at 350ms” in the posterior area of the brain,

In a similar vein, Segalowitz and Lane (2000) tested whether open- and closed-class words had different reading times by asking their participants to read sentences aloud. Their findings indicated that “in meaningful contexts” closed-class words were accessed faster than open-class words (2000:382) and they concluded that this was due to their high frequency and related predictability. However, Segalowitz and Lane state that if the words in their experiment had been presented in isolation, they would have expected slower response time for closed-classes “due to their lack of independent meaning” (2000:337)¹¹. Unfortunately for this research, neither Shapiro and Jensen nor Segalowitz and Lane specifically tested pronouns, but their conclusions still indicate that there is a psychologically real distinction between open- and closed-class words. Indeed, Segalowitz and Lane concluded that “the mental lexicon respects the differences between the two word classes in some way” (2000:377), and Rizzi (2004:440) argues that “it appears that the open and closed lexicons are segregated in the brain, dissociable in acquisition and pathology, accessed differently in language use”¹².

Therefore, having established that there is evidence in the literature which indicates that there are differences between how open- and closed-class items are processed, it is apparent that class membership is highly significant. If, as the literature suggests, the pronouns are a closed-class, then the characteristics attributed to closed-classes are key points for consideration within this research, because the adoption or rejection of any new pronoun as an epicene form will arguably be influenced by the restrictions, which I discuss below, that the literature has shown to apply to closed-classes.

One of the primary characteristics in the literature on the differences between open- and closed-classes (c.f. Herron and Bates 1997) is that the former, such as nouns and verbs, easily admit new members, whilst the closed-classes, or function words, such as prepositions and determiners,

whilst closed-class words caused a negative peak in “the left inferior frontal-anterior temporal area”. Supplementary evidence comes from Pulvermuller (1999) who suggested that “lexical access for function words involves the perisylvian region” which is also in the left hemisphere (Segalowitz and Lane 2000:378).

¹¹ This lack of independent meaning is linked to the position that closed-class words have low semantic value. However, as section 1.2 shows, this may not be the case for the third-person pronouns as they are overtly marked for gender.

¹² I return to the issue of closed-class acquisition in section 1.1.3.

encode grammatical structure and are resistant to change. In the case of pronouns this grammatical structure comes in the form of case marking, which signifies the relationship between the pronoun and the verb. Another equally important characteristic ascribed to closed-classes is that it is generally agreed upon in the literature that closed-class words generally lack the “descriptive content” (Rizzi 2004:438) attributed to open-class words, and subsequently “are not high in information value” (Chafetz 1994:274). I consider this characteristic in more detail in section 1.2, where I review studies that have shown generic *he* to have a default masculine semantic value, but in this section I focus on the issues surrounding changes in a closed-class.

There is evidence in the literature for the argument that characteristically closed-classes are “highly resistant to change”, based, at least partially, on the premise that “pronouns and other function words, are parts of the basic machinery of the language” (Green 1977:152). Considering epicene forms specifically, Segalowitz and Lane (2000:376) argue that although a gender-neutral form would be useful, adding to a closed-class “through cultural change is very slow and open to controversy”. This is most likely due to the structural nature of closed-class words, as Newman (1992:470) likens the changing of a grammatical feature (such as adding a new pronoun or changing a pronoun’s scope of reference) to “altering one part of a whole ecosystem”. In other words, the development of a new pronoun would represent a fundamental change within the system, not simply an addition of another form. Such arguments link to Rizzi’s (2004:439) claim that changes in closed-classes “are rare and involve significant restructurings of the system”. Yet Rizzi does not rule out changes completely, saying that “closed class items are subjected to diachronic changes in the long run” (2004:439).

Importantly, studies have shown that the factors influencing open- and closed-class changes do not always overlap. For example, new inventions may have a heavy influence on English nouns, as forty years ago the nouns *Playstation* and *email*¹³ did not exist, yet there is no evidence in the literature that such neologisms also occur in closed-classes (see Baron 1986

13 In the *OED* entries for the noun *email* begin at 1982 and the first entry for the verb form is 1987.

for a list of epicene neologisms all of which have failed to become part of the closed-class paradigm). However, that does not mean that open- and closed-class changes have nothing in common. According to Yang (2000:237) the “social and cultural factors” which influence the language of a population appear to be an overarching phenomenon that affects both classes of words. For example, Pauwels’ (2001) research indicated that non-sexist language reforms in Australia have caused a decrease in the use of generic *he* on national radio stations. She analysed a corpus of circa half a million words taken from before and after language reform rejecting generic masculines (see section 2.2) and found that the occurrence of generic *he* dropped between the first dataset from the 1960s-1970s and the second taken from the 1990s. Here the influence of a social constraint on a closed-class pronoun is clearly visible¹⁴.

Based on my review of the literature on open- and closed-classes, which highlights the argument that closed-class changes represent structurally significant developments within a language, previous changes within the personal pronoun paradigm are highly important in this research. Any patterns of pronominal change evident in the literature may affect the potential integration of an epicene pronoun into the paradigm. Therefore, in the following section I consider the literature on the development of English pronouns starting from Old English and working chronologically through to the present day.

1.1.2: PREVIOUS CHANGES: TRENDS THROUGHOUT HISTORY

The Old English personal pronoun paradigm, taken from Fennell (2001:68), differs considerably from the present-day standard paradigm, especially in the first and second-persons where the forms are barely recognisable (Table 1.1)¹⁵. However, the forms that look most like their

¹⁴ The influence of social factors on the personal pronouns is an integral part of my discussion of the paradigm’s development (section 1.1.2) and is one of the main focuses of chapter two.

¹⁵ Obviously, there are many differences between Old English and Present-day English, including grammatical changes such as the loss of formal gender marking, and phonological changes such as the Great Vowel Shift. However, the focus here is on the pronoun paradigms.

modern counterparts are the third-person singular masculine forms, suggesting that such forms have been relatively stable throughout the development of English, and providing evidence that the pronouns have resisted change.

TABLE 1.1: THE OLD ENGLISH PERSONAL PRONOUN PARADIGM (500-1100)

	SINGULAR	DUAL	PLURAL
FIRST-PERSON			
NOMINATIVE	IC	WIT	WĒ
ACCUSATIVE	MEC, MĒ	UNC, UNCIT	ŪSIC, ŪS
GENITIVE	MĪN	UNCER	ŪSER, ŪRE
DATIVE	MĒ	UNC	ŪS
SECOND-PERSON			
NOMINATIVE	ÐŪ	GIT	GĒ
ACCUSATIVE	ÐEC, ÐĒ	INC, INCIT	ĒOWIC, ĒOW
GENITIVE	ÐĪN	INCER	ĒOWER
DATIVE	ÐĒ	INC	ĒOW

	MASCULINE	NEUTER SINGULAR	FEMININE	ALL GENDERS PLURAL
THIRD-PERSON				
NOMINATIVE	HĒ	HIT	HĒO, HĪE	HĒO, HĪE
ACCUSATIVE	HINE	HIT	HĒO, HĪE	HĒO, HĪE
GENITIVE	HIS	HIS	HIRE	HIRA, HEORA
DATIVE	HIM	HIM	HIRE	HIM, HEOM

Adapted from Fennell (2001:68)

Alternatively, the reason why such forms have stayed rather constant could be that practically all Old English third-person forms were based on the masculine template, making such forms a fundamental element of the structure of the paradigm. Although not a feature in the wider literature, I argue that the long history of this masculine template can be used as supporting evidence for traditional grammatical prescriptive views (discussed in chapter two) that generic *he* has always been a feature of English, especially as the Old English neuter pronouns (which encoded grammatical gender as opposed to biological sex) were based on the masculine forms¹⁶.

For a review of other fundamental changes in the history of the English language, see Fennell (2001) or for a more historical perspective Skeat (2009 [1912]) is a good resource.

16 One other major difference between Old English and Present-day English is the use of a dual pronoun, which meant that the system had a three-way distinction: singular, two entities, and more than two entities. Yet the dual distinction had been lost by Middle English, and is thus of little consequence to the further development of the pronoun paradigm. However, losing a number distinction completely is clear evidence that a closed-class can change over time. Indeed, such a large change justifies my detailed consideration of the different pronoun paradigms in English history.

As Old English progressed to Middle English grammatical gender marking was dropped from the language. This led to the loss of neuter pronouns in the third-person and, combined with the loss of dual pronouns (Table 1.2), according to Fennell (2001:102) there was “considerably less overlap in the pronouns than... in Old English, and thus less room for confusion”. As the following discussion will show, issues such as pronominal overlap, confusion and comprehension have been key to the development of the paradigm and are found in the literature in relation to the second and third-person forms throughout history.

TABLE 1.2: THE MIDDLE ENGLISH PERSONAL PRONOUN PARADIGM (1100-1500)

	FIRST-PERSON	SECOND-PERSON	THIRD-PERSON
SINGULAR			
NOMINATIVE	I	THOU	HE, SHE, IT
ACCUSATIVE	ME	THEE	HIM, HIR, HIT
GENITIVE	MY, MIN	THY, THYN	HIS, HIR, HIS
DATIVE	ME	THEE	HIM, HIR, HIM
PLURAL			
NOMINATIVE	WE	YE	THEY
ACCUSATIVE	US	YOW	HEM
GENITIVE	OURE	YOURE	HIR
DATIVE	US	YOW	HEM

Adapted from Fennell (2001:102)

The observed loss of grammatical gender across Middle English has still not impacted on the standard present-day third-person singular pronouns, as we still mark gender on present day third-person singular (animate) forms. One explanation for the gender marking, as presented by Howe (1996:63), is that pronouns may “retain or maintain distinctions lost or absent elsewhere in the language”, most likely due to their closed-class nature. However, this does not mean that such gender marking is static, as Howe also notes that pronouns “cannot indefinitely uphold a grammatical category” on their own (1996:63). Indeed, evidence for this position comes from the elimination of the dual person between Old English and Middle English, and the coalescence of the accusative and dative cases between Middle English and Early Modern English (see below).

My Middle English Table 1.2 shows that the first-person forms clearly resemble their present-day equivalents. The nominative first-person singular has lost its Old English unstressed syllable so that it resembles its modern form (Fajardo-Acosta 2002), and apart from the dropping of *e* from *oure* and the addition of *e* to *min*, there is consensus in the literature that the first-

person forms, in their written form at least, are stable from this point forwards. However, the same is not true for the third-person plural forms. The Middle English paradigm illustrates how the *th-* forms of Scandinavian origin, such as nominative *they*, began to infiltrate the system, suggesting that social change and language contact directly influenced the paradigm. Crystal (2004:76) documents that it took 300 years for the Scandinavian *they* forms to completely replace their Old English counterparts and spread throughout the whole country, but the bulk of this change was over by about 1400. However, there is no orthodox account of this change and there is some confusion surrounding its date, as Crystal (2001:45) also argues that in “the 15th century *their* became the norm, and by the beginning of the 16th century *them* had followed it”¹⁷.

Independent of the order in which the Scandinavian forms entered the pronoun paradigm it appears that they were useful to speakers of Middle English. Crystal notes that the *th-* forms “must have been very welcome” in the southern areas where the Old English pronouns *he*, *heo* and *hi* had begun to sound similar (2004:76). In fact, Fennell (2001:124) notes that the southern forms were even more complicated as “the *h-* was often dropped, so that *a* could also be the equivalent of *he*”. Therefore, it appears that alongside the social influences of invasion, there was a language-internal change also affecting the pronouns, where phonological contrast was being lost between forms. The Scandinavian forms appear to have been a good way to avoid the confusion of too many pronouns sounding the same¹⁸.

One other major difference between Old English and Middle English pronouns evident in the paradigms present in the literature is the development of feminine third-person forms with initial [ʃ]. According to Fennell (2001:143), these forms began in the North and East Midlands dialects and “made it easier to differentiate [feminine pronouns] from masculine and plural pronouns”. Significantly, the need to differentiate

17 In any case, neither of Crystal’s claims is borne out in Table 1.2. Fennell (2001:102), where the paradigm is taken from, lists the nominative form *they* as part of Middle English, but not the accusative *them* or possessive *their*. One explanation for this discrepancy is that the two authors may have consulted different dialects before making their assertions. Crystal could have been focused on the northern dialects where the *th-* forms took hold first, whilst Fennell’s paradigm could be based on a southern dialect where the Scandinavian influence was not felt until later in the Middle English period. However, this is speculation as neither author states explicitly which dialects they have consulted for their analysis of Middle English.

18 For a detailed breakdown of the development of *they* see Howe (1996:154-160).

between similar sounding forms parallels the language-internal factors that could have affected the incorporation of the Scandinavian third-person forms discussed above, and suggests that phonological drift, whereby pronouns that were once distinct come to sound the same, could be a key element in pronoun change. However, why the feminine pronouns developed the way that they did, and how Old English *heo* became the present-day form *she*, is not agreed upon by all historians¹⁹, a problem that has been classified by Crystal as “one of the unsolved puzzles in the history of English” (2001:43). What is important however is that, regardless of their development, the [ʃ] forms “had prevailed in general usage” by the time of Early Modern English (Fennell 2001:134-144).

The Early Modern English period was particularly important in the development of standard English, as this period saw the introduction of the printing press and increased access to education (Fennell 2001:156). These factors are discussed further in chapter two, as in terms of epicene pronouns, this period laid the groundwork for the birth of traditional grammatical prescriptivism, which would proscribe singular *they* (section 2.1). A review of the scholarship on Early Modern English pronouns shows that one major change affecting all pronominal forms was the coalescence of the accusative and dative cases; there was no longer formal grammatical differentiation between the direct and indirect object.

TABLE 1.3: THE EARLY MODERN ENGLISH PERSONAL PRONOUN PARADIGM (1500-1800)

	FIRST-PERSON	SECOND-PERSON	THIRD-PERSON
SINGULAR			
NOMINATIVE	I	THOU, YE, YOU	HE, SHE, (H)IT
ACCUSATIVE	ME	THEE, YOU	HIM, HER, (H)IT
POSSESSIVE	MY, MINE	THY, THINE, YOUR YOURS	HIS, HER, HERS, HITS
PLURAL			
NOMINATIVE	WE	YE, YOU	THEY
ACCUSATIVE	US	YOU	THEM, ‘EM
POSSESSIVE	OUR, OURS	YOUR, YOURS	THEIR, THEIRS

Adapted from Nevalainen (2006:77)

In addition, Table 1.3 shows that during this period the feminine [ʃ] and Scandinavian *th*-forms became established and the neutral third-person

¹⁹ See Howe (1996:145-54) for discussion of the development of *she*, or Crystal 2001:43 for a review of the three main theories of its development. Also see Wolfe (1989:87-89) for a discussion on the phonological relevance of the male/female dichotomy and its contribution to the development of the pronoun *she*.

singular *it* came into use. Most importantly to epicene research, the paradigm taken from Nevalainen (2006:77) indicates that the framework for the modern third-person pronouns was established by this period of English development.

However, the same cannot be said for the second-person forms²⁰. In Middle English the *thou/you* distinction was simple, as the former was singular and the latter was plural, but Crystal (2004:307) notes that this distinction was disrupted by the “emergence of *you* as a singular”, which began in the thirteenth century. People were beginning to use the plural form *you* as a singular pronoun (not unlike singular *they*). Whilst this change may have indicated expansion of the possible referents of *you*, Fennel (2001) discusses how the distinction between *thou* and *you* was not merely one of number. The factors affecting pronominal use were “much more complex”, with *thou* seen as “intimate” but *you* considered “polite” (2001:164), paralleling pronominal functions in other languages, for example the *tu/vous* distinction in French²¹.

Nevertheless, by the first half of the 17th century, *thou* had “disappeared from Standard English” (Crystal 2004:310)²². Yet, a survey of research on Early Modern English pronouns indicates that, parallel to some of the arguments that would later be put forward to reject singular *they* (see section 1.3 and chapter two), some groups in society were not happy with singular *you*. Both Fennel (2001:164-165) and Crystal (2004:310) note that in the Early Modern English period, Quakers preferred to use *thou* forms as they were grammatically singular, making them “more exact usage”. This argument was put forward in terms of both number agreement (see section 2.1.2) and a dislike for the social distance that *you* was perceived to create (Crystal 2002:310)²³. Nevertheless, despite such arguments against the pronoun change, the *thou* forms became restricted to archaic usage and *you* took over as both plural and singular.

20 For an analysis of how second-person pronouns were used in the sixteenth century see Brown and Gilman (1989) and Walker (2000).

21 Bodine (1975:141) links the development of the *tu/vous* distinction to the rise of feudalism, thus providing a link between pronoun development and social change.

22 However, Crystal does not define what he means by *Standard English* in relation to this argument.

23 See also Frank (1989:114) for a discussion on Quaker perceptions of *thou*.

The *thou* forms however, were not the only language change casualties in the second-person pronouns. Initially, both *thou* and *you* had morphologically distinct nominative and accusative forms, *thou/thee* and *ye/you* respectively (see Crystal 2004:307). Whilst *thou* was dropped from the paradigm completely, the *you* forms were losing their case distinction and *ye* began to be used less frequently. One explanation for this, proposed by Fennell (2001:142) is that the two forms, *you* and *ye*, sounded similar due to their unstressed vowels, and “could be pronounced almost identically as [jə] or [jʌ]”, meaning that the distinction between them was not salient in speech. This is yet another pronoun change (in this case a written one) which has been influenced by a language-internal phonological change which led to different pronouns becoming indistinguishable in speech.

More evidence for the chronology of this pronoun change comes from Nevalainen (2006:80) who notes that Henry VII, who ruled at the start of the Early Modern English period, “consistently” used *you* as the nominative second-person form “in his personal correspondence”. Whilst this does not mean that the form had any official backing, it was obviously regarded highly enough to be used by the ruling elite²⁴. As Henry VIII ascended to the throne in 1509, this change was clearly already in progress at the very start of the Early Modern English Period, but Crystal (2001:65) notes that *ye* had only “disappeared from Standard English in the late 17th century”. Again, there appears to be little consensus on when such changes occurred, but discrepancies in dates are not my primary concern. The significance of Nevalainen’s argument to the present research is that, as I will discuss in section 2.1, the forms prescribed during the development of traditional prescriptive grammar (see chapter two) which is associated with circa the eighteenth century, were generally based on the forms used by the upper echelons of society. Based on this position, which holds throughout previous scholarship on the development of grammatical norms in English, the King’s pronoun choice is significant and perhaps influenced, in some small way, the adoption of singular *you*. However, there is not enough evidence in the literature to assess the weight, if any, of this proposed influence.

²⁴ Section 2.1 will show why the ruling elite’s choice of language forms is so important.

As a final comment on the Early Modern English pronouns, Nevalainen (2006:82) discusses generic pronominal usage and argues that whilst “the traditional masculine *he* was used throughout the period” for referents of undetermined sex, singular *they* can also be found. Indeed, such usage can also be traced back to Middle English (Nevalainen 2006:82; Curzan 2003:71-72)²⁵. Tracing changes in epicene choice with indefinite pronouns between 1500 and 1800 Laitinen (2007:197) concludes that between Middle and Early Modern English “the plural pronoun extends its semantic sphere to cover the functions carried out by the singular”. Thus, despite the fact that neither generic *he* nor singular *they* appears in the formalised pronoun paradigms presented above, there is research to suggest that both forms were in use, and have been in use for hundreds of years.

Chronologically this brings us to the current standard English pronoun paradigm (Table 1.4). As I suggested at the start of this section, analysing the development of the pronoun paradigm has revealed some useful information about how and why the closed-class pronouns changed. There appears to have been a trend throughout history for pronouns to become phonologically distinct.

TABLE 1.4: THE CURRENT STANDARD ENGLISH PERSONAL PRONOUN PARADIGM (1800-)

	FIRST-PERSON	SECOND-PERSON	THIRD-PERSON
SINGULAR			
NOMINATIVE	I	YOU	HE, SHE, IT
ACCUSATIVE	ME	YOU	HIM, HER, IT
POSSESSIVE	MY, MINE	YOUR, YOURS	HIS, HER, HERS, ITS
PLURAL			
NOMINATIVE	WE	YOU	THEY
ACCUSATIVE	US	YOU	THEM
POSSESSIVE	OUR, OURS	YOUR, YOURS	THEIR, THEIRS

Furthermore, I have shown that the literature points to a consistent interplay between language-internal factors such as phonological change, and language-external factors such as invasions leading to the integration of

²⁵ Curzan (2003:70) takes this argument one step further and claims that “generic *they*” can be found as far back as Old English. However, it is difficult to evaluate this claim, as the *they* forms did not infiltrate the language until Middle English, and so arguably they were not available to the majority of Old English speakers. In addition, the grammatical gender of Old English arguably negates the need for singular *they* because biological sex was second place to grammatical agreement. Thus, it appears that Curzan’s claim would need more research. Nevertheless, showing that singular *they* has been in use since Middle English, is good enough for the scope of this thesis, as it shows that the pronoun was used as an epicene before any debates over sexist language.

new pronouns into the closed-class. Significantly for epicenes there also appears to be parallels between the development of *you* and the development of singular and plural *they*, where a form which was initially restricted to the plural comes to be used as a singular form. These parallels are discussed in more detail in section 1.1.3.

Given the evidence in the scholarship discussed above, it is likely that the pronoun paradigm will continue to change, as there is no evidence to support the argument that the paradigm will remain static from now on. In addition, the literature also illuminates what appears to be a tendency towards decreasing the number of pronouns and simplifying the system as a whole. A comparison of the Old English pronouns and the modern English pronouns shows that the current standard English paradigm has 23 fewer forms than its historical counterpart. The scholarship shows that the only additions to the pronoun paradigm between Old English and Present-day English have been the Scandinavian *th-* forms and the feminine [ʃ] forms, and importantly, these changes appear to be influenced by a combination of language-internal phonological developments and language-external forces.

Thus, the evidence suggests that at least one aspect of the closed-class nature of the pronoun paradigm is extremely robust indeed, as my review of the scholarship on the development of the English pronouns indicates that they are highly resistant to change, but despite this, changes do happen²⁶. Even more important to epicene research, the above discussion has also shone some light on the development of generic *he* and singular *they*. Generic *he* clearly has a strong historical standing, with the origins of the modern form being evident in the Old English paradigm, whilst singular *they*, which was only (potentially) introduced as an epicene candidate with the Scandinavian *th-* forms in Middle English, has parallels with the development of the second-person pronouns and the adoption of *you* as a singular pronoun. In the next section I show the significance of these parallels by addressing them in the context of scholarship on language acquisition.

26 One possible change currently in progress is the tendency in certain dialects to mark plurality in the second-person (see Richardson 1984:58; Maynor 1996; Wales 1996:17).

1.1.3: ACQUISITION AND PRONOMINAL CHANGE

In this section I consider the literature that attempts to account for the apparent slow rate of change in the pronoun paradigm that is evident in the section above. I show that there is support in the wider scholarship for the argument that one of the main factors affecting language change in a closed-class is that, once a paradigm has been built in a person's internal grammar, they are aware that it "is not likely to admit new members" (Chafetz 1994:275). Arguably, this could be either a social barrier, as people may become aware that the pronouns are a definite set, or a psychological barrier linked to how closed-classes are acquired.

According to Yang (2000), despite the closed-class' overall rigidity, there is evidence in the scholarship for the position that it may be possible to influence a paradigm during the acquisition process, whilst it is still flexible²⁷. With this in mind, Yang (2000:231) argues that "ultimately, language changes because learners acquire different grammars from their parents". His support for this viewpoint is that language change can be observed "when a generation of speakers produces linguistic expressions that differ from those of previous generations, either in form or in distribution" (Yang 2000:231), and he thus sees the construction of an internal grammar as "an adaptive response to the linguistic evidence in the environment" (2000:234).

I argue that Yang's arguments can be applied to pronoun change, if it can be shown that each generation acquiring the closed-class pronouns has slightly different input (influenced by language-external forces such as non-sexist language reforms, see chapter two), which could lead to different pronominal forms being incorporated into people's mental pronoun paradigms. Indeed, there is some support for this argument in the current literature, as Egerland (2005:1105-1106) argues that "a child acquiring language has no direct access to his [sic] parents' grammar" and must rely on their linguistic output, therefore "the source of a syntactic change can only lie in the trigger experience, and hence must be external to the grammar itself". However, before addressing this claim, I contextualise my

27 This claim is based on the assumption that children are born without a fully formed paradigm, but do have the innate ability to acquire one, and thus follows the well-established nativist perspective (see Sampson 2005).

argument by briefly reviewing relevant scholarship on the processes by which children acquire closed-classes.

Chafetz (1994:273) argues that during language acquisition children acquire “both the meaning of an individual word and also something about the word’s syntactic category”. According to Rispoli (1994:159), part of this process involves the creation of closed-class paradigms that “express a finite stock of grammatical notions”. However, Rispoli (1994:161) claims that due to the “irregular” case marking on the personal pronouns children must rote learn them, memorising individual forms. In support of this position, Moore (2001:212) argues that “no general rules aid learning across the various pronominal forms”. If the argument that children are not born with a complete pronoun paradigm is accepted, this would suggest that it is possible for children to build different paradigms to those built by the generation before them. This process is explained by Egerland (2005)²⁸:

The first generation has a grammar, G1, which generates an output in the form of a language, L1. The second generation is exposed to L1 and defines its grammar, G2, on the basis of this input. The output of G2, in turn, is a second language, L2

(Egerland 2005:1109).

Evidence for grammatical change through language acquisition, and further evidence that children are not born with complete pronoun paradigms comes from Gerken (1987) who found that children “still in the telegraphic stage” would repeat nonsense closed-class words, such as *na*, suggesting that they were adding to a closed-class (Chafetz 1994:283). Chafetz argues that the children did this “because they had not yet analyzed these items as to function or meaning” (1994:283), and this led her to conclude that “[c]hildren do not easily admit new members to the closed-class vocabulary, *but clearly they are able to do so*” (1994:284, my emphasis).

In addition, Rispoli (1994:158) found that children as old as three years and eight months were still making errors in the third-person pronouns, suggesting that these are the last forms in the paradigm to be fully acquired. I argue that this relatively long timeframe compared to other pronouns - following Owens Jr (2007:225) first-person pronouns are generally acquired

²⁸ See also Pertsova (2011:251).

by 30 months, second-person pronouns by 34 months, and third-person pronouns from 35-40 months - means that there is a larger chance that the third-person forms can be influenced by the linguistic input that children receive (what Egerland 2005, above, refers to as the “trigger experience”). However, Chafetz (1994:275) notes that even children are highly sensitive to the restrictions of a closed-class, and that they “are aware, on some level, that the closed-class vocabulary is not likely to admit new members” (but she does not specify whether this level is psychological, syntactic, etc.).

Applying Chafetz’ claim to epicenes however, I argue that even if the argument holds that children are aware of the closed-class nature of the pronouns, this does not necessarily represent a problem for the integration of an epicene form, especially singular *they*, into their mental grammars. Singular *they* has the same morphological and phonological form as plural *they*, a form which children will be naturally acquiring (as there is no evidence in the literature that normally developing children do not acquire plural *they*). Thus the proposed integration of singular *they* into a developing pronoun paradigm, arguably represents the easiest type of change for a closed-class, because it is simply the expansion of a current form in both meaning (one/more than one) and syntax (singular/plural). Children only have to notice that there is both a singular and a plural form of *they*, a topic I pick up on in section 1.3.3, and assign them individual places in their pronoun paradigms. In addition, as singular *they* is not a neologism it would not be alien to the adults providing the L1 input, and studies such as Stringer and Hopper (1998:209) have indicated that singular *they* is the spoken epicene of choice in “conversational interaction” (see also Newman 1992 and Pauwels 2001) suggesting that the form will be present in L1 input. Similarly Meyers (1990) cites an unreferenced study that showed how a group of children in Minnesota were using singular *they* more than their parents, suggesting that it was part of their pronoun paradigms.

Therefore, based on the current literature, there is evidence for the position that there are parallels between the development of *you* and *they*, such as the expansion of a plural form to the singular, and similarities in arguments opposing such changes (c.f. section 1.1.2 and chapter two), and there is also evidence that singular *they* is common in spoken English. In relation to Moore’s (2001) comment, above, that there are no learning mechanisms for the personal pronouns, I argue that any parallels between

the forms of *they* and *you* may facilitate their acquisition. Based on the identification of parallels between the development *they* and *you*, the acquisition of singular *they* would actually represent a pattern in the usually irregular pronouns (what Harley 2008 refers to as a pronominal syncretism, see section 1.3.3). Thus, as long as singular *they* is received as input, and there are studies which indicate that this is the case, see above, I believe there is no reason why singular *they* cannot be acquired in the same manner as singular *you*.

However, I am aware that a similar argument can also be posited for generic *he*, in that if children receive both masculine *he* and generic *he* as L1 input, then they should be able to include two forms of the pronoun in their paradigm. The problem with this theory is that children do not receive generic *he* as input during the early stages of language acquisition (see Graham 1973, discussed in chapter two). This issue is what Nilsen (1977) touches on when she argues that girls and boys have different acquisition experiences of the pronoun *he*. There is support in wider scholarship for the argument that girls acquiring English learn that the pronoun *she* can refer to them, but the pronoun *he* cannot, yet when *he* is used generically girls must learn “to apply the pronoun... first acquired in its specifically male contexts” to themselves (Martyna 1980:75). Whereas for boys, generic *he* is just an expansion of the form that has always applied to them and Gibbon (1999:44) argues that “they may well not learn to include *she* in their pseudo-generic *he*”. Such arguments are realised in the studies I discuss in section 1.2 which indicate that generic *he* has a default masculine meaning.

In this section I have shown that the English personal pronouns are a closed-class which, according to current scholarship, such as Segalowitz and Lane (2000), will be processed differently in the brain to open-class forms. Furthermore, it is generally agreed upon in the literature that closed-classes display certain characteristics, such as being low in semantic value (see below) and resistant to change. However, the historical development of the pronouns, which I discussed in section 1.1.2, shows that, although infrequent, changes have occurred. Whatever the influencing factors, such as the language-internal and -external factors noted in section 1.1.2, there is evidence in accounts of the development of English that these changes take

place slowly over time. In response to this position I addressed the issue of pronoun acquisition and I argue that there is evidence to support the idea that successive generations acquire different paradigms due to differing L1 input, thus facilitating pronoun change. In terms of epicene pronouns, this argument seems to favour singular *they*, as not only has it been shown to occur in speech (c.f. Newman 1992; Stringer and Hopper 1998; Pauwels 2001), its acquisition also represents a pattern in the normally irregular paradigm, and I argue that it parallels the acquisition of the second-person pronouns.

My review of the literature also indicated that another related characteristic attributed to closed-classes, and thus pronouns, is that they are low in semantic value. Yet there is evidence that pronouns do carry semantic weight, as in Early Modern English *you* had connotations of politeness (section 1.1.2), and Waryas (1973) argues that children learn the semantic value of pronouns before mastering their syntactic forms. Therefore in the following section I address arguments about the semantic salience of gender marking and review studies on epicene comprehension, in order to show that singular *they* and generic *he* are not equally weighted as epicene candidates in the literature on the use and understanding of epicene pronouns.

1.2: THE PSYCHOLOGICAL REALITY OF EPICENE PRONOUNS

A review of the literature indicates that, as well as being characterised as resistant to change, as the above discussion has shown, closed-class items are also perceived to be low in semantic value. For example, Segalowitz and Lane (2000:337) state that closed-classes give structure to language, but lack a level of “independent meaning”. To incorporate this argument into the current discussion of epicene pronouns means that, if this position holds and function words do lack semantic value, then epicene research is inconsequential, as pronominal form would not influence comprehension. It would not matter whether singular *they*, generic *he*, or any other pronoun

was used generically, as neither form would influence the gender (or proposed biological sex) of a generic referent.

In order to address this issue, section 1.2.1 begins with consideration of the arguments surrounding the semantic content of personal pronouns, showing that they are generally perceived in the literature to be grammatical markers with little semantic significance, primarily as a result of their closed-class status. However, a survey the vast literature on the interpretation and comprehension of pronouns in section 1.2.2 indicates that the obligatory gender marking on third-person singular pronouns in the standard paradigm is used in comprehension, and thus the forms carry semantic weight. This section also indicates that research has shown that singular *they* and generic *he* are not processed with equal ease, as the results of reading time studies have indicated that generic *he* can cause a gender mismatch with feminine stereotyped antecedents, which results in additional mental processing. Significantly, the same studies indicate that the same problem does not occur for singular *they*.

1.2.1: THE SEMANTIC SALIENCE OF GENDER MARKING

A review of the literature on English pronouns highlights the position that the semantic contribution of the closed-class forms is low. De Vincenzi (1999:538) suggests that pronouns are merely “noun phrases with little semantic content,” and similarly Newman (1992:453) notes that they are “seen as nominal stand-ins with no real semantic contribution of their own.” The extreme of this view can be attributed to Lakoff (1975), whose views on generic *he* were heavily influenced by the argument that pronouns could not carry semantic (and/or social) weight:

Feminist language reformers who wished to abolish epicene *he* were linguistically naïve because they did not understand that grammatical features, whatever their origins, could not themselves be derogatory

(Lakoff 1975 cited in Newman 1992:450).

Whilst it is true that there is nothing inherently sexist and/or derogatory about the morphological and phonological form *he*, its use as an epicene clearly has a social impact - an argument that I discuss in depth in

chapter two. As I also indicate in the next chapter, the language reformers Lakoff discussed were not rallying against the morphological form *he*, but against the male-as-norm view they associated with its generic usage (see section 2.2.3). This indicates that society, or at least groups within society, have attached a semantic value and a related world-view to the use of *he* in generic contexts. Thus, in relation to the argument that pronouns have low semantic value, the male-as-norm interpretation of generic *he* (evident in the studies discussed below) undermines Lakoff's argument. It also provides evidence for the position, albeit in opposition to general claims in the literature about closed-classes, that the third-person pronouns are not merely replacements for full NPs.

Despite their closed-class status, there is evidence in the literature for the argument that the third-person pronouns cannot be labelled 'low in semantic significance'. Firstly, the English pronouns have overt case marking, and thus carry more information about their place in syntactic argument structure than lexical NPs. Furthermore, this formal and overt gender marking is extremely restricted in English, as Newman (1992:448) notes that "apart from a few suffixes on professions and proper names, morphosyntactic expression of gender in English is a pronominal phenomenon".

There is also evidence in the literature for the argument that a pronoun paradigm where every human antecedent in the third-person is marked for gender must have semantic salience, especially when grammatical gender marking is no longer a core feature of the language. For example, Garnham, Oakhill and Cruttenden (1992:253) state that people notice when "the wrong pronoun is used" and therefore "it is natural to assume" that people will use pronominal gender "to help them distinguish between different possible referents" (1992:232). Similarly, Arnold, Eisenband, Brown-Schmidt, and Trueswell (2000:B14, henceforth Arnold et al.) claim that it is "obvious that pronoun comprehension should be guided by the gender information" provided. It is therefore surprising that some would deny gender-marked pronouns their semantic contribution, when clearly, through its rarity, gender marking must be salient in comprehension. If it was not significant it would not be used, and we would not notice its omission in speech or writing.

The proposed salience of pronominal gender marking is reinforced by cross-linguistic data from existing scholarship, which shows that formal marking of pronominal gender is extremely rare. Bhat (2004:109) noted that in the pronoun paradigms of the 225 languages she studied only 62 marked gender (“or noun class”) on the third-person. There were only ten languages where gender was marked on the second-person, and none marked it on the first-person singular. This finding can account for Corbett’s (1991) claim (discussed in the introduction) that epicene pronouns are rare cross-linguistically. As the vast majority of the pronominal systems studied by Bhat did not mark grammatical gender in the pronouns, their paradigms had no need for an epicene pronoun²⁹.

Contextualising gender marking within the English language as a whole, Kreiner, Sturt, and Garrod (2008:240) note that there are three different ways of marking gender on English open-class words. Their three categories of how gender could be signified are a) words can be morphologically marked for gender, such as *waitress*; b) gender can also be “specified in the lexical entry” as with *boy*; and c) there are also words which are syntactically gender-neutral but stereotyped, such as *soldier*. Although Kreiner, Sturt and Garrod’s classification focused on open-class words only, there is no evidence in the current literature to argue that morphological gender marking is fundamentally different for closed-classes. Therefore, if the premise holds that all of these ways of conveying gender are used in comprehension, and the morphological gender marking on third-person singular pronouns is an example of one of Kreiner, Sturt, and Garrod’s categories, then I argue that there is no reason why pronouns where gender is formally marked should be denied a semantic contribution to comprehension.

Indeed, the salience of pronominal gender marking can be illuminated by research on epicene pronouns; whilst generic *he* is formally syntactically singular, Baranowski (2002:279) notes that arguments levied against it “most often refer to its ambiguity...and equating maleness with humanness”. The sentences in (1) illustrate that pronoun interpretation allows one sentence to have multiple meanings even though the syntax is consistent. In (1a), the

29 This is also the case with BSL pronouns (see Sutton-Spence 1999:385), although signers may add “a spoken component”, where the mouth moves but the voice is not used, to create a distinction between masculine and feminine.

NP and both pronouns are coindexed, whereas the pronoun *his* in (1b) introduces an external referent.

- (1) a) *A nurse_i must believe in his_i training.*
b) *A nurse_i must believe in his_j training.*

The use of *he*, far from being generic, arguably makes it salient that in (1a) the nurse referred to is biologically male (and in (1b) the referent may be assumed to be male). Thus, it is difficult here for generic *he* to select a female referent, and this difficulty results from the semantic specification of *he* as [+MASCULINE]³⁰.

Conversely, singular *they* does not cause this problem, and related arguments that it causes referential ambiguity because it is formally plural (section 1.2.3, section 2.1.2) appear unfounded. Singular *they* does not cause ambiguity in the number of real-world entities it selects:

- (2) a) *A secretary should practice their typing.*
b) *A secretary should practice his typing.*

It is clear that in (2) only one secretary is being referred to by each pronoun. This means that whilst the use of singular *they* may be “*inherently* prone to ambiguity of reference because it increases the number of potential antecedents in the discourse” (Frank and Treicher 1989 in Newman 1992:456), it only selects one real-world entity. Thus, the semantics of singular *they* do not appear problematic for epicene reference, although the sentences in (1) suggest that epicene readings of *he* may be problematic. I discuss this further in the next two sections.

1.2.2: WHAT EPICENES MEAN

Whilst my discussion so far has been largely theoretical, focusing on arguments in the literature over the application of closed-class characteristics to third-person pronouns, in this section I document the results of experimental studies on epicene comprehension and production. The experiments I summarise use a variety of techniques, including sentence

³⁰ Square brackets denote phi-features.

acceptability tests, pronoun matching exercises, reading time tests and measurements of the mental images created by particular pronouns. The results from the studies generally converge to illustrate that the pronoun *he* has a default masculine interpretation, even when it is used generically. My goal in this section is to show that there is evidence in the literature that suggests that singular *they* does not impact upon the presumed biological sex of a gender-neutral referent.

To test how different epicene pronouns are comprehended, Martyna (1980) asked 72 university students to judge whether a set of sentences matched corresponding pictures. The test sentences in her experiment included either generic *he*, singular *they*, or the combined pronoun *he* or *she*. Martyna argued that if *he* was truly generic then pictures that included females only would “always be judged applicable to the sentences with generic *he*” (1980:73). However, her results showed that almost 20% of the students deemed that the female pictures did not apply to sentences including generic *he*, although the results were not uniform, as “the same students often reported contradictory judgements” where they sometimes perceived *he* as generic and sometimes perceived it as masculine (1980:73). To see if this irregularity was a feature of the experimental procedure Martyna questioned the students on whether they had “noticed the *he* in the critical sentences” and if it had consciously affected their judgements. Over 50% said that *he* had gone unnoticed (1980:74).

To reinforce her results, Martyna conducted a second experiment where the test sentences and pictures were presented simultaneously, and told the participants that they would be asked questions following the experiment. In this second experiment the rejection of female images with generic *he* increased, as circa 40% of the participants “reported that the female picture did *not* apply to the sentences with generic *he*” (1980:74). This led Martyna to claim that *he* facilitates a masculine interpretation and excludes women (1980:74). She argues that whether generic *he* is interpreted as masculine 40% of the time, or even “if it occurs only once”, it fails as an epicene pronoun (1980:74).

In a similar study, by Wilson and Ng (1988), participants read aloud a sentence which contained either “masculine or feminine pronouns or nouns”, before they were shown “a subliminal visual image of either a man’s or a

woman's face" and asked to assign a gender to the person (Mucchi-Faina 2005:204). Wilson and Ng (1988:167) were careful to note that their "tachistoscopical procedure" did not measure directly the images the participants created in their minds, but did allow them to "infer beyond reasonable doubt that images might have been induced" by the experimental sentences. The results showed that participants who "read phrases with masculine marks reported that subsequent faces had more male than female features" and the reverse was also true (Mucchi-Faina 2005:204). Therefore, the gender marking on the pronoun influenced how participants saw subliminal images of faces.

More evidence that generic *he* creates masculine images comes from Gastil (1990) who asked 93 participants (48 females and 45 males) to read aloud a set of six target sentences that included generic *he*, *he/she*, or *they*. The participants were then asked to describe any images that the sentences conjured. The results showed that in all conditions *he* evoked the most male images. Gastil supported his findings by noting that other research has shown *he* to cause women to "disappear from the population... in males' minds" (1990:631). However, in Gastil's experiment the same was not true for singular *they*, which was shown to be the most generic form, as it produced the highest number of mixed gender images and evoked the most female images from the male participants (1990:638). These results show that pronominal choice can have non-linguistic consequences; in the work of Gastil (1990), Martyna (1980), and Wilson and Ng (1988), it clearly influenced the way the participants physically saw the world.

This pronominal influence was illustrated in the work of Crawford and English (1984) who investigated whether the use of generic *he* could have an effect on memory. They tested whether 78 college students recalled a psychology essay better if it included generic *he* or combined pronouns. Their findings were that female students performed better on recall if they had read the essay which used combined pronouns, and thus explicitly included them, whilst the males tended to perform better having read the essay which used generic *he* (1984:378). The experiment was repeated with one hundred students reading an essay on "law as a profession" (1984:378), and only the recall of female students was significantly affected by the pronoun condition. Interestingly, when asked which essay they had read, the one with generic *he* or combined pronouns, "very few" of the students could

say (1984:380). This agrees with Martyna's (1980) results and suggests that although the choice of pronouns did not attract the participants' attentions, they were subconsciously affected by the use of generic *he*. Thus, it is arguable that Crawford and English's female participants were excluded from the possible referents of the essays containing generic *he*, and may have assumed that the essays were not relevant to them.

The fact that the participants appeared unaware of the pronouns used could undermine the argument, discussed above, that the gender marking on third-person pronouns is salient enough to have an effect on comprehension. However, the studies discussed above show that whilst pronouns may not always stand out in comprehension, it is clear that their semantic value can affect unconscious processing. One piece of research which clearly showed how pronoun choice can influence processing time is Foertsch and Gernsbacher's (1997) study on the reading times of epicene pronouns. Their research focused on the acceptability of singular *they*, and Foertsch and Gernsbacher assumed that any processing difficulties caused by a gender mismatch between antecedent and pronoun would mean a sentence would take longer to read (and thus longer to process). Their hypothesis was that singular *they* would not cause such processing difficulties.

Foertsch and Gernsbacher gave 87 participants test sentences like that in (3) with antecedents which were either indefinite pronouns, or belonged to one of three types: neutral (*a runner*), male stereotyped (*a truck driver*) or female stereotyped (*a nurse*).

- (3) *A truck driver* should never drive when sleepy, even if *he/she/they* may be struggling to make a delivery on time, because many accidents are caused by drivers who fall asleep at the wheel.

(Foertsch and Gernsbacher 1997:107).

The variables were the pronouns *he*, *she*, and *they*, and each sentence occurred in the experimental data once with each pronoun. Each sentence was presented clause-by-clause on a computer screen, and participants pushed a button to move on to the next clause when they had finished reading. This meant that their reading time for the pronominal clause could be isolated. At the end of each sentence, participants were presented with a yes/no or true/false question to distract them. The theory behind this

method is that it tests pronoun comprehension independent of any conscious opinions on epigenes that the participants may have.

The results support Foertsch and Gernsbacher's hypothesis that if reading times were not significantly affected by the use of singular *they*, then "the argument that singular *they* 'violates the expectations of most readers'³¹ would not be empirically supported" (1997:107). In the sentences with stereotyped antecedents³², the opposite-gender pronouns were read slowest (e.g. *she* used with *truck driver*), however, singular *they* and the pronoun that gender-matched the stereotype "were read with equal facility" (1997:108). In other words, a generic interpretation of *they* was processed just as easily as a gender-specific interpretation of the stereotyped pronoun, whereas opposite-gender pronouns were not processed as easily. In addition, singular *they* was "the pronoun of choice" (1997:108) when read with indefinite pronouns (see section 1.3.1), suggesting that it was the most generic pronoun tested. Foertsch and Gernsbacher then tested how reading times were affected when the antecedent of the pronoun was definite (e.g. *that truck driver*) as such constructions suggest that the speaker/writer knows the sex of the referent. In this experiment, the stereotypical gender-matched pronouns were always read most quickly, but singular *they* was again favoured over the opposite-gender pronoun³³.

In a similar experiment involving pronouns and gender stereotypes Carreiras, Garham, Oakhill and Cruttenden (1996) had thirty participants rate 120 professions on "the likelihood that each would be done by either a man or a woman" (1996:643). Then, based on these stereotype scores they had a further 24 participants read sentences which contained masculine stereotyped (e.g. *the electrician*), feminine stereotyped (*the babysitter*) or neutral roles (such as *the psychology student*) and either a matched or mismatched pronoun. Carreiras et al.'s results showed that test frames including gender-stereotyped roles with mismatched pronouns (*the footballer...she* for example) were read slower than sentences where the

31 This 'violation of expectation' relates to the rejection of singular *they* by traditional prescriptive grammarians, a development which is considered in chapter two.

32 I consider the impact of gender stereotypes on pronominal choice further in chapter three.

33 Foertsch and Gernsbacher conclude by claiming that "singular *they* is an acceptable substitution for gender specific pronouns with nonreferential antecedents" (1997:110). As supplementary evidence they note that 51% of those tested "did not believe using *they* in place of *he* or *she* is ungrammatical" and state that when participants provided pronouns for test sentences "spontaneous use of singular *they* was common" (1997:110).

gender of the pronoun matched the stereotypical gender of the social role. This led Carreiras et al. to conclude that, “we know from the mismatch condition that stereotype information is activated when the pronoun is read” (1996:646).

Similarly, Kennison and Trofe (2003) had eighty participants rate 405 different nouns (and noun compounds) for gender stereotypes on a rank scale from 1.0 to 7.0, which represented a cline from feminine to masculine. Using a reading time experiment, they tested whether the pronouns *he* or *she* could create a gender-mismatch with the stereotyped noun, which would result in longer reading times. Kennison and Trofe (2003:366) found that the gender stereotype did affect reading times, as it took participants longer to read the “two regions following the pronoun *he* or *she* when the gender stereotype and the antecedents mismatched the gender of the pronoun” when compared with the reading times for test frames where there was no gender mismatch. Again these results are similar to Foertsch and Gernsbacher’s (1997) work, and provide more weight for the argument that gender stereotypes, although they are not syntactic, can clash with the syntactic gender marking on third-person pronouns. The results of these three studies illustrate the argument that the overt gender marking, and thus semantic values, of *he* and *she* are used in comprehension, and affect unconscious language processing. However, as singular *they* is not marked for gender, this effect cannot occur.

Indeed, Sanford and Filik (2006:172) argue that “*they* and *them* carry the option of being gender-neutral singular as part of their specification”, which led them to conclude that “a gender-neutral singular is perfectly acceptable from a processing point of view” (2006:172). To test whether this is the case, Sanford and Filik used an eye tracker to see whether their 36 participants had trouble reading *they* with a singular antecedent³⁴. They found that when the antecedent was singular “there was no difference” in processing between sentences containing *he* or *she* and (plural) *they* (2006:174). Thus the processing of *they* “was not reliably affected” when it referred to a singular antecedent (2006:175).

34 See Ehrlich and Rayner (1983) for a discussion of how eye-tracking experiments inform theories about how words are processed and how pronoun/antecedent resolution occurs. See also Choy and Thompson (2005) for discussion of eye tracking and pronominal comprehension.

Sanford and Filik's data does show a small deficit in total reading time when singular *they* is used, but I believe this may be explained by their participants' encounters with the traditional grammatical proscription of singular *they* (see chapter two). In other words, their overall reading time could have been influenced by their perceptions of whether singular *they* is "correct" as by the end of the sentence the participants would have had slightly longer to think about what they had read. In any case, the results led Sanford and Filik to conclude that on initially encountering the pronoun *they* a person will search for a plural antecedent, but due to its common use, singular *they* "is rapidly accommodated as an acceptable deviation" (2006:177)³⁵. Indeed, Sanford, Filik, Emmot and Morrow (2007:373) argue that "*they* seems to be very tolerant" of different antecedents, and my review of the literature on epicene coreference in section 1.3.1 provides more evidence for this claim.

Another study, which links back to the discussion of pronominal acquisition in section 1.1.3, was undertaken by Bennet-Kastor (1996), who recognised that there was very little research concerning children's epicenes³⁶, as the vast majority of empirical studies on this topic deal with adult comprehension. She argues that as a result, the processes surrounding pronominal acquisition were obscured. Yet Fisk (1985) had already found that whilst generic *he* "is always interpreted as masculine", the same is true for singular *they* "prior to first grade in boys and prior to kindergarten in girls" (Bennet-Kastor 1996:287). This led Bennet-Kastor to test 26 children aged between 9 and 12³⁷ on whether sentences including the pronouns *he*, *she* and *they* referred to men, women, or both. She controlled for gender stereotypes using antecedent nouns that were stereotypically masculine (*carpenter*, *minister*), stereotypically feminine (*librarian*, *nurse*) or stereotypically neutral (*child*³⁸, *singer*). Bennet-Kastor also used a mixture of

35 Section 1.3.3 will take issue with the argument that singular *they* is a deviation of the singular form.

36 I address this issue in more detail in chapter six where I analyse a corpus of child language transcripts.

37 In relation to data on pronoun acquisition (section 1.1.3) the children in Bennet-Kastor's study should already have acquired their adult pronoun paradigm as they were beyond the age of pronominal acquisition. I deal with children's epicene production in chapter six, looking at data from much younger children who are within the age boundaries of pronoun acquisition.

38 I argue that the antecedent *child* may not be neutral in this experiment as it could have interacted with the biological sex of the child completing the task, however it does not appear to skew Bennet-Kastor's overall results.

definite and indefinite articles distributed equally across the stereotyped/neutral conditions as a control for reference (see Foertsch and Gernsbacher 1997 discussed above). The children were also tested on their awareness of gender stereotypes, as in a second task they were asked to identify whether the antecedents used in the experiment were “mostly male”, “mostly female” or “either male or female” (1996:289).

The results showed that the children in the study perceived gender-marked pronouns generically in a small minority of cases (4% for *he* and 2% for *she*) - this is what Bennet-Kastor terms “ANAPHORIC GENERIC STRATEGY” (1996:290). The majority however, perceived *he* and *she* as exclusively male and female (53% and 61% respectively) - this is the “ANAPHORIC GENDER STRATEGY” (1996:290). However, the children did not consistently use one strategy or the other across all antecedents. With reference to singular *they* Bennet-Kastor found that the children thought 30.5% of the tokens of *they* could refer to either men or women, and she argued that *they* had the “strongest tendency to induce ANAPHORIC GENERIC readings” (1996:292-293). Bennet-Kastor noted that whilst generic interpretation of *he* was not absent from her results, it was “extremely rare”, whilst *they* was “most likely to be interpreted neutrally” (1996:298).

To support her results, Bennet-Kastor had 27 different children produce stories about stereotypically male antecedents (e.g. *dentist*), stereotypically female antecedents (*secretaries*), or stereotypically neutral antecedents (*artists*, or *children*). The children produced three stories each, based on prompts that contained a definite article, an indefinite article, or a plural antecedent. The results showed that 13.1% of children used *they*, with two thirds using it to refer to singular antecedents (both definite and indefinite), 11.9% used *he* and 5.6% used *she* (1996:295). This indicates that as well as being perceived as the most gender-neutral pronoun, singular *they* was also the most favoured epicene produced by the children.

Based on a similar method, Moulton, Robinson and Elias (1978) asked college students to write a story based on a stimulus sentence that contained either generic *he*, singular *they*, or a combined pronoun. They documented the gender chosen for the main character in 490 different stories and found that only 35% of those written by students given a stimulus including generic *he* had female main characters. This increased to 46%

when singular *they* was used in the prompt, and 56% in stories based on a prompt containing a combined pronoun. These results suggest that using generic *he* blocked a female interpretation of the antecedents and thus, excluded women from being prominent in the stories³⁹.

In some of the most recent work on epicene pronouns Strahan (2008) looked at pronoun choice in the written work of first year university students in Australia. She showed that in the abstract, introduction, and methodology section of seventeen essays on child language acquisition, the students had a slight preference for *they*, with a strong tendency to use it in conjunction with gender-neutral NPs, such as *the child*, even when the students knew the sex of the child that they were writing about. This finding led Strahan to argue that *they* is “not just a third person plural and third person singular ‘indefinite gender’ or ‘general’ pronoun, but it is a third person ‘gender not relevant to discussion’ pronoun” (2008:27), which suggests that the students were using *they* specifically because it was not marked for gender. Strahan’s results reflect much older studies in epicene production, as Green (1977:152) found clear trends in the use of singular *they*, showing that it was “normal usage” for his 184 college-level participants.

The main point to be taken from this review of empirical studies on epicene production and comprehension is that there is a strong body of evidence to support the proposition that gender-marking on third-person singular pronouns is salient in comprehension and affects readers’ perception of possible antecedents. The studies also indicate that generic *he* has a default masculine interpretation but the same is not true for singular *they*, which it appears, can be processed unproblematically with singular antecedents. However, very few of the studies I have evaluated were completed in the last ten years, and thus my review of the literature has also indicated that there is not much data on recent or current epicene usage. One potential reason for this lack of modern data is the overwhelming tendency for studies to provide complementary results.

³⁹ Interestingly, there were two different story themes used in Moulton, Robinson and Elias’ experiment and this also interacted with the gender chosen for the main characters. One set of stories was based on the concept of “the average student” with 39% female main characters (1978:1034), whilst a second theme, involving attitudes to attractiveness, had 52% female main characters (1978:1035). Thus whilst generic *he* played a part in the selection of gender, theme also had an effect. Yet rather than weakening the above conclusions about generic *he*, the interaction of theme and gender suggests that gender stereotypes can influence pronominal choice.

Nevertheless, it is important to collect modern data on epicene usage in order to keep the literature on the epicene debate up to date and to track any changes in pronoun usage.

The results of the studies discussed above indicate that the initial masculine interpretation, and thus semantic contribution, of generic *he* may preclude its use as a gender-neutral pronoun. However, Baron (1986:191) notes that this apparent (semantic) gender violation is “either ignored, or rationalized” in traditional grammar (and beyond) by the claim that the masculine form includes the feminine (see chapter two). Clearly though, the evidence from the studies suggests that “*he* has an inherently masculine meaning” (Newman 1992:453) and plays a semantic role in sentence comprehension. In fact, Gibbon (1999:45) claims that work by Ng (1990) shows “conclusively” that *he* (and related terms like *man*) “are coded in memory as members of the masculine linguistic category”. Indeed, Crawford and Chaffin (1986) note that the semantic value of generic *he* initially decreases the number of possible referents of a pronoun because of its predominant maleness:

When both men and women read the word *he*, a male interpretation (the default value) initially predominates. But if women are not to exclude themselves from what they read, they must do additional mental processing to transform the initial literal interpretation into one that includes them. Thus, they suppress male imagery associated with *he* and avoid its generic use (and the necessity for the transformation process) when writing

(Crawford and Chaffin 1986:16).

The additional processing proposed by Crawford and Chaffin is one possible reason why the use of generic *he* with a stereotypically feminine antecedent results in longer reading times (c.f. Foertsch and Gernsbacher 1997). Furthermore, work by Osterhout, Bersick, and McLaughlin (1997) indicates that generic *he* can trigger brainwaves that parallel those triggered when a person processes a syntactic anomaly. Looking at ERP (Event-Related Potentials) in the brain, Osterhout, Bersick, and McLaughlin showed that there was a processing difficulty when pronominal gender did not match that of stereotypically gendered antecedents. Using reflexive pronouns they found that stereotypical gender disagreements caused a brainwave “similar to the P600 effect” (1997:273).

The P600 (or syntactic positive shift) is a large positive wave, detectable when measuring event-related brain potentials, which appears to be triggered by “syntactic anomalies” (1997:273). Osterhout, Bersick, and McLaughlin showed that the brainwave was also triggered by gender mismatches caused by pronouns, suggesting that the brain was treating gender agreement as a syntactic phenomenon⁴⁰. This means that although gender stereotypes are not syntactic in nature, the brain responds to them as if they are. More importantly, the input that triggers these brainwaves is the semantic data encoded on the pronouns. This supports the claim that “gender marking is a highly salient aspect of a pronoun” (Garnham, Oakhill and Cruttenden 1992:236). However, they did not directly test whether the use of singular *they*, or any other generic pronoun, affected the P600 brainwave.

If Garnham, Oakhill and Cruttenden’s argument holds then when faced with an indefinite referent or generic NP, such as *a person*, any choice of pronominal gender (generic *he* or generic *she*) will arguably influence the gender of the imagined referent. Indeed, such an argument is supported by the findings of Martyna (1980), Wilson and Ng (1988) and Gastil (1990) discussed above. Furthermore, McConnell-Ginet (1979) argues that “it is difficult” for a human being to “imagine persons as unsexed” (Newman 1992:454). Arguably then, if the gender marking on pronouns is of semantic significance, it will influence the sex of an imagined referent, again, as was illustrated by Martyna (1980) and Wilson and Ng (1988), discussed above. Indeed, the experimental evidence does suggest that *he* creates mainly male mental images even when used generically. However, I argue that even if these generic images, or “prototypes” (Newman 1992:454), are imagined as a specific sex, as McConnell-Ginet suggested, if used with singular *they* inferences about the imagined sex of the referent would not, and indeed could not, be determined by the pronoun choice, due to the absence of morphological gender marking on the pronoun *they*.

In this section I have shown that there is evidence in the literature for the argument that third-person pronouns do carry semantic value that is

40 See also Osterhout and Holcomb (1992).

used in comprehension, and most likely in production. The studies I have reviewed show that there is a clear argument that generic *he* encourages a male interpretation. On the basis of this evidence I argue that if pronouns did not have their own semantic value then a masculine interpretation of *he* could not occur. Newman (1998:355) argues that if pronouns are seen as mere stand-ins for full NPs “we are left with an incomplete understanding of how pronouns function as referring expressions”. The view that pronouns merely match their antecedent in terms of syntactic structure “implies [that] pronouns are static elements” and cannot account for such usage as *they* referring to collective nouns or being used as a singular epicene (Newman 1998:355). Newman claims that to be surprised at the masculine nature of *he* means starting from “a theoretical stance that denies pronouns a contribution to meaning beyond mere designation” (1998:365).

The logical implication taken from my synthesis of the relevant literature is that acknowledging the semantics of pronouns appears to favour singular *they* as epicene pronoun, as it is not formally marked for gender. In contrast, the proposed default masculine value of generic *he* does appear to affect the conceptualisation of generic referents. The results from the studies discussed above suggest that singular *they* and generic *he* are not equal epicene candidates in terms of sentence processing and production. However, as I briefly noted above, singular *they* does not seem able to shake the prescription that it is purely plural, and therefore cannot be conindexed with singular antecedents. I explore the language-external reasons for this prescription in chapter two, whilst in the following section I evaluate the literature on the coreference of generic *he* and/or singular *they* with different types of syntactically singular antecedents.

1.3: DIFFERENCES BETWEEN GENERIC HE AND SINGULAR THEY

My review of previous epicene studies has established that generic *he* and singular *they* are the two main candidates for epicene status in English. Based on the above review of current scholarship, in this section I focus on these two forms in detail and assess their viability as epicene pronouns by

focusing on considerations in the wider literature of their syntactic agreement with different types of antecedent. In section 1.3.1 I consider arguments in the literature about whether generic *he* and singular *they* cause referential ambiguity, and I consider the issues of compatibility between singular *they* and generic *he* with different types of antecedents, starting with indefinite pronouns, followed by NPs with quantifiers, indefinite NPs and finally definite NPs. My goal in this section is to compare and contrast the two epicene candidates in terms of syntactic agreement and highlight any differences between the two forms.

Having established that there is evidence in the literature for the use of singular *they* with varying types of antecedent, in section 1.3.2 I address the most prominent theories in the current literature which attempt to account for the syntactic (dis)agreement between singular *they* and formally singular antecedents. My discussion indicates that all of the current theories struggle to account for coreference between singular *they* and singular definite NPs. In response to this I consider one final theory of pronouns, proposed by Whitley in 1978, which has not received much attention in the wider literature, but which, when applied to epicene pronouns may suitably account for the problematic coindexation of singular *they* and all formally singular antecedents.

1.3.1: AGREEMENT WITH DIFFERENT TYPES OF ANTECEDENT

Despite the large body of research providing support for the argument that singular *they* is a viable epicene, the pronoun does not appear able to avoid the prescription (discussed in detail in chapter two) that “it is grammatically incorrect to use plural pronouns to refer to a singular antecedent” (Madson and Hessling 1999:561). Indeed, Weidmann (1984:68) refers to the plural marking of *they* as its “only drawback”. The traditional prescriptive grammarians to be discussed in chapter two who proscribed the use of singular *they* did so by arguing that a pronoun had to agree in gender and number with its antecedent. The literature indicates that these grammarians regarded *they* “as purely plural” meaning that it “disagreed in number” with formally singular antecedents (Baranowski 2002:378).

However, there is evidence in the wider literature that, not only is singular *they* used for singular antecedents, in some cases (at least) it is the preferred pronominal form over generic *he*. To this end, in this section I consider different types of pronominal antecedents, beginning with indefinite pronouns and moving through increased levels of syntactic definiteness to definite singular NPs.

When discussing how coreference is achieved between pronouns and indefinite pronouns, Kolln (1986) used example sentences to show that (in most cases) generic *he* and *everyone* cannot be coindexed (4b). For sentence (4) to be grammatical, *he* (whether it is generic or not) must have an external referent (4a). However, singular *they* does not pose this problem and can be coindexed with the indefinite pronoun (4c).

- (4) a) At first *everyone*_i in the room was singing; then *he*_j began to laugh.
b) *At first *everyone*_i in the room was singing; then *he*_i began to laugh.
c) At first *everyone*_i in the room was singing; then *they*_i began to laugh.

(Adapted from Kolln 1986:101)

This is part of the evidence Kolln used to try and persuade teachers not to ‘correct’ the use of singular *they* with the indefinite pronoun *everyone*, even though it is prohibited by traditional prescriptive grammar (c.f. section 2.3.1). In a similar vein, Balhorn (2004:84) used illustrative sentences to show that generic *he* can actually make “sex a ‘salient property’ of the referent” when used with an indefinite pronoun (5a). Again, this is not a problem for singular *they* (5b).

- (5) a) *Somebody* called when you were out and *he* said *he’d* call back later.
b) *Somebody* called when you were out and *they* said *they’d* call back later.

(Balhorn 2004:84)

Balhorn’s argument corresponds with Wales’ (1996:128) view that when processing generic *he* with indefinite pronouns “the reader or listener might well assume that only groups of men are being referred to”. In any case, Wales also notes that, whether generic *he* is gender-neutral or not, *he* “is remarkably scarce in co-reference with indefinite pronouns” (1996:130).

The sentences (4-5) highlight how indefinite pronouns do not obey traditional prescriptive rules on number agreement, as *somebody* and *everyone*, which are singular, favour the traditionally plural *they* over generic *he*. Newman's (1998:366) explanation for this is that indefinite pronouns have both "singular and plural aspects". He argued that the presence of singular and plural aspects also applied to generic, formally singular NPs (see below), which have plural aspects "because they refer to whole classes by a single example" (1998:366).

In contrast to Newman's arguments, Sklar (1988) does not think that underlying plurality is the overarching factor of pronominal agreement as it cannot account for why *everything* (which is clearly analogous with *everyone*) does not coindex with a plural pronoun, as in (6):

- (6) a) *Everything* looks dirty, doesn't *it*?
b) **Everything* looks dirty, don't *they*?

(Adapted from Sklar 1988:418)

Instead Sklar proposed that when an indefinite pronoun occurs with a plural pronoun "an animate/human substantiative is involved" (1988:418). From her perspective, animate indefinite pronouns, such as *everyone*, involve gender, but inanimate indefinite pronouns, such as *everything*, do not and so "it looks as if the third person plural pronoun occurs in response to gender, rather than number" (1988:418). One interpretation of Sklar's argument is to claim that singular *they* agrees with animate indefinite pronouns as it is specified for gender (as [-MASCULINE], [-FEMININE]) and is also [+ANIMATE]. Yet *everything*, which is inanimate, coindexes with *it*, an indefinite pronoun marked [-ANIMATE] with no gender specification⁴¹.

The discussion of epicene choice with indefinite pronouns also relates to formally singular NPs that are premodified by a quantifier (e.g. *every apple*, *any child*, *no woman*). Although these NPs are singular, and like

41 This complements Weidmann's (1984:62) argument that "the only solution" to grammatical coreference with indefinite pronouns "is to find an existing pronoun that will allow enlargement of its usage". Arguably, singular *they* is this pronoun as it is already specified as gender-neutral, and the use of a plural pronoun as a singular can be found in other areas of the pronoun paradigm (see section 1.3.3). Thus *they* fits Weidmann's criteria that "we should try to find a pronoun whose features come as close as possible to the specifications of our gap" (1984:62).

indefinite pronouns, take a singular verb, they can be coreferenced with either singular *they* or generic *he* as in (7), as long as the noun is [+ANIMATE].

- (7) a) *Each warden* must carry ID with *them* at all times.
b) *Each warden* must carry ID with *him* at all times.

Whilst there is arguably nothing structurally problematic about the coindexation of NPs with quantifiers and singular *they* (7b), Weidmann (1984:63) argued that as “every combines with singular nouns only”, this should “preclude they as an anaphoric pronoun”, I address the fact that it does not in more detail in section 1.3.2. Nevertheless, generic *he* is also problematic. Despite causing no syntactic problems the evidence from the studies discussed above indicates that the pronoun’s semantic value means that it is difficult to avoid a masculine interpretation of *he*, even though the antecedents in (7) are not necessarily gender-stereotyped. In any case, the sentences in (8) suggest that in attested examples - in this case data taken from newspaper articles in the *Daily Mail* and the *Guardian* - the epicene of choice for NPs with quantifiers is singular *they*.

- (8) a) Like *any girlfriend* with someone *they* care about serving on the front line, her emotions were all over the place.
(*Daily Mail*, adapted from Paterson 2011)
b) ...for *any woman*, waiting to hear whether or not *they* have breast cancer is an extremely stressful and worrying time
(Boseley 2008)

What is interesting about these examples is that in both cases, the sex of the pronominal referent is known, as *woman* and *girlfriend* are coded for gender in their lexical specification (see Kreiner, Sturt, and Garrod 2008:240). This is evident in (8a) where singular *they* is used to select any one of the collective term *girlfriend*, but later in the sentence a specific *girlfriend* is coreferenced with *her*. My interpretation of this is that the main factor in using singular *they* in (8a) is its indefinite nature; although *any girlfriend* can only refer to females, the use of singular *they* emphasises that there is no specific real-world referent for the antecedent. Replacing singular *they* with generic *he* in either (8a) or (8b) would create rather odd sentences.

In relation to the studies discussed in section 1.2.2, and more specifically Sanford et al.'s (2007:373) claim that *they* is tolerant of different antecedents, the examples from the literature do indicate that, as with indefinite pronouns, singular *they* is the epicene of choice for NPs with quantifiers. In relation to wider arguments about epicene pronouns, and the apparent acceptability of singular *they* illustrated by my synthesis of the studies on epicene pronouns in the previous section, the use of singular *they* with different forms of syntactically singular antecedents is more evidence for the position that it is the epicene pronoun of choice in English.

Another clear indication of the antecedent tolerance noted by Sanford et al. (2007) comes from Weidmann's (1984) discussion of the use of singular *they* with an indefinite singular NP. Using the example "If there is a Barbara Wassman on board, could *they* make *themselves* known to the cabin?", he argued that even though the sex of the referent could, in this case, be justifiably assumed to be female, what "*they* does is to reiterate the meaning of the indefinite article before the name" (1984:65). This complicates the role of singular *they*, as in this example, where *they* is coindexed with an indefinite, formally singular NP, it is doing more than just being gender-neutral. Weidmann argued that the pronoun allowed the speaker to express "uncertainty about the presence" of any such passenger (1984:65). Indeed, Newman (1998:369) uses Weidmann's example with his claim that "singular *they* cannot be accounted for only through appeals to plural notional number and sex-indefiniteness".

However, arguably Weidmann's use of singular *they* is a manipulation of its singular epicene form, which cannot be directly contrasted with generic *he* as it is performing a different function. This is, in some ways, similar to what Sanford et al. (2007) call 'Institutional They': when *they* (plural or singular) refers to an external referent not coindexed with another NP, as in "*They're* digging up the road again" (2007:378). In the *Barbara Wassman* example, singular *they* still coindexes with the indefinite singular NP, independent from its contribution to the semantics of uncertainty in the sentence. Thus there is still formal grammatical agreement between the NP and the pronoun insofar as the phi-features of the NP 'a Barbara Wassman' and *they* both include a [-PLURAL] specification. In relation to the agreement between singular *they* and indefinite NPs more generally, Weidmann (1984:63) argues indefinite NPs do not refer to one specific entity, but

rather to one entity within a homogenous group of entities, and thus the singularity of indefinite NPs is “superficial”. Similarly, Wales (1996:131) argues that an indefinite syntactically singular NP does not represent an individual entity; rather it is “a grammatical realisation of a concept” which is “a generic representation of a class”.

To illustrate this argument, the sentences in (9) show that using singular *they* with an indefinite NP does not have any apparent effect on the imagined referent, whilst generic *he* can influence meaning:

- (9) a) *A good girl* always says *their* prayers.
**A good girl* always says *his* prayers.
- b) *An actor* wishing to further *their* career must take some challenging roles.
An actor wishing to further *his* career must take some challenging roles.

If generic *he* is used in (9a) it creates a semantically odd sentence; as with the sentences discussed above in (8), the antecedent *girl* is marked as [-MASCULINE] for gender in its lexical entry (c.f. Kreiner, Sturt, and Garrod 2008:240). In (9b) generic *he* arguably changes the meaning of the sentence by indicating through gender marking that the referent is male. In both cases, singular *they* does not cause this problem and thus, based on the example sentences, and the data from Weidmann (1984), it also appears to be the most suitable epicene for indefinite NPs.

However, unlike indefinite pronouns and indefinite NPs, pronouns coindexed with definite NPs do refer to a specific entity, rather than to a generic member of a homogenous group. The use of a definite article suggests that the sex of the referent is known to the speaker/writer (see Foertsch and Gernsbacher 1997, section 1.2.2), and therefore their choice of pronominal gender should match accordingly. In (10a) and the use of singular *they* does not make the sex of the real-world referent explicit. On the other hand, as with (10b), the gender of the third-person pronoun coindexed with a definite NP has to match the biological sex of the referent. Therefore, I argue that this means a generic interpretation of *he* would not be possible in (10c).

- (10) a) *The driver* asked *their* passengers for the fare.
b) *The driver* asked *her* passengers for the fare.
c) *The driver* asked *his* passengers for the fare.

I thus propose, based on Strahan's (2008:27) argument, noted in section 1.2.2, that singular *they* is a "gender not relevant to discussion pronoun" and that the use of singular *they* with definite NPs thus is a stylistic choice. It serves to background the gender of the referent, presumably because their biological sex is not important to comprehension, or is not salient in the wider context. In relation to the studies I discussed in section 1.2.2 it appears that there is evidence for the position that the only epicene form available which will avoid gender marking when coindexed with definite singular NPs is singular *they*, as the alternative, generic *he* has been shown in many empirical studies to carry a default masculine value. The example sentences above, many of which are taken from the wider literature, indicate that singular *they* appears to be the preferred epicene form for all the tested antecedent types. Whilst in some instances, generic *he* is also syntactically grammatical, the semantic value of the pronoun, evident in the large body of literature I discussed in section 1.2.2, may interfere with a gender-neutral interpretation. Thus, singular *they* corresponds with the non-specific nature of generic reference, as it does not convey gender information.

Weidmann (1984:68) suggests that singular *they* gets around the issue of the semantic value of gender marking on pronouns because it "does three jobs" simultaneously, better than any other available form. In Weidmann's view, singular *they* "says nothing about the gender (or sex) of the referent", nor does it mark the referent for number (although it does agree syntactically with singular antecedents due to its phi-feature specification, see section 1.3.3), and also "it does not even specify whether a real referent exists" (1984:68). Weidmann argues that singular *they* can fill "an 'ecological niche' in the English language" as an epicene pronoun, as long as it can adapt to singular usage (1984:68). This argument assumes that singular *they* is a plural form masquerading as a singular form, but in the final two sections of this chapter I show that there is evidence in the literature to suggest that this is not necessarily the case.

1.3.2: NOTIONAL NUMBER AND OTHER EXPLANATIONS

Despite the scholarship that has shown that singular *they* is used as an epicene pronoun in both spoken (Newman 1992, Stringer and Hopper 1998, Pauwels 2001) and written (Baranowski 2002; Laitinen 2007; Balhorn 2009) English, the form *they* is still marked as syntactically plural in the current standard pronoun paradigm (see section 1.1.2). Indeed, there is currently no available scholarship that challenges the position of *they* as a syntactically plural form. In this section I review different explanations in the literature which attempt to account for the coindexation of singular *they* (theorised as a manipulation of the plural form) and syntactically singular antecedents. I evaluate explanations such as notional number, Polysemy Theory, and Homonymy Theory in order to determine whether any particular theory is currently held above any other for explaining the coreference of singular *they* with formally singular antecedents.

To begin with, the most common explanation in the current literature for epicene agreement between *they* and formally singular antecedents is *notional number* (as used by Newman 1992; Newman 1998; Baranowski 2002; Bock et al. 2004; Humphreys and Bock 2005). Notional number is the feature of an antecedent which “specifies how many entities [it]... refers to” (Baranowski 2002:383), and may contrast with its syntactic number marking. According to Newman (1992:458) notional number differs from syntactic number in that it “is probably best viewed as a cline” between the endpoints of singular and plural⁴². A good illustration of notional number is the discrepancy between British and American English over whether collective NPs are singular or plural (see Hundt 2006:207). Another example comes from Bock, Eberhard, and Cutting (2004:253) who noted that although *news* is a singular form, the “typical referent of *news* is notionally plural”.

In reference to pronouns, Baranowski (2002:390) showed the influence of notional number on indefinite pronouns, arguing that in (11) “the semantic plurality [of *anybody*] is so strong that it affects the morphology of

42 Newman (1998:336) likens this distinction to the difference between digital and analogue signals, where the first (syntactic number) is formed of binary distinctions and the second (notional number) is continuous.

the noun.” This is because the plural form *tongues* does not agree in syntactic number with the singular form *anybody*⁴³.

- (11) a) *Anybody* who goes to the Regency Rooms with *their* tongues anywhere other than in *their* cheeks is likely to be sorely disappointed.

(Baranowski 2002:390)

Yet the singularity of the sentence is evident from the verbal morphology which agrees with the subject NP [anybody]. It is not difficult to find complementary examples to support this analysis - e.g. ‘everyone sang their *hearts* out’ - and the example in (11) works for other indefinite pronouns, such as *everybody* and *noone*. Although the meaning of the sentence is changed by the use of *no one* (11b), the number (dis)agreement is constant.

- (11) b) *Noone* who goes to the Regency Rooms with *their* tongues anywhere other than in *their* cheeks is likely to be sorely disappointed.

To illustrate how notional number can influence pronoun choice Baranowski (2002) catalogued the epicene forms used in a corpus of issues of *The Independent* newspaper. He divided their syntactically singular antecedents into three groups: notionally singular, notionally plural and an indeterminate (neutral) group. Whilst the classifications may have been somewhat subjective, Baranowski’s results showed that singular *they* was the epicene of choice for all three groups, holding a 93%, 67% and 51% share for the plural, neutral and singular antecedents respectively (2002:386). Baranowski supported this numerical data by giving examples from the corpus, similar to those in (8), which showed that singular *they* was used even if its antecedent was of lexically-specified sex⁴⁴:

- (12) In *her* late twenties a fairly average PR woman can earn over 50 grand, more than most of the journalists with whom *they* deal

(Baranowski 2002:389).

43 Arguably, the sentence would still be grammatical if it contained the singular *tongue*, but *cheeks* would also have to be singular.

44 It can also explain anomalous results in the data of others, for example Meyers (1993:198) discussed an example where the antecedent of singular *they* was *everybody’s grandmother*, a term clearly marked for gender. One explanation for this is that the use of *everybody* increased the notional number of *grandmother*, meaning that the referent of the NP was merely one person out of a larger group.

Supplementary evidence for how notional number accounts for singular *they* is Weidmann’s (1984) view that singular *they* creates “a less sharp focus on any particular individual” (Newman 1992:453), and one could argue that this is what makes it truly generic. Arguably, according to Baranowski (2002:389) notional number displaces “formal, i.e. syntactic, singularity” and Baron (1986:193) argues that it represents “semantic concord in English overriding grammatical concord”.

Turning this argument on its head, Newman (1992) suggested that the concept of notional number accounts for the perceived semantic singularity of *they*. Thus, not only are indefinite pronouns, NPs with quantifiers, and indefinite singular NPs notionally plural, *they* in its epicene form is also notionally singular. Newman argues that “singularity should not so much be ascribed to the pronominal form itself but that at best it can signal noncanonical coreference relation...with a formally singular NP” (1992:458). So, if notional number holds, whilst the underlying plurality of indefinite pronouns can account for their coreference with *they*, similarly perhaps an underlying singularity of *they* can account for its coreference with definite NPs. Figure 1.1 represents how notional number interacts with indefinite pronouns and formally singular NPs, illustrating the argument that the notional plurality of indefinite pronouns and the notional singularity of *they* are complementary.

HOW NOTIONAL NUMBER ACCOUNTS FOR AGREEMENT:		
	<i>Everyone_i fired their_i gun.</i>	<i>The police officer_i fired their_i gun.</i>
SYNTAX	Syntactically plural <i>they</i> DOES NOT AGREE in number with <i>everyone</i> .	Syntactically plural <i>they</i> DOES NOT AGREE in number with <i>the police officer</i> .
NOTIONAL NUMBER	The notional plurality of <i>everyone</i> AGREES with syntactically plural <i>they</i> . Notionally singular <i>they</i> AGREES with syntactically singular <i>everyone</i> .	Notionally singular <i>they</i> AGREES with <i>the police officer</i> .

FIGURE 1.1: NOTIONAL NUMBER

Yet the concept of notional number is not unproblematic. Conceptualising it as a cline leaves it open to subjective judgement, and if notional number can affect the plurality/singularity of *they* then it must also be able to influence other pronouns, including *he*. Newman (1992:452) argues that when *he* is used with an indefinite pronoun or a NP with a

quantifier, “the singular pronoun *he* can also take on the same plural sense so long as it is also within the scope of the quantifier”. However, based on my review of the literature on epicene coreference and the example sentences in the previous section, Newman’s argument does not appear to hold, as there is evidence for the argument that coreference between generic *he* and indefinite pronouns is problematic. Furthermore, the pronoun is arguably not generic when used when used with indefinite, and definite, singular NPs, due to its well-established semantic value.

Newman’s argument also poses an additional problem for notional number, as if it can affect all pronouns, none can be privileged over any other in terms of epicene status. Based on the examples of antecedent coreference in section 1.3.1, I argue that notional number, if indeed it is the underlying semantic factor allowing singular forms to coreference with singular *they*, does not interact with *he* and *they* in the same way. In relation to indefinite pronouns, there is no suggestion in the wider literature that generic *he* can take on the supposed plurality of such singular antecedents, in the same way that singular *they* can notionally agree with the syntactic singularity of its antecedent (c.f. Figure 1.1)⁴⁵.

Acknowledging the apparent shortcomings of notional number, an alternative explanation present in the literature, which attempts to account for the grammatical acceptability of singular *they*, is that notional number is merely an aid to syntactic agreement. Following Abney (1987) who suggested that “pronouns are generated in the D-position⁴⁶, like definite articles” (Baauw 2002:15), and Déchaine and Wiltschko (2002:419-420) who argued that pronominal *one* is base generated in the NP, Figure 1.2 (overleaf) shows how the notional plurality of indefinite pronouns may be carried by their quantifier element in the same way that the quantifier in NPs like *any girlfriend* suggest plurality. The combination of notional number and syntactic number evident in Figure 1.2 supports Newman’s (1992:452)

45 In the wider literature, Bock, Cutler, Eberhard, Butterfield, Cooper Cutting, and Humphreys (2006) propose that collective nouns have “different lexical specifications” where speakers of different varieties of English “associate different grammatical number values” with collectives (2006:72). Thus in American English it can be posited that collective nouns have a lexically singular specification and are acquired as grammatically singular, with the converse being true for British English. Therefore, according to Bock et al. (2006:72), number does not have to be “notionally controlled but lexically controlled”. In the following section I suggest that the same argument can also be posited for singular and plural *they*.

46 Here D-position refers to the syntactic slot available to determiners in syntactic theory.

analysis that a singular term “can be said to take on a plural meaning because it is in the scope of the universal quantifier every”. Indeed this view accounts for all of the examples above where indefinite pronouns are coindexed with singular *they* (arguably, it also accounts for indefinite singular NPs too).

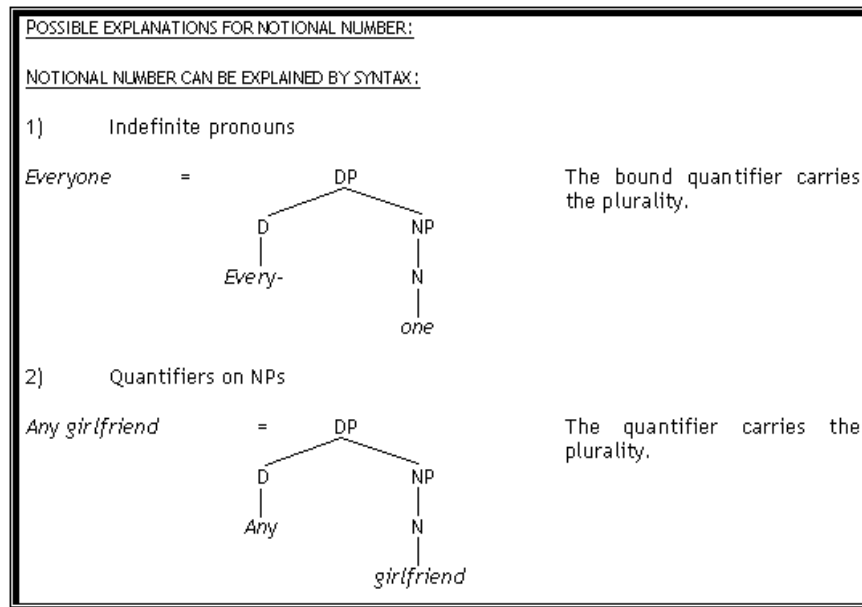


FIGURE 1.2: NOTIONAL NUMBER AND SYNTAX

However, this explanation means that the pronoun would have number concord with the quantifier, but the verb would still be in agreement with the content of the NP, and it seems unlikely that different elements of a sentence would take their number agreement from different elements of the DP, as this is not a general feature of English. In any case however, this explanation cannot account for the coindexation of singular *they* and a formally singular NP where there is no quantifier present in the head of DP. This problem illustrates Meyers’ (1993:185) argument that “it is more difficult to rationalize a plural sense for singular noun phrases than for indefinite pronouns”. Notional number cannot help in this case, as English does not have a null-quantifier that could convey this information⁴⁷. Whilst this lack of agreement initially appears to give weight to arguments against singular *they*, there is another explanation which I evaluate in detail in the following section.

⁴⁷ A null quantifier is a syntactic entity, taking the form \emptyset , which can fill a slot on a syntactic tree and carry key information, but is not spoken or written.

In this section I have evaluated the main theories present in the current literature which attempt to account for epicene coreference between singular *they* and formally singular antecedents. My synthesis of scholarship indicates that, not only is epicene agreement an area of contention in English grammar, no current theory can account for agreement between singular *they* and all the tested types of singular antecedent. However, as I discuss in detail below, the current theories of epicene reference all work with the premise that singular *they* is merely a manipulation of the plural form, yet conceptualising *they* as two forms, one plural and one singular, may hold the key to explaining why *they* can be found coindexed with syntactically singular antecedents.

1.3.3: HOMONYMY THEORY

Having established that there are issues with current explanations of epicene reference, my goal in this section is to address this area of epicene research by adapting an existing theory of pronouns to the problem of singular and plural grammatical coreference with *they*. This alternative theory is based on Whitley's (1978) study of personal and impersonal pronouns, and has not had much, if any, consideration in the wider literature on epicene pronouns. Whitley (1978:31-33) posits two versions of an argument which can potentially be adapted into an explanation for why the current scholarship indicates that singular *they* appears suited to epicene agreement. The first is Polysemy Theory, where singular and plural *they* "are merely alternative semantic interpretations of the same underlying formative" (1978:31), in the same way that the word *wood* can refer equally to a group of trees or to the material obtained from trees. However, like notional number this theory also assumes that *they* is fundamentally plural and therefore coreference with singular antecedents is syntactically impossible.

Alternatively there is Homonymy Theory, where Whitley posits that the impersonal forms of *they* "are distinct from their personal sound-alikes at some underlying level, though phonologically and morphologically" the same (1978:32). However, Whitley explicitly notes that singular *they* is not an

impersonal pronoun, stating that it “is clearly both anaphorical and personal, referring back to NPs that are [+ Ill, + human, - plural] and contain indefinite pronouns or determiners” (1978:28). On the basis of this, under Homonymy Theory, I argue that Whitley’s approach can be adapted from a personal/impersonal distinction to the singular/plural distinction, and as such there can be two distinct forms of *they* in the mental lexicon, one which is marked [+PLURAL] and another which is [-PLURAL]⁴⁸ in its phi-feature specification⁴⁹ (as depicted in Figure 1.3). This hypothesis means that there is no singular/plural disagreement when singular *they* is used as an epicene, as the [-PLURAL] version of *they* is selected for coreference with syntactically singular antecedents⁵⁰.

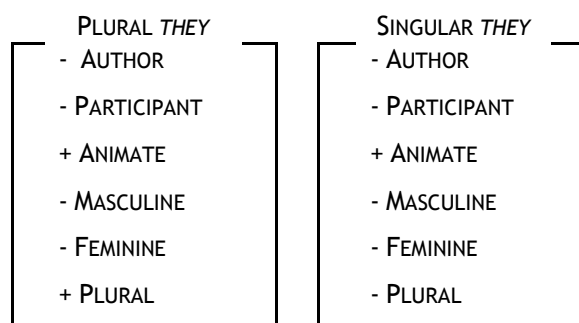


FIGURE 1.3: THE PHI-FEATURE SPECIFICATIONS OF PLURAL AND SINGULAR *THEY*

This single difference in the phi-feature specifications of singular and plural *they* fulfils Whitley’s (1978:34) criteria that Homonymy Theory “requires that the underlying feature make-up of the two sets of pronouns be different”. In addition, my argument that there are two distinct entries for *they* in the mental lexicon represents a syncretism in the pronoun paradigm “where different combinations of morphosyntactic feature values are represented by the same form” (Harley 2008:251). Singular and plural *they* are morphologically identical but distinct in their underlying syntactic

48 In positing Homonymy Theory I am working under the assumption that phi-features must be kept binary and cannot have a null value.

49 Phi-features refer to the underlying syntactic specifications for (in this case) the personal pronouns. Assuming that features such as [NUMBER], [PERSON], and [GENDER] must be binary, each pronoun has a plus or minus value for each phi-feature, and it is the different combinations of these phi-features which makes each pronoun syntactically unique.

50 I am aware that the term *homonymy* may be confusing as it could be taken to imply that the two forms of *they* are unrelated in their development, however this is not the intended meaning as it certainly is not the case. I use the terms as Whitley does, to differentiate between polysemes, which are two meanings developed from one form, and homonyms, which are two syntactically distinct forms. Polysemy refers to a form that has more than one (often-related) meaning, but only has one underlying syntactic specification, and thus meaning is differentiated on a semantic level. Homonymy denotes two forms that are syntactically distinct in terms of phi-features and have different entries in the mental lexicon, but are morphologically and phonologically identical.

specifications, and thus, following Harley (2008:251), the morpheme *they* “‘realizes’ more than one combination of features in a terminal node”.

In addition, Pertsova’s (2011:225) explanation of syncretisms explicitly includes the criterion that, in its technical sense, the term “applies only to those instances of inflectional identity that are judged to be systematic in contrast to accidental homonymy”. The two proposed forms of *they* (which, incidentally, are not considered in the wider literature) are, arguably, systematically different, as the motivation for the singular form lies in the theorised gap for a singular third-person epicene pronoun (as discussed in the introduction). Thus, there is some limited evidence in the literature that such an adaptation of Homonymy Theory could be easily incorporated into wider theories of pronouns. Furthermore, Harley (2008:275) argues that there are three English metasyncretisms in the pronouns, the first is that “gender is not marked in the personal pronouns (first and second person)”, the second is that “gender is not marked in the plural pronouns” and finally “number is not marked in the second person”. I argue that this final metasyncretism can also be applied to singular and plural *they*.

However, working with the same premise, it is also arguable that generic and masculine *he* could be differentiated by phi-features (Figure 1.4), thus creating a set of pronominal homonyms. In this case the two forms are differentiated on their gender-marking values for [FEMININE]. Masculine *he* has the combined [GENDER] specification [+MASCULINE], [-FEMININE], whilst generic *he* is [+MASCULINE], [+FEMININE] - i.e. including both genders but not contradicting the default [+MASCULINE] specification.

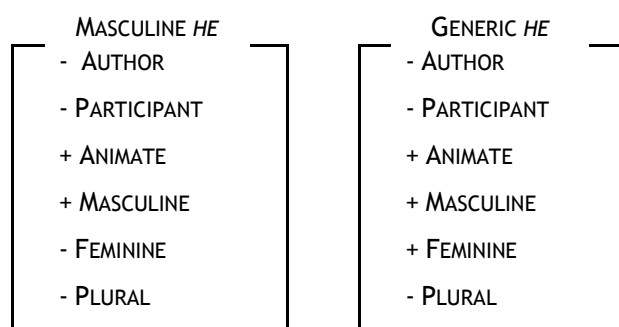


FIGURE 1.4: THE PHI-FEATURE SPECIFICATIONS OF MASCULINE AND GENERIC HE

Yet, such a complex [GENDER] specification is problematic. It is difficult to see how a pronoun with the phi-features [+MASCULINE], [+FEMININE] would coindex with forms such as *doctor*, *child*, *human being*, etc., which are not syntactically marked for gender. A simpler explanation would be that

generic *he* was [-MASCULINE], but the studies reviewed in section 1.2.2 showed generic *he* to have a default masculine interpretation, indicating that *he* is processed as a [+MASCULINE] pronoun⁵¹. Thus, phi-feature specifications cannot account for generic and masculine uses of *he* in the same way that they can neatly illustrate the difference between singular and plural *they*; using the [PLURAL] phi-feature distinction which is used across the whole pronoun paradigm. The proposed utilisation of this common phi-feature is significant, as it represents a pattern in the normally irregular forms, and could arguably aid pronominal acquisition (see section 1.1.3).

Support for this adaptation of Homonymy Theory as an explanation to account for agreement between singular *they* and formally singular antecedents comes from the argument in the literature that when processing a sentence, syntactic information is used before semantics (thus even if notional number did affect pronoun/antecedent resolution it would occur after syntactic processing). For example, Hirst and Brill (1980:174) state that the majority of models depicting how pronouns realise their coreferents “posit that a pronoun triggers a search through the memory representation formed from previous text” and based on this memory each “alternative antecedent encountered... is evaluated on syntactic, lexical... and [then] pragmatic grounds”. Similarly, De Vincenzi (1999:543) claims that “there is an initial stage in which the syntactic processor operates to assign a constituent structure analysis to the input”. Arguably the “incomplete specification” provided by the syntax is then followed by semantic information, and “pragmatic processes” fill in the gaps (Foster-Cohen 1994:248). Thus, if the argument holds that there are two entries for *they* in the mental lexicon and there is not enough syntactic information for either the singular or plural form to be selected, scholarship indicates that pragmatic information can aid this choice. However, it is highly unlikely that

51 Alternatively, it is perhaps possible for masculine and generic *he* to be differentiated on another phi-feature, such as the novel feature [GENERIC]. Masculine *he* would be [-GENERIC], and generic *he* would be [+GENERIC], but by its very nature this hypothetical phi-feature would be in conflict with the already present [MASCULINE]; a pronoun could not refer specifically to males and be gender-neutral at the same time. In any case, it would be unlikely that such a novel phi-feature is acquired, despite the fact that pronouns are rote learned, as studies such as those done by Graham (1973) have shown that children do not receive generic *he* in their L1 input. I argue that this lack of data would thus prevent them from acquiring generic *he* as a form distinct from masculine *he*. Thus, it is not possible that generic *he* is an example of pronominal homonymy.

such pragmatic information would be needed, as selection of the right form of *they* would be constrained by the number marking on its antecedent.

My goal in this final section of the chapter was to illustrate that, despite its classification as a plural pronoun there is scholarship to support the argument that singular *they* coindexes unproblematically with all of the singular antecedent types tested. Having already established in my review of the literature in section 1.2.2, that the gender marking on *he* interferes with an epicene interpretation, my consideration of the scholarship on how epicenes coindex with their antecedents indicates that the use of generic *he* with indefinite pronouns is problematic. Furthermore, the semantic value of *he* influences interpretation when it is coreferenced with other singular NPs, meaning that its use as a gender-neutral pronoun is challenged. However, my review of the scholarship indicates that the same problem does not occur with singular *they*, which has no such coreferent restrictions.

However, there are issues in the literature over exactly how singular *they* syntactically agrees with formally singular antecedents, based primarily on the argument that *they* is fundamentally plural, and as such number concord does not hold between pronoun and antecedent (see Baron 1986:193 and Sanford and Filik 2006:177). In light of this apparent problem with using *they* as a singular pronoun I reviewed the different possible explanations in the current literature for the syntactic agreement of singular *they* with formally singular antecedents. My review of the literature indicates that there are issues with the current most popular explanations for the apparent syntactic acceptability of singular *they*, especially in terms of agreement with definite NPs. There is evidence for the position that notional number cannot completely account for the data in section 1.3.1 which indicated that singular *they* is a viable epicene candidate for all antecedent types.

In response to this finding I adapted Whitley's (1978) theory of personal and impersonal pronouns in an attempt to account for epicene reference with singular *they* and formally singular antecedents. I showed that by adapting Homonymy Theory and positing that there are two forms of *they* which differ on the syntactic level of phi-features, it is possible to account

for the coreference of singular *they* with definite singular NPs⁵². If my adaptation of Whitley's (1978) Homonymy Theory is correct then it can be argued that singular *they* is not a deviance as Sanford and Filik (2006:177) suggested (see section 1.2.2), it is not a manipulation of plural *they* which must be computed each time it is encountered; it is instead its own form.

CONCLUSIONS

In this first chapter I have shown that there is a body of scholarship that supports the argument that the personal pronouns are a closed-class. My evaluation of the literature also indicates that such a classification would predict that the pronoun paradigm would be resistant to change, and the pronouns would carry very little, if any, semantic weight. However, scholarship charting the development of the pronouns from Old English onwards indicates that both internal (phonological) and external factors (such as invasions) have influenced the structure of the pronoun paradigm. This literature review has indicated that such changes have led to an overall decrease in the number of pronouns since Old English, with highly restricted additions. I also showed in section 1.1.2 that there was evidence in the existing literature to suggest that there are parallels in the development of singular and plural *you* and singular and plural *they*, potentially suggesting a regularity in a paradigm which, according to current scholarship, must be rote learned. Furthermore, if the argument that a paradigm becomes fixed once it has been acquired holds, my synthesis of current theories of pronoun acquisition in section 1.1.3 can potentially account for why the literature shows that the pronoun paradigm is resistant to new members.

My evaluation of the literature also showed that there is a body of evidence in support of the position that despite their closed-class status, the third-person singular pronouns carry uncharacteristically high levels of semantic information, as they are formally marked for gender. If this

⁵² I also argued that Homonymy Theory could not apply to generic *he* in the same way, as its application would involve either conflicting phi-feature specifications for [MASCULINE] and [FEMININE], or the adoption of a novel phi-feature.

argument holds then it appears that the pronouns do not possess all of the characteristics of closed-classes that are proposed in the wider literature⁵³. This apparent semantic weight restricts the use of generic *he* as an epicene pronoun, as the studies discussed in section 1.2.2 give rise to the argument that generic *he* has a default masculine interpretation. This view is supported by studies that show that its semantic value makes it difficult to coindex generic *he* with the different antecedent types I discussed in section 1.3.1, especially indefinite pronouns and definite singular NPs. Furthermore, my review of current scholarship also showed that the same body of literature supports the argument that singular *they* on the other hand, does not pose this problem, and is commonly used in production tests as an epicene pronoun.

In the final section of the chapter I focused on scholarship that offers explanations for the coreference between *they* and singular antecedents. Based on my review of the current theories of coreference I proposed that the least problematic explanation of epicene agreement is the application of Homonymy Theory to plural and singular *they*. This theory accounts for why singular *they* can be coindexed with indefinite pronouns, NPs with quantifiers, indefinite singular NPs, and definite singular NPs without causing confusion over the number of entities it refers to. This chapter has shown that whilst certain areas of epicene research are covered extensively in the literature, not least the data on comprehension and production I presented in section 1.2, other aspects of the epicene debate are not so well covered. The above discussion has highlighted several areas where current scholarship cannot fully account for modern epicene usage, both theoretically, in relation to syntactic coreference, and in terms of usage data.

The synthesis of current research presented above indicates that despite any arguments to the contrary, evidence shows that singular *they* is a viable epicene pronoun which does not appear to influence the presumed gender of a generic referent. Also, as I showed in section 1.3.2 it is found in the literature with all types of singular antecedents. Based on this large body of research I hypothesise that singular *they* is used as the epicene of choice in speech and writing in British English today. However, there is very

53 Indeed, Wiese and Simon (2002:2) describe pronouns as having “a borderline status within the linguistic system, between lexical categories like nouns and functional categories like complementisers”.

little available data on epicene usage in the twenty-first century, as the vast majority of studies I have surveyed here are at least ten years old, and do not necessarily focus on written British English. In order to address this lack of modern usage data, and test my hypothesis, I propose the following research questions which I discuss in more detail in chapter three:

- What epicene forms are currently in use in written British standard English?
- Do different epicene forms correlate with different types of singular antecedent?
- Is there an epicene pronoun of choice?

My review of the scholarship also supports Bennet-Kastor's (1996) assertion that there has been very little research on children's use and acquisition of epicene pronouns, despite arguments in the literature that the content of the pronoun paradigm in a person's mental grammar becomes fixed once acquired. Even since Bennet-Kastor's assertion, which is now fifteen years old, my review of existing studies indicates that this area of research is still underdeveloped. Furthermore, the participants in Bennet-Kastor's (1996) study, discussed in section 1.2.2, are at an age where, according to dates in the literature provided by Rispoli (1994) and Owens Jr (2007), their pronoun paradigms are already fixed. Whilst there is evidence in the current scholarship for the argument that it is possible to influence the content of a closed-class whilst it is still being acquired, there have been no studies directly testing the epicene forms children receive as L1 input during this period. Therefore, in relation to my discussion of the acquisition of epicene pronouns, and the implications this may have for changes within closed-classes I pose the following further set of research questions, which are designed to investigate the hypothesis that children will receive singular *they* as L1 input:

- Is there any evidence that children are exposed to either singular *they* or generic *he* when acquiring the personal pronoun paradigm?
- Furthermore, in relation to the historical parallels in the development of the second and third-person pronouns, are there parallels between the development of *you* and *they*?
- If so, can these parallels be used as evidence for regularity in the acquisition of the rote-learned paradigm?
- Is there any evidence that any observed regularity makes acquiring singular *they* easier than acquiring generic *he*?

In chapter three I discuss these research questions in more detail, justifying the methods I use to address each part. However, before I can begin to approach the topics I have set out above, there is another large body of literature on the epicene debate, which has not yet been addressed, and which is primarily concerned with the impact of language-external factors. Elements of this chapter, such as the acknowledgement of the impact of Danish invasions on the introduction of the Scandinavian *th-* forms into the personal pronoun paradigm in section 1.1.2, have already begun to highlight the importance of this aspect of epicene literature. As such, the influence of language-external forces is the primary focus of chapter two.

CHAPTER TWO: SOCIAL MOVEMENTS AND EPICENES

2.1: CONTEXTUALISING EPICENE PRESCRIPTIONS	63
2.1.1: INITIATING FORMAL PRESCRIPTIVE PRACTICE	64
2.1.2: TRACING EPICENE HISTORY	70
2.2: THE EFFECTS OF SEXUAL POLITICS	78
2.2.1: FEMINISM AND LANGUAGE	79
2.2.2: FEMINIST LANGUAGE REFORM AND EPICENES	84
2.3: EPICENES IN EDUCATION	90
2.3.1: "CORRECT" LANGUAGE IN EDUCATIONAL HANDBOOKS	91
2.3.2: MODERN APPROACHES TO EPICENES	96
CONCLUSIONS	103

In this chapter I review the literature on the language-external factors which have played a large part in the epicene debate, with the aim of contextualising the language-internal issues presented in chapter one within wider social issues. My main goal in this chapter is to document the history of epicene prescriptions. Previous research indicates that traditional grammatical prescriptivism, and language reforms born from second-wave feminism, have influenced the prescription of epicene pronouns, and have also been influenced by societal norms and political movements. If this position holds then there is evidence for the argument that epicenes are as much of a social issue as they are a syntactic one.

This chapter is structured chronologically, beginning with the birth of traditional grammatical prescriptivism in the eighteenth century, before moving on to more modern prescriptions, focusing specifically on non-sexist language reforms born out of second-wave feminism. To illustrate the boundaries between the key periods in epicene prescription I have split the chapter into three sections. I begin by addressing the literature on the origins of English standardisation and the codification of grammatical norms, to provide contextual data for my consideration of scholarship surrounding the treatment of epicene pronouns throughout history. Such scholarship indicates that generic *he* was the primary candidate for epicene status

between the eighteenth and twentieth centuries. However, whilst surveying the literature it became apparent that there is evidence in the current scholarship for a shift in momentum in the epicene debate around the time of the rise of second-wave feminism, and based on this observation, the second section of this chapter deals with the treatment of epicene pronouns from circa 1970 onwards. Finally, as the following discussion indicates that grammatical norms are linked throughout history to levels of education, in section 2.3 I consider studies on school grammars and focus on the use of prescriptive materials in education, evaluating the impact of language-external factors on epicene choice in the classroom.

2.1: CONTEXTUALISING EPICENE PRESCRIPTIONS

In this section I focus on the scholarship surrounding the development of traditional grammatical prescriptivism, which I define here, following Yule (2006:77), as the “view of grammar as a set of rules for the ‘proper’ use of language”. I show that the literature indicates that how and why the “correct” forms of language are chosen, and more importantly, who makes these choices is extremely important to the epicene debate. Evidence shows that traditional prescriptive rules of grammar included the promotion of generic *he* as the English gender-neutral pronoun (see Bodine 1975, Stanley 1978, and Baron 1986 for comprehensive examples). Current scholarship also suggests that within traditional prescriptive grammar (which shall be used here to denote grammars that have their foundation in the beginnings of the printed prescriptive movement discussed in section 2.1.1) the use of singular *they* is proscribed. In section 2.1.2 I document studies indicating prescriptive norms up until second-wave feminism, whilst in section 2.1.1 I survey the literature on the development of traditional grammatical prescriptivism and show how there is a large body of scholarship in support of the position that its origins were linked to social change.

2.1.1: INITIATING FORMAL PRESCRIPTIVE PRACTICE

Based on my review of the literature, detailed below, I argue that there is evidence for the position that the beginnings of traditional grammatical prescriptivism coincide with the onset of standardisation, which began in earnest when William Caxton set up his printing press in 1476⁵⁴. Fennell (2001:156) notes that whilst initially, the majority of books were printed in Latin, “by 1640 there were approximately 20,000 titles available in English”. However, literature on the development of the English language indicates that at this time there were still several major dialects of English that were markedly different⁵⁵, and it would have been costly to print copies of books in each dialect, as the master stamps for each page were produced by hand. Arguably, based on this socioeconomic criterion, there was thus a need for a single dialect that could be mass-printed and comprehended by most of the literate population of England.

Clark (2001:30) argued that Caxton’s decision to print in the East Midlands dialect was greatly influenced by “consideration of his likely readership”, and, as the East Midlands area “held a significantly higher proportion of the literate population”, using that dialect was likely “to make publication economically worthwhile” (2001:30). Similarly, Perera (1994:80) claims that “it was an accident of geography that selected the East Midlands dialect as the forerunner of standard English, not any inherent superiority over other fifteenth-century dialects”. Furthermore, in one of the most recent texts on traditional grammatical prescriptivism Tieken-Boon van Ostade (2011:86) highlights the role played by the publishers in the production of eighteenth-century grammars, noting that for Lowth’s (1762) influential grammar, discussed below, even the author considered the text “to be primarily the responsibility of his publishers”. She argues that the success of the text was down to “the publishers’ efforts” and came as a surprise to Lowth (2011:86).

Thus, there is evidence for the position that standardisation was not a neutral progression from the invention of the printing press, it was

54 Febvre and Martin (1958) claimed that printing “was probably the single most important factor in the process of the formation of European languages” (Clark 2001:29).

55 Caxton himself complained about England’s dialects “in his preface to the *Eneydos* in 1490” where he remarked on the difference between the northern form *eggs* and its southern counterpart *eyren* (Shay 2008:138). For a discussion of England’s historical dialects see Skeat (2009).

“motivated in the first place by various social, political and commercial needs” (Milroy and Milroy 1985:22). As printed standardisation began, dialectal differences or “chaotic deviations” (Newman 1992:449) were no longer tolerated, arguably paving the way for traditional grammatical prescriptivism. In terms of epicenes, such “chaotic deviations” referred to the use of both generic *he* and singular *they* as pronouns of generic reference, which can be found throughout the history of English (see section 1.1.2), and as such, only one variant could be endorsed within traditional prescriptive grammar.

Yet, there is evidence in the literature to suggest that the printing press was not the only technological development involved in the rise of English traditional grammatical prescriptivism. Due to mechanical innovations, such as the steam engine, and the start of the Industrial Revolution in the late eighteenth century, the nature of work changed. According to Fennell (2001:156) the overarching influence of industrialisation changed the make-up of society and initiated the rise of the middle class, bringing with it “an increase in leisure time for reading and general interest in education and learning”⁵⁶. Significantly for the epicene debate, Baron (2001:116) documents that this period of social change corresponded with a sharp increase in traditional prescriptivist materials, as in the “second half of the eighteenth century... more than 200 grammars and books on rhetoric appeared”.

The most often-cited example of a traditional prescriptive grammar (c.f. Tieken-Boon van Ostade 2011) is Robert Lowth’s (1762) *A Short Introduction to English Grammar*. According to Aitchison (2001:9), Lowth, who is perhaps the most famous of the traditional grammatical prescriptivists, set out to “lay down ‘rules’ of good usage”, but had no formal linguistic training. He did however have high social status, and, again following Aitchison (2001:12), is characterised as prescribing language norms based on his own “preconceived notions” of what represented “correct” usage. Recently this depiction of Lowth has been challenged by Tieken-Boon van Ostade (2011:8) who argues he could not have had the qualifications of a modern grammarian because he was educated “when English was not part of

⁵⁶ Maynes (2004:51) notes that during this period female employment was on the rise, reaching “astonishingly high” levels in the late eighteenth century, arguing that the growing textile industry “relied on the labor of women and girls” (2004:54).

the school curriculum, and when linguistics as a discipline did not exist". She also heavily criticises Aitchison's characterisation of Lowth, which she considers "far from objective", rejecting the notion that Lowth's selection of prescriptive norms was based on personal choice (2011:6). However, although she questions the traditional depiction of Lowth, Tieken-Boon van Ostade does note that Lowth's text was taken as the foundation for many other prescriptive materials, claiming that there is "a direct continuity between the strictures discussed by Lowth in his grammar, and illustrated by negative examples in his footnotes, and the modern usage guide" (2011:20)⁵⁷. Thus although the characterisation of Lowth is generally negative, and now disputed by Tieken-Boon van Ostade, the impact that his work had on the grammatical norms of English is more clear-cut.

The general aim of traditional prescriptive grammar guides was to regulate "correct" usage and to preserve and protect the language (see Johnson 1747:4 below). Yet there is evidence in the related scholarship for the position that this attempted preservation was underpinned by two main, yet flawed, assumptions. The first premise was that there was a "golden age" of language (Labov 2001:6), that was slowly being forgotten or lost, and which needed to be enforced. This is what Aitchison (2001:13) refers to as a mythical "vintage year where language achieved a measure of excellence". A clear example of how this idea affected practice is that the "chief intent" of Johnson's dictionary, published in 1755, was to "preserve the purity" of English (Johnson 1747:4 in Fennel 2001:149)⁵⁸. Indeed, Burrige (2010:8) refers to the standard that developed through traditional grammatical prescriptivism as "something of a linguistic fantasy", noting that even Lowth "appeared aware that the rules he was laying down belonged to... a more abstract level of language".

However, Halpern (2008) launches a defence of traditional grammatical prescriptivists, arguing that they "never supposed that there was such a

57 Tieken-Boon van Ostade (2010:15) notes that the statement that "Two negatives in English destroy one another, or are equivalent to the Affirmative" is found in versions of both Lowth's and the American Lindley Murray's grammars. Thus a "rule" of grammar tied to traditional grammatical prescriptivism has direct links to Lowth's grammar.

58 This is why early traditional prescriptivism relied heavily on "logic and analogies with classical languages" (Newman 1992:447), as it was assumed that Latin and Greek were in some way more whole than other languages and should be used for reference. As a result English traditional grammatical prescriptive rules were based on a Latin model (Clark 2001:29).

Golden Age of Language” and similarly there was never anyone who suggested that “All change is forbidden!” (2008:130). Halpern argues that prescriptivists “have no personal authority in matters of usage, nor do they seek any”, and claims that, instead, prescriptivists “seek to persuade, not to dictate”. In regards to prescriptivists’ apparent reluctance to acknowledge language change (see below) Halpern claims that “they argue that no change should be accepted until the educated public is aware of the issues it raises” (2008:130). Yet Halpern is a lone voice in the wider literature and his comments go against the grain of the generally accepted characterisation of prescriptivists.

The second assumption ascribed to the traditional prescriptive approach concerns language change. Aitchison (2001:4) states that “large numbers of intelligent people condemn and resent language change” due to the false assumption that change equals “unnecessary sloppiness, laziness or ignorance”. However, my review of the literature on previous pronoun changes in section 1.1.2 illustrates that language contact and its resulting influence on the paradigm can be extremely useful to speakers. According to Fennell (2001), the new forms added to the paradigm throughout history, including the Scandinavian *th*- forms, helped speakers to delineate between pronouns that had become too phonologically similar. Nevertheless, there is evidence for the position that it was the goal of the early traditional prescriptivists to ward off any such changes and maintain the status quo. This is what led Drake (1977:330) to comment that prescriptivism in the 1700s “imposed a static model on linguistic behaviour”. Thus, the inevitability of language change was not acknowledged and traditional grammatical prescriptivists tended to avoid addressing their language’s historical development.

Perales-Escudero (2010:3) claims that the existence of “language guardians, self-appointed and otherwise” is crucial to the standardisation process, as they “ensure that one or another form should fall out of usage by casting it as substandard or deviant”. Poplack, van Herk and Harvie (2002:88) state that grammars “aim to furnish an ordered view of language, in which each form serves a single function and each function is represented by a single form”. Furthermore, they argue that “not only is a particular form to be associated with a particular function, but competing forms are to be eradicated” (Poplack, van Herk and Harvie 2002:89). Although the selection

of “correct” forms was “arbitrary” and other possible variants were “quite serviceable” (Milroy and Milroy 1985:17), scholarship suggests that traditional grammatical prescriptions tended to be conservative and were/are “typically intolerant of innovations in language” (Mesthrie, Swann, Deumert and Leap 2000:14). Thus there is evidence in the literature for the position that, due to the nature prescriptive texts, and the dichotomy of correct and incorrect language, grammarians or “language guardians” as Milroy and Milroy (1985:17) chose to name them, had to select only one “correct” form for each rule they prescribed, to try to avoid contradictions in their work.

Milroy and Milroy (1985:1) link this selection of a single variant to the argument that traditional grammatical prescriptivism was intricately linked to “an ideology” of language, not only that “things shall be done in the right way”, but that there was only one “correct” or “right” way of using language. In a similar vein, Strabone (2010:237-238) argues that in the eighteenth century “[t]he standardizers, whether grammarians, lexicographers, or rhetoricians, generally shared several beliefs about the state of their language”, which included the notion that the language had to be actively “protected and preserved from corruption”. Strabone claims that such corruption was assumed to come from dialectal variants, which the grammarians “deemed barbarous” (2010:238)⁵⁹. Equally, Strabone also claims that these writers assumed that “everyone in Great Britain, regardless of class and birthplace, could learn to speak and write according to the emerging standards”. Yet, Pinker (1999:5) argues that these rules were not open to all, as they were connected with a certain level of education, and thus, conversely to Strabone’s standpoint, they did not act as unifying tools, which would “strengthen the community of English speakers” (2010:238) but rather represented social markers.

Labov (2001:509) notes there is a strong link between the prescription of “correct” forms of language and “the growth of an upwardly mobile merchant class whose early education did not match their newly achieved social status”. Thus, there is evidence for the proposition that prescriptions

⁵⁹ Strabone also claims that the grammarians, as a homogenised group, settled upon dismissing dialectal forms in favour of “following writers of taste” (2010:237). However, Mugglestone (1997:471) notes that grammarians, including Lowth, far from following the usage of respected writers can be found “censuring samples of English from Milton and Shakespeare”.

reflected how the new middle class aspired to gain social prestige, and as a result it was the dialects of the upper echelons that were prescribed as “correct” and not the dialects of the common people. Indeed, Tieken-Boon van Ostade (2011:49) notes that Lowth’s primary purpose for writing his grammar was to provide his son with guidance on language that was appropriate in desired social circles. However, Pinker (1999:5) argues that traditional prescriptions are so “psychologically unnatural” that only people with “access to the right schooling” can follow them. This means that prescriptive rules act like “shibboleths, differentiating the elite from the rabble” (1999:5), separating the “correct” English of the upper classes⁶⁰ from the English used by the middle and lower classes⁶¹.

In terms of epicenes, Laitinen (2007:52) argues that the use of *they* “in anaphora rose to some extent above the level of consciousness at the end of the eighteenth century”, he notes that as a result, “agreement turned into a socially stigmatised linguistic marker in some social circles”. This stigmatisation occurred despite arguments that singular *they* has been “correct for centuries” especially when used with indefinite pronouns (Kolln 1986:100, see section 1.3.1). Indeed, Milroy and Milroy argue that within the traditional prescriptive movement certain usages were “not attacked as non-standard” until they were widespread features (1985:21). This claim provides support for the argument that before it was proscribed, singular *they* was actually the epicene of choice, and can perhaps account for the experimental data in chapter one which showed that in many cases, singular *they* appears to be the most appropriate gender-neutral pronoun.

According to Madson and Hessling (1999:571) singular *they* was the English epicene until traditional prescriptivists “mandated” that generic *he* was the “correct” form. This mandate led to such claims as that of Edward D. Johnson who in 1982 argued that singular *they* “annoys writers, who must

60 In section 1.1.2 I noted that Henry VIII’s choice of *you* as both a singular and plural pronoun may have been significant due to his position in society.

61 Another example of prescriptivism in action would be the creation of language academies, most famously L’academie Francaise which was established in 1635 (see section 2.2.3). Wardhaugh suggests that the primary aim of L’academie was “to fashion and reinforce French nationality” (2006:36), indicating an awareness that language and society/identity are two sides of the same coin. This corresponds with Cooper’s argument that the L’academie is an “excellent illustration” of the links between language prescription and social forces, and shows that “language planning cannot be understood without reference to social context” (1989:3). This supports my argument that the empirical data in chapter one cannot be fully understood without analysis of how language-external influences affect which pronoun is promoted as epicene.

forego the privileges the masculine pronoun has for millennia enjoyed in English and its root languages” (Baron 1986:195). This comment not only overestimates the stability of English, it denies the fact that Old English gender was grammatical, with neuter pronouns forgoing the need for generic *he* (section 1.1.2). To illustrate how generic *he* became the prescribed form, in the following section I review the literature on the epicene debate from its origins pre-1800 up until the middle of the twentieth century.

2.1.2: TRACING EPICENE HISTORY

Having set out the key premises of grammatical prescriptivism identified in the literature I now focus more closely on how the move towards standardisation affected the status of singular *they* and generic *he* as epicene pronouns. Newman notes that during the Middle English period, before the birth of traditional grammatical prescriptivism, although generic *he* was the most common epicene pronoun, “there are cases of pronominal uncertainty” (Newman 1992:448). This is supported by the use of singular *they* in the *OED*, which according to Green (1977:150) included a citation containing singular *them* as far back as 1389, and singular *they* in 1526. It also supports Nevalainen’s (2006) observation that singular *they* was found in Middle English (section 1.1.2). As traditional grammatical prescriptivism is generally considered to begin in earnest circa the eighteenth century, and thus had not begun in Middle English, it is interesting to speculate on the reasons why, according to Newman (1992) generic *he* was the majority epicene at this point in history. One explanation could be the dominance of the masculine template for pronouns in the Old English paradigm (see section 1.1.2).

Alternatively, however, the number of educated, and/or professional men compared with the number of educated women, as well as the relative social standings and social roles of the two sexes, could have influenced epicene choice. Indeed, Schultz (1975:163) suggested that until “recently” (a rather ambiguous date) unmarked masculine nouns such as *congressmen*, *policemen*, and *craftsmen* did refer to men, and were thus not intended to be generic, it is only since “the invasion by women” that a generic reading

of such terms is needed. Changes in the workforce (as noted by Maynes 2004) and the infiltration of women into roles stereotypically performed by men can be seen as support for Labov's claim that "upward social mobility", in this case caused by the changing roles of women in society, is "a primary characteristic" of the forces behind linguistic change (2001:509).

If Schultz' argument holds and unmarked masculine nouns only referred to men, then there is no reason to argue that the same cannot also be said for more general uses of generic *he*. In a related issue, Wolfe (1989:82) argues that the dominance of male generics is understandable in traditional prescriptive grammars as "women were all but excluded from the educated audience" that they were aimed at. Similarly, Stanley (1978:800) suggests that traditional prescriptive texts were "written by men for the edification of other men", and as a direct result such grammars "deal with male concerns from a male point of view".

However, in response to Stanley's argument, Sklar (1983:351) claims that eighteenth-century grammarians were in fact aware of a female contingent in their audience, noting that that there were grammars being written by women as early as 1745. However, Sklar also notes that female grammarians were in the minority and claims that their underrepresentation meant that eighteenth-century and nineteenth-century grammars were "infused... with the social attitudes of their times" which included a negative image of women. This proposition leads Sklar to argue that by following traditional prescriptive rules "we are still serving up eighteenth-century social biases" (1983:348). Yet it is important to note that Sklar's argument implies that a female perspective on grammar would have been different simply *because* it is female, although there is no evidence in the literature to support the premise that a greater number of grammars written by women would have changed the course of traditional grammatical prescriptivism. In any case, it appears highly unlikely that the minority of eighteenth-century female grammarians could have influenced the large body of traditional prescriptive works and rules discussed in the wider literature, even if they did stand against the patriarchal norms of society⁶².

62 In any case, it appears that it was difficult to deviate from the general consensus in the prescriptive movement. For example, Mugglestone (1997) notes that William Cobbett, who published a grammar in 1818, goes against prescriptive convention by not setting out rules of

Such norms are clearly expressed in what Bodine (1975:134) claims is one of the earliest voices in support of generic *he*, where Wilson (1560:189) supports the idea that we should “kepe a *natural order*, and set the man before the woman for manners Sake” (my emphasis). Evidence from Bodine (1975) - an often-cited text on epicene prescriptions - indicates that rules similar to Wilson’s pronouncements on the masculine gender were prescribed as early as the seventeenth century, for example:

The Relative agrees with the Antecedent in gender,
number and person...

The Relative shall agree in gender with the Antecedent
of the more worthy gender...

The Masculine gender is more worthy than the feminine

(Poole 1646:21 in Bodine 1975:134).

It is important to note, however, that Wilson and Poole were writing before the onset of traditional grammatical prescriptivism in the eighteenth century, and as such, at the time of their pronouncements Bodine (1975:135) notes that singular *they* had not yet been openly condemned by traditional prescriptive grammarians, although its rejection was implied by the promotion of generic *he*. According to Bodine, the first real attack on singular *they* appears uncharacteristically early when Kirby (1764:117) claimed “[the] masculine Person answers to the general Name, which comprehends both Male and Female” (Bodine 1975:135)⁶³.

Right at the end of the eighteenth century scholarship shows that in 1795 Lindley Murray’s grammar (a traditional prescriptive volume that would later be heavily criticised, see below) stated that “a pronoun must agree with its antecedent in gender and number” (Baron 1986:191). The rule does not specify how to treat indefinite pronouns, but Stanley (1978:805) notes that Murray’s corresponding “example of a violation of pronominal concord” involves changing “Can *anyone*, on *their* entrance into the world...” to “on *his* entrance”. This example clearly shows that traditional grammatical prescriptivists such as Murray promoted generic *he*, and thus number

pronunciation. Yet when Cobbett’s son published a later edition of the grammar he added a chapter on pronunciation (1997:484), bringing the text in line with the prescriptive conventions of the time; thus indicating that deviating from traditional prescriptive convention was problematic.

63 There is a discrepancy with Bodine’s claim that Kirby (1746) is the first rejection of singular *they*, as Laitinen (2007:50) notes that there is an almost identical quote in Tiekens-Boon van Ostade’s (2006) work which was taken from a grammar published in 1745, which was, incidentally, written by a woman, Ann Fisher.

concord, at the expense of singular *they*, even though their prescriptions may not have reflected usage. Indeed, a review of the literature on prescriptions at the turn of the nineteenth century shows that such prescriptions did evoke some resistance.

Despite the large number of traditionally prescriptive texts published in the eighteenth century (see Bodine 1975 or Stanley 1978 for a review), scholarship indicates that there was no blanket acceptance of the forms they supported. Drake (1977) documents a backlash in the early nineteenth century against traditional grammatical prescriptivism, citing an example from 1825 where Cardell criticises school grammar⁶⁴ as being “opposed to fact, to science, and to common sense” and then claims that traditional grammatical prescriptivist rules are “artificial, perplexing, contradictory, and impracticable” (Drake 1977:326). This viewpoint is supported by “attacks on Murray’s Grammar that occurred especially in the [18]20s and 30s” (1977:327)⁶⁵, which are evidence for the position that throughout history, traditional grammatical prescriptions have been criticised because they do not reflect usage⁶⁶. Indeed, evidence from the literature indicates that this argument is inextricably linked to traditional grammatical prescriptivism, as even at the end of the twentieth century, Zuber and Reed (1993) repeated Cardell’s claim that (modern) grammars do not reflect a student’s experiences (section 2.3).

However, even overt criticisms of traditional prescriptive texts did not stop their publication, as illustrated by the large number of studies on grammar books I evaluate in section 2.3. In the early to mid-nineteenth century, the condemnation of singular *they* by traditional prescriptive grammarians continued, based on the argument that it did not correspond in number with its antecedent (c.f. Baranowski 2002:378). However, those same prescriptivists did not condemn generic *he* even though it does not necessarily match its antecedent for gender. Baranowski (2002:378-379)

64 According to Mesthrie et al. (2000:13) “the traditional approach to the teaching of grammar in English schools” is a clear example of prescriptivism in action, arguing that teachers following a Latin model had to “[uphold] certain forms of language as the norm to be emulated”.

65 These “attacks” included a section of *The American Journal of Education* between July and December 1826 which was “devoted to detailed examination of the inconsistencies of Murray’s rules against what the writer believes is actual behaviour” (Drake 1977:328).

66 Importantly though, there is no evidence to suggest that the criticisms levelled against Murray were based on empirical data. Thus, *The American Journal of Education* may just have been exchanging one form of prescriptivism for another.

argues that this lack of gender concord was ignored, because gender was socially significant and the status quo in the nineteenth century still favoured a male-as-norm attitude. Indeed, the social implications of gender were reflected in an 1850 Act of Parliament which Bodine claims “legally replaced ‘he or she’ with ‘he’” (1975:136, see also Baranowski 2002:379).

The Act stated that “words importing the masculine gender shall be deemed and taken to include females... unless the contrary as to gender... is expressly provided” (Evans and Evans 1957:221 in Bodine 1975:136)⁶⁷. Yet neither Bodine nor Baranowski explicitly notes that the Act only applied to legal documents, and whilst it did graphically replace the form <he or she> with generic *he* in the legal statutes, there was no illegality in using either *he or she* or singular *they*⁶⁸. The reality is less dramatic than Bodine’s interpretation seems to suggest. Nevertheless, the rejection of singular *they* by traditional grammatical prescriptivists continued, as can be seen by Greene’s 1863 comment that singular *they* represents an “irregular use” of the pronoun (Baron 1986:194).

Importantly though, a review of the literature shows that there has never been a blanket rejection of singular *they*, and even after the Act of Parliament of 1850 support for the form can be found. Baron (1986:193) notes that some early grammarians such as Baine (1879 in Jespersen 1922:138-140) thought the use of singular *they* was justified and indeed “not illogical”. Such opinions support the argument that singular *they* was still used as an epicene despite its continuing proscription. A good example of support for singular *they* comes from Gould Brown, who, despite rejecting singular *they* in 1825 on the grounds that it “is of the plural number and does not correctly represent its antecedent noun” (Brown 1825:142 in Baron 1986:194), advocated the adoption of “descriptive standards” as part of the backlash against prescriptivism in 1832 (Drake 1977:334).

Arguably, Brown’s changing of allegiances indicates that even advocates of traditional norms may have struggled with endorsing generic *he*, even though Drake (1977:334) notes that Brown returned to his original view

67 Importantly, the Act also designated that “the singular [is] to include the plural” (Evans and Evans 1957:221 in Bodine 1975:136). This clearly parallels the prescription of generic *he*. However, the Act also states that the plural is to include the singular, but not that the feminine will include the masculine; therefore, number and gender are treated differently.

68 Importantly, there is no evidence to suggest an overt sexist motivation for the Act of Parliament; rather the Act just follows the traditional prescriptive norms of the time.

in 1851, once again condemning singular *they*. Brown's contradictory statements make it clear that prescription of the "correct" third-person epicene pronoun has long been a point of contention in the English language. Furthermore, this contention implies that prescriptions of generic *he* may never have accurately represented generic pronoun usage.

Following on from the epicene debates of the nineteenth century, a review of the literature shows that the 1900s brought rather radical views on grammatical gender, as Weseen (1928 in Baron 1986:192) claimed that "the feminine is the only gender that actually expresses sex". In other words, this view meant that the use of feminine pronouns and other grammatical markings such as the suffix *-ess* actually drew attention to gender, whilst also implying that the same is not true for masculine pronouns and generic masculine references, as they are the unmarked forms. Whilst initially this claim may seem unusual, Madson and Hessling (1999) investigated whether alternating generic pronouns in a text (one paragraph using *he*, and the next *she*) would effectively make a text gender-neutral and their results indicate that *she* was more salient to their participants. They gave 114 student participants a questionnaire and asked them to read either a text that alternated between pronouns, or a control text which used *he/she* forms. Participants were asked questions about how often they perceived *she* to be used in the text, and whether they thought the text they read was biased in some way to one gender.

As Madson and Hessling had originally hypothesised, the results showed that within the alternating text female pronouns "were perceived to occur more frequently than masculine pronouns" (1999:565-566)⁶⁹. The alternating text was perceived as most sexist, with a bias favouring women (1999:566) leading arguably to the conclusion that not using generic *he* is sexism against men (thus turning the gender-neutral argument on its head). Also, the author of the alternating text was perceived to be a woman by over half of the participants⁷⁰. Their results led Madson and Hessling to the conclusion that even though *he* may not be generic "readers are at least accustomed to

69 However, Madson and Hessling may have inadvertently drawn attention to the feminine forms as the alternating text used generic *he* first, which would make the switch to generic *she* in the second paragraph extremely salient.

70 This finding relates to arguments over non-sexist language and the impact of political allegiance and biological sex investigated by Harrigan and Lucic (1988), which I discuss in detail in section 2.2.2.

encountering generic masculines” and therefore generic *he* may not have been as salient in the texts as generic *she* (1999:562). Importantly though, Madson and Hessling could not tell from their experimental data whether the participants were “consciously aware of the male-bias generic masculines create in their minds” (1999:562)⁷¹.

However, rather than prove the generic nature of *he*, Weseen’s observation about feminine grammatical markings attracting attention only serves as further evidence for the claim that *he* cannot have a generic interpretation. Although Madson and Hessling’s (1999:569) data suggests that, whilst *she* draws attention to gender, “unmarked terms like *he* blend into [a] text, allowing readers to pass over them without much thought”, the unconscious processing of *he* does not necessarily allow a generic reading. If anything, the unconscious processing and acceptance of the default masculine value of *he* evident in the literature I evaluated in section 1.2.2 draws attention to why the masculine form cannot be a true epicene.

Nevertheless, according to Baranowski (2002:379, based on Stanley 1978), by 1906 generic *he* was considered the only epicene form by American grammarians, and from this point forwards it was the form prescribed in the majority of grammar books up until the 1970s. Baranowski (2002:379) argues that grammarians had a “fixation with number concord” and as such, other epicene candidates such as singular *they* were ignored, or dismissed as incorrect. However, Baron (1986:194) is quick to point out that in 1929 Leonard noted that despite traditional prescriptive grammarians’ views on epicene pronouns, their prescriptions on this matter “did not greatly influence nineteenth century usage”. Grammars can only go so far (see section 2.3.3) and Zuber and Reed claim that even in the twentieth century, in the 1930s, “panels of educated writers and editors” did consistently accept certain uses of singular *they* (1993:521)⁷². This finding reinforces the notion that prescriptions on epicene pronouns may never have reflected usage.

71 Expanding on their previous work Madson and Hessling (2001) looked at how readers perceived the use of singular *they*. Using an essay reading exercise they tested the perception of combined pronouns, singular *they*, and pronoun alteration between paragraphs (as with their 1999 study). However, despite directly testing singular *they*, the authors have very little to say on the pronoun, noting only that “the use of ‘they’ in singular contexts may not be ideal as readers perceived this strategy as low in overall quality” (2001:175).

72 In chapter six I show that the same limited acceptance is still a feature of grammarians’ views on singular *they*.

Also, much like Gould Brown in the nineteenth century, grammarians in the twentieth century changed their minds on epicene prescriptions. For example, in 1931 George O. Curme conceded that singular *they* occurred in “popular speech” (1931:557-558 in Baron 1986:193), but sixteen years later he ignored his previous comments and endorsed generic *he*:

The masculine pronouns and possessives are usually employed for persons without regard to sex wherever the antecedent has a general indefinite meaning

(Curme 1947:221 in Jochnowitz 1982:198).

However, by recognising that singular *they* was used in speech, Curme was actually taking a step towards descriptive analysis and away from simply following traditional grammatical prescriptivist norms. Comments, such as those made by Curme indicate that arguments against the use of generic *he* as epicene were gaining momentum in the mid twentieth century. In the following section I continue the chronology of epicene prescriptions, and indicate that current scholarship supports the position that the rise of second-wave feminism had a discernible impact on the use of generic masculines.

My review of the literature on the development of grammatical prescriptivism and the related grammar “rules” on epicenes, indicates that there is evidence for the argument that from the outset of printed prescriptive norms, generic *he* has traditionally been promoted as the English epicene pronoun. However, based on scholarship considering the history of grammatical conventions it seems that descriptive studies of language were not used as support for the prescription of generic *he*. Furthermore, the creation of language “rules” was clearly influenced by, and interacted with, industrial developments and societal norms, such as the differing roles of men and women and the rise of the middle class.

My evaluation of the literature also shows that prescriptions on epicene pronouns appear to have been a point of contention even in the eighteenth century, with discrepancies between grammarians (and even between different grammars written by the same grammarians) suggesting that traditional prescriptive grammarians’ pronouncements on *he* may never have reflected epicene usage. There is a large body of literature in support of this argument, indicating that gender-neutral pronouns are not merely a

syntactic phenomenon, as there is evidently influence from language-external factors on the epicene debate. Indeed, despite the rejection of singular *they* evident in scholarship on traditional grammatical prescriptivism, my review of the literature has shown that there is evidence indicating that singular *they* was still used in the eighteenth and nineteenth centuries.

In the twentieth century, debates over English epicene pronouns were ongoing and Baron (1986) claims that the rise of second-wave feminism and the non-sexist language reforms that would follow it impacted upon the epicene debate. Support for singular *they* was based on arguments that it had been “standard English for the word *everybody* for more than four hundred years” (Evans and Evans 1957 in Baron 1986:195), combined with research including Bodine (1975) and Stanley (1978), and the studies I cited in chapter one, which illustrate the non-generic nature of generic *he* (Baron 1986:196). I deal with the literature on this period of epicene prescriptions in the following section.

2.2: THE EFFECTS OF SEXUAL POLITICS

In this section I consider the state of the epicene debate around the onset of second-wave feminism in the 1960s and 1970s. I use this historical base to draw on relevant material to argue that there was disillusionment with generic *he*, highlighted by groups of feminist language reformers. I briefly summarise and examine the origins and development of second-wave feminism in section 2.2.1 before focusing on feminism-based views on language, including a “feminist attack” on generic *he* which “began in force about 1970” (Bodine 1975:130). Baranowski (2002:279) states that opponents of prescriptions of generic *he* claim that it violates gender concord, and results in “ambiguity, [and] exclusion of women”, whilst “equating maleness with humanness”. Arguably this position depends upon accepting the argument that generic *he* is masculine by default (see chapter one). In section 2.2.2 I evaluate the literature on the adoption of non-sexist language reforms surrounding generic *he* and note that alternative epicene candidates, such as singular *they* have come under criticism. I also begin to note the ties

between epicene prescriptions, standard English, grammar handbooks, and education, which I focus on more closely in the final section of this chapter.

2.2.1: FEMINISM AND LANGUAGE

There is a general consensus in the literature that the seeds of second-wave feminism were planted in the late 1960s. Meehan (1990:190) argues that “women’s activism, whether in avowedly feminist politics or not, has risen substantially” from this point onwards. She proposes that the commemoration of “the fiftieth anniversary of women’s suffrage in 1968” was the main “catalyst” for the emergence of the second-wave movement (1990:193). Similarly, Rowbotham (1989:3) claimed that women’s liberation “erupted in the late 1960s”, suggesting that it was born from disillusionment with the realities of being a woman within the family dynamic. These two differing opinions on the motivations behind second-wave feminism illustrate the argument that no single factor launched the women’s movement. Indeed, Platt (2007:962) argues that second-wave feminism in Britain “cannot be given a precise chronology”.

One reason why there is no clear-cut developmental history is because there is evidence in the literature for the argument that second-wave feminists were not a homogenous group. Just some of the issues highlighted by second-wave feminism included a focus on getting the Equal Pay Act passed in the UK in 1970 (Dale and Foster 1986:131), and the establishment of campaign groups with specific focuses, such as WOAR (Women Organized Against Rape). Yet Whelehan (1995:26) argues that despite these differing approaches and goals, second-wave feminists tended “to foreground the same substantive issues”. Indeed McLaren (2002:19) states that there was an overarching belief that united all feminists; they recognised “that women have been subordinate to men” and McLaren claims that, as such, “the primary aim of feminism is to overcome this subordination”⁷³. Scholarship indicates that some feminists directly attributed part of this subordination to language, and Sontag claimed in 1973 that generic *he*, and other generic

⁷³ See also Pauwels (1998:92) for the argument that the common goal underlying feminism is the elimination of “all forms of discrimination against women”.

masculine forms were “the ultimate arena of sexist brainwashing” (Martyna 1980:75).

According to McLaren (2002:8), it was radical feminists who began to draw attention to the power of language, and believed that words were “potent transmitters of social and cultural values”. She claims that one aspect of radical feminism was to “urge feminists to reclaim and revalue words that have derogatory connotations and devalue women” (2002:8). Similarly, Crawford (2004:228) argues that “as long as feminism has been a social movement, language has been a battleground”, and according to Cooper (1989:62) the creation of language policies influenced by feminism (section 2.2.2) shone a light on “sex bias in language” which up until that point in history had been “part of the grounds of everyday life”. Cooper directly attributes a “heightened awareness of such bias” to the women’s liberation movement (1989:62), and Pauwels (1998:17) notes that such a focus on “the linguistic treatment and representation of women is said to be characteristic” of second-wave feminism. In terms of epicenes, Jochowitz (1982:200) notes that some feminists such as Miller and Swift (1976) endorsed singular *they* instead of the form *he or she*, as Magner (2002:272) argues that feminists “came to see language itself as a major roadblock” in the way of equality. Indeed McLaren suggests that for these feminists, “language not only describes, but also creates, reality” (2002:8).

According to Shute (1981:31) justification for linguistic action within a feminist framework is based on the argument that because sexist language⁷⁴ “is just another instance of sexism”, it should be treated like any other form of sexist discrimination. Shute argues that a non-sexist society is impossible if sexist language exists, seeing it as a device that specifically “serves to limit the activities of people of one sex but not those of the other” (1981:31). This view is crystallised by Jacobson and Insko (1985:1) who claim that accepting the idea that “societal sexism influences linguistic sexism is not enough”. They argue that the relationship between language and sexism works both ways, and that “sexism in language helps to perpetuate sexist attitudes” (1985:1). Linking back to the discussion of acquisition in section 1.1.3, Parker (1977) argues that if sexist language occurs in input to children

74 I adopt Baker and Ellece’s (2011:77) definition that sexist language is “language which discriminates on the basis of gender or biological sex”, whereas non-sexist language is characterised as “a deliberate attempt to avoid using gender discriminatory words”.

they may “perceive [it]... to be part of the core of English” and as a result it may become entrenched in their grammars (Wolfe 1989:81). Indeed, Pauwels notes that as language “[shapes] people’s views of reality”, it is possible to “see a direct, even causal, link between women’s subordinate status in society and the androcentrism in language” (1998:xi-xii).

As a result of such opinions, Simpson (1993:160) argues that the worldview supposedly perpetuated by sexist language has been heavily investigated and “occupies a central role in feminist linguistics”. Equally, Cameron (1998:150) notes that it was feminists in the 1970s who argued that “languages which mark gender assiduously in their grammars, and treat the masculine as the unmarked gender, will lead their speakers to perceive the world in gender-polarised and androcentric ways”. This is a key argument in the feminist rejection, not only of generic *he*, but of terms overtly marked for gender such as *chairman* and *stewardess*, which supports the more general claim that he/man language excludes women (see Crawford 2004:238-241 for a summary). Therefore, there is support in the literature for the argument that in the 1970s the use of generic *he* was classified as an example of sexist language and opposed on the grounds that its masculine value meant that it could not be used gender-neutrally.

According to Stanley, masculine generics only became “correct” because of “male control of the educational establishment in England” (1978:800). Indeed, my review of the literature on the development of traditional prescriptive grammar does seem to indicate that there is support for this position, as the majority of grammars produced in the latter half of the nineteenth century were written by men. There is no evidence to suggest that Stanley’s argument cannot also be applied to the prescription of generic *he*, and thus it is arguable, following Cook and Suter (1980 cited in Wolfe 1989:85) that traditional grammatical prescriptivism has created a “sexually biased” but standard English pronoun in generic *he*.

Having analysed the language and identified examples, such as generic *he*, which they believed to be instances of sexism, groups of feminist language reformers had to decide how to publicise the changes they wished

to propose⁷⁵. However, there is a body of scholarship supporting the argument that certain avenues of implementation were simply not open to feminist ideas. Formal language bodies “were often among the most vehement opponents of the proposed reforms” (Pauwels 1998:143). One explanation for this is that such institutions “regard themselves as the guardians of language” taking a conservative stance towards change (1998:143)⁷⁶. According to Pauwels (1998:101) restrictions on access to formal language bodies meant that feminist language reformers had to find alternative ways to promote change and expose “the discriminatory treatment of women in language”. Pauwels (1998:144) argues that by far the “most widespread and popular strategy” chosen for this purpose is “the formulation and distribution of language guidelines”.

According to Cooper (1989:38) feminist language reform, that is, the identification of sexist language and promotion of alternative forms, was “an instance of language planning... adopted at high levels of authority”. Similarly, Cameron (2006:20) claims that the promotion of non-sexist language is “one of those feminist ideas that has somehow managed to achieve the status of orthodoxy” in society in general, not just in feminist circles. If the consensus in current scholarship holds then, following Crawford (2004:240), the current situation is that advice on non-sexist language has been consistently in publication from the 1970s onwards in a variety of institutions, including government agencies, educational bodies, and professional organisations (see also Madson and Hessling 1999).

However, scholarship also indicates that there has not been a uniform adoption of feminism-based proposals for language change, and Cooper (1989) documents one of the main arguments against the promotion of non-sexist language:

75 Feminist language reforms do not just involve “the purging of certain words”, the focus is more on drawing attention to problematic constructions and, according to Vetterling-Braggin (1981:54) “halting the *use* of such words” because of the implicatures they arguably carry about “conceptions of women and men or about women’s and men’s roles in society”. Feminist language reform has also led to additions to language as it facilitated the creation of terms like *sexual harassment*.

76 However, this may not be the whole picture; to take an international example, L’Academie Francaise only elected its first female member in 1980 (Noordenbos 2002:132). This gender imbalance, and thus lack of formal female influence, could partially explain why such institutions are reluctant to consider proposals of linguistic change based on feminist ideas, although the election of a female *Immortel* has little impact on issues of non-sexist language if she is not concerned with such issues.

It is one thing to demonstrate that people tend to envisage males rather than both males and females when encountering androcentric generics. It is quite another to demonstrate that the use of androcentric generics contributes to sexual discrimination

(Cooper 1989:18).

The underlying issue here is to what extent one assumes a causal relationship between language norms and societal norms. Rather than being solely concerned with the linguistic effects of second-wave feminism, the argument highlighted by Cooper really questions the extent to which language shapes people's experience of "reality" by influencing their perception of the world⁷⁷.

Despite arguments against language reform based on the concept of sexist language, there is evidence in the literature that second-wave feminism, and the language reforms that it facilitated (see below), did have "an observable impact on the written usage of publications addressed to the general audience" (Cooper 1989:20). Indeed, Pauwels' (2001) study, discussed in section 1.1.1, on the epicenes used on radio stations indicates that language policies on non-sexist language can directly affect usage. Thus, language has changed because groups of second-wave feminists drew attention to usage they found problematic, usage which, significantly for this research, included generic *he*.

Based on the preceding review of scholarship I argue that there is weight for the position that there are parallels between traditional grammatical prescriptivism and feminist language reforms. Although Frank (1989:133) argues that where traditional grammatical prescriptivists were "basically elitists" who believed that "their rules embodied the ultimate authority on linguistic 'purity'", those concerned with feminist language reform are instead "conducting a popular movement" with core values of "fairness and equity". However, as Frank's other works include *Sexism, Grammatical Gender, and Social Change* (1978) and *Language and the Sexes* (Frank and Anshen 1983) it is possible that her claim contains some biases, yet I do not intend to imply that Frank's claims should be discounted. What is significant here is that, in line with Cameron (2006:23), "it seems self-

⁷⁷ This conceptualisation of language reflects the Sapir-Whorf hypothesis that "our grammar shapes our thought" (Whorf 1956 in Gastil 1990:630), meaning that language perceived as male-biased or sexist influences "sexist attitudes and behaviors in a... subtle, psychological manner" (Gastil 1990:630).

evident” that both the traditional norms discussed in section 2.1 and non-sexist language guidelines are both prescriptive.

The guideline approach appears to have been (partially) successful for some feminist language reforms, such as the adoption of *Ms* (see Eckert and McConnell-Ginet 2003:53 and Mühlhäusler and Harré 1990:247). Such changes led Crystal (1984) to claim that “the feminist campaign against sexist language was amongst the most successful instances of prescriptivism in living memory” (Cameron 1995:118). However, Crystal’s statement implies that feminist language reform is complete, yet a review of the literature indicates that this is not the case, and in the following section I show that there is a large body of literature in support of the position that there has been no blanket acceptance of feminism-based linguistic reforms.

2.2.2: FEMINIST LANGUAGE REFORM AND EPICENES

Despite current scholarship indicating that the rejection of generic *he* has had an observable impact on epicene use (c.f. Pauwels 2001), there is evidence in the wider literature suggesting that the influence of feminist language reforms on epicene pronouns is debatable. For example, Crawford and Fox (2007:483) argue that feminist initiatives on epicene pronouns are an example of language reform which has “failed spectacularly”. Similarly, in 1978 Stanley argued that “male grammarians have succeeded in their efforts to promote number concord as their primary issue” (1978:810)⁷⁸. Thus Stanley’s claim suggests that traditional prescriptive grammarians’ promotion of number concord outweighs feminist language reformers’ promotion of gender neutrality (if indeed the two groups can be polarised as such). Whilst the date of Stanley’s work is significant, as it represents an interpretation of the state of affairs around the time of second-wave feminism, more recent work still suggests that feminism-based rejections of generic *he* have not been universally accepted. According to Pauwels (1998:181) singular *they*, which is arguably the main alternative epicene to generic *he*, has “attracted severe criticisms... on the grounds that it is

⁷⁸ However, as the discussion in section 2.1 showed, it is not necessarily the sex of the grammarian that is important, but rather their position in relation to social norms.

linguistically incorrect". Pauwels' view gives weight to Stanley's (1978:800) argument that authors using generic *he* are "appealing for authority to the men who have gone before [them]".

Stanley argues that even those "modern grammarians who do mention the problem of pronoun reference" see the male dominance that feminist language reformers have attached to generic *he* as "hardly worth mentioning"(1978:810), despite the large body of evidence of its default masculine value (see chapter one). Even when singular *they* is used, Jacobson and Insko argue that "it is not always clear whether such usage stems from a desire to be non-sexist or from an ignorance of proper grammar" (1985:1). In this case, "proper grammar" must assumedly refer to traditional prescriptive grammar. Thus, there is evidence of epicene conflict in the wider literature on non-sexist language, supporting the position that any changes instigated by language reform policies are far from complete.

Interestingly, some feminist language reformers have noted that many of those who oppose non-sexist guidelines, especially on the grounds "that they impose censorship, are often the ones who are strong advocates of... other language prescriptions" (Pauwels 1998:183). Thus, they are opposed to censorship based on the criterion of non-sexist language but accept traditionally prescribed norms. This cannot be explained linguistically as there is no evidence in the literature for the argument that a particular form of language prescription is superior to another. Therefore, the choice to adhere to certain linguistic "rules" must be analysed from a social perspective, and research, which I detail below, has shown that judgement of non-sexist language reform, and its related guidelines, is linked to wider language-external factors such as political allegiance and biological sex.

Using the Attitudes to Women Scale developed by Spence and Helmreich (1972), Jacobson and Insko (1985) found a positive correlation between allegiance to feminist views and the use of combined pronouns (and generic *she*) with gender-stereotyped antecedents. Despite the fact that Jacobson and Insko say the correlation cannot be the basis for strong, general conclusions (1985:5), the result still indicates an interaction between language-external factors and language, insofar as an awareness of feminist politics influenced the pronominal choice of those in the study. Additionally, Harrigan and Lucic (1988) investigated whether the adoption of

gender-neutral pronouns interacted with wider contextual factors such as group membership. They used a questionnaire to examine the attitudes to gender-neutral language, and more specifically pronouns, of five groups: members of a NOW branch, university faculty, medical students, English graduate students, and psychology graduate students. When asked whether “English should have a sex-neutral pronoun”, the members of NOW and the psychology students agreed, the university faculty members replied *maybe*, and the English and medical students were *not sure* (1988:135). However, significantly for this research, Harrigan and Lucic also note that singular *they* was the form “most likely used to avoid gender-biased language” (1988:137).

Interestingly, Harrigan and Lucic reported that the English students (along with the medical students) “were less likely to shift to gender-neutral alternatives for pronoun selection” (1988:137). The authors proposed that this is because the English students “had difficulty” using pronouns which violated number concord and suggested this could be because the students “may wish to keep the language ‘pure and proper’ in a manner similar to the French” (1988:137). Although Harrigan and Lucic gave no evidence for the argument that the English students aimed to preserve language, I argue that their participants may have been influenced by the promotion of number concord associated with traditional prescriptive grammars (and their modern counterparts, see section 2.3). For example, in Adamsky’s (1981) study of generic *she* one participant claimed that “It would be difficult for me to use it [*she*] because of what has been ingrained in me previously” (1981:778). I argue that such comments provide evidence for the impact of traditional rules of grammar and the related promotion of generic *he*.

In a similar vein to the studies on political allegiance, noted above, research has also shown that there is correlation, between sex and the uptake of feminist language reforms. For example, Rubin and Greene (1991) analysed how 128 undergraduate students and 119 college graduates between 30 and 45 perceived a selection of feminist language reforms. Their results show that whilst 77.2% of the participants agreed that English included some sexist elements (1991:398), Rubin and Greene argued that it was the women in the study who had “most to gain by language-inclusive reforms, and their attitudes appear to be correspondingly more positive than men’s” (1991:404). Indeed, similar findings were also borne out in studies by

Parks and Robertson (2005, 2008). The results of such studies show how political allegiances (which, in this case, may be influenced by biological sex) and “the desire for language change” interact with both “psycholinguistic factors governing the best change to make” and other social factors concerning how changes are implemented (MacKay 1980:364).

More recent data comes from Laitinen’s (2007:255) corpus-based study of the BNC which shows that “women consistently use the plural *THEY* as the epicene anaphor more than men”, and that “women rarely resort to epicene *HE*”. One explanation for this finding is that Laitinen’s results could relate back to arguments about how men and women have different experiences with learning the value of generic *he* (see the arguments of Nilsen 1977, Martyna 1980, and Gibbon 1999 in section 1.1.3). Thus, the men in the BNC sample, having perhaps never adopted a generic interpretation of *he* may not be using the pronoun gender-neutrally.

Similar recent evidence for the non-generic nature of *he* comes from Balhorn’s (2009:403) corpus-based analysis of the pronoun uses of male and female authors and journalists, which shows that “neither male nor female writers can get an epicene reading or interpretation for the male third-person pronoun”. Significantly, Balhorn concludes that, as an alternative, singular *they* “with its nonspecification for sex and one-syllable phonemic realization rolls off the tongue and out of the unedited pen” (2009:406). Data such as this is evidence for the argument that non-sexist language reform has “become more firmly entrenched” in society and “influential social institutions”, arguably making non-sexist language “more normative” (Rubin and Greene 1991:394).

Therefore, based on the literature I have surveyed it appears that Crawford and Fox (2007:483) have overstated their claim, noted above, that feminist initiatives on epicene pronouns have “failed spectacularly”. There is no evidence in wider research to suggest that feminist language reform is a completed process, and as I noted in section 1.1.2, the closed-class nature of the pronoun paradigm and its apparent resistance to change is arguably evidence for the position that it may take time for a pronominal change to infiltrate the paradigm. Based on the above discussion of current scholarship it is clear that second-wave feminism, feminist language reforms, and non-sexist guidelines have had a notable impact on the epicene debate, and the

following anecdote from Pauwels (1998) illustrates just how feminist language reforms on epicenes have directly influenced society:

An airline company was taken to court in Berlin for using the words “jobholder...he...” in its English version of a job advertisement. In its defence the company claimed that they had used the terms “jobholder” and “he” in a gender-neutral manner. This argument was rejected by the court on the grounds that “he” could not be interpreted as gender-neutral in this context

(Pauwels 1998:147).

Evidence such as Pauwels’ example supports the argument that generic *he* is out of favour as the English third-person singular epicene. Furthermore, McConnell-Ginet (1989) argues that using generic *he* has now “become more difficult because more people have become aware of its controversial nature as generic” (cited in Pauwels 1998:219). Indeed, Hellinger and Schräpel (1983:183) oppose “having to wait for (con)textual or situational signals” to determine whether women are included or excluded from an expression (Pauwels 1998:178) and Martyna draws attention to how it is difficult to interpret whether an instance of *he* is meant to be generic or masculine (1978:131-132).

However, there are clear indicators in the literature to suggest that there has not been a blanket acceptance of feminist language reforms on the rejection of generic *he*. At Harvard in 1971 protests rejecting generic *he* evoked a surprisingly negative response from the linguistics faculty (see Cameron 1992:93-94). Seventeen members of the linguistics staff endorsed the following letter in the campus newspaper (Figure 2.1).

EXTRACT FROM THE *HARVARD CRIMSON* 16TH NOVEMBER 1971, PAGE 17:

MANY OF THE GRAMMATICAL AND LEXICAL OPPOSITIONS IN LANGUAGE ARE NOT BETWEEN EQUAL MEMBERS OF A PAIR BUT BETWEEN ONE OF WHICH IS MORE ‘MARKED’ THAN THE OTHER... FOR PEOPLE AND PRONOUNS IN ENGLISH THE MASCULINE IS THE UNMARKED AND HENCE IS USED AS A NEUTRAL OR UNSPECIFIED TERM. THIS REFLECTS THE ANCIENT PATTERN OF THE INDO-EUROPEAN LANGUAGES... THE FACT THAT THE MASCULINE IS THE UNMARKED GENDER IN ENGLISH (OR THAT THE FEMININE IS UNMARKED IN THE LANGUAGE OF THE TUNICA INDIANS) IS SIMPLY A FEATURE OF GRAMMAR. IT IS UNLIKELY TO BE AN IMPEDIMENT TO ANY CHANGE IN THE PATTERNS OF SEXUAL DIVISION OF LABOR TOWARD WHICH OUR SOCIETY MAY WISH TO EVOLVE. THERE IS REALLY NO CAUSE FOR ANXIETY OR PRONOUN ENVY ON THE PART OF THOSE SEEKING CHANGES.

(PAUWELS 1998:69)

FIGURE 2.1: HARVARD CRIMSON

The fact that so many members of staff endorsed the letter suggests that epicene pronouns are a point of contention in language, but the way the faculty belittled arguments against generic *he*, by coining the term

“pronoun envy”, indicates that the faculty portrayed the issue as unimportant⁷⁹. Perhaps most interestingly, the reason the staff gave for their condemnation of the protests was that “the English generic masculine is simply a feature of grammar” (Cameron 1992:95). Arguably, this view portrays language as an independent entity, a view which both illustrates and denies the effects of traditional grammatical prescriptivism. Yet my review of studies on the history of traditional grammatical prescriptivism and its social motivations (section 2.1) show that generic *he* is not just a feature of grammar. As the empirical evidence I evaluated in chapter one has qualified, “the psychological impact” of generic *he* “involves much more than what the Harvard linguists once termed ‘pronoun envy’” (Martyna 1980:70).

In a similar example, Bate (1978:145) interviewed twenty university faculty members and found that eleven were opposed to singular *they*, preferring instead *he or she* (16 subjects) or generic *he* (10 subjects). However, six of the eleven people opposed to singular *they* used the form in the interview section of Bate’s study. Bate’s results provide more evidence in support of the position that singular *they* is the spoken epicene of choice (see Stringer and Hopper 1998) and adds to the body of literature already discussed which showed that prescriptions did not necessarily reflect language use. These two examples, the letter in the *Harvard Crimson* and the study of university faculty undertaken by Bate (1978), lead into the final section of this chapter, in which I survey the literature on epicene use in education and show that there is evidence for the position that (traditional or feminist) prescriptive grammar can influence classroom usage.

In this section I have established that there is support in the literature for the position that second-wave feminism has raised awareness of, and instigated the challenging of certain linguistic norms which could be classed as *sexist* due to the way in which men and women are represented asymmetrically by them. Scholarship shows that one such norm which was

⁷⁹ One way of interpreting this clear link to Freud is that, as penis-envy (the clear parallel of pronoun envy) was deemed by Freud to be “the crux of female development” (Burke 1998:xiv), it is therefore arguably implied that those campaigning against generic *he* have not yet developed to the level of the authors. Interestingly, it is also not without note that Freud argued that penis envy “becomes the necessary pre-condition for the achievement of the form of ‘femininity’ required by patriarchy” (Minsky 1996:50).

challenged was the use of generic *he* as an epicene pronoun, a practice which, as I have shown in section 2.1, is heavily associated with and perpetuated by traditional grammatical prescriptivism. My review of the literature has shown that up until the start of the twenty-first century the rejection of generic *he*, on the grounds that it cannot include women in its possible referents, appears to have been relatively successful, although it has not necessarily corresponded with a related endorsement of singular *they*.

As I noted above, there is evidence in the literature to suggest that one area where the impact of epicene prescriptions can be seen is within the realms of education. Therefore, in order to investigate the continuing influence of grammatical prescriptions on epicene pronouns, in the final section below, I evaluate current scholarship on the impact of language prescriptions in the classroom. Furthermore, I also consider the large body of literature concerned with the continuing discussion of epicene pronouns in grammars and language handbooks.

2.3: EPICENES IN EDUCATION

My main goal in the final section is to complete the chronology of epicene prescriptions by evaluating the literature on their perpetuation in grammars and handbooks after the initial onset of second-wave feminism. I also explore the links between epicenes and education, as my review of the literature in the preceding section suggests that any observable connections would be significant to the epicene debate. Therefore, in section 2.3.1 I synthesise the results of studies on epicene prescriptions that have used relatively modern school and college text books as their primary sources. Scholarship shows that although generic *he* may now be out of favour, grammar texts published during and after second-wave feminism do not endorse singular *they*. However, my review of the literature in section 2.3.2 indicates that the most recent published scholarship in this field is now almost twenty years old. I conclude this section by evaluating research that

has focused on investigating how much influence grammars and traditional prescriptive rules actually have over attested epicene usage.

2.3.1: “CORRECT” LANGUAGE IN EDUCATIONAL HANDBOOKS

A survey of the scholarship on grammar guidelines in textbooks indicates that the importance of the link between grammatical prescriptivism and education cannot be underestimated. Pauwels (1998:23) claims that school grammars and handbooks have an “alleged importance as repositories of linguistic knowledge” and their analysis provides an insight into the linguistic norms of education. For example, when working on the corpus of school textbooks used to create *The American Heritage School Dictionary*, Graham (1973:58) found that the ratio of masculine to feminine pronouns “was almost four to one”, and she argued that the majority of the occurrences of *he* did not allow a generic interpretation. Graham found that in an “experimental sampling” only 32 out of 940 occurrences of *he* “referred to the unspecified singular subject” (1973:58). The overwhelming majority referred to male humans, male animals, or indefinite referents which were “assumed to be male” due to their social roles/occupations, e.g. *sailor* (1973:58)⁸⁰. To tie this in with the discussion of acquisition in section 1.1.3, Graham’s findings support the claim that children do not receive generic *he* as L1 input in educational materials.

Graham’s data supports the idea that “human beings were to be considered male unless proven otherwise” (Bodine 1975:133), at least within educational texts. This argument can be tied in to Weseen’s argument that only the feminine gender expresses sex (1928, see section 2.1.2), insofar as it appears that the masculine is the default sex, whilst the feminine is atypical. According to Gibbon (1999:42) the apparent “slippage”, where generic *he* is used gender-specifically, facilitates an “invisibility of women” where “human experience [is categorised] in terms of males’ experience”. Indeed, this argument supports Graham’s (1973) depiction of how

⁸⁰ Importantly, Graham also claimed that the dictionary was the first of its kind “to define sexism, to include the phrase liberated women, and to recognize Ms”, yet it was “a wordbook for children” (1973:57).

asymmetric linguistic representation of men and women creates two unequal subgroups of a whole:

If you have a group half of whose members are A's (sic) [men] and half of whose members are B's [women] and if you call the group C [people] then A's and B's may be equal members of group C. But if you call the group A [or use generic masculine forms], there is no way that B's can be equal to A's within it. The A's will always be the rule and the B's will always be the exception - the subgroup, the subspecies, the outsiders

(Graham 1974 [1973] in Treichler and Frank 1989:146).

Linking to education, Silveira (1980:174) argued that students can “develop people=male bias in...[their] thinking in part because...[they] are rewarded for doing so and punished for not doing so”. She notes that using generic *he* may warrant an A+ grade as it is “correct” grammar, whilst the use of singular *they* may be marked as an error, or bad grammar, worthy of only a C- (1980:171). This position is supported by Cook and Suter (1980) who note that whilst generic *he* is “perhaps ‘sexually biased’ [it] is acceptable in Standard English” (Wolfe 1989:85). Cook and Suter’s argument is that if a person uses generic *he*, they may be considered sexist, or rather “chauvinistic or socially unaware” but they will not be deemed unintelligent or illiterate (1980:147 in Wolfe 1989:85).

Such viewpoints correspond with Sunderland’s (1986) analysis of 22 school grammar textbooks, which showed that although there had been an increase in discussion of sexist language between 1975 and 1985, “in most cases the innovations or non-sexist practices... [were] given a negative evaluation” (Pauwels 1998:207). Similarly, Bodine (1975) analysed thirty-three American high school grammars that were being used in the US in the 1970s with publication dates ranging from 1958 to 1967 (the majority being from the mid-to-late 1960s and thus pre-second-wave feminism). She found that twenty-eight of these books (an overwhelming 85%) advised against using either singular *they* or combined pronouns, such as *he or she*, and only three of the twenty-eight books gave an “adequate explanation of the use of ‘they’”, although what is meant by *adequate* is not explicitly stated. Her findings led Bodine to call the authors of the grammars “the docile heirs to the androcentric tradition of the prescriptive grammarians” (1975:139).

These results of investigations into school grammars occur despite the large body of scholarship in support of the argument that generic *he* is only

favoured in standard English because of its selection as the epicene pronoun of traditional grammatical prescriptivism. For example, Cameron (1992:95) argues that “grammarians intent on prescribing rules of correct usage preferred *he* over *they* and stigmatised the latter as incorrect”. In Valian’s (1977) view, generic *he* “was not really a natural part of the pronominal system” but was rather “a prescriptive imposition” with sexist undertones (Newman 1992:450). Thus, Mühlhäusler and Harré (1990:232) argued that it is “indisputable” that generic *he* is, “if not exactly a grammarians’ invention, a usage pressed on a *they*-using public”. This assessment agrees with Bodine’s (1975:131) ideas on how backlash against generic *he* is simply “a counter-reaction to an attempt by prescriptive grammarians to alter the language”⁸¹.

However, Pauwels (1998:163) notes that within grammars and textbooks it appears that there are discrepancies between grammar guidelines on “non-sexist alternatives for generic pronouns”. If Pauwels’ argument holds then this is evidence for the position that epicene pronouns are still a point of contention in grammars (arguably due to the fact that prescriptions may never have reflected usage). Furthermore, Gershuny (1989:98) argues that grammars “frequently include short discussions of... the problem of pronoun agreement when the antecedent is of unspecified sex”. For example, whilst recognising that 1970s grammars “struggled with the fact that singular *they* seems more natural than generic *he*” (at least in some contexts), Foertsch and Gernsbacher claimed that in the 1990s “most begrudgingly allow writers to use *they* as a pronoun for two limited classes of singular antecedents”, these are indefinite pronouns, and “corporate nouns” or collectives (1997:106)⁸². Similarly, Meyers notes that Quirk, Greenbaum, Leech and Svartvik included limited uses of singular *they* in their 1985 grammar, having “questioned them in 1972” (1993:187). However, Foertsch and Gernsbacher also note that some modern grammars, i.e. grammars written after the initial influence of second-wave feminism on language, were opposed to singular *they* as it “violates the expectations of

81 Bodine’s epicene of choice is singular *they*, and she claims that it standard form until it came “under attack” from a group of unnamed “prescriptive grammarians” and was subjected to over two hundred years of “vigorous attempts to... regulate it out of existence” (1975:131). She notes that whilst singular *they* is “alive and well”, its current usage is “remarkable” as “virtually the entire educational and publishing establishment has been behind the attempt to eradicate it” (1975:131).

82 In chapter five I show that such limitations are still placed on singular *they*.

most readers” (Fowler and Aaron 1983:195 in Foertsch and Gernsbacher 1997:106). Such changes in prescriptions can be used as evidence for the argument that grammarians’ rulings on epicenes are not stable.

Supporting evidence for this position comes from Zuber and Reed (1993) in their analysis of 1980s and early 1990s grammar handbooks, (which is, unfortunately, the most-recent grammar book epicene study before this work, see chapter five). They claim that grammarians see generic *he* as “a tradition ‘rooted in the beginnings of the English language’” (1993:523), and Zuber and Reed argue that some grammar books published as late as the early 1990s had “actually returned to a more prescriptive view” of singular *they* (1993:525). This is substantiated by Madson and Hessling’s (1999:571) claim that “some style guides... continue to prohibit” singular *they* because “some readers” see it as “inappropriately informal or grammatically incorrect”. According to Zuber and Reed this means that by privileging number concord, grammarians had “reverted to the traditional rational established by the 1850 Act of Parliament”, (see section 2.1.2), which in effect hid the “linguistic discrimination” of generic *he* (1993:526).

Zuber and Reed’s claims are supported by their analysis of “the addition or deletion of sections on nonsexist language” in multiple editions of six American college-level grammars (1993:526). Their research identified three main approaches to singular *they* taken by grammarians in the late twentieth century; the authors a) ignored or denied singular *they*, b) restricted it to spoken language, or c) “tentatively” conceded its written use (1993:524). However, Zuber and Reed did not document how many of the grammars adopted each approach, and gave no indication of which approach was most popular, or how approaches changed between different texts. Documentation of this could have supported their claim that the relatively frequent revisions of the grammars indicate “conflict between forces for language change and the tradition of authority” (1993:517). Despite finding that some of the teachers’ editions did consider singular *they*, the students’ editions did not (1993:525), and Zuber and Reed concluded that grammar writers ignored “the professional literature in linguistics... which for several

decades has indicated the discrepancy between actual linguistic practice and the handbook proscriptions” (1993:527)⁸³.

The argument that prescriptions do not reflect usage is not a new criticism, as whilst discussing the difference between grammar books from the 1900s and the 1940s Dawson (1956:36) questioned why “textbook authors [are] not willing to accept the changed usage and give it space without parenthesis”. Dawson argued, independently from feminist language reforms, which had not yet made their impact, that if “Everyone can bring *their* own lunch to the picnic”, grammatical rules about pronominal agreement should be revised (1956:37). Dawson’s argument corresponds well with MacKay’s comments about late twentieth-century prescriptivism, providing evidence for the position that prescriptions do not reflect usage:

By ignoring linguistic knowledge, prescriptivism has remained narrow, uninformed, and unprincipled, following arbitrary, unconscious or poorly formulated criteria and biases rather than general rules or principles

(MacKay 1980:350).

Evidence for the argument that grammar writers ignored linguistic knowledge comes from Meyers’ (1991:343) study of the glossaries of fifty different grammars, which he found to be “riddled with practices that descended in an unbroken line from the grammarians of the eighteenth century”. His study also illustrated an overwhelming inconsistency between the contents of grammar handbooks, as only one grammatical feature was common to all fifty books he analysed, although this was not epicene pronouns. This suggests that the consistency of mention of epicenes in Bodine’s (1975) and Sunderland’s (1986) samples, and Zuber and Reed’s set of grammars, is rare and singles out the epicene debate as a key point of grammatical contention.

The above collection of studies indicate that, despite attempts within the scope of feminist language reform to highlight the non-generic nature of generic *he*, and subsequently reject the prescription of it (section 2.2.3), current scholarship suggests that this has had little impact on (relatively modern) grammar handbooks. Also, my reviews of studies such as Zuber and

83 See also Amare’s (2007) study on sexist language in online grammar guides supported by American academic institutions, and Schaffer (2010) provides a list of online grammar blogs where usage is debated and prescribed.

Reed (1993) indicate that singular *they* is still out of favour in school/college grammars and language guidelines. However, the literature review also shows that prescriptions are not uniform. I explore this issue in more detail in the following section.

2.3.2: MODERN APPROACHES TO EPICENES

Having established that there is evidence to indicate that epicenes are still discussed in relatively modern grammars, based on the data from Bodine (1975), Sunderland (1986) and Zuber and Reed (1993), my goal in this section is to analyse such prescriptions in more detail. I thus review the literature on the epicene avoidance tactics and techniques found in grammars and guidelines published towards the end of the twentieth-century. For example, Jochowitz (1982) examined grammars from the 1970s and 1980s and catalogued how traditional prescriptive grammarians responded to the general rejection of generic *he*. He claims the most popular guideline in the grammars was for readers to avoid the need for epicenes all together (1982:200). This result is complemented by the work of Pauwels (1998:128), and Madson and Hessling (1999:560), which also showed that recasting and pluralisation were promoted in grammars and usage guidelines⁸⁴.

Thus there is evidence for the position that although some grammar authors acknowledge the dissatisfaction with generic *he* emphasised by feminist language reforms, their choice to promote avoidance tactics means they do not have to endorse an alternative epicene candidate. One of the most comprehensive lists of grammarians' advice on epicene pronouns comes from Treichler and Frank (1989), which I have reproduced in Figure 2.2 (overleaf)⁸⁵.

84 Avoidance tactics are not employed in all grammars and singular *they* did receive some support from bodies such as *The Chicago Manual of Style*.

85 This schema is the framework for my avoidance tactic analysis in chapter five.

GUIDELINES ON ALTERNATIVES TO GENERIC HE:

1. RECAST THE SENTENCE IN THE PLURAL.
CHANGING SINGULAR NOUN PHRASES AND PRONOUNS TO PLURALS IS OFTEN THE MOST PAINLESS WAY TO AVOID MALE "GENERIC" PRONOUNS IN EVERYDAY WRITING.
 - a. IN MOST WRITTEN CONTEXTS, AVOID *THEY* WITH A SINGULAR ANTECEDENT.
BECAUSE THE USE OF THEY WITH A SINGULAR ANTECEDENT IS WIDELY CONDEMNED WITHIN THE SCHOLARLY COMMUNITY, WE DO NOT ADVOCATE ITS USE IN WRITING.
 - b. TAKE CARE IN USING SINGULAR *THEY* COLLOQUIALLY.
 2. SHIFT THE PERSON OF THE PRONOUN TO THE FIRST PERSON OR TO THE SECOND PERSON.
 3. USE *HE OR SHE* OR *SHE OR HE*, BUT AVOID REPEATING SUCH PHRASES.
WHILST THE OCCASIONAL USE OF HE OR SHE IS GENERALLY ACCEPTED, ITS REPEATED APPEARANCE ON A PAGE IS ALMOST ALWAYS OBJECTIONABLE...THE REPEATED USE OF HE OR SHE OR SHE OR HE IN A LONGER PASSAGE IS OBVIOUSLY INTOLERABLE.
 4. ALTERNATE MASCULINE AND FEMININE PRONOUNS IN APPROPRIATE CONTEXTS.
 5. AVOID ALTERNATIVE-GENDER FORMS REQUIRING SLASHES OR PARENTHESES.
 6. USE "GENERIC" *SHE* IN SPECIAL CIRCUMSTANCES.
 7. EDIT OUT THE PERSONAL PRONOUN.
 8. USE *ONE*.
 9. AVOID REVISIONS THAT DISTORT THE ORIGINAL MEANING.
 10. PRESERVE THE FLAVOUR OF THE ORIGINAL.
 11. AVOID INTRODUCING STYLISTIC FLAWS.
 - a. TAKE CARE IN SHIFTING TO THE PASSIVE VOICE.
IN GENERAL...WE DO NOT SUGGEST USING THE PASSIVE VOICE TO ACHIEVE NONSEXIST LANGUAGE IF A MORE CONCISE AND GRACEFUL SOLUTION IS AVAILABLE.
 - b. AVOID NEEDLESS REFLEXIVES.
 12. AVOID NEEDLESS CORRECTION OF APPROPRIATELY USED SEX-SPECIFIC PRONOUNS.
 13. AVOID INCONGRUITY AND INCONSISTENT CORRECTION.
 14. AVOID AMBIGUITY; CLARIFY THE TERMS OF ANALYSIS.
REMOVING THE SURFACE "SEXIST LANGUAGE" FLOWS EASILY FROM MAKING AN EFFORT TO THINK GENUINELY ABOUT THE REFERENT.
- (ADAPTED FROM TREICHLER AND FRANK 1989:153-181)

FIGURE 2.2: GUIDELINES FOR THE AVOIDANCE OF GENERIC *HE*

An example of how some of these guidelines/prescriptions work in practice comes from Nilsen (1984) who analysed her own editing practices, comparing the unedited manuscripts of the November 1982 and April 1983 issues of the *English Journal* with their published counterparts. Based on this comparison, and arguing that pronouns were "not discussed when most of us were in high-school and college grammar and usage classes" Nilsen offered four guidelines on how to be non-sexist with pronouns (1984:152). The first guideline was that people should be consistent in their usage, whilst her second and third guidelines parallel the tactics noted in Figure 2.2. For example, Nilsen suggests that where possible writers should use words which are not marked for gender, and that dual pronouns should only be used

“when they contribute to clearer understanding” (1984:155), although what constitutes “clearer understanding” is not defined.

However, Nilsen’s final guideline is interesting, as she states that writers should “make the surface structure match the deep structure” (1984:155). Her argument is that, based on the premise that a generic referent can be of either gender, it is possible that in the deep structure, its notional plurality (section 1.3.2) is reflected. Thus, according to Nilsen, when people sense this underlying plurality, they will use a plural pronoun “without bothering to change the surface structure to plural” (1984:156). This argument is not without support, as Whitley (1978:32) claims that a pronoun can have “one set of referential possibilities in deep structure” but they may not match the eventual set of referents “after syntactic rules have stirred up the dust”. However, Nilsen’s overall advice on this matter was to avoid such practice and pluralise sentences (the first criterion in Figure 2.2).

What is clear about Nilsen’s approach is that she is reaffirming a set of prescriptive norms. Based on the guidelines in other prescriptive texts summarised by Treichler and Frank (1989), I argue that Nilsen’s paper is a clear example of the rejection of singular *they* based on rules established within traditional grammatical prescriptivism. Nilsen’s paper thus provides further evidence that, despite any rejections of generic *he*, there was not a widespread acceptance of singular *they* in the 1970s and 1980s. Indeed, a 1985 survey of editors discussed by Kingsolver and Cordry (1987) “found them overwhelmingly opposed to singular *they*” (Meyers 1990:228)⁸⁶. Based on such data it is therefore possible to argue for the position that, as editors are in a relative position of power, it is unlikely that they would publish a text on grammar which advised readers to use an epicene pronoun that they did not condone.

In a similar vein, Gastil (1990:629) notes that some writers of style guides, such as Strunk and White (1979:60), were vehemently in support of generic *he*, as like the Harvard faculty noted above, they claimed that it “is a simple, practical convention rooted in the beginnings of the English language”. Strunk and White, who are heavily criticised by Pullum (2010),

⁸⁶ Also, Kingsolver and Cordry (1987) noted that “only about half” of the editors covered by the survey “favoured *he or she*” (Meyers 1990:228). This means that for the other half, if tactics such as pluralisation etc. are not considered, generic *he* was their only option for an epicene pronoun.

went on to suggest that because of its historical status, generic *he* has “lost all suggestion of maleness... has no pejorative connotations [and] it is never incorrect” (Strunk and White 1979:60 in Gastil 1990:629). Their position agrees with McCawley (1974:103) who argued that generic *he* “carries no overtones of its primary masculine meaning” when used consistently (cited in Bodine 1975:138). However, such views deny the findings in the literature I discussed in chapter one, which lean strongly towards the consensus that generic *he* is not a true generic and has a default masculine meaning.

Nevertheless, there is evidence in the literature to suggest that even some feminists were indirectly in support of generic *he* (although not for the same reasons as Strunk and White), illustrating that feminist political allegiance does not demand subscription to feminist language reforms. As already discussed in section 1.2.1, Lakoff (1975) argued that “feminist language reformers who wished to abolish epicene *he* were linguistically naïve” (Newman 1992:450). Yet Lakoff’s position denies the socio-cultural element of the development of traditional grammatical prescriptivism discussed above, as the prescription of any form carrying semantic value (section 1.2.1) encodes a particular view of society. The acknowledgement of the impact of language-external factors on language prescriptions corresponds with the views of feminists Sniezek and Jazwinski (1986:643), that generic *he* “and similar words ‘not only reflect a history of male domination’ but also ‘actively encourage its perpetuation’” (Gastil 1990:630).

However, Green (1977:152) argues that women are “unlikely to tolerate the slippery exclusivity of *he*, which, even when officially generic, suggests that women need not apply”. Thus there is evidence that scholars have argued for an interaction between societal norms and this grammatical feature. In a similar vein, Sklar (1983:415-416 in Meyers 1993:184) argues that the prescription that number concord is senior to gender concord “has never reflected usage” despite its apparent influence on educational grammars. Cameron (1992:96) claimed that because generic *he* “was originally prescribed for sexist reasons” (an idea that may not be 100% accurate), “feminists who find it sexist are hardly projecting some novel and bizarre interpretation onto an innocent and neutral rule”. Moreover, Black and Coward (1988:105) drive home the point that traditional grammatical prescriptions “do *not* constitute evidence of the structural and systematic properties of syntactic structure”. In other words, such prescriptions are

down to the choices of those writing the grammars, as illustrated by Nilsen's (1984) guidelines.

Whilst the above discussion illustrated the findings of studies on the types of prescriptions present in educational handbooks (up until the early 1990s), what is not clear is just how much influence such grammars actually have⁸⁷. One interesting way of approaching this topic was adopted by Pauwels and Winter (2006) who investigated how, or rather if, traditional prescriptive grammar affected the epicene pronouns chosen by teachers. They argued that a language teacher had to be both a "guardian of grammar" and, in terms of non-sexist language, an "agent of change" (2006:129), and Pauwels and Winter assumed that in the case of epicene pronouns the two roles would be in conflict. To test this hypothesis they collected questionnaire responses from 182 teachers from primary, secondary, and tertiary education in Australia and supplemented this data with extended interviews of twenty participants. The object of the study was to investigate how the teachers perceived their own epicene usage and to gauge their opinions on students' choices of epicene forms.

The teachers had to indicate which pronominal form they would use in relation to a set of five different antecedents (including indefinite pronouns, and heavily gender-stereotyped forms, such as *prisoner*) both inside and outside the classroom. Pauwels and Winter found that whilst there was a strong tendency to use singular *they* outside the classroom, there was a decrease in the teachers' perceived use of it within the classroom (2006:131). Yet this did not correspond to a marked increase in generic *he* within the classroom, it instead correlated with a reported increase in combined pronouns.

Key to the debate on whether prescriptive grammars influence epicene usage, Pauwels and Winter noted that a subset of the younger teachers in their study, those below the age of thirty, "seemed to be unaware of the grammatical correctness argument" concerning singular *they* (2006:131). They claimed that two of the younger teachers "did not grasp the issue of

⁸⁷ It is important to note that the influence of prescriptive rules may not be uniform across modes. Pauwels and Winter (2006) argued that whilst traditional prescriptivists had most success making generic *he* "dominant as a generic pronoun up until the 1970s" the same prescriptive success cannot be attributed to spoken epicenes (2006:129). Similarly, Mülhäusler and Harré (1990:236) noted that "for 400 years spoken English has resisted" the prescription of generic *he*.

linguistic prescriptivism surrounding singular *they* despite clarifications from the interviewer” (2006:131). Interestingly this links to the comments of Foertsch and Gernsbacher, who suggested that it was “unclear whether many of the people who now choose to use singular *they* realize that it is ‘ungrammatical’” (1997:106). It is Pauwels and Winter’s interpretation that for these teachers “the controversy surrounding singular *they* is a non-issue” (2006:131). If this position holds then it appears that traditional prescriptive grammar has had little-to-no effect on the younger teachers in the study, or alternatively, as I have argued elsewhere, these teachers may not have had much contact with traditional prescriptive norms during their education (see Paterson 2010).

In contrast, the older teachers in Pauwels and Winter’s study, although they “expressed a greater sense of the grammatical correctness argument” justified the use of singular *they* by stating, “*I know this is supposed to be incorrect but everyone uses it*” (2006:132), highlighting again that prescription does not necessarily reflect usage. However, this was not the universal response from this group of teachers, and some still stated that they preferred *he/she* forms. One interesting finding that Pauwels and Winter (2006:133) note is that some of the female teachers made direct reference to breaking traditional grammatical prescriptive rules, stating that they used singular *they* “to ensure that children and students are aware of the issue” of gender-inclusive language. Pauwels and Winter argue that to these teachers singular *they* was “more radical” than combined pronouns as “it signalled a deliberate breach of a prescriptive norm” (2006:133). This finding suggests that some of the female teachers were not only aware of prescriptive norms surrounding epicenes, but actively chose to reject what they had been taught.

I argue that the teacher’s overt rejection of prescriptive norms supports the position that second-wave feminism and non-sexist language reforms have had a direct impact on the epicene debate. Indeed this correlates with the fact that 60% of the teachers in Pauwels and Winter’s study stated that they would correct any students use of generic *he* (2006:136). Another 14% stated that they would disapprove of the form but would fall short of correcting it (2006:136), whereas only 8% were disapproving and would correct a student’s use of singular *they* (2006:137). Thus, Pauwels and Winter conclude that despite any prescriptions stating

the opposite, “[g]ender inclusivity is the preferred reported practice for educators” and they expected this practice “to spread to successive generations” (2006:139)⁸⁸.

However, the results from just one study are not conclusive evidence that traditional prescriptive grammars go unnoticed in the classroom. In their study of American college-level grammars, Zuber and Reed (1993:518) go as far as to claim that “some English teachers appear to forget that grammars, too, are written by human beings”. However, the large number of respondents and the high level of consensus in the questionnaire data in Pauwels and Winter’s study suggests that their sample of teachers is representative⁸⁹. In any case, Zuber and Reed argue that “[t]oo often handbooks promote rules of standardization outside the students’ linguistic experience” (1993:518) and indeed the data from Pauwels and Winter (2006) seems to support this assertion.

Therefore, if the argument that traditional prescriptive rules do not affect the epicene pronouns used by teachers holds, then it is arguably unlikely that teachers will instruct their pupils to use generic *he*, or to avoid singular *they*. However, I am aware that neither Zuber and Reed (1993) nor Pauwels and Winter (2006) test this hypothesis directly. Nevertheless, the evidence presented above concerning the teacher’s use of pronouns, supports the position that “descriptivism can rule, and usage can overturn prescriptivism” (Joseph 2003:1). However, this does not mean that school grammars will suddenly endorse singular *they*, as there has been resistance to the pronoun since the eighteenth century. Therefore, any grammarian wishing to endorse singular *they* “would have to challenge another tradition of authority to sanction its use in writing” (Zuber and Reed 1993:522)⁹⁰. Unfortunately though, the data on what types of epicene prescriptions occur in modern grammars stops with Zuber and Reed’s (1993) study, and therefore, based on the current literature it is not possible to see what

88 This argument provides another link to pronoun acquisition (section 1.1.3).

89 However, the sample was comprised of Australian teachers and responses from UK teachers may differ due to differences in their education (see Paterson 2010).

90 This argument has parallels with my discussion of Stanley’s (1978) claim about female grammarians in the eighteenth century (section 2.1) in that arguably the same social norms which would have influenced traditional grammatical prescriptivism at its birth still now have some effect on epicene prescriptions.

epicene forms were being prescribed in school grammars at the time of Pauwels and Winter's more recent study.

In this final section I have addressed the literature on the state of epicene prescriptions after major influences from traditional grammatical prescriptivism and language reforms born out of second-wave feminism. My analysis of post-1970 studies on language guidelines shows that there is evidence for the position that although generic *he* may no longer be prescribed as the "correct" form for indefinite reference, due to its classification as an instance of sexist language, studies have shown that grammar authors are reluctant to endorse an alternative epicene candidate - especially singular *they*. The studies I noted in section 2.3.2 provide evidence for the position that epicene avoidance tactics, and not the endorsement of a particular epicene candidate, predominate as the current approaches to epicene usage.

CONCLUSIONS

In this chapter I have evaluated scholarship on the standardisation and codification of British English, which appears to indicate that there is support for the position that traditional epicene prescriptions have their roots in the onset of standardisation. Furthermore the promotion of generic *he* as the "correct" epicene pronoun can be linked to the differing social positions of men and women in the eighteenth century. My review of current scholarship supports the position that up until the late twentieth century, generic *he* was the prescribed English epicene pronoun, and its position was justified by traditional grammatical prescriptivism's promotion of number concord before gender concord. Previous research on grammar texts show that this was the status quo until the second-wave feminist movement bore non-sexist language reforms under which a masculine generic was an oxymoron, and thus it was not possible for a masculine pronoun to be used for both male and female referents.

In addition, there is scholarship to support the position that in reaction to the classification of generic *he* as an example of sexist language, feminist language reform guidelines were produced condemning the pronoun, leading to overt rejections of the traditionally prescribed pronoun. The studies I discussed in section 2.3.2 indicate that from circa 1970 onwards, generic *he* began to fall out of favour, and as a result traditional prescriptive grammars leant towards the prescription of avoidance tactics, thus negating the need for an epicene pronoun altogether.

Whilst this chapter has provided evidence for the position that language-external factors have impacted upon the epicene debate, the influence of traditional prescriptive grammars and feminist language reforms is hard to judge. Although Pauwels and Winter's (2006) study of Australian teachers suggests that traditional prescriptive grammar has had little influence on the epicenes used in classrooms, the adoption of *Ms* as a title for women, and Pauwels' anecdote about the job application (discussed in section 2.2.2) suggests that language prescriptions rooted in feminist language reform have had an impact on society. However, my literature review of the language-external elements of the epicene debate has also highlighted the fact that the most recent study of grammar books, undertaken by Zuber and Reed in 1993, is now almost twenty years old. Furthermore, Zuber and Reed's study focused on American college texts only, and thus the data on British English grammars used throughout this chapter is even older.

Based on my synthesis of previous research I argue that there is support in the literature for the position that the promotion of generic *he* has decreased since the beginnings of second-wave feminism, combined with an observable increase in the prescription of epicene avoidance tactics. If this trend holds then I hypothesise that more modern grammar handbooks will continue this trend in prescriptions and reject generic *he* in favour of epicene avoidance tactics. Therefore, to move the analysis of the language-external side of the epicene debate forward I propose the following sets of research questions, which I shall address alongside those already posed at the end of chapter one, beginning with a focus on addressing the lack of empirical data on current epicene prescriptions.

- Which epicene forms (if any) are being prescribed in modern grammars and handbooks aimed at native speakers of British English published in the last ten years?
- Is there any evidence that generic *he*/singular *they* are promoted in such grammar books as the ‘correct’ epicene form?
- Is there any evidence that epicene avoidance tactics predominate?
- If particular epicenes are endorsed, are they restricted in the antecedents they can take?
- Does the data on grammar prescriptions correlate with the data on epicene usage collected in response to the research questions in chapter one?

In addition, throughout this chapter my synthesis of the literature has indicated that epicene prescriptions have been consistently a point of contention in British English from the initial endorsement of generic *he* in the eighteenth century. Furthermore, the literature on the history of traditional grammatical prescriptivism, which I evaluated in section 2.1, supports the position that the prescription of this particular pronoun was not based on descriptive linguistics, but was rather intertwined with wider social norms and values. My consideration of the changes in epicene prescriptions, and the influence of feminist language reform and the classification of generic *he* as an example of sexist language in section 2.2, reinforces the argument that language-external factors have influenced the epicene debate. Thus, the chronology I have presented here suggests that there is no evidence to contradict the hypothesis that modern prescriptions will be influenced by social forces. In response to this proposition I pose the following final set of research questions designed to test whether grammatical prescriptions are explicitly linked to language-external factors.

- How might language-external factors, such as political movements and language campaigns, interact with and/or be manifest within epicene prescription and epicene usage?
- Are any of the following terms used in modern grammar books: ‘political correctness’ and/or ‘politically correct’, ‘sexist language’ and/or ‘non-sexist language’, ‘traditional’ and/or ‘out-dated’, ‘feminism’ and/or ‘feminist’?
- Are the antecedents of generic *he*/singular *they* in usage data, and in the modern grammars, examples of gender stereotypes?

These questions are designed to bring research on epicene pronouns up to date in order to establish current debates and trends in epicene

prescription. The results of these questions can then be compared with the results from the research questions posed in chapter one, which are designed to investigate epicene usage. However, before I begin my investigation I set out and justify my chosen methodology for this thesis in the following chapter, focusing on how corpus linguistics can be used to investigate both the language-external, and the language-internal, elements of the epicene debate.

CHAPTER 3: INVESTIGATING EPICENES

3.1: CORPUS LINGUISTICS: A PROBLEMATIC METHODOLOGY?	108
3.1.1: CORPUS LINGUISTICS AND GRAMMATICAL FEATURES	109
3.1.2: INTRODUCING THE CORPORA	113
3.2: WHAT THE RESEARCH QUESTIONS COVER	119
3.2.1: ANALYSING EPICENE USAGE AND PRESCRIPTIONS	120
3.2.2: INVESTIGATING LANGUAGE-EXTERNAL INFLUENCES	123
3.2.3: EPICENES AND PARADIGM ACQUISITION	125
CONCLUSIONS	127

My literature review has led me to identify four main issues in current research on epicene pronouns that cannot be fully explained by the existing scholarship. In chapter one my synthesis of the literature indicated that there was very little data on epicene usage in the twenty-first century, as the vast majority of research currently available is at least ten years old. Furthermore, I also highlighted that at present there is very little data available on the epicene pronoun use of children, and despite evidence to suggest that the pronoun paradigm is somewhat flexible whilst being acquired, and then becomes fixed, there have been no studies directly testing what epicene forms occur in the L1 input to language-acquiring children. In chapter two, my evaluation of the literature indicated that studies of the epicene prescriptions in grammar handbooks appear to stop in the early 1990s, meaning that the current available data in this area of epicene research is almost twenty years old. A further implication of this lack of data is that there is no substantial body of literature documenting the impact, if any, of language-external factors on the epicene debate in the twenty-first century.

Based on the identification of these four main points of interest in the literature I formulated the following four hypotheses:

- a) Investigation of current usage will indicate that singular *they* is the epicene of choice in speech and writing in British English today;

- b) As there is support for the position that singular *they* is the spoken epicene of choice, children who are acquiring their pronoun paradigms will receive it as L1 input;
- c) Modern grammar texts will continue the trend evident in the literature and reject generic *he* in favour of prescribing epicene avoidance tactics; and
- d) Modern grammar prescriptions will be influenced by language-external factors.

In order to address these hypotheses, I also proposed four different sets of research questions, which are designed to begin to address the areas of contention that I have identified in the literature and thus move the epicene debate forward. These specific research questions are the focus of the following three chapters on epicene usage (chapter four), epicene prescriptions (chapter five), and epicene acquisition (chapter six). I introduce the resources I used to inform my answers to these sets of questions in section 3.1 and discuss the sets of questions in more detail in section 3.2.

3.1: CORPUS LINGUISTICS: A PROBLEMATIC METHODOLOGY?

The aim of this first section is to introduce the sources I have chosen for this research. In order to investigate epicene usage and prescription in current British English, I elected to use corpus linguistics as my primary methodology. This selection was based on the position that corpora are categorised in the literature as rich sources of attested language use, which provide data spanning “the language use of many different speakers and writers” (Stubbs 2001:154). Thus, any patterns of pronoun use found in the corpora would be indicative of wider epicene use (and not individual choice). Indeed, Stubbs (2001:168) notes that multiple occurrences of a phenomenon in a corpus show “that meanings [and/or usage] are not personal and idiosyncratic” but rather, representative of a wider scope of language.

By selecting corpus-linguistics as my primary methodology, I follow a long line of research using corpora to investigate epicene usage (c.f. Graham 1973; Bodine 1975; Jochowitz 1982; Sunderland 1986; Newman 1992; Zuber

and Reed 1993; Newman 1998; Stringer and Hopper 1998; Pauwels 2001; Laitinen 2007; Baranowski 2002; Balhorn 2009). However, within the literature discussing the pros and cons of corpus linguistics there are two particular repeated issues that apply specifically to this research: the argument that corpora cannot provide qualitative data, and the claim that the occurrence of a form in a corpus is not evidence of its grammaticality. I address these issues in detail in section 3.1.1 and I provide evidence for the argument that corpora are suitable resources for collecting both quantitative and qualitative information on the use of singular *they* and generic *he*. Having established that there is evidence that the methodology I have chosen is suitable for epicene research, in section 3.1.2 I detail the three main corpora I use throughout the rest of this thesis to address the research questions I proposed at the end of the previous two chapters.

3.1.1: CORPUS LINGUISTICS AND GRAMMATICAL FEATURES

A review of the literature on corpus linguistics indicates that the methodology centres on the idea that a subset of texts from a representative sample can indicate linguistic norms for a larger whole, meaning that any claims made about the results of corpus analyses are applicable to more than just the corpus itself. The notion of representativeness, where a given sample matches the criteria of a larger set of texts, forms the backbone for the position that “statements derived from the analysis of... [a] corpus will be largely applicable to a larger sample or to the language as a whole” (Tognini-Bonelli 2001:57). For example, section A of the BE06 corpus which I use in chapter four (see below), includes texts from national UK newspapers published circa 2006, yet the texts are arguably representative of a wider range of UK newspaper texts. Therefore, as the corpus is representative of this given text type the results of analyses of the texts in BE06 are to a large extent attributable to other texts of the same form and distribution⁹¹. Thus, a sample from any

⁹¹ The composition of BE06 also facilitates diachronic analysis as it mirrors the Brown and LOB corpora and can thus be directly compared with its predecessors.

representative corpus will, in theory, indicate linguistic norms general to the text types and/or genres analysed⁹².

However, as I noted above, there are two key issues in the literature on corpus linguistics that need to be considered within the scope of this research. The first is expressed by Meurers' (2006:1621) claim that corpora do "not provide grammaticality judgements", in the sense that there is no way to tell whether a native speaker would find a construction found in a corpus acceptable, and thus there is no reason to argue that the occurrence of an utterance is "a proof of the grammaticality of that utterance". Framing this claim within the epicene debate, Meurers' argument can be recast as suggesting that the occurrence of singular *they* (or indeed generic *he*) in a corpus does not mean that it is grammatically acceptable (especially from a prescriptive perspective).

I accept that finding singular *they* in a corpus does not guarantee its syntactic acceptability let alone its social acceptability, yet Baker (2006:48) argues that if a specific grammatical construction occurs in a corpus it illuminates "something about [the] intentions" of the speaker/writer. Thus, using Baker's argument, the occurrence of singular *they* or generic *he* in a corpus of texts is not merely a matter of grammatical agreement and the selection of a pronoun to coindex with an antecedent. The language-external forces operating on epicene endorsement and choice, which I documented in my literature review in chapter two, indicate that there is evidence for the argument that there is more to epicene pronouns than syntactic coreference. Therefore, the occurrence of either epicene form in a corpus is worth consideration, irrespective of whether its use signals grammaticality (in the form of formal coreference between pronoun and antecedent, as discussed in section 1.3.1). In any case, I posit that the acceptability of a grammatical form, such as singular *they*, can be assumed if it consistently occurs in (a sample of) a corpus⁹³.

92 Biber, Conrad, and Reppen (1994:171) argue that this core belief "allows us to test assumptions about language use against patterns found in naturally occurring discourse", and note that evidence from corpus research "shows that the actual patterns of function and use in English often differ radically from prior expectations".

93 Another positive attribute of using corpus analysis is that the volume of texts in a corpus means that authorship and individual, idiosyncratic usage is not a major problem as they are statistically negligible (as noted by Sinclair 1987:81). According to Kennedy (1998:9), corpus analysis allows the extraction of "linguistic information from texts on a scale previously undreamed of".

Another argument that supports the use of corpus linguistics for epicene research comes from Teubert and Čermáková (2007:63), who note that “nearly all” of the top one hundred most frequent words in English “are function words such as pronouns and prepositions”. This suggests that there should be enough tokens of the grammatical forms *they* and *he* within a corpus to make the search for epicene usage viable (and the results in chapters four through six illustrate that this is the case). Therefore, arguments about corpus analysis not being suitable for analysing an element of grammar are not problematic for research on epicenes⁹⁴. Furthermore, corpus data is arguably indispensable when researching a grammatical phenomenon, as Teubert and Čermáková’s data supports the position that corpus analysis is a good way to ascertain large totals of the target forms⁹⁵.

The second major issue in the literature on corpus linguistics, which applies to using corpora for epicene analysis, is the argument that corpora cannot provide qualitative data. It is arguable that the statistics provided by quantitative analysis of corpora may only indicate the occurrence of a particular form, and not any nuances in its use. In a rather bold statement, Aarts (2000:7-8) argues that “in and of themselves statistics are of no interest whatsoever”, and therefore “statistical data... should be the starting point for qualitative research questions” (2000:9). Thus, there has to be qualitative data to support my numerical analyses, as merely documenting the use of epicenes is not sufficient for answering my research questions, and such practice sheds little light on how different epicene candidates are actually used.

Nevertheless, I propose that there is scholarship to support the position that corpora can also provide the qualitative data needed for research on epicenes. Kennedy (1998:9) rightly notes the focus of corpus linguistics is not just on how many times certain constructions occur, but rather it is also concerned with “how particular forms are used”. In relation to this research, my focus is not only on how often singular *they* and generic *he* are used, but

94 Perales-Escudero (2010) also uses corpus linguistics to investigate a language norm associated with traditional prescriptive grammar. He uses the Corpus of Contemporary American English to search for occurrences of the split infinitive.

95 It is worth noting also, that a descriptive trend in grammatical analysis, based upon the results of corpus analysis highlighting trends in epicene usage may (eventually) influence language prescriptions. However, my discussion in chapter two about the non-acceptance of singular *they* in prescriptive grammar books suggests that this influence may be minimal.

also on how, when, and why, certain epicene pronouns are used. For example, in the following three chapters I not only document the relative occurrences of singular *they* and generic *he* I also show what types of antecedent are most commonly used with epicene pronouns, and document the influence of gender stereotypes, and levels of definiteness, on epicene choice.

Despite criticisms, I argue that there is support in the literature for the argument that corpus analysis is suitable for research on epicene pronouns, as statistical analyses of corpora can highlight changes and indicate trends in epicene usage. According to Mair (2004:234) corpora are also suited to research on grammatical features as they can “show how innovations spread slowly and gradually, and at different rates in different text types”. Indeed Mair argues that as “almost all grammatical change will manifest itself in shifting statistical preferences in usage” corpora are a necessity (2004:234).

In addition, McEnery and Wilson (2001:114-115) argue that the identification of a set of features within “a given sample of texts” could lead to questions about “whether these features are actually tied to the specific social practices concerned or whether they arise through more general social practices” (2001:114-115); a topic pertinent to this research. Indeed, Mautner (2007:54) claims that the “awareness” of using corpora to generate qualitative data “has been increasing, over the past ten years or so” and now corpora may be used “for uncovering relationships between language and the social”. Furthermore, Holmes (2000:142) argues that corpora should be used as sources for the production of grammars, and “provide insight into the sexist and non-sexist usages currently available”. She argues that in order to avoid being prejudiced on such matters, the analysis of “current usage” is paramount, and as such “corpus analysis is indispensable” (2000:142).

This brief discussion has shown that there is evidence for the position that corpora (and the methods of analysis associated with corpus linguistics) are suitable resources for research on epicene pronouns. Having established that my chosen methodology is suitable for the present research, in the following section I provide more detail on the specific corpora used in the following three chapters of this thesis.

3.1.2: INTRODUCING THE CORPORA

There are three primary corpora used in this research, each of which is used to investigate a different aspect of epicene research, as identified by my review of scholarship in chapters one and two. The first is the BE06 corpus (of which I use two sub-corpora) which is used in chapter four to document current epicene usage in written British English. Secondly, there is the grammar corpus, which I compiled as a new resource in order to investigate modern prescriptions on epicene pronouns in chapter five. Finally, I also use the CHILDES corpus (which I sampled from the larger CHILDES database) in order to analyse the pronouns children receive as input as part of my discussion of pronominal acquisition in chapter six. As I sampled (and/or constructed) each corpus in a slightly different way, I present each one in turn below.

THE BE06 CORPUS

The BE06 corpus, which is held at Lancaster University, follows the corpus-construction pattern of the LOB and Brown family of corpora, and consists of British English texts with an average publication date of 2006. Specifically, according to Baker (2009:316, who constructed the corpus) “82% of the texts collected were published between 2005 and 2007”. This resource is currently the most recent reference corpus of its kind and by using it I provide the most recent data on epicene usage in written British English⁹⁶. The corpus itself consists of a million words, which by today’s standards may seem relatively small, especially considering the availability of online texts and the option to use the web as a corpus (see Renouf 2003; Fletcher 2004). However Baker justifies its construction not only in terms of its compatibility with previous corpora, but also argues that “a million words is probably acceptable for examining usage of high frequency words”, specifically noting the high frequency of “most grammatical words”

⁹⁶ Whilst the comparative construction and content of the LOB, Brown, and BE06 corpora facilitates diachronic analysis, this is presently beyond the scope of this research. However, it may be interesting to do such a comparative study of epicenes at a later date as initial research on epicenes in the LOB corpus (Paterson 2011) indicates a dominance of generic *he*. Indeed, Baker (2009:323) notes that in terms of frequency the pronoun *he* has dropped “four rankings down” from 12th to 16th place between LOB (and other similar corpora) and BE06.

(2009:314). Thus, the corpus is clearly suitable for this research; a statement which is supported by the fact that *he* is the 16th most frequent word in the corpus, with *his* following in 19th place, and Baker's acknowledgement that "the most frequent 20 items account for about 30% of the whole of the BE06" (2009:321).

It was not necessary to document epicene usage in the whole corpus for two reasons. Firstly, running the queries I detail in section 4.1.2 on the whole corpus gave extremely large data sets, with the total number of concordance lines for *they* and *he* reaching almost twenty-three thousand. Thus, without creating a subset of concordance lines using the sampling methods discussed below, the data would be unmanageable. Secondly, I did not wish to look at all the text types in the corpus, as my primary focus was investigating non-specialised, and "everyday" uses of epicenes, and as such I restricted my BE06 samples to the types of text that are most accessible to members of the general public. Thus, newspaper and general prose texts were arguably more suitable for this research than text types such as Academic Writing (section J of the corpus), which I chose not to consider. I also discounted the Fiction sections (K-R) to avoid any instances where pronominal choice may have been expressly considered for stylistic effect. For example, there would be no way to determine whether an author had chosen to use generic *she*, or pair all female-stereotyped roles with masculine pronouns, as a political or literary device, and thus I excluded the fiction texts to avoid skewing my data.

To narrow down the data and obtain my samples I constructed two smaller sub-corpora of BE06. The first, which I labelled *pronouns_press* (99635 words over 44 texts), consists of all the texts in section A (Press: Reportage) of BE06⁹⁷, which are newspaper articles taken from both local and national publications such as *The Mirror*, *The Scotsman*, and the *Yorkshire Evening Post*. In this sub-corpus there were 851 concordance lines including *they* and 1125 concordance lines for *he* (which are discussed in more detail in the sections 4.2.1 and 4.2.2). However, whilst O'Halloran (2007:8) notes that using newspaper "texts can create a sense of what regular readers of news texts are conventionally exposed to", Gilquinn and

⁹⁷ Sections B and C are also press texts (Editorials and Reviews respectively), but they are not included in *pronouns_press*.

Gries (2009:7) argue that “newspaper articles are a very particular register”. They highlight the fact that such texts “are created much more deliberately and consciously than many other texts”, are restricted by word length and may have been heavily edited (Gilquinn and Gries 2009:7).

Taking this point on board, and not wishing to skew my analysis to only one text type, I decided to supplement *pronouns_press* by also looking at more general texts which were not restricted to one genre. I therefore created a second sub-corpus of BE06 entitled *pronouns_general_prose*, which was larger than *pronouns_press* (463, 930 words in 206 texts) and included 3522 tokens of the lemma *they* and 3720 instances of *he*. This sub-corpus included text types ranging from auto/biographies and books about religion, to magazines and parliamentary reports, meaning that the texts in this sub-corpus were representative of wider styles of writing. As the queries I ran on both sub-corpora gave more manageable sets of data I decided not to sample any further and analysed all of the total 9218 tokens of *they* and *he* in both *pronouns_press* and *pronouns_general_prose*. The results of my analysis of the BE06 sub-corpora are documented in chapter four of this thesis.

THE GRAMMAR CORPUS

Whilst I was able to use a pre-existing corpus to tackle the research questions concerned with modern epicene usage, there was no suitable corresponding corpus of grammar textbooks, which I could use to analyse modern epicene prescriptions. Therefore, I constructed a new specialist corpus of grammars published between 2000 and 2010⁹⁸. The texts span different levels of language proficiency, insofar as they are aimed at different audiences with different levels of metalinguistic knowledge, but they are all written primarily for native speakers of English, and thus they do not focus on second language learning, or English for specific purposes (ESP).

98 The reason why I chose a ten-year period for the grammar corpus is because restricting the publication dates even further gave an extremely small number of texts which would not make analysis viable. In any case, there is no reason to assume that grammars published in 2000 would be markedly different from texts published in 2010.

In order to compile the modern grammar book corpus I began by constructing a pilot corpus⁹⁹, in order to test my method, based on a sample of the top 35 bestselling books on grammar at waterstones.com (31 of which were published post-2000). I collected a sample of these thirty-five texts based on their availability in a medium-large sized Waterstones store¹⁰⁰. Fifteen of the top thirty-five grammars were available, but three were different issues of the same book (with or without answers/CD-Roms) and so the analysis covered 13 texts. Although the sample collection was dependent entirely on the stock levels at a particular Waterstones store, this method of sampling is not problematic. The texts in the pilot study are a true representation of the grammar books available to members of the public in the store in question when the sample was collected.

Of this small sample of thirteen texts, seven books covered epicene pronouns, indicating that, although the topic was well discussed, it only appears in just over half of the grammars in the pilot study. Nevertheless, the pilot study indicated that, like previous studies such as Bodine (1975) and Zuber and Reed (1993), it is possible to find notes on epicene usage in modern grammar handbooks, and the corpus was thus expanded to include more texts in order to form a more complete picture of current epicene prescriptions. Therefore the pilot study does not stand as its own entity, as rather than compile a completely new corpus I included 12 texts from the Waterstones sample (as one was eliminated as it was published in 1990) in the larger grammar corpus, which I discuss in detail below.

To compile the final grammar corpus I used the copyright library at the University of Cambridge. Using the library's online catalogue, I searched for the terms "English grammar", and restricted the results to works published in English between 2000 and 2010. The search returned 95 results in the catalogues of University Library and its Dependent Libraries. However, although this large number of texts looks promising not all the results were applicable. I thus had to eliminate the following types of text:

- books and bilingual dictionaries aimed at English learners of other languages (*Acholi for beginners: Grammar: Acholi-English, English-Acholi*),

⁹⁹ This initial pilot corpus is used in Paterson (*forthcoming b*) to investigate whether four different traditional grammatical rules are still present in modern grammars.

¹⁰⁰ The data was collected in the Nottingham store on the 25th August 2010.

- books targeted specifically at second-language learners of English (*Grammar of Spoken English and EAP teaching*),
- translations (*Book of Intimate Grammar* [translated from Hebrew]),
- books on other languages (*Basic introduction to Biblical Hebrew*),
- linguistics textbooks, especially if they concerned syntax (*Grammar of the English Verb Phrase*),
- books on linguistic theory (*Systemic Functional Grammar of Spanish: A Contrastive Study of English*),
- books focusing on international varieties of English (*Comparative Studies in Australia and New Zealand English Grammar and Beyond*),
- dialectology texts (*Comparative Grammar of British English Dialects: Agreement, Gender, Relative Clauses*),
- historical books (*Imagining an English Reading Public, 1150-1400*),
- and CD-Roms and electronic resources (*Collins COBUILD* [electronic resource]), as I restricted my corpus to the written mode.

After narrowing down the search, I was left with a corpus of 42 texts (including the twelve from the pilot study). I manually inspected each text in turn in order to ascertain whether or not the texts covered the topic of pronouns by searching the contents pages, indexes, and glossaries for terms including ‘pronoun(s)’, ‘they’ and ‘he’. Elimination of those texts that did not directly reference pronouns led to a final sample of twenty grammar books. I found in my analysis of the 42 texts that the vast majority of those aimed at younger children did not cover pronouns in any detail. In addition, some of the texts in the original sample were teacher’s issues and only contained answers to exercises from a different volume, and therefore it was not possible to see whether the texts included tests of generic pronominal reference. These two factors primarily account for why less than 50% of the 42 texts made it into the final corpus.

The final grammar corpus, which is used in chapter five, is comprised of twenty texts and includes school course books (*Key Grammar Book 1*), general guides (*English Grammar Workbook for Dummies*), and popular reference books aimed at a wider adult audience (*My Grammar and I, or Should that be Me?*). There are also books associated with publishing houses that have a long-standing tradition of producing usage guides (e.g. *Oxford A-Z of Grammar and Punctuation*, and *Collins’ Improve Your English*). I used

the corpus to document and analyse how epicene pronouns are treated in modern texts, in order to compare prescriptions to the usage data that I collected from the BE06 sub-corpora. I also use this corpus to investigate whether the grammar authors explicitly note the impact of any language-external factors on their approaches to epicenes.

THE CHILDES CORPUS

My third and final corpus, which I use in chapter six of this thesis, relates to my finding that, in line with Bennet-Kastor's (1996) assertion, there is very little previous research on pronoun acquisition. However, the primary focus of this thesis does not lie in the field of child language acquisition, and as such, a longitudinal study on pronoun development was not within the boundaries of this research (not least for reasons including the difficulties of sourcing child participants, selecting a suitable recording method, then collecting, transcribing, and analysing the data).

However, collecting new data on children's pronoun use was not necessary. In order to obtain suitable data that would facilitate an analysis of children's pronoun production and input I used the CHILDES database (MacWhinney 2000), which is a vast online data store of transcripts of attested child language, hosted by the Psychology Department at Carnegie Mellon University. The CHILDES website allows researchers to upload transcripts of child language to a central store, which can then be accessed by others interested in their data. The CHILDES project is based on the principle of copy-left (as opposed to copyright) and one of its aims is to make large amounts of raw data on child language available for use (with permission) by others doing research on child language¹⁰¹. Therefore, in the same way that I did not have to construct a new corpus for analysing written British standard English in chapter four, as the BE06 corpus was available, the same is true for the corpus of acquisition transcripts.

I restricted my search of the CHILDES database to the British English section, in order to make the data as comparable as possible with the above

¹⁰¹ Although here I focus only on the transcripts in the database, there are also supporting audio and video files for some datasets.

corpora, and selected my final dataset based on six criteria (which are discussed in detail in section 6.1.1). The final CHILDES corpus consists of four sets of caregiver/child recordings from the Manchester Dataset (Theakston, Lieven, Pine and Roland 2001), with twelve transcripts each per pair, a total of twenty-four hours' worth of child language recordings. The recordings include a total of 1193 occurrences of *he* and 985 occurrences of *they* (as well as 4413 occurrences of *you* which I analysed in order to address the issues of parallels between the acquisition of second- and third-person pronouns).

In this section I have justified my selection of a corpus-based methodology for my analysis of epicenes in British English, showing that there is evidence in the wider literature to support the argument that corpora are suitable for analysing grammatical features. I also noted that the current research follows a long line of previous corpus-based studies on epicene pronouns. Having introduced the three corpora that form the backbone of my analysis of epicene pronouns, I now move on to address each research question in detail, noting what is covered by each part and which corpus-tools I will use to provide new data on each point of contention I noted at the end of the previous two chapters.

3.2: WHAT THE RESEARCH QUESTIONS COVER

Having introduced the resources I will use to address the areas of the epicene debate which cannot be fully explained by the current literature, in this section I repeat the research questions proposed at the end of chapters one and two in the order in which I will address them. I have numbered each question and its parts for ease of reference throughout the rest of the thesis. In section 3.2.1 I document how I will collect and analyse new data on epicene use and epicene prescriptions in twenty-first century British English, focusing primarily on the written mode. In section 3.2.2 I take a closer look at the research questions I proposed which focused on measuring whether or not language-external forces still influence epicene prescriptions (and

epicene use). In this section I also introduce the framework that I use for measuring the use of stereotyped antecedents with epicene pronouns. Finally, in section 3.2.3 I focus on the research questions designed to address epicene acquisition, focusing on L1 input of generic *he* and singular *they*. I also cover the related topic of second-person pronoun acquisition posing questions designed to analyse the use of the pronoun *you* in the CHILDES corpus.

3.2.1: ANALYSING EPICENE USAGE AND PRESCRIPTIONS

In chapter one I surveyed the literature on the language-internal elements of the epicene debate and illustrated that there was evidence for the argument that singular *they* could be found used with all types of singular antecedents. This argument stands, independent of whether or not current theories of epicene coreference can explain the use of the pronoun with formally singular antecedents. My review of previous studies on epicene comprehension also gave weight to the position that singular *they* was the most gender-neutral option available for epicene reference, and the results of the studies I evaluated in section 1.2.2 suggested that it was used unproblematically in this manner. Based on this evidence I hypothesised that in current British English singular *they* would be the epicene of choice in both speech and writing (with my focus primarily on the latter). Furthermore, my literature review of the language-external factors influencing the epicene debate reinforces this hypothesis insofar as there is evidence to support the position that the traditionally prescribed form, generic *he*, does not, and never has reflected epicene use. In order to investigate whether this first hypothesis holds I posed the following set of research questions (Figure 3.1), henceforth referred to as research question one.

1. WHAT EPICENE FORMS ARE CURRENTLY IN USE IN WRITTEN BRITISH STANDARD ENGLISH?
 - a. DO DIFFERENT EPICENE FORMS CORRELATE WITH DIFFERENT TYPES OF SINGULAR ANTECEDENT?
 - b. IS THERE AN EPICENE PRONOUN OF CHOICE?

FIGURE 3.1: RESEARCH QUESTION ONE

The purpose of this initial question was to bring research on epicene usage up to date so that my research claims are based on modern data, as my review of current scholarship indicates that the bulk of previous studies on epicene usage, such as Newman (1998) and Baranowski (2002) are now circa ten years old. I do acknowledge a limited amount of more recent research, including Balhorn (2009), who focuses on epicene usage in US newspapers, but there is not enough research on current epicenes available in the existing literature to address my first hypothesis which refers specifically to British English. Furthermore, in previous research epicene interaction with syntactic category of antecedent has gone largely unacknowledged, or is restricted to indefinite pronouns (e.g. Kolln 1986; Balhorn 2004). Therefore, by specifically investigating the relationship between pronoun choice and types of antecedent I am focusing on an untapped element of research on epicene pronoun/antecedent resolution¹⁰².

Following on from my documentation of early 21st century epicene usage in written British English, my second research question (Figure 3.2) relates to my identification of the lull in research on epicene prescriptions over recent years. Based on the chronology I presented in chapter two, where I surveyed the literature on the development of epicene prescriptions, I hypothesised that modern grammar books will continue the trend observed in the literature and reject generic *he* in favour of epicene prescriptions.

- | |
|--|
| <p>2. <u>WHICH EPICENE FORMS (IF ANY) ARE BEING PRESCRIBED IN MODERN GRAMMARS AND HANDBOOKS AIMED AT NATIVE SPEAKERS OF ENGLISH, PUBLISHED IN THE LAST TEN YEARS?</u></p> <p>a. IS THERE ANY EVIDENCE THAT GENERIC <i>HE</i>/SINGULAR <i>THEY</i> PROMOTED IN SUCH GRAMMAR BOOKS AS THE “CORRECT” EPICENE FORM?</p> <p>b. IS THERE ANY EVIDENCE THAT EPICENE AVOIDANCE TACTICS PREDOMINATE?</p> <p>c. IF PARTICULAR EPICENES ARE ENDORSED, ARE THEY RESTRICTED IN THE ANTECEDENTS THEY CAN TAKE?</p> <p>d. DOES THE DATA ON GRAMMAR PRESCRIPTIONS CORRELATE WITH THE EPICENE USAGE DATA FROM QUESTION 1?</p> |
|--|

FIGURE 3.2: RESEARCH QUESTION TWO

The research questions I posed in response to this hypothesis have two main purposes, the first of which is to provide data on newer grammar texts, thus updating epicene research on grammatical prescriptions. In addition,

¹⁰² This matter has received some limited investigation, for example, Bennet-Kastor (1996) used both definite and indefinite articles in her study, and Foertsch and Gernsbacher (1997) tested both definite and indefinite antecedents. However, this work is over ten years old.

the collection of new data will facilitate diachronic analysis of epicene guidelines between my data and previous studies. Secondly, as the data I used to address research question one is arguably representative of British English circa 2006, it is logical to do a comparative analysis of grammar textbooks, handbooks, and usage guides, published within the same time period - and this matter is explicitly noted in question 2d). The comparison of current prescriptions on epicene usage with data from question one is designed to bring into view any differences between the advice given to native speakers of British English and their actual language use. The results of my analysis of the grammar corpus are presented in chapter five of this thesis.

Using the grammar corpus detailed above in relation to question two, I obtained relative frequency figures for the texts that did and did not reference epicenes. I also noted the amount of space given to the issue of epicenes (in terms of word count), and catalogued the antecedents of the example sentences used in the grammars used (section 5.2.2). In order to be able to address question 2a) I collected frequency data for those grammars which did and/or did not endorse generic *he* and singular *they* respectively. In order to address question 2b), the results of which are presented in section 5.1.1, I catalogued any alternative prescriptions not involving generic *he* or singular *they*, based upon the general schema of avoidance tactics summarised by Treichler and Frank (1989, section 2.3.2), using key words such as *pluralisation*, *recast* and *avoid*. In addition, to address question 2c) I classified the antecedents of epicene pronouns in terms of their syntactic category (using the same method as for question one) in order to determine whether the examples given in the grammars (if any) indicated a preference for a particular antecedent class. Finally I address question 2d) in chapter five and in the general discussion in chapter seven, in light of the results presented in chapters four and five. Thus, I directly compared modern prescriptions on epicene pronouns with the language use attested in the sub-corpora of BE06.

3.2.2: INVESTIGATING LANGUAGE-EXTERNAL INFLUENCES

Having addressed current epicene prescriptions with research question two, above, my focus in research question three, is the investigation of how, or rather if, epicene prescriptions, and epicene usage, interact with language-external factors, and wider social norms (Figure 3.3). Starting from my literature review, which showed that there was little-to-no data on epicene prescriptions in the past ten-to-twenty years, I argued that there was another resulting area of interest involving the impact of language-external social forces on the epicene debate in the twenty-first century. Based on the evidence in the scholarship I surveyed in chapter two, which indicated support for the position that prescriptions are both tied to, and examples of, language-external influences on epicene pronouns, I hypothesised that, given the long history of impact of language-external forces, modern prescriptions would not be immune. To address this hypothesis I posed the following research questions:

3. HOW MIGHT LANGUAGE-EXTERNAL FACTORS, SUCH AS POLITICAL MOVEMENTS AND LANGUAGE CAMPAIGNS, INTERACT WITH AND/OR BE MANIFEST WITHIN EPICENE PRESCRIPTION AND EPICENE USAGE?
 - a. ARE ANY OF THE FOLLOWING TERMS USED IN MODERN GRAMMAR BOOKS: “POLITICAL CORRECTNESS” AND/OR “POLITICALLY CORRECT”, “SEXIST LANGUAGE” AND/OR “NON-SEXIST LANGUAGE”, “TRADITIONAL” AND/OR “OUT-DATED”, “FEMINISM” AND/OR “FEMINIST”?
 - b. ARE THE ANTECEDENTS OF GENERIC *HE*/SINGULAR *THEY* IN USAGE DATA, AND IN THE MODERN GRAMMARS, EXAMPLES OF GENDER STEREOTYPES?

FIGURE 3.3: RESEARCH QUESTION THREE

In order to address these questions, and thus test my hypothesis, I primarily focus on the grammar corpus, but compare my results with an analysis of the BE06 sub-corpora in order to compare and contrast data on prescriptions with data on epicene usage. For question three, I analysed the data in order to find out whether different epicene candidates in the grammar corpus are associated with specific political allegiances, using frequency counts, concordance analyses, and corpus comparative statistical keywords (CCSKs). Basing my selection of keywords on the reoccurrence of themes throughout my literature review, for question 3a) I catalogued any references to traditional feminism, sexist language, etc. using simple corpus queries for keywords, details of which are given in section 5.3.1.

However, there is evidence in the literature for the argument that epicene prescriptions are not the only language-external force that may

influence epicene choice. Studies such as Bennet-Kastor (1996), Foertsch and Gernsbacher (1997), and Osterhout, Bersick and McLaughlin (1997), all of which are discussed in chapter one, indicated that gender stereotypes could also impact upon the processing of different epicene candidates. In order to see if stereotypes affect the modern data from the BE06 corpus (and to a lesser extent the antecedents chosen for example sentences in the grammar corpus) I posed research question 3b). This question is designed to investigate whether the antecedents of generic *he* and singular *they* showed tendencies to be gender stereotyped as either masculine or feminine.

In order to determine whether the antecedents in my data were stereotyped, I compared the list of condensed antecedents for each pronoun (in both the BE06 and the grammar corpus data) to two sets of data already available in the literature on gender-based stereotypes of social roles. As collecting new data on the gender stereotype values of particular NPs was beyond the scope of this thesis, I found two existing studies - Kennison and Trofe (2003) and Carreiras, Garnham, Oakhill and Cruttenden (1996) - where participants were asked to rate nouns in terms of their applicability to men and women.

Although these two studies (the results of which are discussed in section 1.2.2) are not the only pieces of research testing gender stereotypes in this way, what made them different to other works, and suitable for this thesis, was that both of them included a comprehensive list of the terms tested in appendices. Kennison and Trofe explicitly state that they provided this data “as a way to facilitate future research on the use of gender stereotype information during language comprehension” (2003:365). As such, there was no need for me to conduct a similar experiment provide gender-stereotyping information for this research, as both Kennison and Trofe, and Carreiras et al. had made their own data available for comparison¹⁰³. Therefore, in order to gauge the level of stereotyping (if any) attached to the antecedents of generic *he* and singular *they* from the BE06 sub-corpora, I cross-referenced my data with the appendices provided in these studies.

I am aware that the two sets of data are not as up-to-date as would be liked, especially considering that the Carreiras et al. data is now fifteen

103 Carreiras et al. do not provide a complete list, instead listing the top-rated masculine nouns, the lowest-rated female nouns.

years old. However, for the purpose of this thesis the data is sufficient, as it is not within the remit of this research to collect new information of the stereotypical values associated with nouns. In any case, as is shown in section 4.3.3 and section 5.3.2, the Kennison and Trofe data was the most useful in this analysis, as there were more matches in their dataset with the antecedents from BE06, and thus the older data from Carreiras et al. was relegated to a supporting role. The Kennison and Trofe data is from 2003, and therefore it was collected close enough to the publication dates of the texts in BE06 to make it comparable, and thus useful for this research, allowing me to address the final part of research question three.

3.2.3: EPICENES AND PARADIGM ACQUISITION

Finally, research question four is designed to address the lack of material in the current literature on the epicene input and production of children still acquiring their pronoun paradigms. My synthesis of the literature in section 1.1.3 indicated that there is evidence for the position that the pronoun paradigm is still flexible during acquisition. If this position holds, when combined with research indicating that singular *they* is the spoken epicene of choice (c.f. Stringer and Hopper 1998), I hypothesised that children within the age range of pronoun acquisition will receive singular *they* as L1 input (and potentially be able to acquire the form as part of their pronoun paradigm).

To test this hypothesis I proposed the following set of research questions (Figure 3.4). To address research question four I used the CHILDES corpus and catalogued all occurrences of generic *he* and singular *they* in both the input given to the children by their primary caregivers, and in the children's output. The results of this analysis are presented in chapter six of this thesis. However, the research questions I set out in response to my final hypothesis also focus on another issue raised by my literature review: the significance of parallels between the development of the second- and third-person pronouns.

4. IS THERE ANY EVIDENCE THAT CHILDREN ARE EXPOSED TO EITHER SINGULAR *THEY* OR GENERIC *HE* WHEN ACQUIRING THE PERSONAL PRONOUN PARADIGM?
 - a. IN RELATION TO THE HISTORICAL PARALLELS IN THE SECOND AND THIRD-PERSON PRONOUNS, ARE THERE PARALLELS BETWEEN THE DEVELOPMENT OF *YOU* AND *THEY*?
 - b. IF SO, CAN THESE PARALLELS BE USED AS EVIDENCE FOR REGULARITY IN THE ACQUISITION OF THE ROTE-LEARNED PARADIGM?
 - c. IS THERE ANY EVIDENCE THAT ANY OBSERVED REGULARITY MAKES ACQUIRING SINGULAR *THEY* EASIER THAN ACQUIRING GENERIC *HE*?

FIGURE 3.4: RESEARCH QUESTION FOUR

As a preface to question 4a) I have already shown in section 1.1.2 that there is support in the literature for the argument that there are parallels in the historical development of *you* and *they*. For example, both forms began as plurals and but are now used with singular antecedents, and both have come up against similar arguments (c.f. the Quaker's views on singular *you* in section 1.2.2 and the traditional grammatical prescription of singular *they* in section 2.1.2). In addition, support for singular *they* has come from its analogy with the second person forms, for example Nagle, Fain, and Sanders (2000:269) note that it is "no more or less logical or illogical" than singular *you*. Similarly, Laitinen (2007:167) argues that the rejection of generic *he* in favour of singular *they* is similar to the replacement of *thou* with *you* (as discussed in section 1.1.2).

Another area in the literature where the two forms appear to be complementary, which I have not previously discussed in this thesis, concerns verbal agreement. The pronoun *they*, whether it is singular or plural, always takes a plural verb form, and there is no indication in the literature of any arguments contesting this statement. Arguably, the consistent use of *they* with a plural verb gives weight to the argument that it is fundamentally plural and its singular use is merely a deviation (see section 1.3.3). However, the second person also always takes a plural verb form, yet there is no opposition in current scholarship to its classification as a singular pronoun. Significantly for this research, Tieken-Boon van Ostade (2011) claims that one reason why the second person form does not currently co-occur with both singular and plural verb forms is due to the impact of traditional grammatical prescriptivism.

Tieken-Boon van Ostade (2011:225-226) notes that Robert Lowth explicitly proscribed the form *you was* in his grammar, although its usage had been on the rise in the 1750s. She argues that "a sharp decrease in

usage immediately after the publication of the grammar” is evidence of the impact and popularity of Lowth’s text (Tieken-Boon van Ostade 2011:226)¹⁰⁴. Whilst Tieken-Boon van Ostade does emphasise that *you was* did not just disappear, she notes that “its relegation to the non-standard registers seems the direct result of its stigmatization by Lowth in his grammar” (2011:226). Thus, I argue that this example illustrates the sizable impact that traditional grammatical prescriptivism can have on language use, as there is no reason evident in current scholarship why prescriptions should have differing levels of influence on second and third-person pronouns.

In order to expand on this previous research, which has already drawn attention to the historical parallels between *they* and *you*, I opted to analyse whether there were any further similarities between the forms by looking at their acquisition. My choice to focus on acquisition data relates to the fact that I showed in section 1.2.2 that there has been no substantial research on the epicene pronouns of children still developing their pronoun paradigms. The result of having a dual purpose for the final set of research questions means that the focus of my analysis of the CHILDES database is two-fold. Section 6.2.1 and 6.2.2 concern the parallels between *they* and *you*, in response to question 4b), whilst section 6.2.3 provides additional data on whether generic *he* is received as L1 input, facilitating a comparison of children’s epicenes (section 6.3.2), which informs my response to question 4c).

CONCLUSIONS

In this chapter I have presented the methods I use in this thesis, justifying my choice of corpus linguistics as my primary methodology on the grounds that this research follows a long line of corpus-based studies on epicenes. I have also addressed the issues in the literature on the suitability

¹⁰⁴ Wales (1996:127) even notes that singularity was once morphologically marked on the reflexive singular *yourself*, which was used in the fourteenth century, but is now non-standard. Significantly, she also states that the analogous third-person form *themselves* predates the standard *themselves* in the OED, although interestingly it was allegedly a plural form.

of using corpora to analyse grammatical phenomena and shown that the high number of pronoun tokens which can be expected in my three corpora means that I will have enough data to make justifiable conclusions. Having justified my choice of methodology I also introduced the three corpora that I analyse in detail over the next three chapters. In the final section of this chapter I considered each of my hypotheses, which I formulated based on my literature review in the previous two chapters, and the related sets of research questions, detailing what each was intended to cover.

Finally, there is one other issue that I raised in my literature review, which does not feature in my hypotheses but is still important to the current research, yet as my brief discussion below will show, directly testing Homonymy Theory is beyond the scope of this thesis. In chapter one I proposed that my adaptation of Homonymy Theory can account unproblematically for epicene coreference between singular *they* and formally singular antecedents, including definite NPs. However, the vast majority of current scholarship on pronouns and the brain does not necessarily consider their neurological representation, although there are a limited number of works on the effects of pronominal processing, such as Osterhout, Bersick and McLaughlin (1997), and Streb, Rösler and Hennighausen (1999)¹⁰⁵. Current literature is of little help in distinguishing whether it is possible that there are two forms of *they* in the mental lexicon,¹⁰⁶ and whilst further research on Homonymy Theory is needed, directly testing if the brain processes singular and plural *they* differently is beyond the scope of this thesis.

Nevertheless, I return to the issue of Homonymy Theory in the general discussion in chapter seven in order to assess the theory in relation to the results I present the following three chapters in order to see whether it can account for my findings on epicene pronouns. Therefore, to return the focus

105 Works such as Greene, McKoon and Ratcliff (1992), Kennison and Trofe (2003), and Kennison (2003), are focused on more pragmatic features, gender stereotypes, or features of discourse cohesion.

106 Although there are studies on how and where language is stored in the brain (see section 1.1.1) such research is not targeted towards specific word types, and any consideration of the differences between open- and closed-classes is minimal. In addition, mentions of open- and closed-class words in the mental lexicon literature tend to treat the two classes as binary categories, which are homogenised, yet, as I showed in section 1.2.2, the semantic value of gender marking on the third person pronouns means that they are atypical of more general closed-class paradigms. More importantly for this research, none of the brain imaging studies discussed in section 1.1.1 specifically looked at pronouns, and thus the data from the literature is extremely limited.

to my current sets of research questions, I have reiterated my hypotheses and set out the chosen method for the following three chapters of corpus analysis. I consider epicene acquisition in chapter six, and epicene prescription in chapter five, but I begin with a focus on current epicene usage in written British standard English in chapter four.

CHAPTER 4: EPICENES IN THE 21ST CENTURY

4.1: STANDARD ENGLISH CORPORA	131
4.1.1: OCCURRENCES OF SINGULAR <i>THEY</i>	131
4.1.2: OCCURRENCES OF GENERIC <i>HE</i>	137
4.2: COMPARING <i>THEY</i> WITH <i>HE</i> IN BE06	143
4.2.1: QUANTITATIVE ANALYSIS	143
4.2.2: QUALITATIVE ANTECEDENT ANALYSIS	149
4.2.3: GENDER STEREOTYPED ANTECEDENTS	156
CONCLUSIONS	164

My focus in this first corpus-based chapter is to bring data on epicene usage up to date. I use the BE06 sub-corpora introduced in the previous chapter to compare and contrast singular *they* and generic *he* in written British English, using both quantitative and qualitative methods of analysis. To this end, my research brings the literature on epicene usage up to date, as I noted in chapter one that there is currently insufficient data in the wider scholarship to make any firm conclusions about modern epicene usage. Furthermore, I am also testing the hypothesis that singular *they* will be the epicene of choice in written British English. I also consider whether the pronouns are restricted to particular antecedent types, as my synthesis of current literature indicated that this topic has received very little coverage in the previous scholarship.

In section 4.1 I present the raw quantitative data for my analyses of *they* and *he*. My results in this section show that, numerically at least, singular *they* is more popular in both sub-corpora than generic *he*, indicating that the former is the epicene of choice in written British standard English. I then compare and contrast the data for *they* and *he* in section 4.2, using both the quantitative data on the occurrence and distribution of the two pronouns, and qualitative data on the semantic values of the pronominal antecedents. In addition, in response to research question three, I perform a stereotype analysis of the antecedents, using the data from Kennison and Trofe (2003) and Carreiras et al. (1996) which I discussed in section 3.2.2,

which shows that generic *he* is marginally more likely to occur with masculine stereotypes than singular *they*.

4.1: STANDARD ENGLISH CORPORA

In order to get a clearer picture of current epicene usage, and ascertain whether generic *he* or singular *they* is favoured in written British English - thus addressing research question one - I elected to use an existing online reference corpus; the BE06 corpus based at Lancaster University¹⁰⁷, which is accessible, with permission, using the CQP Web interface (see Hardie *forthcoming*). The method I used in my analysis was to query the two sub-corpora of BE06 (see section 3.1.2) to find all tokens of *they* and *he* in `pronouns_press` and `pronouns_general_prose`. I then processed each concordance line generated by my queries by matching the pronoun to its antecedent. I present the analysis of *they* in section 4.1.1 before moving onto *he* in section 4.1.2 (comparing the results in section 4.2).

4.1.1: OCCURRENCES OF SINGULAR *THEY*

In the case of *they* I was only interested in tokens where the form was used as a singular pronoun. Therefore, I analysed the tokens of *they* selected by the queries detailed below, and codified each coindexed antecedent as either 'singular' or 'plural'. All classifications were based on syntactic data only. Thus, for a token to be classified as singular *they* its referent had to be syntactically singular in form, being further classified as one of the following: indefinite pronoun, NP with quantifier, indefinite NP, or definite NP. To avoid confusion both collective nouns and institutional uses of the pronoun (c.f. Sanford et al. 2007) were classified as plural, as these two varieties of *they* are outside the scope of the singular/plural dichotomy I focus on in this research. This classification system was unproblematic for both sub-corpora, as all tokens of *they* fit neatly into these categories.

¹⁰⁷ See Baker 2009:312-320 for a discussion of how the corpus was constructed.

Using the query [they,they*,them,their,theirs,themselves*,themselves] (which was designed to select each occurrence of the lemma *they*)¹⁰⁸ in sub-corpus pronouns_press returned a total of 858 concordance lines consisting of 395 tokens of *they*, 130 tokens of *them*, 326 tokens of *their*, one *theirs*, and six tokens of *themselves*. However, seven of these tokens were duplicates and were omitted from the rest of the study. Of the 851 remaining tokens 26 (3.06%) were instances of singular *they* which were distributed across antecedent types as shown in Table 4.1.

TABLE 4.1: DISTRIBUTION OF TOKENS OF SINGULAR *THEY* IN PRONOUNS_PRESS

	INDEFINITE PRONOUN	NP WITH QUANTIFIER	INDEFINITE NP	DEFINITE NP	TOTAL
No.	7	2	7	10	26
%	26.92	7.69	26.92	38.46	

In the corresponding sub-corpus, pronouns_general_prose, the same query gave a larger number of overall tokens but a similar percentage value for singular *they*, as well as a comparable distribution pattern across antecedent types. In this second corpus there were a total of 3557 instances of the lemma *they*, divided by case as 1466 tokens of *they*, 663 tokens of *them*, 1286 tokens of *their*, five tokens of *theirs*, 101 tokens of *themselves*, and *themselves* occurred once. There were also 35 problematic search results where the query picked up other lexical items, such as *theme(s)* and *thematic*, which were omitted from the final study, giving a total of 3522 tokens of *they*. Singular *they* occurred 154 times (4.38%), distributed across antecedent types as in Table 4.2.

TABLE 4.2: DISTRIBUTION OF TOKENS OF SINGULAR *THEY* IN PRONOUNS_GENERAL_PROSE

	INDEFINITE PRONOUN	NP WITH QUANTIFIER	INDEFINITE NP	DEFINITE NP	TOTAL
No.	38	14	39	63	154
%	24.68	9.09	25.32	40.91	

The similarities between the antecedents of singular *they* across both sub-corpora is better represented visually. The graphs in Figure 4.1 illustrate that the antecedents for *they* for both pronouns_press and pronouns_general_prose are comparable in their syntactic distribution.

¹⁰⁸ Including all case forms and instances where *they* was part of a longer contraction, such as *they're*.

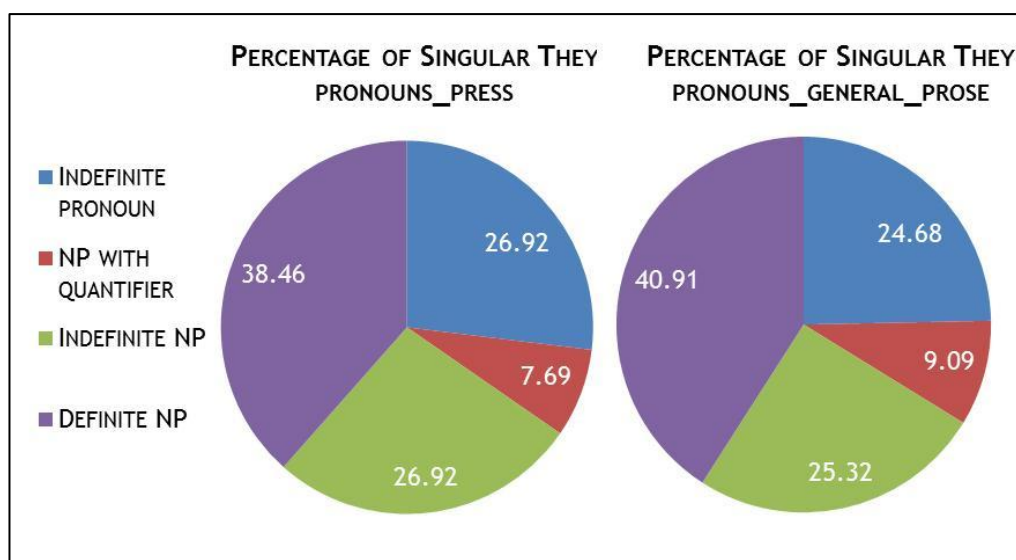


FIGURE 4.1: PERCENTAGE COMPARISON OF DISTRIBUTION OF ANTECEDENTS OF SINGULAR *they*

These results indicate general trends in the use of singular *they* and show that, although across both sub-corpora an average of 25% of antecedents of singular *they* are indefinite pronouns (see Table 4.3), the largest majority of antecedents are definite NPs, which account for 40.56% of tokens in the combined sub-corpora. The results also indicate that use of singular *they* with NPs with quantifiers is rare, accounting for only circa 8.89% of tokens in both sub-corpora. However, the results do show that singular *they* is used with all the different syntactic classifications of antecedents, much in line with the discussion of singular reference in section 1.3.1.

TABLE 4.3: DISTRIBUTION OF TOKENS OF SINGULAR *they* IN BOTH SUB-CORPORA

	INDEFINITE PRONOUN	NP WITH QUANTIFIER	INDEFINITE NP	DEFINITE NP	TOTAL
NO.	45	16	46	73	180
%	25.00	8.89	25.56	40.56	

As well as having similar distribution in types of antecedent, the distribution of case forms across the tokens of singular *they* were also comparable. In both *pronouns_press* and *pronouns_general_prose* the nominative and possessive forms of *they* predominate across all antecedent types, with only a minority occurrence of accusatives (and only two reflexives in the whole data set)¹⁰⁹.

¹⁰⁹ I compared the distribution of case forms for singular *they* with the distribution of all 4370 tokens of *they* across both sub-corpora (851 from *pronouns_press* and 3519 from *pronouns_general_prose*), but the results were not drastically different. Therefore, I condensed the results for all cases.

On inspection of the complete list of antecedents of singular *they* it became apparent that certain texts within the sub-corpora used singular *they* a large number of times to refer to the same antecedent, which could skew the results. For example the antecedent *your child* occurred six times in a single text in *pronouns_press* (Text A24), accounting for 23.08% of the total 26 tokens of singular *they* in this sub-corpus, but the form does not occur in any other text in *pronouns_press*. Thus it would be misleading to claim that singular *they* has a high collocation rate with *your child* as this is clearly not the case.

To eliminate this problem, I followed the methodology used by Newman (1992:456) and condensed multiple references to the same antecedent into one token of singular *they*. If the same antecedent occurred in different texts I counted these as different instances of singular *they*. For example, the antecedent *the person* occurred in both Text A15 and A28 in *pronouns_press* therefore *the person* appears twice in the results as it represents two different, independent uses of singular *they*. This procedure meant that the 26 tokens of singular *they* in *pronouns_press* were condensed to 18 unique antecedents and the results from *pronouns_general_prose* were condensed from 154 tokens to 90, giving a total of 108 tokens of singular *they* across the sub-corpora (a drop of 40%), as summarised in Table 4.4.

TABLE 4.4: DISTRIBUTION OF ANTECEDENTS OF SINGULAR *THEY* AFTER THE CONDENSATION OF DUPLICATES

	INDEFINITE PRONOUN	NP WITH QUANTIFIER	INDEFINITE NP	DEFINITE NP	TOTALS
PRONOUNS_PRESS	6	2	5	5	18
PRONOUNS_GENERAL_PROSE	30	10	22	28	90
BOTH SUB-CORPORA	36	12	27	33	108

In order to compare the distribution of the complete list of antecedents from the raw corpus data, with the modified results where duplicate references to the same antecedent were condensed, Tables 4.5 and 4.6 show the relative percentages for each different syntactic classification of antecedent before and after the condensation process.

TABLE 4.5: PERCENTAGE DISTRIBUTION OF ANTECEDENTS OF SINGULAR *THEY* **BEFORE** THE CONDENSATION OF DUPLICATES

	INDEFINITE PRONOUN	NP WITH QUANTIFIER	INDEFINITE NP	DEFINITE NP
PRONOUNS_PRESS	26.92	7.69	26.92	38.46
PRONOUNS_GENERAL_PROSE	24.68	9.09	25.32	40.91
BOTH SUB-CORPORA	25.00	8.89	25.56	40.56

TABLE 4.6: PERCENTAGE DISTRIBUTION OF ANTECEDENTS OF SINGULAR *THEY* **AFTER** THE CONDENSATION OF DUPLICATES

	INDEFINITE PRONOUN	NP WITH QUANTIFIER	INDEFINITE NP	DEFINITE NP
PRONOUNS_PRESS	33.33	11.11	27.78	27.78
PRONOUNS_GENERAL_PROSE	33.33	11.11	24.44	31.11
BOTH SUB-CORPORA	33.33	11.11	25.00	30.56

The tables show that, when duplicate references to a single antecedent are condensed to just one token, the syntactic distribution of antecedents is different. For example, there is an increase of seven percentage-points in the relative occurrence of indefinite pronouns, with a corresponding two percentage-point increase in NPs with quantifiers. Conversely, the definite NPs dropped by 10 percentage-points. The differences between the two sets of data are best illustrated visually and Figure 4.2 shows the overall distribution of tokens in both sub-corpora before and after the duplicate antecedents were merged.

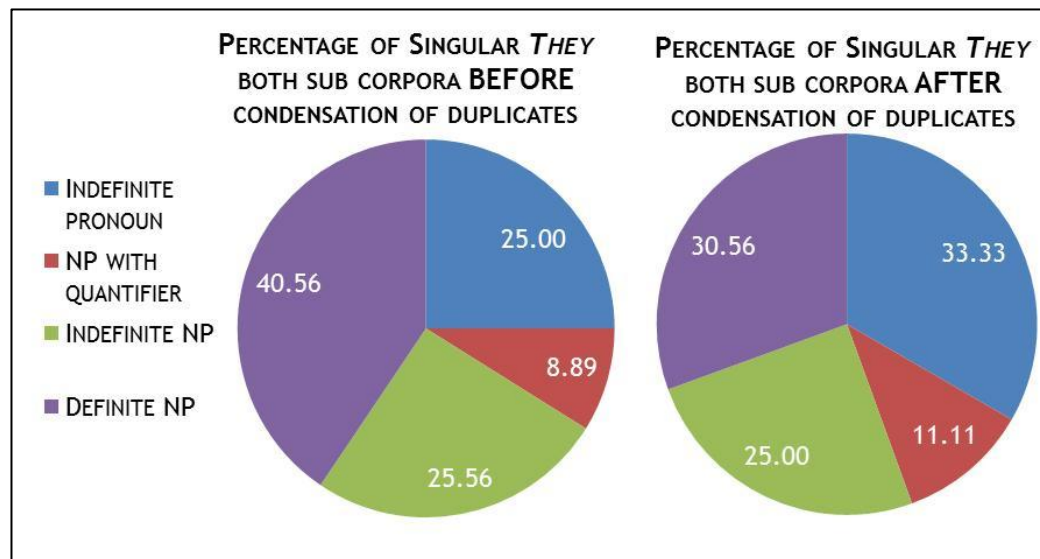


FIGURE 4.2: PERCENTAGE COMPARISON OF ANTECEDENT DISTRIBUTION **BEFORE** AND **AFTER** CONDENSATION

The change in distribution evident in Figure 4.2 confirms my intuitions, noted above, that analysing only the raw data for singular *they* could lead to over representation of certain antecedents, and thus of syntactic categories. The distribution of syntactic forms is visually different when duplicate references to one antecedent are condensed to just one token of singular *they*. Therefore, having analysed the raw number of occurrences of singular *they* I proceeded using the condensed set of antecedents only.

In this section I have documented the clear parallels between both sub-corpora for the occurrence of singular *they* in both overall percentage value (which is an average of 4.12% before condensation) and distribution of antecedent forms (as illustrated by Figure 4.1). I acknowledged that some of the antecedents of singular *they* were duplicate references within the same text, and so, in order to avoid skewing the results and over-representing my findings I condensed such occurrences to just one antecedent token. This changed the antecedent distribution as the percentage of definite NPs coindexing with singular *they* dropped by ten percentage-points whilst the number of indefinite pronouns rose eight percentage-points overall, (c.f. Table 4.5/4.6; Figure 4.2) confirming that I was right to assume that working from the raw data could over-represent certain syntactic categories of antecedents.

Although my investigation is still incomplete, this concludes my initial analysis of singular *they*. I have collected and analysed data on singular *they*, which I will use to inform my responses to research question one, and which can be directly compared with similar data for *he* below. But also, in a more general sense, I have begun to illustrate that corpus linguistics is a suitable methodology for research on epicene pronouns, providing quantitative data on the use, and coreference patterns of such pronouns, thus justifying my methodological choices for this thesis. Now however, I turn my attention to the corresponding analysis of *he* in both sub-corpora in order to facilitate a comparative analysis in section 4.2.

4.1.2: OCCURRENCES OF GENERIC HE

In order to analyse the tokens of *he* in the two sub-corpora I classified the antecedents of the pronoun as either ‘masculine’ or ‘generic’, with the generic tokens being divided by antecedent form in the same way as the tokens of singular *they*. Obviously, syntactic classification by number would not have aided the labelling of tokens of *he* as either ‘masculine’ or ‘generic’, so in this case I was careful to check whether or not each token corresponded to an individual real-world entity. Antecedents that were classified as ‘masculine’ included (in the main) proper nouns, NPs which referred back to previously specified male/masculine referents, and indefinite antecedents where gender was coded in the lexical category, insofar as the term could only be used as masculine (e.g. *father*). For an antecedent to be part of the ‘generic’ class there had to be no discernable specific male referent in the surrounding context of the token. Although again institutional uses of the pronoun were classified as ‘masculine’, more conventional generic uses of *man* and *the human being* were classified ‘generic’ *he* (although *a man* was classified as ‘masculine’ as a generic reading was difficult).

Using the query [he,he*,him,him*,his,himself,hissself] to select all case forms of *he* (and any relevant contractions) in sub-corpus pronouns_press returned a total of 1125 concordance lines (occurring over 43 of the 44 texts) consisting of 633 tokens of *he*, 116 tokens of *him*, 361 tokens of *his*, and 15 tokens of *himself*. Of the 1125 tokens of *he* in all case forms only nine (0.8%) were instances of generic *he* and they were distributed across antecedent types as shown in Table 4.7. These figures are much lower than their corresponding values for singular *they* in Table 4.1.

TABLE 4.7: DISTRIBUTION OF TOKENS OF GENERIC HE IN PRONOUNS_PRESS

	INDEFINITE PRONOUN	NP WITH QUANTIFIER	INDEFINITE NP	DEFINITE NP	TOTAL
No.	0	0	4	5	9
%	0.00	0.00	44.44	55.56	

As with singular *they*, the same query for *he* was run in the corresponding sub-corpus pronouns_general_prose¹¹⁰, again resulting in a

¹¹⁰ In this case, 39 texts, out of the total 206, contained no tokens of any of the search terms.

larger number of tokens than in *pronouns_press*, with a total of 3724 concordance lines. Of the tokens, four were erroneous results for the acronym HE (Higher Education) which were eliminated, leaving a total of 3720 tokens of pronominal *he* (in all case forms). The tokens consisted of 1661 occurrences of *he*, 427 concordance lines including *him*, a rather large 1522 occurrences of *his*, and a much lower figure of 110 tokens of *himself*. Importantly, when *he* and *she* occurred simultaneously, as in *his or her* (which occurred nine times in *pronouns_general_prose*, alongside two occurrences of *he or she* and one *himself or herself*), the pronoun was considered masculine, even though the overall effect may have been to produce a combined epicene pronoun¹¹¹. Also of note were two institutional tokens of *he* where the referent was external and no direct antecedent could be found - in order to avoid skewing the data these tokens were classed as masculine as their generic nature could not be verified.

Again, as with singular *they*, there were more overall tokens in the second sub-corpus, but unlike the above analysis, the percentage value and distribution patterns for the different types of antecedent differed for generic *he* between the sub-corpora. Out of the 3720 tokens of *he* in *pronouns_general_prose* 147 (3.95%) were generic, distributed by antecedent type as in Table 4.8. This figure is markedly different from the corresponding value of 0.8% for *pronouns_press*.

TABLE 4.8: DISTRIBUTION OF TOKENS OF GENERIC *HE* IN *PRONOUNS_GENERAL_PROSE*

	INDEFINITE PRONOUN	NP WITH QUANTIFIER	INDEFINITE NP	DEFINITE NP	TOTAL
No.	2	8	54	83	147
%	1.36	5.44	36.73	56.46	

The differences between the distribution of generic tokens of *he* across both sub-corpora are represented graphically in Figure 4.3 which shows that the two sets of tokens are somewhat comparable in their distribution between antecedents, although not to the same extent as the results for singular *they* (c.f. Figure 4.1). Interestingly, neither indefinite pronouns nor NPs with quantifiers occurred with generic *he* in *pronouns_press*, and the forms were not well represented either in *pronouns_general_prose*. This is a

¹¹¹ In *pronouns_press* this phenomenon only occurred once, there were also two occurrences of *he/she* in *pronouns_general_prose* (which were not picked up by the original query).

marked contrast to Figure 4.1, which showed a relatively high level of both of these types of antecedent for singular *they*. Such differences are expanded upon in section 4.2 where I compare and contrast the results for *he* and *they* in more detail.

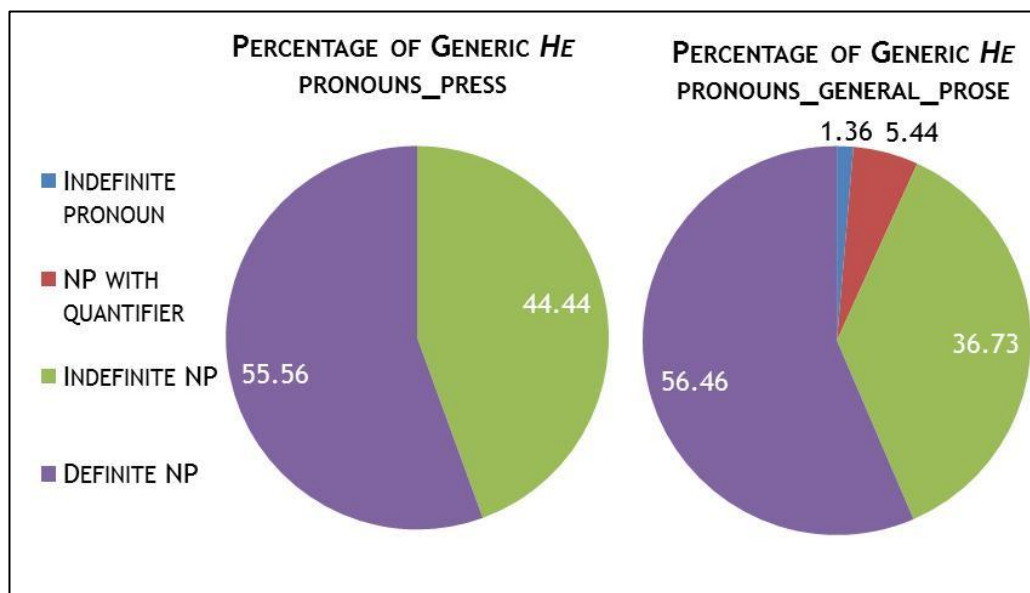


FIGURE 4.3: PERCENTAGE COMPARISON OF DISTRIBUTION OF ANTECEDENTS OF GENERIC *HE*

However, the two sub-corpora are generally comparable in their distribution of occurrences of generic *he*, and the lack of indefinite pronouns and NPs with quantifiers coindexed with generic *he* in *pronouns_press* may simply reflect the fact that there were less tokens of generic *he* in this first sub-corpus than in *pronouns_general_prose*, which is bigger. However, even in *pronouns_general_prose*, indefinite pronouns and NPs with quantifiers combined only occurred with generic *he* only 6.8% of the time, suggesting that their occurrence is rare overall.

When the total number of 156 tokens of generic *he* are combined for both sub-corpora, as in Table 4.9, it is clear that neither distribution of antecedents from the individual sub-corpora is anomalous to the group result. That is, the distribution of antecedents illustrated in Figure 4.3 is not too far removed from the pooled data in Table 4.9. The predominant syntactic category is the definite NPs (a similarity with the data for singular *they* in Figure 4.1), which account for almost 60% of antecedents of generic *he*,

followed by indefinite NPs, which represent well over a third of the antecedents across both sub-corpora¹¹².

TABLE 4.9: DISTRIBUTION OF TOKENS OF GENERIC *HE* IN BOTH SUB-CORPORA

	INDEFINITE PRONOUN	NP WITH QUANTIFIER	INDEFINITE NP	DEFINITE NP	TOTAL
NO.	2	8	58	88	156
%	1.28	5.13	37.18	56.41	

In order to make my analyses of generic *he* and singular *they* parallel I used the same method as described in section 4.2.1 to condense all the duplicate references of generic *he* to a single antecedent into one token. The results are shown in Table 4.10. The condensing process decreased the total number of generic tokens dramatically from 156 to 47, with the occurrences in *pronouns_press* declining from nine to four (a 55.56% reduction) and those in *pronouns_general_prose* dropping from 147 to just 43 (a 70.75% decrease).

TABLE 4.10: DISTRIBUTION OF ANTECEDENTS OF GENERIC *HE* AFTER THE CONDENSATION OF DUPLICATES

	INDEFINITE PRONOUN	NP WITH QUANTIFIER	INDEFINITE NP	DEFINITE NP	TOTALS
PRONOUNS_PRESS	0	0	2	2	4
PRONOUNS_GENERAL_PROSE	2	4	9	28	43
BOTH SUB-CORPORA	2	4	11	30	47

The overall decrease from duplicate tokens to single references was 69.87%, meaning that each antecedent was referred to by 3.32 pronouns on average, compared with a figure of 1.67 for singular *they*. Although there are almost two extra pronominal references to each antecedent of generic *he* than singular *they*, this could be a feature of pronouns *per se*, as opposed to a feature of epicene reference. Indeed, Baker (2009) notes that across the whole of BE06 masculine personal pronouns are more frequent than forms of *she* or *they* (although he does not look at how many antecedents the pronouns refer to as this would be infeasible without manual analysis, due to the current limits of corpus tools and tagging systems).

112 Although the syntactic distributions of antecedents across both sub-corpora are comparable, the distribution of case forms for generic *he* is not so uniform. Yet, although the sub-corpora do not appear comparable, when they are combined and compared with all the tokens of *he* in the sub-corpora, the distributions are not too dissimilar. Therefore, as with singular *they*, I do not differentiate my results by case.

Alternatively, the higher ratio of pronouns to antecedents for generic *he* could be a feature of this particular corpus, or more specifically, of the text types within it. For example, the texts in section H of pronouns_general_prose are legal texts (which are discussed in more detail in section 4.3) and involve complex descriptions of the roles of particular parties in an agreement, as in Figure 4.4. The example shows that pronouns are used four times for one repeated antecedent.

PART 1 PERSONS WHO LACK CAPACITY THE PRINCIPLES 1 THE PRINCIPLES (1) THE FOLLOWING PRINCIPLES APPLY FOR THE PURPOSES OF THIS ACT.

(2) [A PERSON] MUST BE ASSUMED TO HAVE CAPACITY UNLESS IT IS ESTABLISHED THAT [HE] LACKS CAPACITY.

(3) A PERSON IS NOT TO BE TREATED AS UNABLE TO MAKE A DECISION UNLESS ALL PRACTICABLE STEPS TO HELP [HIM] TO DO SO HAVE BEEN TAKEN WITHOUT SUCCESS.

(4) A PERSON IS NOT TO BE TREATED AS UNABLE TO MAKE A DECISION MERELY BECAUSE [HE] MAKES AN UNWISE DECISION.

(5) AN ACT DONE, OR DECISION MADE, UNDER THIS ACT FOR OR ON BEHALF OF A PERSON WHO LACKS CAPACITY MUST BE DONE, OR MADE, IN [HIS] BEST INTERESTS.

TAKEN FROM TEXT H11 - MENTAL CAPACITY ACT PART 1.

FIGURE 4.4: EXAMPLE OF LEGAL TEXT

Whatever the reasons for the difference between pronoun-to-antecedent ratio, the 47 different antecedents of generic *he* represent a fundamentally smaller number than the 108 antecedents of singular *they*. This result occurs despite the fact that the number of concordance lines for *he* was higher in both sub-corpora than for *they*, even though the total numbers of uncondensed tokens were somewhat comparable for the two forms (156 for *he* and 180 for *they*). The differences between the two tested pronouns are considered in more detail in section 4.3 (c.f. Figure 4.9). In terms of generic *he* however, Tables 4.11 and 4.12 illustrate the different percentage values of antecedent distribution before and after the condensation of antecedents.

TABLE 4.11: PERCENTAGE DISTRIBUTION OF ANTECEDENTS OF GENERIC *HE* BEFORE THE CONDENSATION OF DUPLICATES

	INDEFINITE PRONOUN	NP WITH QUANTIFIER	INDEFINITE NP	DEFINITE NP
PRONOUNS_PRESS	0.00	0.00	44.44	55.56
PRONOUNS_GENERAL_PROSE	1.36	5.44	36.73	56.46
BOTH SUB-CORPORA	1.28	5.13	37.18	56.41

TABLE 4.12: PERCENTAGE DISTRIBUTION OF ANTECEDENTS OF GENERIC *HE* **AFTER** THE CONDENSATION OF DUPLICATES

	INDEFINITE PRONOUN	NP WITH QUANTIFIER	INDEFINITE NP	DEFINITE NP
PRONOUNS_PRESS	0.00	0.00	50.00	50.00
PRONOUNS_GENERAL_PROSE	4.65	9.30	20.93	65.12
BOTH SUB-CORPORA	4.26	8.51	23.40	63.83

Condensing the tokens changed the distribution of antecedent types slightly for generic *he*. In the condensed set of antecedents indefinite pronouns are much better represented, increasing from 1.28% to 4.26%, with a similar increase for NPs with quantifiers. The largest difference, which is better shown graphically as in Figure 4.5, is that indefinite NPs drop by almost fourteen percentage-points, whilst definite NPs account for over 60% of pronominal referents.

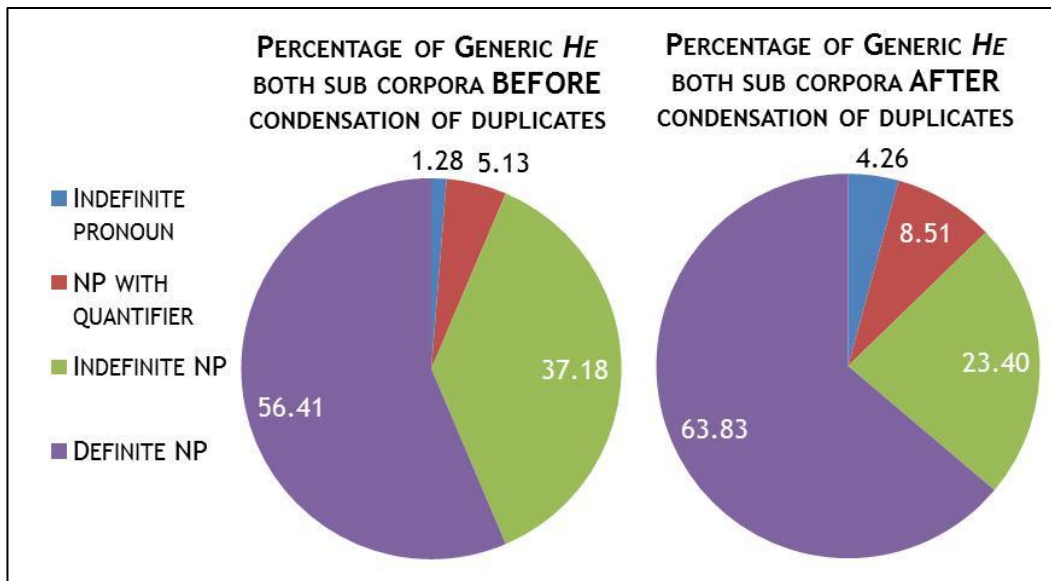


FIGURE 4.5: PERCENTAGE COMPARISON OF ANTECEDENT DISTRIBUTION **BEFORE** AND **AFTER** CONDENSATION

Compared with Figure 4.2, Figure 4.5 shows that the distribution of the type of antecedents used with generic *he* is completely different from the distribution for singular *they*, and I explore possible reasons for this in section 4.2.1.

In this section, I have presented the results for my initial analysis of *he* in both sub-corpora. I have shown that, conversely to singular *they*, which I analysed in the previous section, the results for *he* across the sub-corpora

are different on many levels: from the initial percentage value of generic tokens (0.8% and 3.95% respectively), to the distribution of antecedents by syntactic classification (Figure 4.3). The average percentage value of generic *he* for the sub-corpora stands at 3.22% (156 out of 4845 concordance lines) before condensation of antecedents. In addition, generic *he* is used more with indefinite and definite NPs than with NPs with quantifiers or indefinite pronouns (Figure 4.3).

Again, as with singular *they*, the condensation of antecedents, where multiple references to an antecedent were classed as one token of generic *he*, changed the syntactic distribution of antecedents, as shown in Figure 4.5, but the dominance of indefinite and definite NPs in the dataset for generic *he* is still evident. Using this data as a starting point for analysis of how generic *he* is used within the sub-corpora of BE06, in the next section I compare it with the data I presented in section 4.2.1 for singular *they*.

4.2: COMPARING THEY WITH HE IN BE06

In order to compare the results for singular *they* with the results for generic *he* in both *pronouns_press* and *pronouns_general_prose* I begin by quantitatively comparing the figures presented in the previous two sections in section 4.2.1, before moving on to a more qualitative analysis in section 4.2.2. I compare and contrast the antecedent types for each pronoun, drawing on specific examples and highlighting the differences and similarities between the entities referred to by *they* and *he*, before performing stereotype analysis on the antecedents in section 4.2.3 in order to address research question three.

4.2.1: QUANTITATIVE ANALYSIS

Running the queries detailed in the sections above returned 851 concordance lines for *they* in *pronouns_press* and 3552 lines in *pronouns_general_prose* of which 26 (3.06%) and 154 (4.38%) were instances

of singular *they*. This means that across the sub-corpora there were 180 occurrences of singular *they* in 4370 concordance lines; a percentage value of 4.12%. Correspondingly the query for *he* resulted in 1125 concordance lines in pronouns_press and 3720 concordances in pronouns_general_prose, nine (0.8%) and 147 (3.95%) of which were instances of generic *he*. Thus there was a total of 156 tokens of generic *he* in 4845 concordance lines, which is 3.22%, making the occurrence of generic *he* in these sub-corpora 0.9 percentage-points lower than the occurrence of singular *they*.

TABLE 4.13: DISTRIBUTION OF TOKENS OF SINGULAR *THEY* AND GENERIC *HE* IN PRONOUNS_PRESS AND PRONOUNS_GENERAL_PROSE

		INDEFINITE PRONOUN	NP WITH QUANTIFIER	INDEFINITE NP	DEFINITE NP	TOTAL
PRESS	SINGULAR <i>THEY</i>	7	2	7	10	26
	GENERIC <i>HE</i>	0	0	4	5	9
GENERAL _ PROSE	SINGULAR <i>THEY</i>	38	14	39	63	154
	GENERIC <i>HE</i>	2	8	54	83	147
BOTH SUB-CORPORA	SINGULAR <i>THEY</i>	45 (25.00 %)	16 (8.89%)	46 (25.56%)	73 (40.56%)	180
	GENERIC <i>HE</i>	2 (1.28%)	8 (5.13%)	58 (37.12%)	88 (56.41%)	156

Based on Table 4.13, which shows the numerical values for singular and generic uses of *they* and *he*, the total number of tested epicene forms occurring in the sub-corpora is 336, of which 53.57% are singular *they* and 46.43% are generic *he*. This initially appears to be a relatively even split, yet (as shown in section 4.2.2) the queries for *he* returned 475 concordance lines more than the query for *they*. Therefore, I hypothesise that if the sample sizes had equal numbers of concordance lines the normalised ratio of epicene usage would be 200:156 for *they* and *he* respectively¹¹³. Simplified to its smallest form this ratio is 50:39 in favour of singular *they*, which in percentage terms is 56.18% to 43.82%. This is a 12.36 percentage-points difference, indicating that in raw figures singular *they* is favoured as an epicene form over generic *he*.

¹¹³ This figure was calculated using the relative percentages of singular *they* (4.12%) and generic *he* (3.22%) on the largest total of concordances (4845).

Moving onto the syntactic distribution of antecedents, as illustrated by Figure 4.6, it is clear to see that the two epicene forms co-occur with antecedent types differently. The most popular antecedents are definite NPs followed by indefinite NPs, with the combined totals for the two forms being 66.12% for singular *they* and 93.19% for generic *he*.

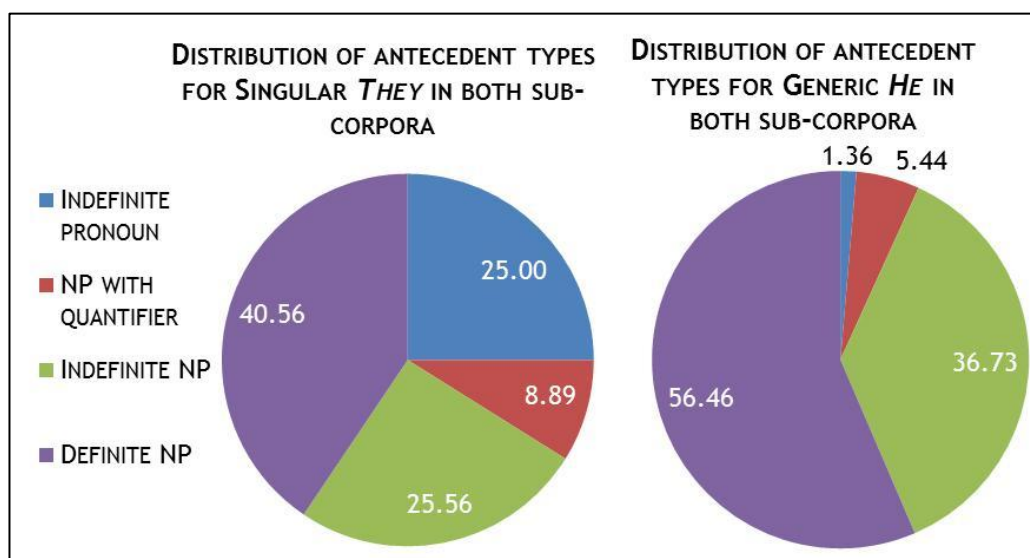


FIGURE 4.6: PERCENTAGE COMPARISON OF ANTECEDENT DISTRIBUTION FOR SINGULAR *THEY* AND GENERIC *HE* ACROSS BOTH SUB-CORPORA

The main difference between the two distributions is the relatively high use of indefinite pronouns with singular *they*, which contrasts with its low occurrence in the antecedents of generic *he*. This findings links to Wales' (1996) argument, noted in section 1.3.1, that the use of generic *he* is rare with indefinite pronouns. A similar pattern of distribution also occurs for NPs with quantifiers, with respective values of 8.89% and 5.44%, but this is not as marked as the 25.00% to 1.36% for indefinite pronouns. The differences between how each syntactic form correlates with each pronoun are better illustrated using the raw figures, as in Figure 4.7 (overleaf), which shows that different antecedents clearly favour one epicene form or the other (the condensed antecedents are dealt with in Figure 4.9).

Both indefinite pronouns and NPs with quantifiers are more likely to occur with singular *they*, with values of 95.74% and 66.67% respectively, whilst indefinite NPs and definite NPs co-occur more frequently with generic *he*, although these results are only marginal, with values of 55.78% and 54.66%. This result may be because of issues associated with notional number (section 1.3.2) or it could be a feature that interacts with the

stereotypical value and imaginability of specific antecedents - a concept I explore in chapter seven.

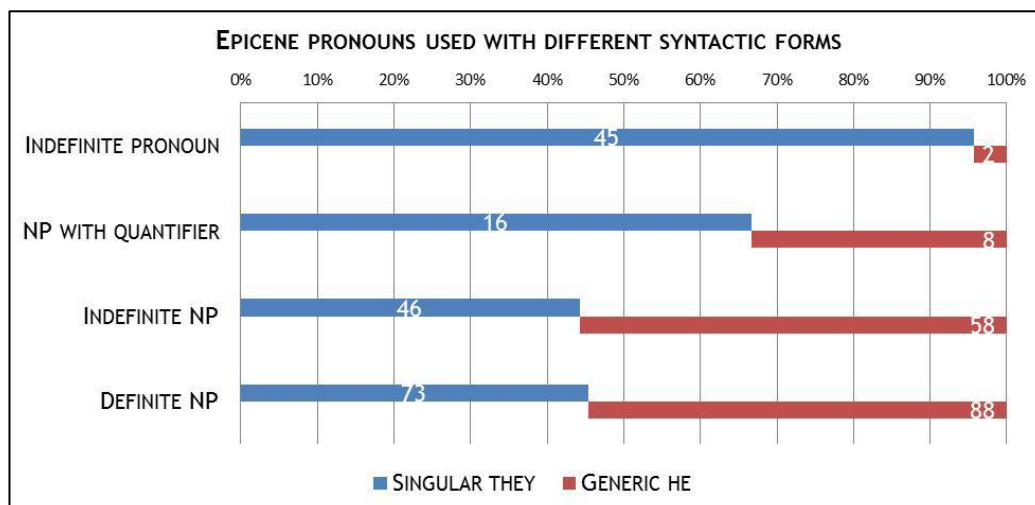


FIGURE 4.7: EPICENE PRONOUNS USED WITH DIFFERENT SYNTACTIC FORMS

The contrast between the data for singular *they* and generic *he* is even more marked after the duplicate references were condensed to single tokens. Table 4.14 includes a summary of the distribution of antecedents for both tested epicenes after the condensation process; compared with Table 4.13 the decrease in the number of individual antecedents is rather large. The total for singular *they* drops from 180 to 108 (a 40% decrease) and the figure for generic *he* changes from 156 to only 51 (a much higher 71% drop).

TABLE 4.14: DISTRIBUTION OF TOKENS OF SINGULAR *THEY* AND GENERIC *HE* IN PRONOUNS_PRESS AND PRONOUNS_GENERAL_PROSE AFTER CONDENSATION OF ANTECEDENTS

		INDEFINITE PRONOUN	NP WITH QUANTIFIER	INDEFINITE NP	DEFINITE NP	TOTAL
PRESS	SINGULAR <i>THEY</i>	6	2	5	5	18
	GENERIC <i>HE</i>	0	0	2	2	4
GENERAL PROSE	SINGULAR <i>THEY</i>	30	10	22	28	90
	GENERIC <i>HE</i>	2	4	9	28	43
BOTH SUB-CORPORA	SINGULAR <i>THEY</i>	36 (33.33%)	12 (1.11%)	27 (25.00%)	33 (30.56%)	108
	GENERIC <i>HE</i>	2 (4.26%)	4 (8.51%)	11 (23.40%)	30 (63.83%)	47

Thus, the ratio between unique uses of singular *they* and generic *he* is 108:47 in favour of singular *they*, with percentage values of 69.68% to

30.32%. Compared with the analysis of Table 4.13 the use of singular *they* and generic *he* is much less evenly split after multiple references are condensed, again indicating that singular *they* is favoured over generic *he*. However, the apparent preference for singular *they* is not uniform across all antecedent types. As discussed above with Figure 4.6, Figure 4.8 shows that the distribution of forms used with each pronoun are different.

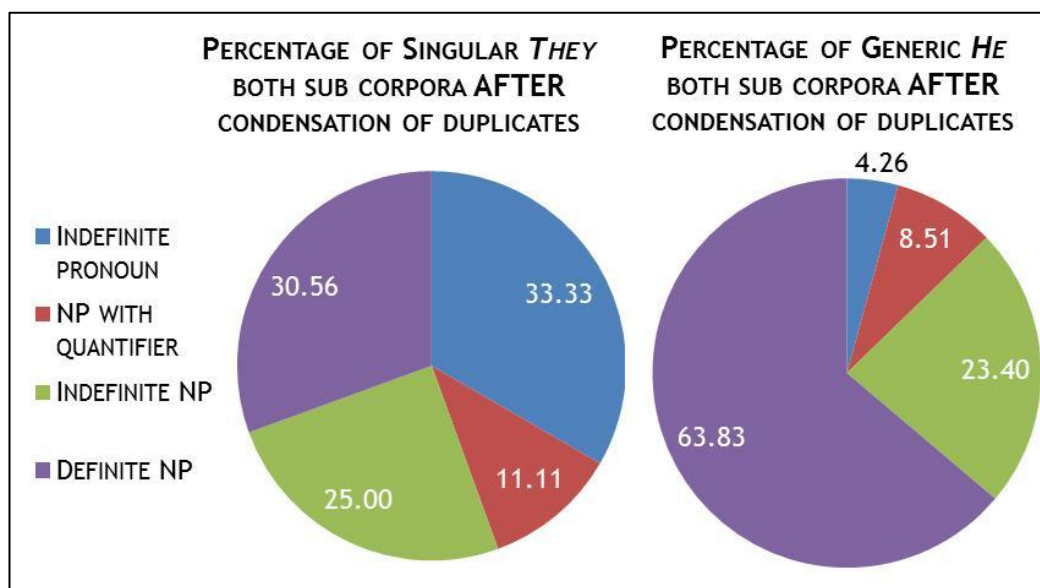


FIGURE 4.8: PERCENTAGE COMPARISON OF ANTECEDENT DISTRIBUTION FOR SINGULAR *THEY* AND GENERIC *HE* ACROSS BOTH SUB-CORPORA **AFTER** CONDENSATION PROCESS

The clear preference for using definite NPs with generic *he* is even more marked here, as such antecedents account for double the percentage of generic *he* compared with singular *they*. However, the differences between the use of indefinite NPs has balanced out and the difference between NPs with quantifiers is relatively small. Yet there is a noticeable difference between the relative use of indefinite pronouns, with the form occurring almost eight times more with singular *they* than generic *he* (33.33% to 4.26%). This suggests that reference between singular *they* and indefinite pronouns is common.

When dealing with the condensed tokens, as in Figure 4.9, the indefinite pronouns across all of the data occur with singular *they* 94.74% of the time, indicating a clear trend for this particular epicene to coindex with this type of antecedent. A similar pattern is also true for NPs with quantifiers where, in this data at least, 75% of the tokens coindex with singular *they*. Furthermore, the corresponding figure for indefinite NPs is 71.05%, thus indicating that these three antecedent types show a high level

of preference for singular *they*. None of the antecedent types favour generic *he*, with singular *they* accounting for a total of 69.68% of all the epicene pronouns in the corpus after the antecedents have been condensed. Thus, the preference for singular *they* in this data set (at least) is clear.

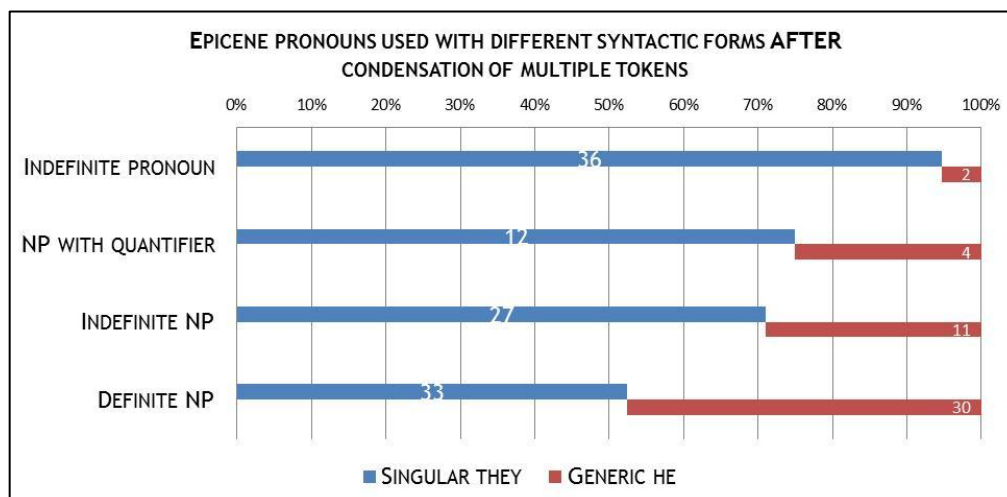


FIGURE 4.9: EPICENE PRONOUNS USED WITH DIFFERENT SYNTACTIC FORMS **AFTER** THE CONDENSATION OF ANTECEDENTS

Interestingly though, there is less of a gap between singular *they* and generic *he* when they are used with definite NPs, although the former is still the most common. As is shown in Figure 4.9 the two forms are used fairly equally (with 52.38% and 47.62% shares respectively). I hypothesise that the relatively high number of uses of generic *he* with definite NPs is twofold: (1) because such antecedents are structurally definite, the authors of the BE06 texts in the sample may have had specific referents in mind, and (2) using these mental images (see Gastil 1990) the authors were then led to choose the masculine pronoun based on the gender stereotyping of the antecedents. This hypothetical mental image of an individual may not have been as dominant for the other antecedent types as they decrease in their levels of definiteness from indefinite NPs to NPs with quantifiers to indefinite pronouns (as discussed in section 1.2.3, c.f. McConnell-Ginet 1979). Indeed, the values in Figure 4.9 show that this decrease in definiteness correlates well with the decrease in use of generic *he* and the corresponding increase in singular *they*. This change in pronoun preference in relation to the definiteness of antecedents is discussed in more detail in chapter seven.

4.2.2: QUALITATIVE ANTECEDENT ANALYSIS

In the above section I compared the numerical data presented in section 4.1, highlighting the differences between the tokens of singular *they* and generic *he* in pronouns_press and pronouns_general_prose. However, the raw figures themselves cannot illuminate the nuances of how each pronoun is used. Whilst the quantitative data has shown that, especially when multiple tokens are condensed, there is a clear preference for singular *they* over generic *he* in the BE06 data, it cannot show what specific antecedents are used with either pronoun. Therefore, I now move towards a more qualitative approach, listing the antecedents for both generic *he* and singular *they*, in order to highlight their semantic similarities and differences. Table 4.15 lists the 47 different antecedents used with generic *he* across both sub-corpora organised by syntactic category.

TABLE 4.15: ANTECEDENTS OF GENERIC *HE*

ANTECEDENTS OF GENERIC <i>HE</i> (BOTH SUB-CORPORA)			
INDEFINITE PRONOUNS	INDEFINITE NP	DEFINITE NP	
ANYONE X2	A SCHOOL PRINCIPAL	BRITISH NEOLITHIC MAN	THE IFA [INDEPENDENT FINANCIAL ADVISOR]
	A FAN	D [PERSON]	THE INDIVIDUAL POET
	A FRIDAY NIGHT POKER PLAYER	MAN X4	THE INITIAL ASSESSOR
	A HUMAN BEING	MY CORRESPONDENT	THE MODERN POET
	A MEMBER OF ANY SUCH COMMITTEES	P [PERSON]	THE PERSON CONCERNED
	A PATHOLOGIST	THAT PERSON	THE PERSON MAKING THE DETERMINATION
	A PERSON WHO IS EXTREMELY SHY	THE ART HISTORIAN	THE PERSON X2
	A PERSON X2	THE CHILD	THE POLICE OFFICER
	A SOLDIER	THE CLERK OF THE HOUSE	THE SLAVE
	AN AUTHOR	THE CUSTOMER	THE SUBJECT
		THE EDITOR	THE UK'S AMBASSADOR TO IRAQ
		THE FARMER	YOUR DOCTOR
		THE HUMAN BEING	
		THE HUMAN PERSON	

On initial inspection of the condensed antecedents, the patterns between them are not immediately obvious. However, there are several antecedents which are masculine-stereotyped¹¹⁴, such as *doctor*, *farmer*, and *soldier* (see section 4.2.3), whilst feminine stereotypes do not occur at all, further supporting the position that generic *he* has a default masculine value.

¹¹⁴ Based on the classifications in Kennison and Trofe (2003), discussed in section 3.2.2.

Yet I am aware that individual antecedent analysis can only go so far in illuminating which types of antecedents are used with generic *he*. Therefore, in order to show patterns within the antecedents, I grouped them under the following super-ordinates:

- **employment:** for antecedents that referring to specific job titles;
- **service users:** antecedents such as *customer* or *user*;
- **generic person:** including *somebody*, *person* and the generic *man*;
- and finally an ‘**other**’ category for all anomalous antecedents.

These four categories were initially chosen *post-hoc* for an analysis of generic *he* in the spoken section of the BNC, but, as Figure 4.10 shows, the same super-ordinate categories also hold for the BE06 data (see Paterson *forthcoming a*); illustrating how effective the categories are for identifying semantic parallels between antecedents of generic *he*.

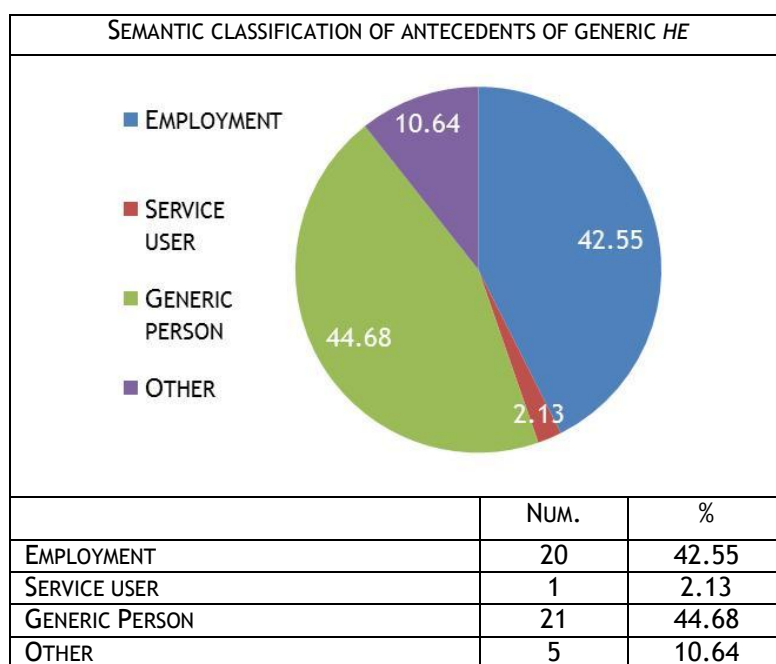


FIGURE 4.10: SEMANTIC CLASSIFICATION OF ANTECEDENTS OF GENERIC HE

The only original category that was underrepresented in the BE06 data was the ‘service user’ category, where there was only one applicable antecedent (*customer*). However, I have chosen to include it here because this category is relatively well represented in the data for singular *they* which I discuss below. As Figure 4.10 shows, the use of generic *he* correlates very well with antecedents in the ‘generic person’ category. Indeed, this is the most represented category, accounting for 44.68% of antecedents.

Whilst the consistent use of generic *he* in the ‘generic person’ category appears to give weight to the argument that it is used as a generic pronoun in British English, a closer look at specific texts in pronouns_general_prose supports an alternative explanation. Before the condensation process, Text H11, which is a section from the *Mental Capacity Act part 1* (see Figure 4.8), included 27 tokens of generic *he* coindexed with *the/a person (concerned)*. Significantly, Text H11 is a legal text, and therefore I suggest that these 27 tokens (which represent four of the 21 occurrences of the ‘generic person’ in Figure 4.10) are examples of legislated use of generic *he*, where no other pronoun would be accepted and published (c.f. 1850 Act as discussed in section 2.1.2). In addition, eight of the 21 antecedents in the ‘generic person’ category were traditional and conventional uses of the generic masculine, with antecedents such as *man*, *human being*, and *British Neolithic man*. The impact of such conventionalised, or indeed mandatory use of generic *he*, is considered in more detail below.

But first, the other key category in Figure 4.10 is the ‘employment’ group of antecedents, which mostly occur in the form of job titles. This is the second most popular group, accounting for 42.55% of antecedents. The vast majority (40%) of the ‘employment’ antecedents referred to legal and political jobs, including *each member on a committee on private bills*, *the clerk of the house*, and *the UK’s Ambassador to Iraq*. As above, I argue that the relatively high number of such antecedents may have influenced the number of uses of generic *he* overall, as the vast majority of these job titles also occurred in the legal texts in Section H of pronouns_general_prose. Thus, such examples are actually mandated uses of generic *he*, and there is no aspect of pronominal choice about them.

The seven Section H texts in pronouns_general_prose account for 94 (60.26%) occurrences of generic *he* overall, and 16 (34.04%) unique uses after condensation. The majority of these texts are from legal documents such as the *Drugs Act 2005*, the *House of Commons Staff Book*, and the *Parliamentary and Health Service Ombudsman: Annual Report*. The inclusion of these texts in the second sub-corpora could account for the fact that generic *he* occurred much more frequently in pronouns_general_prose, than in pronouns_press (section 4.1.2). If these enforced uses of generic *he*, where no alternative is permitted, are removed from the sample, then the total figure for unique antecedents of generic *he* across the sub-corpora falls

to just 31, and the quantitative gap between generic *he* and singular *they* would widen further. What this close analysis of the texts in the corpus has shown is that there are some problems with the figures for generic *he* when mandatory uses of the form are removed from the sample, a finding which was not evident in the quantitative data. Thus, the results provide even more evidence that the epicene of choice in the BE06 data is not generic *he*.

TABLE 4.16: ANTECEDENTS OF SINGULAR *THEY*

ANTECEDENTS OF SINGULAR <i>THEY</i> (BOTH SUB-CORPORA)			
INDEFINITE PRONOUNS	NP WITH QUANTIFIER	INDEFINITE NP	DEFINITE NP
ANYBODY X2	EACH [VOTER]	[CANDIDATE] 1	[THEIR] CHILD
ANYONE NAMED BY THE PERSON AS SOMEONE TO BE CONSULTED ON THE MATTER IN QUESTION OR ON MATTERS OF THAT KIND	EACH OWNER	A BAILIFF	MY NOMINATED NEIGHBOUR
ANYONE WHO CANNOT GIVE A REASONABLE EXPLANATION	EACH PARTICIPANT IN NEST BOX CHALLENGE	A CHILD OF EITHER GENDER	THE ATHLETE
ANYONE X6	EVERY CANDIDATE	A CHILD X3	THE BRITISH NEOLITHIC MAN
EVERYBODY	EVERY CHILD ON THE SWINGS OR MONKEYBARS	A CLIENT	THE CHILD X2
EVERYONE INVOLVED	EVERY CITIZEN IN A POLITICAL ORDER OF SUITABLE AGE , ETC	A CONSUMER	THE CUSTOMER
EVERYONE INVOLVED IN THE CITY'S REGENERATION	EVERY PARTICIPANT	A FRIEND	THE FELLA CONCERNED
EVERYONE X8	EVERY RUNNER	A MAN	THE HOLIDAYMAKER
NO ONE X3	EVERY SELF-RESPECTING TRAVELLER	A MEMBER OF THE PUBLIC	THE INDIVIDUAL
NOBODY X3	ONE IN NINE WOMEN	A PARENT	THE LOCAL REPORTER WHO HABITUALLY COVERS YOUR PATCH
SOMEBODY	ONE IN THREE AMERICANS	A PERSON X2	THE MUSLIM OF TODAY
SOMEONE AS INTERNATIONALLY FAMOUS AND SUCCESSFUL AS SERENA	ONE OF MY ANCESTORS	A SITTING PRIME MINISTER	THE PERSON X4
SOMEONE WITH HIV	SOME PHANTOM	A SQUADDIE	THE PHOTOGRAPHER
SOMEONE X5		A SQUAWKING TODDLER	THE PROTESTER
WHOEVER		A THEOLOGIAN	THE RECIPIENT
		A TYPICAL FIRST-TIME BUYER	THE SENDER
		A VULNERABLE HOUSEHOLDER	THE SPEAKER
		AN EDUCATIONAL PSYCHOLOGIST	THE SUBJECT
		AN INDIVIDUAL X2	THE TEACHER
		KID'S	THE USER X2
		V2 [VOTER 2]	THE VICTIM
		V3 [VOTER 3]	THIS OFFICER
			THIS PERSON
			YOUR CHILD X2
			YOUR EX-PARTNER
			YOUR LOCAL MP
			YOUR PARTNER

Moving on to the data for singular *they*, I have listed the 108 unique antecedents of singular *they* in Table 4.16. What becomes clear when comparing this table with Table 4.15 is that after the condensation process the antecedent types are more balanced for singular *they* than for generic *he*, as each syntactic category is well represented. Also in contrast with generic *he* there are only limited uses of masculine-stereotyped antecedents (see section 4.3.3 for more details). Nevertheless, there are some parallels between the two sets of data.

There are eight different antecedents common to both sets: *anyone*, *a man*, *a person*, *British Neolithic man*, *the child*, *the customer*, *the person*, and *the subject*. Although they do not all occur in the same texts with both pronouns simultaneously, this did happen with *a person* which is used in Text H23 once with generic *he* and once with singular *they*. The same phenomenon occurred for *British Neolithic man* in Text F12 even though, as discussed above, this is a conventional generic, which more often than not takes generic *he* as a pronoun. Significantly, the use of *one in nine women*, which is lexically feminine, with singular *they* parallels those examples in Baranowski (2002) and Paterson (2011), discussed in section 1.3.2, where the use of a quantifier influenced pronominal choice even when sex was lexically specified.

A similar interference of syntactic definiteness occurs in Text G05 where the indefinite NP *a fan* is used with generic *he*, whilst the indefinite pronoun *everybody* coindexes with *they're*, suggesting again that choice of pronoun may be affected by the definiteness (and arguably visibility) of a discreet antecedent (see above and McConnell-Ginet 1979). However, an unusual example similar to this appears in Text G43 where two indefinite NPs, *a person who is extremely shy* and *a person*, coindex with generic *he* and singular *themselves* respectively. In this case, there is no difference in the syntactic classification of antecedents; arguably, they are both as definite as each other. Although it is not possible to find out what motivated these particular pronoun choices it is plausible that the author(s) had a definite singular referent in mind and overtly did not want to express gender, which may help to explain the use of the morphologically singular gender-neutral *themselves*.

Alternatively, working with McConnell-Ginet's (1979) argument that it is difficult to imagine a referent without assigning biological sex, it could be argued that *a person who is extremely shy* is more easily visualised in the human brain as it is post-modified by certain personality traits. On the other hand, *a person* does not have this extra information and is therefore easier to visualise as a sex-less entity. However, both explanations are practically impossible to verify, and as there are no other similar contrasts across the data, there is not enough data to support either hypothesis.

Perhaps the most interesting individual comparison of antecedents is the use of *a squaddie* and *a soldier* in Text A10. Both antecedents are indefinite NPs yet the first coindexes with singular *they* and the latter is used with generic *he*. These two antecedents are both syntactically and semantically similar, and thus there is no apparent reason why both pronominal forms are used, especially considering that singular *they* is used alongside calls to "Back Our Boys" (Text A10, a NP which simply cannot have a gender-neutral reading). This is most likely an idiosyncratic use of generic *he* and singular *they*, which has no real parallels in the data.

There are however, some other intriguing fine-grained contrasts between the two sets of antecedents. For example, there is an interesting difference in the use of *the teacher* and *a school principal*, where, in contrast with the usage noted above, the indefinite NP is coindexed with generic *he* whilst the definite NP is an antecedent of singular *they*. Although these two forms do not appear in the same text, I argue that the choice of pronoun here may have been influenced by social roles and gender stereotypes (see section 4.3.3). The higher position of authority of school principals compared with teachers may mean that the author using the former visualised the referent as masculine, and thus used (pseudo)generic *he* with *a school principal*. A similar phenomenon occurs in Text D03 with *the local reporter who habitually covers your patch* and *the editor*. The more senior position of editor is coindexed with generic *he* whilst the reporter coindexes with singular *they*. These two examples do not represent a general trend in the data, but they are perhaps worth bearing in mind.

However, as with the quantitative analysis in the first part of this section, the real similarities and differences between the two sets of data cannot be ascertained through the analysis of individual examples. Therefore,

in Figure 4.11 I grouped the antecedents of singular *they* according to the categories for antecedents of generic *he* shown above (c.f. Figure 4.10).

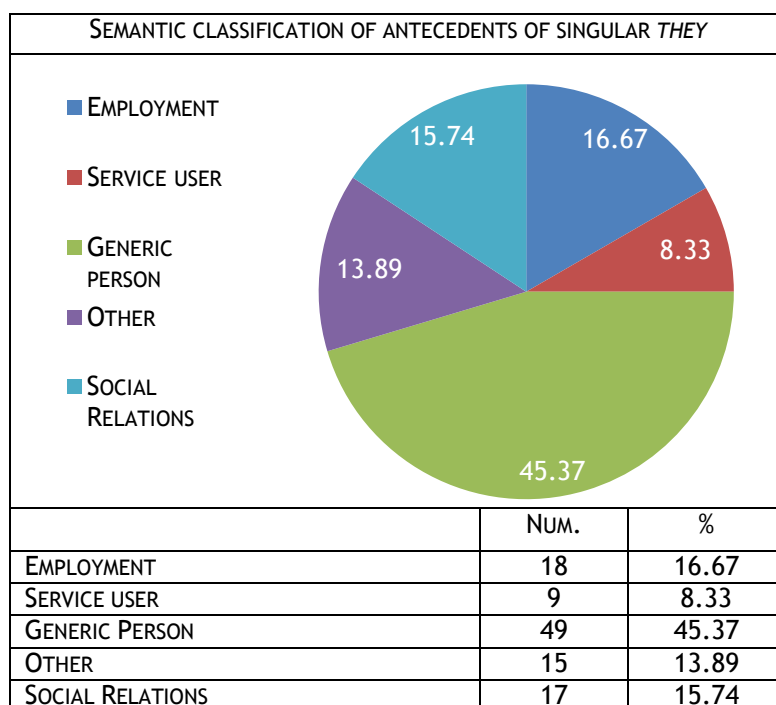


FIGURE 4.11: SEMANTIC CLASSIFICATION OF ANTECEDENTS OF SINGULAR *THEY*

The chart shows that the antecedents for singular *they* do not have the same semantic distribution as generic *he*. The higher overall total of indefinite pronouns, combined with other antecedents, such as *a person*, means that the “generic person” category represents almost half (45.37%) of the antecedents. Also, on a par with the BNC data discussed in Paterson (*forthcoming a*), the “service user” category is better represented in the antecedents of singular *they*. Whilst the “employment” category, which was a major component of the antecedents of generic *he*, only accounts for 16% of the singular *they* antecedents.

In addition, there were seventeen semantically similar antecedents of singular *they* which could not be classified under the current categories. As Figure 4.11 shows, I have added a category labelled “social relations” for antecedents such as *a parent*, *your child*, *my nominated neighbour* and *your ex-partner*. Thus, it is clear that, although the categories used above to classify the antecedents of generic *he* can be applied to singular *they*, they are not the best fit. However, there are some similarities, for example, as with generic *he*, the legal and political job-related antecedents are prominent in the data for singular *they*, suggesting that this type of job is

not exclusively referenced using generic *he*, and showing overlap between the two sets of antecedents.

Finally, there is one significant nuance in the usage of singular *they*; there are nine instances of the pronoun in the legal texts of Section H. Although seven of these antecedents are in public information leaflets or company reports, which are obviously not legal texts, the other two tokens are noteworthy. One occurs in Text H11 (*Mental Capacity Act part 1.*) with *anyone named by the person...*, and the other token is in Text H23 (*Parliamentary and Health Service Ombudsman: Annual Report*) with the antecedent *a person*. The occurrence of singular *they* in these two texts goes against the assertion I made above that it is standard convention to use generic *he* in legal texts. However, these two instances appear to be anomalies, as in terms of raw frequency (i.e. before the condensation of antecedents) generic *he* occurs in Text H11 no fewer than 50 times, whilst singular *they* only occurs once. Therefore, this occurrence of singular *they* may just have slipped through the editing process before the document was published. The very small percentage value for singular *they* in Text H11 (1.96%) suggests that its use is not significant. However, it potentially indicates that the original author of the text used singular *they* but the pronouns were changed to generic *he* during the editing process, but there is no way to justify this claim.

Having established that a qualitative comparison of my datasets is fruitful for establishing the nuances of epicene usage, the final part of my analysis of the BE06 data concerns the gender stereotyping attached to the antecedents of both singular *they* and generic *he*. I deal with this topic in the section below, comparing my data to the lists of stereotyped nouns provided by Kennison and Trofe (2003) and Carreiras et al. (1996) which I discussed in section 3.2.2.

4.2.3: GENDER STEREOTYPED ANTECEDENTS

My goal when conducting this stereotype analysis was to see whether there is a tendency for a particular epicene to occur with either masculine or feminine stereotypes. If a pattern of stereotype usage is evident, then

the results would indicate that stereotypes can influence epicene choice. Such a result would provide more evidence that the epicene debate involves both language-internal and language-external factors, as stereotypes are not a syntactic phenomenon (c.f. Osterhout, Bersick and McLaughlin (1997)). In order to compare my antecedents of generic *he* and singular *they* with the stereotype data from Kennison and Trofe (2003) and Carreiras et al. (1996) I removed the articles and quantifiers attached to the nouns in the BE06 data, as well as any adjectival pre or post-modification. Thus, the final lists of antecedent forms discussed below were bare NPs only, which made them directly comparable to the lists of stereotype values. In addition, antecedents such as *person*, which occurred in the data as both definite and indefinite NPs, were condensed to just one token to avoid repetition of analyses.

Whilst Kennison and Trofe used a seven point scale which ranged from feminine-stereotyped at 1.0 to masculine-stereotyped at 7.0, Carreiras et al. used an eleven-point scale from masculine to feminine, where the lower numbers corresponded to masculine stereotypes and the higher numbers were associated with feminine stereotypes. As the two scoring systems are different I have provided the raw figures for each antecedent tested in the two tables below. In all cases, values from Kennison and Trofe that were higher than 5.0 were classified as masculine-stereotyped, and values less than 3.0¹¹⁵ were classed as feminine-stereotyped. In the case of Carreiras et al., the lower the value, the more masculine-stereotyped the noun was perceived to be, with the higher values corresponding to feminine stereotypes.

In some cases, the antecedents in the BE06 data did not map directly onto the forms tested in previous research and, for the most part, I eliminated such antecedents from this consideration of gender stereotypes. However, there were limited cases where, although the antecedents did not match the tested forms exactly, there were clear near-synonyms in the lists of nouns tested by Kennison and Trofe and/or Carreiras et al. For example, the antecedent *school principal* was covered by both *elementary school principal* and *high school principal* in Kennison and Trofe's data, whilst I used values for *government official* and *diplomat* to estimate a stereotype

115 The scale started at 1.0 (not at zero).

value for *ambassador*. In cases where substitute nouns were used every care was taken to avoid subjective judgements that could have skewed the data, and classification as a synonym was as strict as possible.

For example, whilst *card player* was included as a comparable form to *poker player*, the antecedent *pathologist* was eliminated as the only tested forms in the previous data which could have been used included *allergist*, *physician*, and *research scientist*, none of which accurately and adequately covered the job done by a pathologist. In addition, although arguably *doctor* could have been used as a coverall term for *pathologist* I did not make this comparison, as *doctor* already occurred in the original list of antecedents and I did not want to repeat data. Yet, in any case, the results given for the antecedents that did not match the tested forms exactly are only a guide to the stereotypical value of the actual antecedent form, and thus they are of less weight than the results for those forms that matched exactly.

The results for the antecedents of generic *he* are presented in Table 4.17, and instances where substitute nouns were used are highlighted in grey. As Table 4.17 shows, the majority of the results correspond to data from Kennison and Trofe¹¹⁶, but where possible the Carreiras et al. data is also used. Interestingly, where there was data available from both studies, there tends to be a high level of consensus in whether the participants in the studies perceived the nouns as masculine or neutral, for example Table 4.17 shows that participants in both studies deemed the antecedent *farmer* to be highly masculine-stereotyped. This correlation suggests that the stereotypes for these particular nouns are relatively fixed, but confirmation of this would need more research. Nevertheless, it is interesting that the stereotypes hold across the seven years between the two studies.

Eighteen antecedents of generic *he* occurred in the lists of previously tested nouns, and seven of these antecedents (which is almost half) were classified as masculine-stereotyped. Ranked in order of most-masculine (the highest values from the Kennison and Trofe data and the lowest values in the Carreiras et al. data) to least-masculine, these were *solider*, *farmer*, *police officer*, *ambassador*, *magazine proprietor*, *poker player* and *school principal*.

116 Kennison and Trofe (2003) give stereotype values for male and female participants separately. Whilst the figures have been combined for the purpose of this thesis I have provided the individual figures for completeness, to indicate any points where participant sex may have interacted with stereotype values.

TABLE 4.17: STEREOTYPES WITH GENERIC HE

ANTECEDENT	SOURCE	TESTED FORM	RESULTS	STEREOTYPE
AMBASSADOR	KENNISON AND TROFE	DIPLOMAT	FEMALE: 4.80 MALE: 5.30	MASCULINE
		GOVERNMENT OFFICIAL	FEMALE: 5.25 MALE: 5.30	
FARMER	KENNISON AND TROFE	FARMER	FEMALE: 6.20 MALE: 6.20	MASCULINE
	CARREIRAS ET AL.	FARMER	2.27	
MAGAZINE PROPRIETOR	KENNISON AND TROFE	COMPANY PRESIDENT	FEMALE: 5.05 MALE: 5.45	MASCULINE
		BOSS	FEMALE: 5.10 MALE: 5.50	
POKER PLAYER	KENNISON AND TROFE	CARD PLAYER	FEMALE: 4.80 MALE: 5.25	MASCULINE
POLICE OFFICER	KENNISON AND TROFE	POLICE OFFICER	FEMALE: 5.55 MALE: 5.45	MASCULINE
SCHOOL PRINCIPAL	KENNISON AND TROFE	ELEMENTARY SCHOOL PRINCIPAL	FEMALE: 4.40 MALE: 4.55	MASCULINE
		HIGH SCHOOL PRINCIPAL	FEMALE: 5.15 MALE: 5.15	
SOLDIER	KENNISON AND TROFE	SOLDIER	FEMALE: 6.15 MALE: 6.21	MASCULINE
	CARREIRAS ET AL.	SOLDIER	1.90	
ART HISTORIAN	KENNISON AND TROFE	ARTIST	FEMALE: 3.80 MALE: 4.05	NEUTRAL
		HISTORIAN	FEMALE: 4.30 MALE: 4.45	
	CARREIRAS ET AL.	ART HISTORIAN	5.33	
AUTHOR	KENNISON AND TROFE	AUTHOR	FEMALE: 3.90 MALE: 3.95	NEUTRAL
CHILD	KENNISON AND TROFE	CHILD	FEMALE: 4.00 MALE: 4.10	NEUTRAL
CLERK OF THE HOUSE	KENNISON AND TROFE	CLERK	FEMALE: 3.80 MALE: 3.75	NEUTRAL
CUSTOMER	KENNISON AND TROFE	CUSTOMER	FEMALE: 3.75 MALE: 4.05	NEUTRAL
DOCTOR	KENNISON AND TROFE	DOCTOR	FEMALE: 4.65 MALE: 4.60	NEUTRAL
EDITOR	KENNISON AND TROFE	EDITOR	FEMALE: 4.15 MALE: 4.95	NEUTRAL
FAN	KENNISON AND TROFE	FAN	FEMALE: 4.00 MALE: 4.40	NEUTRAL
IFA [INDEPENDENT FINANCIAL ADVISOR]	KENNISON AND TROFE	ACCOUNTANT	FEMALE: 4.25 MALE: 4.55	NEUTRAL
PERSON	KENNISON AND TROFE	PERSON	FEMALE: 4.00 MALE: 4.10	NEUTRAL
POET	KENNISON AND TROFE	POET	FEMALE: 3.90 MALE: 3.90	NEUTRAL
ANTECEDENTS NOT TESTED: ANYONE, ASSESSOR, BRITISH NEOLITHIC MAN, CORRESPONDENT, D [PERSON], HUMAN BEING, MAN, MEMBER OF A COMMITTEE, P [PERSON], PATHOLOGIST, POET, SLAVE, SUBJECT				

As the male and female responses to the tested forms were separated in the list given in Kennison and Trofe (2003:366-374) when the values for the sexes differed (as with *card player*) I took an average to determine the level of gender stereotyping. The lowest value was 3.90 for *poet*, which is still

well within the middle range on a scale from 1.0 to 7.0 and thus generic *he* did not occur with any feminine stereotypes.

Following on from the antecedents of generic *he*, Table 4.18 (overleaf) shows the corresponding data on stereotypes for the antecedents of singular *they*. There were twenty antecedents in the BE06 data that had been tested in previous research, and four of these were masculine-stereotyped. They rank from the most-masculine-stereotyped *squaddie*, through *prime minister*, and *local MP*, to *officer*, which is the least-masculine-stereotyped antecedent. However, looking more closely at these masculine-stereotyped antecedents, none of them were direct matches between the antecedent and the forms tested by Kennison and Trofe and/or Carreiras et al. As noted above, in cases where there was no direct match between an antecedent and the previous research near synonyms were used to calculate the stereotype value. However, this method can only result in estimated stereotype values that do not have as much weight as those values that represent a match between the antecedent and the tested form.

Table 4.18 also shows that two antecedents of singular *they* were feminine-stereotyped; these were *teacher* and *victim*. The results for *teacher* were not based on a direct comparison between the antecedent and the corresponding tested form, instead, there were four comparable forms used to estimate a stereotype value: *elementary school teacher*, *high school teacher*, and two occurrences of *kindergarten teacher*¹¹⁷. What is interesting, is that the stereotype values change for each form, with teachers working with younger children being perceived as relatively more female than teachers working with older children. Indeed, the value attributed to *kindergarten teacher* by the female participants in the Kennison and Trofe data is the lowest value across all the antecedents of both generic *he* and singular *they* from the BE06 sub-corpora. The form *kindergarten teacher* was also tested by Carreiras et al. and their participants marked the form as highly feminine-stereotyped, giving it an overall value of 9.57 out of 11.0, illustrating the consistency between both previous pieces of research.

¹¹⁷ Taking an average for all of the forms gives a value of 2.51, which is below the threshold of 3.0, and therefore makes the antecedent feminine-stereotyped.

TABLE 4.18: STEREOTYPES WITH SINGULAR *THEY*

ANTECEDENT	SOURCE	TESTED FORM	RESULTS	STEREOTYPE
LOCAL MP	KENNISON AND TROFE	POLITICIAN	FEMALE: 5.10 MALE: 5.80	MASCULINE
OFFICER	KENNISON AND TROFE	POLICE OFFICER	FEMALE: 5.55 MALE: 5.45	MASCULINE
PRIME MINISTER	KENNISON AND TROFE	POLITICIAN GOVERNMENT OFFICIAL	FEMALE: 5.10 MALE: 5.80 FEMALE: 5.25 MALE: 5.30	MASCULINE
SQUADDIE	KENNISON AND TROFE CARREIRAS ET AL.	SOLDIER SOLDIER	FEMALE: 6.15 MALE: 6.21 1.90	MASCULINE
ATHLETE	KENNISON AND TROFE	ATHLETE	FEMALE: 4.65 MALE: 5.20	NEUTRAL
CHILD	KENNISON AND TROFE	CHILD	FEMALE: 4.00 MALE: 4.10	NEUTRAL
CLIENT	KENNISON AND TROFE	CLIENT	FEMALE: 4.15 MALE: 4.50	NEUTRAL
CONSUMER	KENNISON AND TROFE	CUSTOMER	FEMALE: 3.75 MALE: 4.05	NEUTRAL
EDUCATIONAL PSYCHOLOGIST	KENNISON AND TROFE	SCHOOL PSYCHOLOGIST	FEMALE: 3.40 MALE: 3.50	NEUTRAL
HOLIDAYMAKER	KENNISON AND TROFE	HOTEL GUEST CAMPER	FEMALE: 3.95 MALE: 4.20 FEMALE: 4.60 MALE: 5.35	NEUTRAL
KID	KENNISON AND TROFE	KID	FEMALE: 4.05 MALE: 4.25	NEUTRAL
NEIGHBOUR	KENNISON AND TROFE	NEIGHBOUR	FEMALE: 3.95 MALE: 3.95	NEUTRAL
PARENT	KENNISON AND TROFE	PARENT	FEMALE: 3.90 MALE: 3.90	NEUTRAL
PERSON	KENNISON AND TROFE	PERSON	FEMALE: 4.00 MALE: 4.10	NEUTRAL
PHOTOGRAPHER	KENNISON AND TROFE	PHOTOGRAPHER	FEMALE: 4.00 MALE: 3.80	NEUTRAL
PROTESTER	KENNISON AND TROFE	PROTESTOR	FEMALE: 4.25 MALE: 3.90	NEUTRAL
REPORTER	KENNISON AND TROFE	REPORTER	FEMALE: 3.60 MALE: 3.50	NEUTRAL
TODDLER	KENNISON AND TROFE	TODDLER	FEMALE: 4.00 MALE: 4.05	NEUTRAL
TEACHER	KENNISON AND TROFE CARREIRAS ET AL.	ELEMENTARY SCHOOL TEACHER HIGH SCHOOL TEACHER KINDERGARTEN TEACHER KINDERGARTEN TEACHER	FEMALE: 2.30 MALE: 2.20 FEMALE: 3.20 MALE: 3.00 FEMALE: 1.95 MALE: 2.40 9.57	FEMININE
VICTIM	KENNISON AND TROFE	VICTIM	FEMALE: 2.95 MALE: 2.95	FEMININE
<p>ANTECEDENTS NOT TESTED: AMERICAN, ANCESTOR, ANYBODY, ANYONE, BAILIFF, BRITISH NEOLITHIC MAN, CANDIDATE, CITIZEN, EVERYBODY, EVERYONE, EX-PARTNER, FELLA, FIRST-TIME BUYER, FRIEND, HOUSEHOLDER, INDIVIDUAL, MAN, MUSLIM, NO ONE, NOBODY, OWNER, PARTICIPANT, PARTNER, PHANTOM, PUBLIC, RECIPIENT, RUNNER, SENDER, SOMEBODY, SPEAKER, SUBJECT, THEOLOGIAN, TRAVELLER, USER, VOTER, WHOEVER, WOMEN</p>				

Comparing the results for singular *they* with those discussed above for generic *he*, what is clear is that the vast majority of antecedents in this data were stereotypically neutral. As such, any arguments about the gender stereotypes associated with either generic *he* or singular *they* are limited to those few antecedents that had clear stereotyping. Yet comparisons can be made between the seven occurrences of masculine stereotypes with generic *he* and the four masculine stereotypes used with singular *they*. The most masculine-stereotyped antecedent was *soldier*, which had a value of 6.21 out of 7.0 when assessed by the male participants in Kennison and Trofe's research, and occurred with both pronouns. In addition, the average masculine-stereotype values¹¹⁸ for both sets of antecedents are highly comparable; masculine-stereotyped antecedents of singular *they* have an average value of 5.57, compared with a 5.56 for generic *he*. These figures suggest that although there were more masculine-stereotyped antecedents used with generic *he* than singular *they*, there was no preference for generic *he* to co-occur with antecedents which were more heavily masculine-stereotyped (which would have values closer to 7.0) than singular *they*.

However, looking at the values for all of the antecedents covered in Table 4.17 and Table 4.18 there is a slight discrepancy between the overall average stereotype values, which were 4.69 for generic *he* and 4.15 for singular *they* based on the Kennison and Trofe data. This suggests that the pool of antecedents used with singular *they* are less masculine-stereotyped than the antecedents of generic *he* which have an average stereotype value 0.54 higher. Importantly, the average figures for both sets of antecedents are still within the neutral range of between 3.0 and 5.0, although both figures tend towards the masculine end of the spectrum. The lower average stereotype value for singular *they* can be explained by the fact that two of its antecedents were feminine-stereotyped whilst generic *he* had no feminine-stereotyped antecedents at all.

Although this stereotype investigation is limited, and it is based on data from other studies, the results of my analysis of the co-occurrence of stereotypes with singular *they* and generic *he* has shown that, in the data from the BE06 sub-corpora, generic *he* has more masculine-stereotyped

118 Average values were calculated by adding together all the Kennison and Trofe values for the masculine stereotyped antecedents of each pronoun (both male and female values) and dividing by the total number of pieces of data entered.

antecedents than singular *they* in a ratio of 7:4. In addition, none of the masculine-stereotyped antecedents of singular *they* were directly tested by either Kennison and Trofe or Carreiras et al. and a substitute form had to be used. As such, the stereotype values given for these antecedents are only estimations. In comparison, although four of the seven masculine-stereotyped antecedents of generic *he* were also calculated in this way, there were three other masculine stereotypes that matched the forms tested in previous data exactly.

In terms of feminine stereotypes, none of the antecedents of generic *he* had values below the 3.0 threshold, although there were two antecedents of singular *they* which were classed as feminine stereotypes, with one of these antecedents matching forms tested in previous research exactly. Thus the ratio of feminine-stereotyped antecedents is 2:0 in favour of singular *they*. However, perhaps most interestingly, the vast majority of antecedents included in Table 4.17 and Table 4.18 have neutral stereotype values. Thus, the participants in Kennison and Trofe's study, and in Carreiras et al.'s research, did not strongly attribute a gender to most of the social roles covered by the antecedents of the BE06 sub-corpora. It therefore appears that there is little interaction between the selection of a generic pronoun and the gender stereotyping associated with a particular antecedent. However, the raw figures for the number of stereotyped nouns used with each pronoun do suggest that there is a slight preference for generic *he* with masculine-stereotyped antecedents, and a larger preference for singular *they* with feminine-stereotyped antecedents.

In this section I compared and contrasted the results for singular *they* and generic *he* that I obtained as part of my analysis of the BE06 sub-corpora pronouns_press and pronouns_general_prose (as described in section 3.1.2). I began with quantitative analyses, looking at frequency counts for both pronouns, and showing how they are distributed across different types of antecedent (an area that has not been given much consideration in previous research). However, noting the limitations of a purely quantitative approach I also included some more-qualitative elements in my analysis, looking at how the text types in the sub-corpora may have influenced the totals for generic *he*, and comparing and individual examples of antecedents common

to both pronouns. Finally, in relation to research question three, I analysed the antecedents in my BE06 data in terms of gender stereotypes, directly comparing my research with previous research on stereotype values. I summarise my findings, which strongly support the conclusion that singular *they* is the most common epicene in written British standard English, below.

CONCLUSIONS

In this chapter I have shown that singular *they* is the preferred epicene form in the BE06 data. As the BE06 corpus was constructed to reflect, and thus be representative of, modern written British English usage I can thus argue that singular *they* appears to be the widely-used epicene of choice, thus supporting my hypothesis that it is the most-used epicene in current written British English. This assertion is supported by the vast majority of the data from my analysis:

- The distribution of singular *they* across sub-corpora is relatively uniform (3.06% and 4.38%) whilst the same is not true for generic *he* (0.8% and 3.95%).
- The overall percentage value of singular *they* (4.12%) across both sub-corpora is greater than the corresponding figure for generic *he* (3.22%).
- The ratio of usage of singular *they* to generic *he* when the sample sizes when normalised is 50:39 in favour of singular *they*.
- When the antecedents are condensed unique occurrences of singular *they* outnumber generic *he* at a ratio of 108:47.
- Singular *they* is distributed more evenly than generic *he* across different antecedent types (Figure 4.8).
- Singular *they* is the most-used epicene with all syntactic classifications of antecedents (Figure 4.9).
- Singular *they* is even used with traditional generics, such as *British Neolithic man*.

- Singular *they* is used with more qualitatively (semantically) different classes of antecedents (Figure 4.10; Figure 4.11).
- Singular *they* is used with masculine and feminine-stereotyped nouns, whilst generic *he* is only used with the former.
- The figures for generic *he* are inflated somewhat by the inclusion of legal texts in pronouns_general_prose.

The results also showed that the use of singular *they* across the different syntactic categories of antecedent increases rather uniformly from the most definite antecedents, definite NPs, to the most indefinite antecedents, indefinite pronouns. In fact, generic *he* only rarely coindexes with indefinite pronouns, with singular *they* accounting for over 95.74% of tokens before condensation.

However, it is not the case that the antecedents of generic *he* and singular *they* are exclusive; there is overlap. What is apparent is that singular *they* occurs with all the different syntactic classifications of antecedent in both sub-corpora, whilst generic *he* is more restricted, occurring with very few indefinite pronouns, for example. This suggests that the range within which singular *they* is used as an epicene pronoun in this data is greater than the range for generic *he*. Thus I have shown that numerically, syntactically, and to some extent qualitatively, singular *they* is the epicene of choice in this BE06 data, and thus arguably in wider instances of written British English.

CHAPTER 5: STANDARD ENGLISH AND PRESCRIPTIONS

5.1: ANALYSING PRESCRIPTIONS	167
5.1.1: ENDORSING EPICENE CANDIDATES	168
5.1.2: ANTECEDENT ANALYSIS	174
5.1.3: DEALING WITH STEREOTYPES	179
5.2: SOCIAL INTERFERENCE?	182
5.2.1: CORPUS COMPARATIVE STATISTICAL KEYWORDS	182
5.2.2: KEYWORDS IN THE CORPUS	186
CONCLUSIONS	194

Having established current epicene usage in the previous chapter, and shown that singular *they* appears to be the preferred epicene form in written British standard English, I now turn my attention to epicene prescriptions and research questions two and three, which involve language-external influences (section 3.2). My aim in this chapter is to bring research on epicene prescriptions in grammar books up to date, as my review of the current scholarship in section 2.3 indicates that there has been little research on this type of data in the past ten to fifteen years. I thus address the hypothesis that grammars published at the start of the twenty-first century will continue the trend evident in the wider literature to reject generic *he* in favour of epicene avoidance tactics. In addition, I also focus on the related hypothesis that language-external forces will affect language prescriptions.

In section 5.1 I detail how I structured the analysis of the grammar corpus, presented in section 3.1.2, and provide initial quantitative data on the relative mentions of generic *he* and singular *they*. The results show that in this grammar corpus, generic *he* is categorised as *traditional* usage, whilst there is much more discussion of singular *they* than in previous similar corpora (c.f. Zuber and Reed 1993). I also document the epicene avoidance tactics proposed by the grammar authors in order to avoid full endorsement of either epicene candidate. I analyse the antecedents used in the illustrative examples provided by the grammar authors, showing that there is an avoidance of syntactically definite antecedents. I also show how the

authors of the texts in the grammar corpus avoid using heavily stereotyped antecedents, thus distorting the issue of epicene reference, and making it appear as though the choice of a third-person singular gender-neutral pronoun can be explained relatively simply and unproblematically. This position thus ignores the large body of literature I addressed in chapters one and two, which indicates that epicene reference is a complex issue influenced by both language-internal and language-external forces.

In section 5.2 I investigate whether the keywords set out in research question three occur in the grammar corpus, thus making explicit reference to language-external factors. I compare the grammar corpus with a standard reference corpus to obtain the Corpus Comparative Statistical Keywords and show how, although the occurrences of the keywords associated with research question three are not statistically significant, a closer look at their use within the grammar corpus and the BE06 data indicates that this is a feature of the CCSK method. A more fine-grained analysis of the keywords highlights the fact that they are pertinent to the corpus, as the grammar authors use the terms in relation to prescriptions and agency.

The results from the analysis I present in this chapter bring the data on epicene prescriptions in grammar handbooks up to date, thus adding new data to the wider literature on epicene prescriptions. In addition, the results are directly comparable to the usage data from the BE06 sub-corpora I analysed in chapter four due to their similar publication dates. I compare and contrast the two sets of results in more detail in chapter seven.

5.1: ANALYSING PRESCRIPTIONS

The focus of this chapter is to provide data on modern language prescriptions on epicene usage. I began my analysis by manually checking whether or not the twenty grammars in the corpus (see section 3.1.2) explicitly covered the topic of epicones, and how much space was given to

these discussions¹¹⁹. Thirteen out of the twenty texts covered epicenes, with the other seven only focusing on the standard pronoun paradigm; classing *they* as a plural pronoun and not discussing singular generic references¹²⁰. The total number of words in the grammar corpus is 3874, which consists of the thirteen sections of text devoted to epicene pronouns. In comparison with the BE06 sub-corpora this number is relatively small and thus the grammar corpus does not lend itself to the primarily quantitative analysis I performed in the previous chapter.

Therefore, as well as using some methods of analysis common to chapter four, such as documenting the syntactic distribution of antecedents of generic *he* and singular *they* I also had to use more qualitative methods in order to categorise the approaches the grammar authors took to epicene pronouns (section 5.1.1). In section 5.1.2 I take a closer look at the antecedents the grammar authors selected for their example sentences, where they illustrate the use of their chosen epicene, subjecting these antecedents to stereotype analysis in section 5.1.3. However, before the numerical analysis of the grammar, I begin with a note on the graphology of the grammar books and the visual salience of epicenes in such texts.

5.1.1: ENDORSING EPICENE CANDIDATES

All thirteen texts in the grammar corpus that covered epicene pronouns dealt with the issue as a separate point of English grammar, that is, consideration of epicenes was delineated in some way from other more general discussions of pronouns. Most grammars included a separate heading for epicenes, whilst three texts made the topic graphologically salient by placing comments on epicenes in a special text box, separate from the rest of the text (as illustrated in Figure 5.1). I argue that making the topic visually salient has two implications; firstly, the authors have deemed

119 Whilst the dataset is not overly large, the number of texts in my corpus is comparable to previous studies of this type of data (c.f. Bodine 1975; Sunderland 1986; Zuber and Reed 1993) and whilst a small dataset cannot be taken as conclusive evidence that all grammars will follow the trends in my results, my corpus is sufficient for bringing to light common issues and/or key themes in modern grammar prescriptions.

120 The average amount of space devoted to the topic of third-person epicene reference was 298 words, ranging from an upper limit of 641 words (Text 1) down to 75 words (Text 7).

epicenes important enough to warrant their own section, and secondly, using such graphological features makes the issue stand out to readers, suggesting that choice of epicene pronoun is an important issue in English grammar.

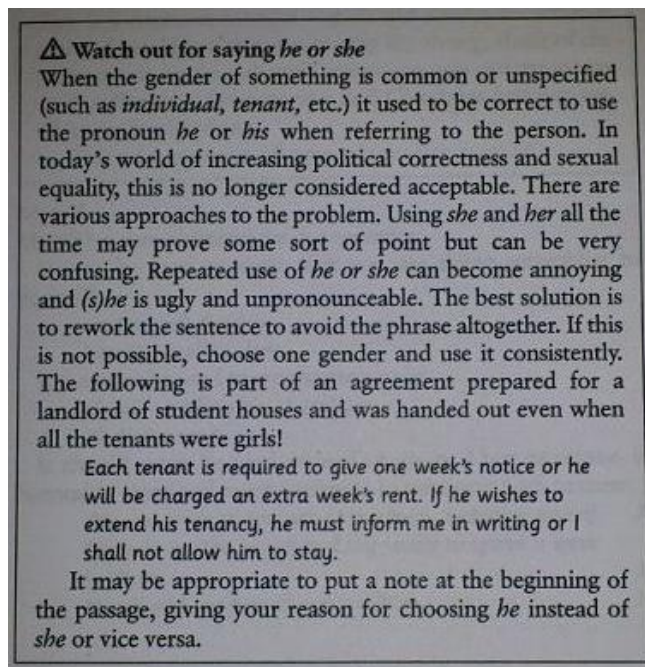


FIGURE 5.1: GRAPHOLOGICAL SALIENCE OF DISCUSSIONS OF EPICENE PRONOUNS (TEXT 8)

The fact that all of the texts dealt with epicenes under a separate heading from other discussions of personal pronouns was not a feature of the manual search, as my method involved reading any notes on pronouns which occurred in each text. Therefore, if epicenes had been mentioned only in passing I would still have picked up the occurrences when reading longer sections concerned with other pronouns. Thus, the data strongly suggests that, if epicene pronouns are discussed in modern grammar books, they are separated from longer sections on pronouns by graphological features ranging from a simple heading, in line with the rest of the formatting of the text, to special text boxes. I argue that such practices suggest that, although not all grammars include a consideration of epicene pronouns, when they are mentioned, they are treated as a special topic, independent from the general discussion.

Having established which of the twenty original texts discussed epicene pronouns, I then categorised the positions that the grammar authors took on epicenes. To do this I transcribed the relevant sections in the grammars and counted and categorised the approaches the authors took to epicenes into

three groups, depending on whether the authors *endorsed*, *dismissed*, or *did not consider*, generic *he* or singular *they* (each category is defined below). Although this process was arguably subjective, and due to the nature of the task electronic analysis was not possible, I was as conservative as possible when allocating each discussion of epicenes in the corpus to a category. In order to avoid any incorrect classifications, each text could be allocated to more than one category. For example, if there were instances in a given text of both acceptance and rejection of generic *he* then the text was classified as both *endorsing* and *dismissing* the candidate. As the vast majority of grammars did not completely endorse a particular pronoun, allowing each text to be allocated to multiple categories more accurately reflects the content of the grammars, when compared with a system where each grammar could only be allocated to one group. For example, a single-entry tripartite system would not cover instances where grammars endorsed the use of singular *they* but only in speech, and rejected the pronoun in its written form.

I classified a text under the endorsement category if the use of a particular pronoun was mandated or expressed as an imperative, for example, there is clear endorsement for generic *he* in Text 12¹²¹ which states that readers should “use *he* throughout”. Similarly for singular *they*, texts endorsing the form contained phrases such as “increasingly common” and “generally accepted” (Text 4), or “you have our permission to use it” (Text 19). Conversely, a text was deemed to *dismiss* an epicene candidate if a form was mentioned, but not formally endorsed. Examples for this approach for generic *he* include the statements that “many people dislike this” (Text 2) or statements that the form is “no longer appropriate” (Text 17) or “no longer considered acceptable” (Text 8). Interestingly, there were fewer examples of imperatives in this category (and I discuss the use of the passive in more detail in section 5.2.2). For singular *they* the same criteria applied, and authors were deemed to have dismissed the form if they included negative evaluations of the pronoun, such as “In serious writing, however, it is often avoided” (Text 2), or explicit statements including “This is not recommended” (Text 11).

121 A list of the texts in the grammar corpus is included in Appendix A.

The final category of *did not consider* was the easiest to codify, and the least subjective, as it involved checking whether or not the grammars explicitly mentioned generic *he* and/or singular *they*, or whether only other forms (or epicene avoidance tactics) were considered instead. A form was either mentioned, which meant that the text was then classified either endorsing or dismissing the form (or indeed both), or it was not mentioned, and thus a text was assigned to the *did not consider* category. If there was endorsement for a particular pronoun in a text (either generic *he* or singular *they*), I looked at what reasons the grammar writers had used to justify their prescription. Conversely, if there was explicit mention of generic *he* (or singular *they*) in order to dismiss it, and/or advise against such usage, I documented the reasons why this was the case. Finally, if either generic *he* or singular *they* was rejected or not considered at all I noted what alternative form(s) if any were prescribed as good usage. The results of this initial analysis are documented in Table 5.1.

TABLE 5.1: SINGULAR *THEY* AND GENERIC *HE* IN THE GRAMMAR CORPUS

	NOTED	ENDORSED	DISMISSED
GENERIC <i>HE</i>	11	1	11
SINGULAR <i>THEY</i>	12	10	7

The results for generic *he* showed that eleven of the thirteen texts noted the traditional position of *he* as epicene pronoun, but importantly only one grammar specifically endorsed its use, and even then it was only a tentative endorsement as it was suggested that readers could “use *he* throughout and apologize to the reader” (Text 12). In fact, all of the texts which mention generic *he*, including Text 12, explicitly dismiss the pronoun, categorising it as traditional, older convention, or past usage, which is now “no longer considered acceptable” (Text 8). The results indicate a trend for modern grammars to move away from the traditional grammatical prescription of generic *he* found in previous studies of grammar texts (see section 2.3).

Interestingly, twelve of the thirteen texts acknowledge the use of singular *they* - one more than for generic *he* - with only Text 8 not mentioning *they* as an epicene pronoun at all. A total of ten grammars formally endorsed singular *they*, although there are limitations imposed on its usage. The vast majority of those grammars which acknowledge singular

they argue that it is either restricted to speech (Text 2; Text 3), or associate its use with indefinite pronouns (Text 4). However, the authors of Text 10 note that they use singular *they* for generic reference throughout their grammar, but they are in the minority.

Conversely, seven texts acknowledge but dismiss singular *they*, at least in formal written contexts, or “serious writing” as specified in Text 3. Yet interestingly, the majority of texts opposing the form blame the opinions of people external to the grammars. That is, the prescriptions against singular *they* are not necessarily attributed to the authors, instead they use passive sentences, such as “using the plural instead of the singular is usually considered <incorrect>” (Text 2), “it is often avoided” (Text 3), or “less widely accepted” (Text 4), which mean that the form is rejected without explicit agency. I expand on this finding in section 5.2.2, alongside my discussion of the keywords set out in research question three.

Based on the numerical figures only, singular *they* is the epicene of choice for the grammar corpus, being endorsed in 76.92% of cases (ten out of thirteen grammars). This is compared with only one endorsement for generic *he* (7.69%), which, as I have noted above, comes with a caveat. Equally, in terms of what pronouns were dismissed, singular *they* is again favoured, with only seven rejections (53.85%) compared with eleven rejections for generic *he* (84.62%). It is therefore clear from Table 5.1 that the results from this particular sample of grammars go against the grain of previous similar research (section 2.3.1), such as Bodine (1975), whose study showed a preference for the endorsement of generic *he*. The relatively equal mentions of both generic *he* and singular *they* in the modern grammar corpus indicate that both pronominal forms are contenders for epicene status, thus indicating that these two pronouns are still the two main epicene candidates in British English, and continuing the trend found in previous research (c.f. section 1.2.2, section 2.3).

However, the results also indicate that the rejection of one particular pronoun does not necessarily entail the endorsement of another. My review of the literature on epicene prescriptions indicates that other approaches, such as pluralisation, using combined pronouns, or recasting are all alternatives to prescribing singular *they* having proscribed generic *he*. In light of this acknowledgement, I also documented how many of the texts in

the grammar corpus endorsed alternatives to generic *he* and singular *they*, focusing mainly on whether or not the grammars advised the avoidance tactics noted by Treichler and Frank (1989) and Pauwels (1998), discussed in section 2.3.2.

There were four different avoidance tactics present in the grammar corpus. These were pluralisation, combined pronouns, generic *she*¹²², and reworking/recasting the sentence in some way to remove the need for a third-person pronoun. These tactics occurred twenty-two times across the thirteen texts (as in some cases more than one alternative was offered, perhaps suggesting uncertainty on the part of the authors). The use of avoidance tactics were distributed as in Table 5.2.

TABLE 5.2: ALTERNATIVES TO EPICENES

	PLURALISATION	COMBINED PRONOUNS	GENERIC <i>SHE</i>	REWORK SENTENCE	TOTAL
No.	6	10	2	4	22
%	27.27	45.45	9.10	18.18	

Combined pronouns such as *he or she* or *s/he* were mentioned in ten of the thirteen texts (76.92%) and account for almost half of the epicene avoidance tactics in the corpus. Interestingly however, such forms are not always endorsed, with the authors of Text 4 claiming that “He or she can become annoying” and “(s)he is ugly”. Similarly such combined constructions are also labelled as “clunky” (Text 19), “long-winded and clumsy” (Text 12), “tedious” (Text 15), and “cumbersome” (Text 16). In much the same way, although generic *she* is mentioned twice, it is not positively reviewed. In Text 4 the authors state that “using *she* and *her* all the time may prove some sort of point but can be very confusing”, whilst in Text 15 the form is described as “the opposite extreme”. Therefore, out of the twenty-two proposed avoidance tactics, twelve are treated negatively, again illustrating that the epicene debate is a point of contention in English grammar.

However, both pluralisation and recasting, which account for 45.45% of the alternatives to epicenes (occurring in six and four texts respectively), tend towards more positive reviews. Instructions to recast a sentence are generally written as imperatives, such as “avoid the phrase altogether”

122 Although this is technically an epicene endorsement, the treatment of generic *she* is limited, and predominantly negative (as is discussed below).

(Text 4) and “turn the sentence into a passive” (Text 12), whilst in Text 11 pluralisation is classified as “the easiest” way to avoid using a singular pronoun. The authors of Text 1 explicitly note that recasting allows a writer to be “grammatically correct and semantically accurate without violating traditional conventions of good use or using sexist forms of expression”. However, even though this tactic is classed as “acceptable and neat” in Text 12, the authors note that it is “sometimes not possible”.

What these alternative approaches to epicene reference indicate is that, even though most of the texts in the corpus endorse either singular *they* or generic *he* (at least in some limited capacity), there is clearly still controversy surrounding epicene prescriptions in modern grammar handbooks. The high rate of dismissal of the two main epicene candidates, supported by advice to pluralise or recast sentences, suggests that there is no consensus over which pronoun is the standard epicene of British English (both in modern grammar books, and arguably in a more general sense too). However, the data presented above cannot necessarily indicate how epicene prescriptions interact with epicene usage, and thus, in the next section I analyse the antecedents of generic *he* and singular *they* in any examples or illustrative sentences in the grammar corpus in order to facilitate a comparison with the BE06 data.

5.1.2: ANTECEDENT ANALYSIS

Each of the thirteen texts in the grammar corpus that included a discussion of epicene pronouns also included illustrative examples, or example sentences, presenting the authors’ points on epicene usage. In order to analyse the example sentences separately from the rest of the grammar corpus, I extracted them using manual analysis, reading every text and documenting each sentence that illustrated the use of an epicene pronoun. All of the antecedents in the example sentences were animate, but I excluded sentences with plural antecedents, which accounted for fourteen out of seventy-one sentences. I also eliminated sentences where the pronoun did not refer back to a particular antecedent, so called ‘institutional’ examples (see Sanford et al. 2007), such as “He who hesitates

is lost” (Text 11). Once these sentences were eliminated there were 56 eligible sentences left in the grammar corpus.

Having compiled a list of example sentences I then codified the antecedents used with singular *they* and generic *he* using the same syntactic criteria as in chapter four. Antecedents were thus classified as indefinite pronouns, NPs with quantifiers, indefinite NPs, or definite NPs. Table 5.3 shows the distribution of antecedent forms across the example sentences in the grammar corpus, delineated by text. It also shows which pronouns the antecedents co-occurred with. In addition, as my initial analysis had indicated that avoidance tactics were popular in the grammar corpus, I also provide the data for sentences including generic *she* and combined pronouns.

TABLE 5.3: PRONOUN CHOICE AND ANTECEDENT CHOICE IN THE GRAMMAR CORPUS EXAMPLES

TEXT	PRONOUN				CLASSIFICATION			
	HE	SHE	THEY	COMBINED	INDEFINITE PRONOUN	NP WITH QUANTIFIER	INDEFINITE NP	DEFINITE NP
TEXT 01	1	1	7	7	13	2	1	0
TEXT 02	1	0	0	0	1	0	0	0
TEXT 03	1	0	3	3	3	2	0	2
TEXT 04	0	0	3	0	1	1	1	0
TEXT 07	0	0	1	0	1	0	0	0
TEXT 08	1	0	0	0	0	1	0	0
TEXT 10	1	0	1	1	1	1	1	0
TEXT 11	6	0	1	2	2	5	1	1
TEXT 12	1	1	1	1	0	0	0	4
TEXT 15	0	0	1	0	0	0	0	1
TEXT 16	0	0	0	1	1	0	0	0
TEXT 17	1	0	7	0	7	1	1	0
TEXT 19	0	0	1	0	1	0	0	0
TOTAL	13	2	26	15	31	13	5	8

Looking at the totals, the lowest number of example sentences included generic *she*, a figure that corresponds well with the limited discussion of this epicene alternative in the grammar corpus (as shown in Table 5.2). Interestingly however, generic *he* is in third place, as there were more combined pronouns used in the examples than the traditionally-endorsed form. This does not support the dominance of generic *he* found in previous grammar book studies on epicene pronouns (section 2.3.1; section 2.3.2) and is yet more evidence that there has been a move away from the

form's endorsement. Conversely, singular *they* came out on top, with twenty-six occurrences (46.43%) in the example sentences; double the amount for generic *he*.

However, this figure is problematic for two reasons, and the dominance of singular *they* may not be as clear-cut as it seems. Firstly, there is no indication whether the sample sentences were used to support the use of the pronoun as a standard English epicene, or whether examples were used merely to be illustrative, and then to be dismissed by the grammar authors. Secondly, there may be interference from the types of antecedents chosen by the grammar authors for their example sentences. There were only five indefinite NPs and a further eight definite NPs in the example sentences. NPs with quantifiers were relatively well represented, with thirteen sentences, but most common, with 31 sentences (55.36%), are indefinite pronouns. In order to see whether or not this dominance of indefinite pronouns affected the authors' choice in epicene pronoun, I eliminated the antecedents of combined pronouns, and of generic *she* in order to compare the different antecedents of generic *he* and singular *they* only. This left me with 39 example sentences, with antecedents distributed as in Table 5.4.

TABLE 5.4: SYNTACTIC ANTECEDENT CLASSIFICATION FOR GENERIC *HE* AND SINGULAR *THEY* IN THE GRAMMAR CORPUS EXAMPLES

	INDEFINITE PRONOUN	NP WITH QUANTIFIER	INDEFINITE NP	DEFINITE NP	TOTAL
GENERIC <i>HE</i>	3 (23.08%)	6 (46.15%)	2 (15.38%)	2 (15.38%)	13
SINGULAR <i>THEY</i>	19 (73.08%)	2 (7.69%)	3 (11.54%)	2 (7.69%)	26

The table shows again that there were double the number of occurrences of singular *they* in example sentences in the grammars, compared with the number of examples including generic *he*. Most of the example sentences, 22 to be exact (56.41%), which included either generic *he* or singular *they* had indefinite pronouns for antecedents. The other syntactic categories were much less well represented; NPs with quantifiers accounted for only eight examples, indefinite NPs for five examples, and definite NPs for only four examples. These figures suggest that the grammar authors are avoiding discussing epicene pronouns with definite referents, preferring instead to use examples including indefinite pronouns. This is

most likely for two reasons; firstly, indefinite pronouns are not marked for gender, nor do they carry any stereotypical gender (section 5.2.3), and secondly, the indefinite pronouns seem to be the exception to the rule for number concord and have been accepted (with limitation) in other, older prescriptive grammars (see chapter two).

Table 5.4 also shows the differences in how the antecedent types are distributed between the two pronouns. Singular *they* is overwhelmingly more likely to occur with indefinite pronouns than generic *he*, by a ratio of 19:3 or 86.36% of the time. Across the rest of the antecedents the distribution between pronouns is much more balanced, although generic *he* does account for 75% of the NPs with quantifiers (even though there are only eight in total). In order to make the distribution of antecedent types clearer, the numbers are represented graphically in Figure 5.2.

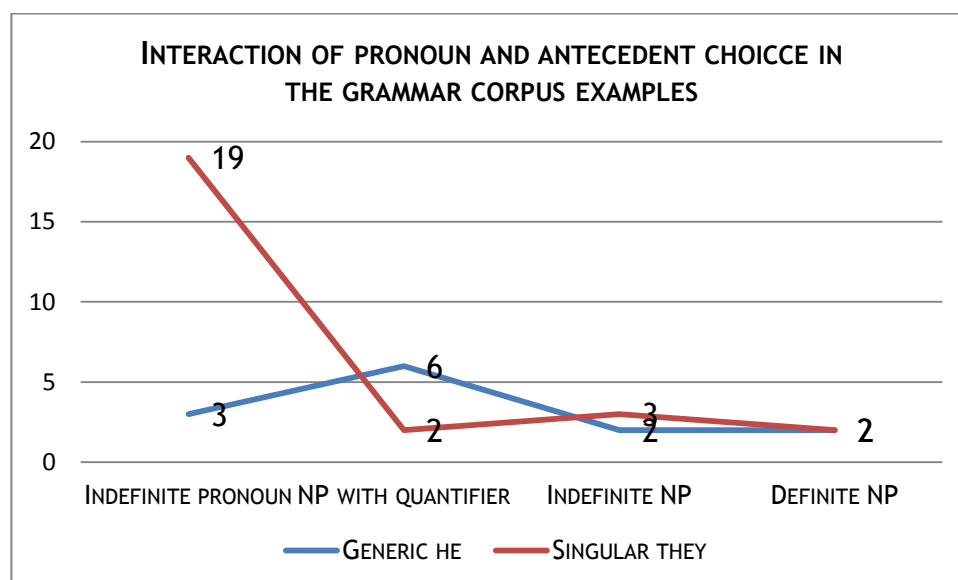


FIGURE 5.2: DISTRIBUTION OF GENERIC *HE* AND SINGULAR *THEY* ACROSS ANTECEDENT TYPES

The graph highlights the trend for indefinite pronouns to occur with singular *they*, whilst the pronoun use across the other syntactic categories is more comparable, if not extremely limited. Interestingly the pattern of generic *he* increasing with definiteness, as I found in the BE06 data, discussed in section 4.2.1, is not realised in these examples, as the pronoun is most popular with NPs with quantifiers. Also, the pattern whereby occurrences of singular *they* decrease as the definiteness of antecedents increases does not occur either, as all syntactic forms other than indefinite pronouns are equally represented, albeit rather poorly. These figures suggest that possibly, although singular *they* is endorsed in ten grammars

(see Table 5.1), the endorsements are restricted to example sentences including (rather unproblematic) indefinite pronouns.

A closer look at the antecedents of generic *he* and singular *they* (as shown in Table 5.5) shows that, whilst generic *he* only occurs with two different indefinite pronouns, singular *they* occurs with the whole range, including *somebody* and *no one*. One interpretation of this difference is that the grammar authors saw singular *they* as the epicene of choice for indefinite pronouns, insofar as the range of indefinite pronouns in the example sentences illustrate the versatility of singular *they* with this particular antecedent type.

TABLE 5.5: ANTECEDENTS OF GENERIC *HE* AND SINGULAR *THEY* IN THE GRAMMAR CORPUS EXAMPLES

INDEFINITE PRONOUN	NP WITH QUANTIFIER	INDEFINITE NP	DEFINITE NP
ANTECEDENTS OF GENERIC <i>HE</i>			
ANYONE EVERYONE X 2	ANY RUNNER EACH APPLICANT EACH NOVELIST EACH PASSENGER EACH STUDENT EACH TENANT	A DOCTOR A PUPIL	THE APPLICANT YOUR EMPLOYEE
ANTECEDENTS OF SINGULAR <i>THEY</i>			
ANYBODY ANYONE X 2 EVERYBODY EVERYONE X 8 NO ONE X 2 NOBODY SOMEBODY X 2 SOMEONE X 2	EACH [PERSON] EACH PASSENGER	A FRIEND A PERSON X 2	THE JUDGE YOUR EMPLOYEE

Nevertheless, the distribution of antecedents in the example sentences of the grammar corpus does not reflect their distribution in wider epicene usage, in relation to the data I presented in section 4.2.1. Thus, on this measure of analysis at least, the grammars do not reflect language use. In the 39 example sentences which included singular *they* or generic *he*, only 22 different antecedents were used, as 17 antecedents, many of them indefinite pronouns, were repeated throughout the corpus. These figures offer further evidence that the grammar authors tend to use unproblematic examples in their texts, in order to avoid a more-detailed discussion of epicene reference.

Table 5.5 also shows that many antecedents were used with both generic *he* and singular *they*, but unfortunately, it was uncommon to find the same antecedent used with both forms in a single text. There were only two instances where a particular antecedent was used with both singular *they* and generic *he*, including *each passenger* in Text 11. However, neither pronoun is endorsed as readers are told “it is best not to say” the example including generic *he*, whilst singular *they* “is not recommended”. Thus although both forms are illustrated using the same antecedent, neither is presented as the “correct” form, and the authors of Text 11 instead support the notion that “It is far better to rephrase your statements completely”. In Text 12 the opposite occurs with *your employee*, as both the sentence including generic *he* and the corresponding sentence with singular *they* receive limited endorsement. Readers are advised to “Use he throughout and apologise to the reader”, but told that “Many people find this approach unacceptable”, whilst on the other hand, readers are told that they “is increasingly used, but traditionalists disapprove” and as such the form “should be avoided in more formal writing”.

What these two examples serve to illustrate is that, across the grammar corpus, it is rare to find explicit endorsement for either generic *he* or singular *they*, and to that end, very few of the grammars overtly prescribe a particular usage. Thus, in contradiction to previous research on similar texts (section 2.3.1 and section 2.3.2), there appears to be an air of caution in the sections on epicene prescriptions in modern grammar texts. My antecedent analysis also indicated that the coreferents the grammar authors selected for illustrative sentences including singular *they* or generic *he* were not representative of the figures for epicene usage I presented in the previous chapter. In the section below I expand on this finding, looking at the gender-stereotype values for the antecedents in the grammar corpus, in order to compare them with their BE06 counterparts.

5.1.3: DEALING WITH STEREOTYPES

Using the same method of stereotype analysis I used to analyse the antecedents in the two sub-corpora of BE06 (section 4.2.3) I cross-

referenced the stereotype data provided in Kennison and Trofe (2003) and Carreiras et al. (1996), discussed in section 3.2.2, with the antecedents used in the example sentences in the grammar corpus. The criteria for judging whether an antecedent was gender stereotyped were as follows: in the Kennison and Trofe data, values below 3.0 were feminine and above 5.0 were masculine, whilst in the data provided by Carreiras et al., the higher the value, the greater the feminine stereotype. As before, I eliminated the indefinite pronouns from further analysis, as these were not tested in either of the previous studies. This process decreased the number of antecedents dramatically, as 22 out of 39 example sentences used indefinite pronouns. Nevertheless, after I also eliminated the quantifiers/articles pre-modifying the antecedents, there were still fourteen usable tokens for analysis.

Although using the same method of analysis made the grammar corpus data directly comparable to the data from the sub-corpora of BE06, the results for the grammar corpus are extremely limited. As shown in Table 5.6 and Table 5.7 only five antecedents in the grammar corpus data were listed in the stereotype data provided by Kennison and Trofe and/or Carreiras et al. Only two of the testable antecedents occur in the BE06 data, *doctor* and *person* (both of which occur with generic *he*; section 4.2.3), and thus there is very little overlap between the two datasets.

TABLE 5.6: STEREOTYPES WITH GENERIC *HE*

ANTECEDENT	SOURCE	TESTED FORM	RESULTS	STEREOTYPE
DOCTOR	KENNISON AND TROFE	DOCTOR	FEMALE: 4.65 MALE: 4.60	NEUTRAL
NOVELIST	KENNISON AND TROFE	NOVELIST	FEMALE: 3.75 MALE: 3.65	NEUTRAL
	CARREIRAS ET AL.	NOVELIST	6.03	
STUDENT	CARREIRAS ET AL.	STUDENT	5.60	NEUTRAL
<u>ANTECEDENTS NOT TESTED:</u> APPLICANT, EMPLOYEE, PASSENGER, PUPIL, RUNNER, TENANT				

TABLE 5.7: STEREOTYPES WITH SINGULAR *THEY*

ANTECEDENT	SOURCE	TESTED FORM	RESULTS	STEREOTYPE
JUDGE	KENNISON AND TROFE	JUDGE	FEMALE: 4.75 MALE: 4.65	NEUTRAL
PERSON	KENNISON AND TROFE	PERSON	FEMALE: 4.00 MALE: 4.10	NEUTRAL
<u>ANTECEDENTS NOT TESTED:</u> EMPLOYEE, FRIEND, PASSENGER				

None of the antecedents of either generic *he* or singular *they* in the grammar corpus were stereotypically marked for gender, with average

(neutral) values of 4.16 for generic *he* and 4.38 for singular *they* (based on the figures from Kennison and Trofe). Interestingly, in contrast with the BE06 data, generic *he* is slightly more gender-neutral than singular *they*, although it is impossible to gauge the significance of this result as the stereotype data for the grammar corpus is so limited¹²³. I argue that the stereotype data provides more evidence that the grammar authors simplify their discussions of epicene reference, cherry picking their examples and avoiding heavily gender-stereotyped antecedents.

In this section I have shown that, in the majority of cases the grammar authors both endorsed and dismissed singular *they*, with only five of the twelve texts supporting the form without limitations, whilst only one text (Text 12) rejected the form without limited acceptance. Similarly for generic *he* none of the texts explicitly endorsed the form without censure, whilst ten rejected it completely. These results indicate that the figures in Table 5.1 simplify how the pronouns are treated within the grammars. If neither form is formally endorsed without caveat, the authors must give their readers some alternative advice, and indeed I showed that epicene avoidance tactics predominate in the grammar corpus, with 22 occurrences across the thirteen texts.

Following on from my initial analyses of the texts I also looked at the types of antecedents used in the grammar corpus to illustrate epicene reference and found that indefinite pronouns were overrepresented when compared to the relative usage values presented in chapter four. In addition, I also showed that the grammar authors avoid choosing heavily gender-stereotyped antecedents for their example sentences. I expand on this finding in section 5.2.1 where I discuss the grammar authors' simplification of the epicene debate. My results indicate that my hypothesis that current epicene prescriptions will continue the trend in endorsing avoidance tactics, furthering the findings of Zuber and Reed 1993, does hold. However, the overwhelming result from this first section is that the data in the modern grammar corpus does not correspond to the results of previous studies on corpora of grammar texts (such as Sunderland 1986; Bodine 1975), and

123 However, the average values for both generic *he* and singular *they* are comparable with their corresponding figures of 4.69 and 4.15 in the BE06 data (section 4.2.3).

singular *they* is much more prominent in this data set of modern grammars when compared with previous works. In the next section I investigate whether there is any explicit mention of the reasons why epicene prescriptions appear to have changed.

5.2: SOCIAL INTERFERENCE?

In order to determine whether the reasons for such diachronic changes in prescriptions are explicitly mentioned in the grammar corpus I uploaded the texts to Sketch Engine (Kilgarriff, Rychly, Smrz and Tugwell 2004), an online corpus-analysis tool where users can upload their own data and construct their own corpora. I used this software to generate the Corpus Comparative Statistical Keywords for the grammar corpus (section 5.2.1), comparing it with the most recent available reference corpus of British English, the BNC, in order to determine if the occurrence of any of the keywords set out in research question three were statistically significant. Following on from this initial quantitative analysis I searched for the keywords directly, taking a closer look at their concordance lines and collocates in section 5.2.2.

5.2.1: CORPUS COMPARATIVE STATISTICAL KEYWORDS

In order to investigate the possible reasons for changes in epicene prescriptions between previous grammar book studies and the present corpus, and in order to highlight whether such changes are acknowledged within the grammars themselves, I used Corpus Comparative Statistical Keyword (CCSK) analysis. This process involved comparing the statistical frequencies of words in the grammar corpus with their relative counterparts in a larger reference corpus of British English. Following O'Halloran (2009:35) CCSKs are words that are “statistically more frequent in a text or set of texts than in a large reference corpus”, and can provide “a rough snapshot

of salient topics in a corpus” (2009:29)¹²⁴. However, Tyrkkö (2010:80) notes that statistical keywords in a corpus may not correlate with “items perceived by human readers as key”, that is, the CCSKs may not reflect the keywords set out in research question three. As is shown below, this distinction between statistical significance and topic-pertinence is evident in the grammar corpus.

Ideally, the grammar corpus, which is arguably representative of grammars published between 2000 and 2010, would be compared with the BE06 data representative of British English circa 2006, but the resources needed to do this were not available. My access rights to the BE06 data allowed me to query the corpus only, and I was not able to move the sub-corpora over to the Sketch Engine software I used to analyse the grammar corpus (nor could I do vice versa). Therefore, it was not possible to use the BE06 corpus as the standard reference corpus for creating a list of CCSKs for the grammar corpus. I thus had to choose the most suitable corpus available to me using the Sketch Engine programme. This was the British National Corpus (BNC), which includes examples of British English taken from the 1990s.

The BNC is a standard reference corpus of British English and to that extent it is comparable with the texts in the grammar corpus¹²⁵. Using the Sketch Engine software, I searched for the top CCSKs in the grammar corpus. This process highlighted the words that were used with greater frequency in the grammar corpus than they were in the reference corpus. As the grammar corpus is rather small, with a total of only 3874 words, the results were grouped into lemmas and the search criteria stipulated that forms had to occur with a minimum frequency of five words per million. The search returned fifty-one lemmas that were key in the grammar corpus, when compared to the frequency values for the BNC. Thirty-five were open class-words, such as *gender* or *usage* (Table 5.8), whilst ten were closed-class words (Table 5.9), which unsurprisingly included eight personal pronouns. There were also six erroneous results in the CCSK search, which related to punctuation and graphology, which I eliminated from further study.

124 Importantly for this research O’Halloran notes that “corpus-comparative statistical keywords can be both lexical and grammatical words” (2009:35), and the two word classes are separated accordingly below.

125 Unfortunately, it was not possible to eliminate the spoken data of the BNC in this query.

TABLE 5.8: OPEN-CLASS CCSK

	LEMMA	CCSK SCORE		LEMMA	CCSK SCORE
1	PRONOUN	4906.7	11	MALE	44.8
2	SINGULAR	2016.8	12	FEMALE	38.3
3	PLURAL	1666.3	13	AVOID	31.6
4	ENGLISH	1225.1	14	CORRECT	26.2
5	NOUN	684.6	15	INSTEAD	15.9
6	GENDER	316.7	16	PERSON	12.8
7	USAGE	199.8	17	USE	11.5
8	SENTENCE	86.2	18	WORD	9.3
9	REFER	61.6	19	PROBLEM	6.8
10	SEX	46.7	20	EXAMPLE	6.6

Beginning with the top twenty open-class words in Table 5.8, it is no surprise that the highest scoring CCSK is *pronoun*, followed by *singular* and *plural*, then *English*, and *noun*. All of the top five words are directly connected to metalanguage, which is of course, a key topic of the grammar corpus, and therefore the occurrence of such terms is to be expected. Similarly, it is not particularly surprising to find *gender*, *sentence*, and *word* in texts concerned with grammar, and I argue that the occurrence of such terms in the grammar corpus is not exclusively related to discussions on epicene pronouns.

What is interesting is that terms such as *avoid*, *correct*, *instead*, and *problem* occur in the top twenty open-class CCSKs. Using my evaluation of the literature in section 2.1.1 as a foundation, I argue that terms like these are negatively loaded when applied to grammatical issues. For example, using the term *correct* implies that other forms are *incorrect*, whilst the term *avoid* may suggest that other alternative forms are more highly valued. Arguably, these terms illustrate the use of a prescriptive view of grammar in the corpus (see chapter two), as they highlight the view that certain forms of language are somehow more acceptable than others.

However, the data for the open-class words is largely unremarkable, as the CCSKs contain terms that one would expect to occur in grammar books. Interestingly, in relation to research question three, none of the keywords identified, including *sexist language*, *political correctness*, etc., appeared in the CCSK list. However, although this quantitative analysis does not highlight such terms, this does not mean that they are not significant at all, and in section 5.2.2 I look at the keywords from research question three in much

more detail. In any case, the open-class CCSK do not really tell us much about epicene prescriptions in the grammar corpus, but there is more to say about the closed-class items, which I deal with in Table 5.9.

TABLE 5.9: CLOSED-CLASS CCSK

	LEMMA ¹²⁶	CCSK SCORE		LEMMA	CCSK SCORE
1	THEM	2851.1	6	SHE	3.8
2	EVERYONE	2392.7	7	THEIR	3.2
3	HIM	2012.6	8	HE	2.4
4	HER	6.0	9	HIS	2.4
5	OR	4.6	10	YOUR	2.2

The most important results are the top three, with extremely high CCSK scores of over two thousand. It is not unexpected that two of these top three words are instances of the two pronouns under scrutiny in this thesis. These results clearly show that, although *she* is also in Table 5.9, both *he* and *they* are the most discussed pronoun forms in the grammar corpus. In addition, as the discussion in section 5.1.2 showed, it is not surprising that the indefinite pronoun *everyone* is highly significant (even though there were only eight raw tokens in the grammar corpus).

The most significant closed-class word in the CCSK analysis was *them*. Although it is not possible to delineate between plural and singular uses of the form in the corpus, its high occurrence rate illustrates that the pronoun is discussed within the grammar corpus with a relatively high frequency. The high statistical significance of *them* is interesting for two reasons. Firstly, *them* is the accusative form of the pronoun, whilst the nominative does not occur in Table 5.5. I argue that this result highlights a flaw in the CCSK method in relation to epicene research. In section 4.1.1 I showed that there were more tokens of the lemma *they* than any other case form of the pronoun across both of the BE06 sub-corpora, and thus it was more frequent in language use than other case forms. Therefore, if the distribution of case forms in the BE06 data holds across the BNC data, and there is no argument in the literature suggesting that it should not, for *they* to be a CCSK it would have to occur much more often than other forms, including *them*. Thus whilst the relative frequencies of *they* and *them* may have been similar in

126 It was not possible to conflate case forms using the Sketch Engine software, and this leads to some anomalies in the CCSK results (see below).

the grammar corpus, only one form is deemed to be statistically significant in relation to the reference corpus, yet this does not mean that the form was any more salient to the epicene debate.

Secondly, the occurrence of *them* is more significant than *him*, which comes in third place in Table 5.5. This pattern is consistent throughout the table, as *their* comes ahead of both *he* and *his* in the CCSK results. Whilst these figures may look like support for singular *they* as the epicene pronoun of choice in the grammar corpus, I argue that the results are skewed by the fact that masculine pronouns tend to be more common than forms of *they* in general. This argument is reinforced by the fact that in both BE06 sub-corpora (section 4.1) there were more concordance lines for *he* than *they*, and Baker's (2009) analysis of the whole BE06 corpus shows that *he* is more frequent than other personal pronouns. Therefore, as with the discussion of case forms, above, for the occurrence of *he* to be statistically significant in the grammar corpus, it would have to occur many more times than other pronouns.

Whilst it is clear that there are problems with using CCSK to investigate epicene prescriptions, what is important for this study, is that, knowing the relatively high frequency of *he* in general, it is interesting that the nominative form of *he* occurs in Table 5.5, although its counterpart *they* does not. Potentially then, there is an argument that the use of *he* is more significant than *they* in the grammar corpus, and in terms of raw figures (45 for *they* compared with 69 for *he*) this appears to be true. However, the CCSK data for the closed-class words is limited, as it is greatly affected by the overall high frequencies characteristic of closed-class words (section 1.1.1) and thus any conclusions based on Table 5.5 must be tentative.

5.2.2: KEYWORDS IN THE CORPUS

Having highlighted the problems with the CCSK approach to the grammar corpus data, I move on to a closer analysis of the keywords set out in research question three, in order to investigate whether possible reasons for the lack of overt endorsement of either singular *they* or generic *he* are acknowledged within the grammars. I generated concordance lines and

frequency counts for the keywords, searching the grammar corpus for each of the following key terms (discussed in section 3.2.2), *political correctness*, *politically correct*, *sexist language*, *non-sexist language*, *traditional*, *out-dated*, *feminism*, and *feminist*. I again used the Sketch Engine software and searched the grammar corpus using the following set of lemma queries: [political.*]; [.*sexist]; [traditional]; [out dated]; [feminis.*]. As is shown in Table 5.10 below, the terms *politically correct* and *out dated* did not occur, although the other keywords all occurred in a limited capacity.

TABLE 5.10: KEYWORDS IN THE GRAMMAR CORPUS

SEARCH TERMS	OCCURRENCES
POLITICAL CORRECTNESS	1
POLITICALLY CORRECT	0
SEXIST LANGUAGE	2
NON-SEXIST LANGUAGE	1 (NON-SEXIST CONSTRUCTION)
TRADITIONAL	5
OUT DATED	0
FEMINISM	1
FEMINIST	2

The raw data does not initially look significant, with only twelve hits across the whole corpus, and many keywords only mentioned once. However, I rearranged the results into four larger categories (as shown in Table 5.11), grouping together terms such as *sexist language* and *non-sexist construction*, as they clearly referred to the same overarching topic. This second representation of the data makes it clearer that the keywords from research question three do play a role in the grammar corpus.

TABLE 5.11: GROUPED KEYWORDS IN THE GRAMMAR CORPUS

GROUPED TERMS	OCCURRENCES	TEXT(S) WHERE KEYWORDS OCCURRED
TRADITIONAL	5	TEXT 1, TEXT 3, TEXT 4, TEXT 10 (2)
FEMINISM	3	TEXT 15, TEXT 17 (2)
SEXIST LANGUAGE	3	TEXT 11 (3)
POLITICAL CORRECTNESS	1	TEXT 8

Not only does Table 5.11 give the overall figures for the occurrence of the keywords, I have also listed the different texts in which the words occurred in order to illustrate that the keyword tokens were relatively well distributed throughout the corpus. Only five of the thirteen texts did not include any of the keywords, and, perhaps more importantly, none of the grammars included more than one of the keywords (although some of them,

such as Text 17, included multiple occurrences of one particular keyword). The significance of the fact that no grammars mentioned more than one keyword is discussed below in relation to the agency attributed to the rejections of generic *he*. However, the data in Table 5.11 alone cannot shed light on how significant the occurrence of the keywords may be.

TABLE 5.12: KEY TERMS IN THE BE06 SUB-CORPORA

	PRONOUNS_PRESS	PRONOUNS_GENERAL_PROSE
POLITICAL CORRECTNESS	0	1
POLITICALLY CORRECT	2	2
SEXIST LANGUAGE	0	0
NON-SEXIST LANGUAGE	0	0
TRADITIONAL	9	60
OUT DATED	0	0
FEMINISM	0	7
FEMINIST	0	10

Therefore, in order to see how important the overt mentions of the keywords are I also conducted the same search in the two sub-corpora of the BE06 corpus. In this instance I used the CQP software (Hardie *forthcoming*) and the following set of queries: [political* correct*]; [*sexist language]; [traditional]; [out dated]; [feminis*]. The results, as depicted in Table 5.12, showed that the terms occurred more frequently in the BE06 data than in the grammar corpus, and this is most likely due to the comparable sizes of the two resources. However, interestingly, the majority of these terms did not occur in pronouns_press, and the limited occurrences of *politically correct* and *traditional* did not relate to discussions about grammar in any way. Similarly, in pronouns_general_prose five of the search terms occurred, but none of the instances were in reference to epicenes or grammar. Therefore, in this sense there is no overt interaction with the use of singular *they* or generic *he* with mentions of language-external elements of the epicene debate in the BE06 sub-corpora.

Thus, comparing the data from the grammar corpus, and treating the sub-corpora of BE06 as a reference for standard written British English, the occurrence of twelve keywords across only 3874 words (when compared with the total word count for the BE06 sub-corpora) is significant. Such a high concentration of keywords in the grammar corpus indicates that the grammar authors are at least aware that there is some interaction between the nuances of language they are discussing, and wider social debates on the

inclusivity of language. This awareness is perhaps best highlighted by the five occurrences of the term *traditional*, which are used in relation to generic *he*. The pronoun is also referred to as “older conventional” (Text 1) and “used to be correct” (Text 8), which suggests an awareness that generic *he* has a historical pedigree within language prescriptions, but that it has been cast aside by arguments concerned with non-sexist language.

However, although the numerical data for the occurrence of the keywords can indicate that certain terms were more likely to occur in the grammar corpus than in other reference corpora, such as the BE06, the numbers alone are limited. Therefore, in Table 5.13 I list the KWIC concordance lines for each of the twelve instances of the keywords in the grammar corpus, in order to take a more qualitative look at the data. I compiled the list of concordance lines using the Sketch Engine software, expanding the number of words to each side of the node in order to make the context clear. Table 5.13 also has three extra concordance lines than the original twelve occurrences of the keywords because I have added the search results for *sexist* (without language) as they are clearly relevant to this discussion.

Starting at the top of Table 5.13 (overleaf) it is clear that the use of the term *traditional* always refers to generic *he*, and the authors of the third concordance line (Text 4) explicitly note their awareness of issues over the applicability of generic *he* to both sexes. Although these are the only authors to explicitly note the historical development of generic *he*, I argue that most of the references to *traditional* are implicitly negative, especially when juxtaposed with the other concordances in Table 5.13.

Moving on to the second section of Table 5.13, the authors of the three concordance lines in the *feminism* category all attribute the opinion that generic *he* is not gender-neutral to (undefined groups of) feminists. I argue that this explicit mention of feminists, both selects an agent for those in opposition to the *traditional* approach, whilst equating those with a particular political view with people who are anti-generic *he*. What is interesting though, is that none of the concordances in the *sexist language* category (the third section in Table 5.13) explicitly mention feminists. Instead, those opposing generic *he* on sexist grounds are referred to in the grammars only as “many people” (Text 11) or “many writers and speakers”

TABLE 5.13: CONCORDANCE LINES FOR KEYWORDS IN THE GRAMMAR CORPUS

TRADITIONAL	
he/him and she/her are only used to refer to humans or animals. In anyone who believes he has been injured by a senator 's words can ask to have a response written into the Senate 's records. (In the late 20th century, as the	formal usage, he/him may occur with reference to both sexes . formal usage). Any student who thinks(s) he has been unfairly treated can appeal.
When there is no need to refer to 'male or female', the	use of he to refer to a person of either sex came under scrutiny choice, again, is in favour of the masculine: using he to refer to both sexes
be grammatically correct and semantically accurate without violating	conventions of good usage or using sexist forms of expression .
FEMINISM	
gender-neutral singular pronoun that includes both men and women, an omission that hasn't gone unnoticed by the	movement. Feminists argue that current usage
omission that hasn't gone unnoticed by the feminist movement .	argue that current usage is not only biased against women but is also illogical .
'the masculine was deemed to embrace the feminine ', but 1960 s	put a stop to that sort of hanky-panky and has left us with a grammatical problem
SEXIST LANGUAGE	
he was an acceptable pronoun for both sexes, but nowadays many people label this as	. So now it is best not to say: Each passenger must ensure that he has
. Pronoun gender ,	and generic 'he' he/she his/hers him/her himself/herself
the judge (who may be a man or a woman , we don't know and it would be	to assume either) adjusted their wig.
many writers and speakers wish to avoid the possibly	practice of using a masculine pronoun to refer to girls and women
you can be grammatically correct and semantically accurate without violating traditional conventions of good usage or using	forms of expresion
or they use a rather clumsy but	which is now widespread : Passengers must ensure that they have
POLITICAL CORRECTNESS	
it used to be correct to use the pronoun he or his when referring to the person . In today 's world of increasing	and sexual equality , this is no longer considered acceptable

(Text 1). Therefore, the agency of opposition linked to feminism is not carried through to mentions of sexist language. My interpretation of the effect of this is that, whilst the grammar authors do not align themselves with the feminist movement, which is treated separately from the issue of sexist language, they appear to suggest a general consensus that generic *he* has been rejected by generic “people” without overt political opposition to the form.

Indeed, the final concordance, for which the node is *political correctness*, shows that this lack of agency is taken one step further. There is no specific opponent to the use of generic *he*, as rejection of the form, which is deemed to be “no longer considered acceptable” (Text 8), is a passive construction. The significance of these different levels of agency links back to the fact that none of the grammars included more than one of the keywords¹²⁷. Therefore, in each of the eight texts that includes a keyword the authors take one of the following stances on generic *he*:

- a) the pronoun is classed as *traditional* and generally reviewed negatively, or
- b) it is opposed based on arguments associated with feminism and feminists, or
- c) it is rejected on the grounds of sexism in terms of general agency, through references to generic groups of “people”,
- d) or finally, it is rejected under consideration of political correctness, using a passive construction, and thus eliminating the agency of those opposing the form.

What is surprising is that singular *they* is not mentioned in any of the concordance lines in Table 5.13 and thus does not occur in close proximity with any of the keywords listed as part of research question three. This illustrates that whilst the rejection of generic *he* is generally associated with language-external forces born out of second-wave feminism, the corresponding promotion of singular *they* does not have such links. Indeed, as I noted in section 5.1.1 the proscription of generic *he* does not necessarily entail the prescription of singular *they*, and the discussions of modern

127 There is a small amount of overlap when the results for *sexist* are included, as one of the concordance lines is taken from Text 1, which includes the term *traditional*.

approaches to epicenes in section 2.3 indicated that there is still resistance to the form.

In order to take a closer look at the terms used in relation to singular *they* I used the Sketch Engine software to search the grammar corpus for the term [singular *they*] and significantly, there were no hits, meaning that the term itself never occurs in the grammar corpus. Similarly, I ran the same query for the term [generic *he*] and this term does not occur either. On finding that the grammar authors do not specifically use the terms most associated with these two epicene candidates (c.f. chapters one and two) I decided to investigate the wider collocates of *they* and *he*. Using Sketch Engine I searched for [they] and [he] and created a list of collocates occurring plus or minus ten words of the node. That is, I had the software generate a list of all lemmas occurring within ten words either side of the target form *they* or *he*, which I then organised by frequency. I then eliminated erroneous results involving punctuation, and other closed-class words that were not personal pronouns. The results are presented in Table 5.14¹²⁸.

TABLE 5.14: TOP TEN COLLOCATES OF *THEY* AND *HE*

COLLOCATES OF <i>THEY</i>			COLLOCATES OF <i>HE</i>		
	LEMMA	OCCURRENCES		LEMMA	OCCURRENCES
1	THEIR	13	1	SHE	39
2	SINGULAR	10	2	HIS	18
3	HE	10	3	USE	17
4	THEM	9	4	HIM	11
5	PLURAL	8	5	PRONOUN	11
6	SHE	8	6	THEY	11
7	PRONOUN	7	7	HER	10
8	NOUN	6	8	NO	8
9	REFER	6	9	CAN	7
10	WORD	6	10	SEX	6

The frequent occurrence of *she* as a collocate of *he* perhaps indicates the prevalence of discussions of combined pronouns in the grammar corpus, as discussed in section 5.1.1, and the use of *plural* in relation to *they* can be related back to the association between singular *they* and number concord

128 The results listed in the table are based on the nominative case forms only. I did perform the analyses for other case forms, but the results were extremely similar and not noteworthy. The collocates common to both pronouns are highlighted in grey.

(section 2.1.2). Furthermore, the term *singular* does occur in close proximity to *they* ten times in the grammar corpus, suggesting that, although it is never its immediate collocate, the grammar authors note its use as a singular pronoun. However, overall there are no particular collocates of note, and perhaps the most significant finding is that none of the keywords from research question three occur frequently in close proximity to either *they* or *he*. Indeed, on analysis of the full lists of collocates, I found that *he* only collocated with *generic*, *masculine* and *gender* three times each, and neither *he* or *they* were within ten words either side of any keywords.

However, although numerically these results look poor, I return to the initial finding in this chapter: that the discussions of epicene pronouns in the grammar corpus were always delineated in some way from the rest of the text. I argue that this graphological marking (as illustrated in Figure 5.1) suggests that epicene pronouns are their own issue in English grammar, separate from more general discussions of pronouns, and thus the sections are to be taken as a whole. That is, it does not necessarily matter how close to *they* and *he* the keywords occur, they are arguably significant to the reader, and in this analysis, as long as they are in the same section of the grammar.

The final section of this chapter began by focusing on the Corpus Comparative Statistical Keywords for the grammar corpus, comparing it to the standard reference British National Corpus. My results indicated that the terms which were statistically significant did not match the keywords I set out in research question three. However, my CCSK results confirmed that *they* and *he* were the most frequent pronouns in the grammar corpus, and there were indications of both the prescriptive approach of the authors and the overrepresentation of indefinite pronouns identified in section 5.1.2. Yet it was clear that, due to interferences of general pronoun frequency, the CCSK method is not entirely suitable for analyses concerning epicene pronouns, and I thus decided to directly search the corpus for the keywords from research question three.

My results for this part of my analysis were much more fruitful, and I showed that, even though there were more tokens of the keywords in the sub-corpora of BE06, none of the hits were explicitly linked to discussions of

English grammar. Therefore, I argued that the occurrence of the keywords in the grammar corpus was significant, and my analysis of the concordance lines including the keywords indicated that modern prescriptions on epicene pronouns (more specifically the rejection of generic *he*) interacted with agency. I showed that the grammar authors took one of four approaches to generic *he* depending on whether they mentioned particular aspects of the epicene debate, such as *feminism*, *non-sexist language*, or *political correctness*.

Finally I took a closer look at the collocates of *they* and *he* in the grammar corpus, with the results showing that neither pronoun occurred within ten words either side of any keyword. Nevertheless, I argued that this finding was not problematic, as the graphological delineation of discussions on epicene pronouns in the grammars meant that the sections should be taken as wholes. Thus, the distance between the tokens of *they* and *he* and the tokens of the keywords was not too important; the significant finding was that the majority of the keywords occurred in the grammar corpus, and were thus used in discussions of epicene pronouns.

CONCLUSIONS

In response to research question two, I have shown that the discussion of epicene pronouns is relatively common to modern grammars, occurring in thirteen out of twenty texts, and importantly, my research indicates that the modern grammars I have analysed do reflect the language-external pressures on the epicene debate. The overwhelming rejection of generic *he*, combined with the predominance of avoidance tactics, and the explicit mention of feminism and non-sexist language, illustrates that, in line with my second hypothesis (c.f. section 3.2.1), modern language prescriptions have continued the trend evident in previous studies.

In section 2.3.2 I argued that there was support in the literature for the argument that the current most popular outlet for language prescriptions was school grammars and textbooks, following Mackay's (1980) claim that many such texts are still influenced by grammatical norms set out

in the eighteenth and nineteenth centuries. Previous studies on educational materials had shown the dominance of both the prescription, and use, of generic *he* (c.f. Graham 1973; Bodine 1975; Sunderland 1986). However slightly more recent work had shown that, although support for generic *he* was still present, grammars tended to endorse epicene avoidance tactics instead (see the discussion of Jochowitz 1982 and Treicher and Frank 1989 in section 2.3.2). This promotion of avoidance tactics was also evident in my grammar corpus, yet the primary candidate to replace either generic *he* or singular *they* was the use of combined pronouns, despite the fact that combined pronouns are relatively uncommon (see section 4.1.2).

However, what was not hypothesised based on previous scholarship is the increase in discussions of singular *they* in modern grammars. This finding illustrates that the lack of data on twenty-first century prescriptions evident in my review of the literature has concealed a development in the prescription of epicene pronouns. Based on the above data it is clear that, not only has there been a movement away from endorsing generic *he*, combined with discussions about epicene avoidance tactics, there has also been an increase in the consideration of singular *they*. For example:

- Only one text in the grammar corpus endorses the use of generic *he*, compared with ten texts (76.92%) for singular *they* (Table 5.1).
- Eleven texts (84.62%) dismiss generic *he* but only seven reject singular *they*, but this is still over half (53.85%).
- Generic *he* is only the third-most used pronoun in the example sentences behind singular *they*, which is the most popular (Table 5.3).
- The explicit endorsement of an epicene candidate is extremely rare (occurring only once in this corpus for singular *they*).

However, the rejection of generic *he* did not mandate the endorsement of singular *they*. Rather, epicene avoidance tactics predominate in the grammar corpus, with twenty-two occurrences across the thirteen texts. The fact that not all of these avoidance tactics are reviewed positively is further evidence that epicene pronouns are still a point of contention in English grammars.

In addition, I also highlighted several key findings that indicated that the grammar authors only present a simplified version of the epicene debate.

Significantly, the terms singular *they* and generic *he* did not occur in the grammar corpus, and the authors avoided areas of contention such as antecedent definiteness and gender stereotypes:

- Indefinite NPs are overrepresented as antecedents of epicene pronouns in the grammar corpus (Table 5.4) when compared with the usage data from the BE06 sub-corpora.
- Singular *they* is six times more likely to coindex with an indefinite pronouns than generic *he* (Figure 5.2).
- None of the antecedents in the example sentences are gender-stereotyped.

My research also showed that although six out of the eight keywords in research question three occurred in the grammar corpus, their use is not statistically significant in terms of CCSK, nor did they occur as collocates of the nodes *they* or *he*. However, when compared with the BE06 sub-corpora, which are much larger but did not contain any relevant uses of the keywords, the pertinence of the occurrence of the keywords in the grammar corpus is highlighted. Thus, in relation to my third hypothesis, such occurrences signify the grammar authors' explicit awareness of the language-external factors influencing the epicene debate.

CHAPTER 6: ACQUIRING EPICENES

6.1 THE CHILDES CORPUS	198
6.1.1: SELECTING THE DATASET	198
6.1.2: SAMPLING THE DATASET	202
6.2: PRONOUNS PRESENT IN ACQUISITION	206
6.2.1: SINGULAR AND PLURAL <i>YOU</i>	208
6.2.2: SINGULAR AND PLURAL <i>THEY</i>	209
6.2.3: MASCULINE AND GENERIC <i>HE</i>	211
6.3: ANTECEDENTS IN ACQUISITION	214
6.3.1: SYNTACTIC CLASSIFICATION OF ANTECEDENTS	215
6.3.2: PROBLEMS WITH PERSONIFICATION	217
CONCLUSIONS	220

This final corpus-based chapter in the thesis focuses on addressing the issue I identified based on my literature review, that there has been very little research undertaken on children's acquisition and use of epicenes. My final set of research questions, detailed in section 3.2.3, focus on the input given to children during the period that they are acquiring their closed-class personal pronoun paradigms. In addition, the questions also address whether there are any parallels in the acquisition of *they* and *you*. To this end, I analyse the occurrences of *you*, *they*, and *he* in acquisition data taken from the CHILDES database (section 3.1.2).

In section 6.1 I document how I selected a suitable dataset from the extremely large CHILDES database, based on six selection criteria. I then discuss the sampling method I used to narrow down the set of transcripts to create the corpus of child language data used in this study. In section 6.2 I present the results of my analysis for the three tested pronouns, comparing the occurrences of singular and plural forms of *they* and *you* in both L1 input and the children's output. This is followed by an analysis of the tokens of generic *he* in the corpus, allowing me to compare the relative weightings of epicene candidates as I have done in the past two chapters.

Finally, in section 6.3 I take a closer look at the antecedents of the pronouns produced by the children in the CHILDES corpus. I begin, as in

previous chapters, with a syntactic analysis of antecedent types, which shows that the occurrence of generic *he* increases with the definiteness of the antecedent. This is followed by a qualitative consideration of the antecedents, where I highlight the fact that the apparent predominance of generic *he* in the CHILDES corpus is problematic due to issues of animacy and personification.

6.1 THE CHILDES CORPUS

In chapter three I introduced the CHILDES corpus that I use throughout this chapter, but I did not detail exactly how I selected my dataset and chose my samples. This is because, as is shown below, the process of data selection was based on several criteria, each of which warrants extended consideration. Therefore, before presenting the results of my CHILDES corpus analysis, I detail how I selected and sampled a particular dataset for this thesis. I discuss the set of transcripts that I chose for the present study in section 6.1.1, noting that the child language data from the Manchester Dataset (Theakston et al. 2001) is arguably representative of everyday interactions between a child and their primary caregiver. Thus, in the same way that the BE06 data was representative of standard written British English, and the grammar corpus is representative of modern grammar books, so too is the CHILDES data representative of normal parent/child interaction. However, because the Manchester Dataset is so large I had to take a sample from it for this research and I detail the sampling process in section 6.1.2.

6.1.1: SELECTING THE DATASET

In the British English section of the CHILDES database there are transcripts from eleven different studies of child language dated from 1978 through to 2009. The majority of the studies are diachronic, with researchers observing children (or occasionally a single child) at various periods and at differing intervals. The children in the studies range in age from infants at six weeks old to an upper boundary of seven years of age.

Each set of transcripts were collected for particular studies with different agendas, and therefore not all of the datasets are potential candidates for the CHILDES corpus. For my research, I needed data that met the following six criteria, each of which is discussed in detail below:

- (a) The children in the study had to be within the age range usually associated with pronoun acquisition,
- (b) The study had to focus on more than one child,
- (c) The participants in the study could not be related,
- (d) The studies had to avoid strong regional dialects,
- (e) The recordings had to be diachronic, and
- (f) The method of recordings had to be continuous (i.e. not a compilation of snapshot recordings).

In relation to criterion (a), according to Owens Jr (2007:225) the average age range for acquisition of second and third person pronouns is between 30 and 40 months (excluding reflexives), otherwise expressed as 2;6-3;4¹²⁹, whilst Rispoli (1994:158) notes that children can still make mistakes in the third-person at 3;8¹³⁰. As my discussion of pronoun acquisition in section 1.1.3 has indicated, there is evidence for the position that a child's pronoun paradigm becomes fixed relatively quickly, and therefore in order to analyse pronouns as they are still being acquired the children in the CHILDES corpus had to be within this age range. Thus, I excluded four of the eleven British Datasets in the CHILDES database on criterion (a), as the age range of the children tested did not fall within these boundaries (two stopped before 2;6, and two began well after 2;6).

Secondly, I needed a dataset that included more than one child, as the results from a single case study would not provide generalisable results. Data from a single child would only provide an insight into idiosyncratic usage. Also, by using a corpus that includes the language of more than one child this chapter is comparable to the other two corpora of British English used in the previous chapters, keeping the method for the whole thesis uniform. I

¹²⁹ I use the standard notation for expressing children's ages throughout this chapter, where the first figure corresponds to how old the child is in years, with any additional months listed as the second figure, for example 2;6 equates to two and a half years old.

¹³⁰ Similarly, Ricard, Girouard and Decarie (1999:683) note a wealth of studies indicating that "the third animate person pronoun is mastered after the first person pronoun and either later than or simultaneously to the second person pronoun".

took the position that, if the data was to be representative of trends in epicene usage, it should include tokens from more than one individual in order to show that trends in epicene usage can be generalised. Thus under criterion (b) I eliminated a further three studies from the British English section of the CHILDES database.

Similarly, in order to gain the most input data from adults I avoided datasets where caregivers were responsible for more than one child in the study. The British English section of the CHILDES database includes a study of twins (Cruttenden 1978), but this was not suitable for the present research as I wished to have a 1:1 ratio of children and caregivers. In addition, I wanted to avoid issues of cryptophasia or twin language (Mogford 1993), and I am not looking at child-child interaction in this thesis. Thus, I eliminated the twins' dataset based on criterion (c). Furthermore, in order to avoid dialectal usage and pronominal input associated with a particular geographical area, neither of which are key points of focus in this thesis, I eliminated one study based solely on Belfast English in line with criterion (d). Justification for this comes from the fact that any results restricted to a particular dialect would not be generalizable. Indeed, Wales (1996:17) notes the use of *youse* as a non-standard second-person plural in other varieties of Irish English, which, if they occurred in the Belfast data would skew the results of my analysis.

These eliminations left me with two possible datasets, both of which conformed to criteria (a) through (e). In relation to criterion (e), the recordings had to be diachronic because I wanted to record epicene usage across the window of pronoun acquisition noted by Owens Jr (2007). Single recordings for multiple children could not guarantee that each child was at the same stage in pronominal acquisition. In addition, there is no guarantee that a single recording would be representative of a child's pronominal competence, as for instance, a child may not produce a particular pronominal token within a recording session, but may use it consistently in everyday speech. Diachronic recordings, therefore, are much more likely to capture a more accurate representation of the language a child actually produces. Also, it is possible to argue that children subjected to multiple recordings would become accustomed to the recording equipment and procedure. In the Manchester Dataset - detailed below - the children were recorded from a very early age, and were used to the presence of the

researcher, arguably limiting the potential impact of the Observer's Paradox on the language produced by the children.

The first of the two remaining datasets was a study done by Wells (1981) which included 32 children between the ages of 1;6 and 5;0. Initially this dataset looked promising as the children all fell comfortably within the age range of pronominal acquisition (see above); however, on closer inspection the data had two large drawbacks. Firstly the Wells dataset is thirty years old, and thus reflects pronoun acquisition during (and towards the end of) the second-wave feminist movement (section 2.2.1). Therefore, the impact (if any) of the heightened awareness of sexist-language and feminism-based language reform discussed in chapter two may not be evident in this data. Secondly, the method of data collection for this study was to record children for 90-second intervals throughout the day, and thus the recordings do not represent continuous stretches of interaction. Working with the assumption, as stated in criterion (f), that it would be more likely to find pronouns in continuous stretches of discourse I thus eliminated Wells' data¹³¹.

This left me with the Manchester Dataset (Theakston et al. 2001) which consists of 401 hour-long recordings of twelve children from Britain (half from Manchester and half from Nottingham) taken over a period of twelve months when the participants were between 1;8 and 3;0 years old.¹³² Although the Manchester Dataset is ten years old, it is the most suitable batch of files in the CHILDES database for this particular project, as it corresponds to all criteria from (a) to (f). The Manchester Dataset is suitable for this research because the children in the study were within the age range for second and third-person pronoun acquisition (as shown above). In addition, the study included twelve children and their caregivers (although I had to take a sample, see section 6.1.2), meaning that the results of my

131 The method of data collection used for the Wells' data does have some positive aspects, such as, the caregiver/child would not know when they were being recorded and thus their speech would not necessarily be affected by the knowledge that they were being observed. However, for this research at least, the negative points outweigh the positives. For example, a caregiver/child may not speak during the 90-second recordings, or they may speak but not produce any pronouns. Where pronouns did occur, a 90-second snippet may not be enough to retrieve the contextual information needed to determine the antecedent of the pronouns.

132 I am aware that the upper boundary of 3;4 is not reached in this data, and thus there may be limited data on the production of third-person pronouns. However, as the transcripts include the adult speech directed at each child, they should be a good source of input data if the child is not producing third-person pronouns, and as such, the data will suffice.

analysis would not be skewed by idiosyncratic usage. In addition, the participants were recorded with a 1:1 ratio of child to caregiver¹³³.

Furthermore, the children in the study were taken from two different areas of the country, Nottingham and Manchester, meaning that my results on epicene usage are not restricted to one particular geographical area, limiting any potential impact of local dialect features. Any correlation between the participants from both locations will suggest patterns of pronoun usage that are arguably generalisable to a wider geographical area. In terms of data collection, the recordings are diachronic, and taken at fairly regular intervals (see below), as well as being continuous and uniform in terms of format: half an hour of free play followed by half an hour of manipulated play (see below for more details). The age of the data is not problematic as the child participants will now be between the ages of twelve and thirteen, and will be some of the potential targets of the texts in the grammar corpus of chapter five.

Having identified a suitable dataset to address research question four, a closer look at the Manchester Dataset showed that it contained 401 different sets of transcripts, equating to almost seventeen days' worth of recordings. The fact that pronominal tokens occur frequently in language (c.f. Baker 2009:314), meant that looking at all the data would not necessarily yield different results from looking at a much smaller sample. Thus, in the following section I detail how I narrowed down the Manchester Dataset to a manageable number of transcripts for the CHILDES corpus.

6.1.2: SAMPLING THE DATASET

In order to make working with the Manchester Dataset feasible, I decided to look at only four of the twelve children. I selected the oldest girl and oldest boy from the Manchester half of the study who had a complete set of transcripts and the corresponding boy/girl pair from the Nottingham half of the study. These were Becky (2;0) and Joel (1;11) in Nottingham, and

133 There were limited occurrences when other caregivers spoke in the data, but these were in the minority and, in any case, only the primary caregiver's speech is included in this analysis.

Liz (1;11) and Warren (1;10)¹³⁴ in Manchester. I took the oldest children because they spent the most time during the recording period within the pronoun acquisition boundaries set out by Owens Jr (2007). Therefore my selection maximised the possible number of pronominal tokens (especially for the third-person pronouns) produced by the children. The recordings for each child are separated into two sections - part a) half an hour of the children playing with their own toys, and part b) the children playing with specially selected toys. I analysed only the first parts as they more accurately represent the children's daily interactions independent of researcher interference evident in the part b) recordings.

There are 34 part a) data files for each of the four selected children - a total of 136 files and 68 hours of recordings. To address the feasibility of looking at all this data, I began my analysis by documenting the frequency of occurrence of the following five pronominal forms in each of the Becky files: *you, they, them, he, and him*¹³⁵. To analyse the CHILDES files, which are based on the CHAT transcription format (MacWhinney 2000:16-20), I used the CLAN software (MacWhinney 2000: vol. 2), which I downloaded from the CHILDES website. This allowed me to isolate the frequencies for the target pronouns in each of the 34 Becky files.

Initially I focused on the child's utterances (signified by [+t*CHI]), using the query [FREQ +t*CHI +f @] where the @ symbol denotes a selection of files (in this case becky01a through to becky34a). Having documented these figures, I then isolated the utterances of the caregiver (signified by [*MOT]), and repeated the process using the following amended query: [FREQ +t*MOT +f @]. The frequency counts for each conversational participant are shown in Table 6.1.

TABLE 6.1: FREQUENCY COUNTS FOR TESTED PRONOUNS IN BECKY MANCHESTER DATA

	YOU	THEY	THEM	HE	HIM	TOTALS
*CHI (BECKY)	698	180	66	225	65	1234
*MOT (CAREGIVER)	3121	364	236	539	200	4460
TOTALS	3819	544	302	764	265	5694

134 Although Warren does not have a complete recording set, the session that he missed (session 3b) was one of the half hours of manipulated play and thus it was not relevant to this analysis. He does have a full set of part a) recordings.

135 I focused on nominative and accusative forms only, as possessives and reflexives are acquired within a different timeframe (see Rispoli 1994:158, and Owens Jr 2007:225).

The figures illustrate that, on the one hand, there is clearly enough pronoun data to analyse in this dataset, with a total number of 5694 pronominal tokens for just one of the four child/caregiver pairings. However, although I am aware that individual children may produce different numbers of pronouns, extrapolating the Becky data to the other three subjects could have led to an overall total of 22, 776 pronominal tokens. Analysing such large numbers of tokens is unnecessary as in many cases the recordings were only taken days apart¹³⁶ and there is no reason to believe that pronominal usage (either by the child or the caregiver) would change vastly between such recordings. Therefore, I sampled the data further, taking only the first recording of each month¹³⁷ (thus leaving me with a total of twelve files per child). This led to totals of 457 child pronouns, and 1501 caregiver pronouns for Becky. Extrapolated to all four children, this gave a much more manageable predicted total of 8072 pronominal tokens.

When the actual frequencies were calculated, after analysing the data for all four children, the total number of pronouns was 6591 (1065 for the children, and 5526 for the caregivers). The figures for both children and caregivers are presented together in Table 6.2¹³⁸. In all cases, the speech of the investigator/researcher was ignored; only the child and caregiver utterances were analysed in order to record pronoun input that would be representative of the child's day-to-day routine, and to keep the sample to a manageable size. Looking across Table 6.2 what also becomes clear is that the files for two children from the Nottingham half of the research, Becky and Joel, contain almost a third more pronouns than the corresponding files for Warren and Liz. However, this is not necessarily problematic, as both of the latter sets of data still contain over 1200 pronominal tokens each.

136 The shortest distance between two recordings was one day, and the longest was 28 days. However, most of the larger gaps in recording came during the within-month period and thus the initial monthly recordings are not affected.

137 These were not calendar months, they were months of the child's life. For example, the first Becky recording was taken when she was 2;0.07, so the next recording in my sample was when she was 2;01.11, then 2;2.15, and so on, until she was 2;11.01. There were twelve recordings for each child.

138 As with the BE06 data in chapter four the different case forms of *they* and *he* were conflated, in order to make the figures comparable to the second person forms where nominative and accusative pronouns share the same morphological form.

TABLE 6.2: FREQUENCY COUNTS FOR ALL CHILDREN AND CAREGIVERS¹³⁹

	YOU	THEY	HE	TOTALS
BECKY	1307	267	384	1958 (29.71%)
JOEL	1407	253	334	1994 (30.25%)
WARREN	851	184	172	1207 (18.31%)
LIZ	848	281	303	1432 (21.73%)
TOTALS	4413	985	1193	6591

Looking down the columns of Table 6.2 shows that the majority of pronouns in the data were second-person forms, with the occurrence of *you* at 66.95%, with the highest figures coming from the Becky and Joel datasets. The predominance of this pronoun is explained by three factors. Firstly, in the second person, one morphological (and phonological) form represents both nominative and accusative usage, and therefore it can be expected that this pronominal form would occur more frequently than forms of pronouns that differ across case. Secondly, due to the nature of the data in which, for the most part, caregivers are speaking directly to the children, the high occurrence of *you* is to be expected. A potential third reason is that, following the arguments of Rispoli (1994) and Owens Jr (2007), the second-person forms may be acquired before the third-person pronouns, meaning they would be more frequent in the children's output.

Across the sample (henceforth the CHILDES corpus) the adults accounted for between 76% and 89% of the pronouns in each child/caregiver pairing, compared with their own children. Interestingly, Becky is proportionately the child who produces most pronouns relative to her caregiver (as shown in Table 6.3). This is most likely due to the fact that she is the oldest child, and thus her recordings start slightly later in her development than the other children. However, the correlation with age and pronoun production does not hold across the sample, as Table 6.3 illustrates the between-child variation and the individuality of language acquisition (thus providing further justification for criterion b), insofar as individual

¹³⁹ The figures in this table have been modified to reflect the results from the KWAL queries, which, in some cases, did not match the results from the initial frequency counts. This was due to two main factors. Firstly, repetitions and false starts were counted as just one pronominal token, but had been counted as separate occurrences in the frequency queries. Secondly, there were several discrepancies between the results of the FREQ queries and the corresponding data for the KWAL queries.

usage cannot be generalised to all children). Joel produces almost twice as many tokens as Warren and Liz, despite his similarity in age with the latter.

TABLE 6.3: DISTRIBUTION OF PRONOMINAL TOKENS ACROSS CHILDREN AND CAREGIVERS

	TOTAL	%		TOTAL	%
BECKY (2;0.07)	457	23.34	CAREGIVER BECKY	1501	76.66
JOEL (1;11.01)	298	14.94	CAREGIVER JOEL	1696	85.06
LIZ (1;11.09)	154	12.76	CAREGIVER LIZ	1053	87.24
WARREN (1;10.06)	156	10.89	CAREGIVER WARREN	1276	89.11

In order to illuminate any patterns in pronominal acquisition between the second and third person pronouns and to document epicene usage I analysed the sample of transcripts for each child/caregiver pair in order to log all the nominative and accusative occurrences of *you*, *they* and *he* in adult speech and in the children’s speech. One additional useful feature of the Manchester Dataset is that only child-directed speech was transcribed for the recordings (except where other information was contextually relevant), meaning that there is no interference from adult-adult communication in the data.

In this section I have provided more detail about how I constructed the CHILDES corpus (introduced in section 3.1.2) based on the British English section of the CHILDES database. Focusing on one set of transcripts from the database, I documented how I sampled the Manchester Dataset in order to obtain a manageable corpus of child language transcripts for the present research. In the following two sections I detail the different aspects of my analyses of these transcripts, beginning with frequency data for *you*, *they*, and *he*.

6.2: PRONOUNS PRESENT IN ACQUISITION

In order to document pronoun usage I analysed each child’s set of recordings individually. In all cases, the children’s pronouns were analysed

and processed first, followed by the adult pronouns. I used the KWAL (Key Word and Line) function of the CLAN software to isolate occurrences of the target pronouns. Using the basic query [KWAL +t*CHI +syou -w2 +w2 +f @], which was amended accordingly depending on whether I was looking at production or input data, I searched for each pronominal form in turn, using the [+s] element of the query. Thus, there were five searches - +syou, +sthey, +sthem, +she, and +shim - performed on each file for each participant in the interaction¹⁴⁰.

The pronouns were shown in a context of two lines either side, as signified by the [-w2] and [+w2] aspects of the query¹⁴¹. The transcript lines were isolated from the morphological, grammatical, and error analysis lines of the transcription so that, in the majority of cases, the query results included only turns taken by the child, signified as *CHI, or the caregiver, signified as *MOT¹⁴². In each case the @ symbol in the queries referred only to those files which represented the first recording for each month for each child. Unfortunately, a major drawback of the query system was that it did not recognise terms such as [they'*] designed to pick up constructions such as *they're* or *they'll* which had been included as tokens in Table 6.1 to Table 6.3. Therefore, the queries had to be modified slightly, although this did lead to some irrelevant data in the query results - such as the inclusion of words such as *help* which had been picked up by the query [he*] - which had to be eliminated during the manual analysis.

In the sections below, I discuss the results of the queries, detailed above, initially looking at the distribution of singular and plural *you* in section 6.2.1. Following on from this I present the results of my analysis of epicene pronouns, beginning in section 6.2.2 which focuses on whether or not the children produce singular and plural *they* (section 6.2.2), and whether they receive it as input. I then document any cases where generic *he* is used by either the child or their primary caregiver in section 6.2.3.

140 Five queries for each child on each text (5 x 12) plus five queries for each adult in each text (5 x 12) led to a total of 120 queries across the dataset.

141 Whilst this level of context was sufficient for determining the antecedents of most pronominal tokens, in the few cases where there was not enough data, I simply expanded the number of lines presented by the query in order to obtain the correct antecedent value.

142 The [+t] element of the query refers to the tiers of the transcript, so [+t*CHI] includes all the child's utterances, and [+t*MOT] includes all the caregiver utterances.

6.2.1: SINGULAR AND PLURAL *YOU*

Beginning with the second person pronouns, the total number of singular and plural tokens of *you* in the final dataset is 4352, which is less than the initial total of 4413 (as shown in Table 6.2). Sixty-one of the tokens were either occurrences of institutional *you* (see Sanford et al. 2007), or their antecedent was not recoverable and the token was thus eliminated from the study. The vast majority of second person pronouns in the data were singular in both child production and caregiver input, accounting for 97.79% and 99.92% respectively. Arguably, this result was expected, given the dialogic nature of the recordings, and the fact that the participants were speaking directly to each other. However, as shown in Table 6.4, there was limited use of *you* in the plural with a total of thirteen plural tokens across the whole dataset (0.3%).

TABLE 6.4: DISTRIBUTION OF SINGULAR AND PLURAL *YOU*

	PRODUCTION				INPUT		
	SINGULAR	PLURAL	TOTAL		SINGULAR	PLURAL	TOTAL
BECKY (2;0.07)	222 (100%)	2 (0.89%)	224	CAREGIVER BECKY	1056 (100%)	0 (0%)	1056
JOEL (1;11.01)	113 (99.76%)	0 (0%)	113	CAREGIVER JOEL	1265 (99.76%)	3 (0.24%)	1268
LIZ (1;11.09)	91 (100%)	7 (7.14%)	98	CAREGIVER LIZ	747 (100%)	0 (0%)	747
WARREN (1;10.06)	17 (94.44%)	1 (5.56%)	18	CAREGIVER WARREN	828 (100%)	0 (0%)	828
	443 (97.79%)	10 (2.21%)	453		3896 (99.92%)	3 (0.08%)	3899

What is significant here is that the children produced an average of 2.21% of plural pronouns, compared with a much lower figure than the adults at 0.08%. Although this data is only a snapshot of second person pronoun acquisition, and the children in the study will obviously have other forms of exposure to occurrences of *you* (from other caregivers, from books and other media, etc.), arguably their primary caregiver is one of their main sources of language input. Thus, based on this data, it appears that children are able to acquire both singular and plural *you* based on limited instances of the plural form in input.

Interestingly, Liz and Warren, who were part of the Manchester side of Theakston et al.'s (2001) study, both have relatively high values for plural *you* compared with the other children (although Warren only produces one

raw token), perhaps suggesting that the form is more prevalent in their shared dialect (although it is still a standard English form). However, neither child receives the form as L1 input from their caregiver. In fact, only one caregiver in the whole sample uses plural *you* in the CHILDES database, yet three of the children produce the form. This result can potentially be explained by two factors. Firstly, and most importantly, it is unlikely that the caregivers who do not use plural *you* in the CHILDES corpus do not use the form at all; that is, they will use the form, they just did not use it during the recording sessions that generated the transcripts in the corpus. Therefore the children's use of the form indicates that they are not just repeating a stimulus, as they have acquired the form and can use it appropriately.

Secondly, although the adults in the corpus are the primary caregivers of the children involved, they are obviously not their only source of linguistic input, and each child will be acquiring their pronouns based on language data from several sources. Therefore, even in the highly unlikely case that Becky, Liz, and Warren's caregivers never use plural *you*, they will no doubt receive examples of it from other adults. The low value for occurrences of the plural form of *you* is most likely because in child-directed speech, and in this data specifically (see above), the majority of the occurrences of *you* have a singular referent because interaction is between a single child and a single caregiver.

6.2.2: SINGULAR AND PLURAL *THEY*

Moving on to the data for singular and plural *they*, as with the figures for *you* there were less tokens of *they* in the final sample when institutional and problematic tokens were eliminated. Thus the original number of 985 dropped to 866 (a decrease of 119 tokens). The reason the tokens of *they* are fewer than the total number of tokens of *you* can be attributed to two main factors; the first is that the children do not reach the upper bounds of the age of acquisition associated with third person pronouns (see Rispoli 1994 and Owens Jr 2007, discussed above). The second is that there are very few external referents in the CHILDES corpus. The majority of non-present

referents are family members, the investigators, pets, or specific, and gender-specified referents.

The figures for the distribution of singular and plural occurrences of *they* are presented in Table 6.5. When compared with the values in Table 6.4, the distribution of singular and plural forms is reversed. It is the plural form of *they* which is predominant in the CHILDES corpus, yet there are a total of 14 tokens of singular *they*, which accounts for 1.58% of the pronouns in this dataset. This figure is 2.54 percentage points lower than the 4.12% figure for singular *they* across the BE06 data. However, the figure for just the children’s production is much closer to the BE06 value at 3.93% (only 0.19 percentage points lower). Considering the relative number of pronouns analysed in both corpora and the fact that the children in the CHILDES data are still acquiring language, this level of occurrence for singular *they* is notable.

TABLE 6.5: DISTRIBUTION OF SINGULAR AND PLURAL *THEY*

	PRODUCTION				INPUT		
	SINGULAR	PLURAL	TOTAL		SINGULAR	PLURAL	TOTAL
BECKY (2;0.07)	1 (1.39%)	71 (98.61%)	72	CAREGIVER BECKY	2 (1.22%)	162 (98.78%)	164
JOEL (1;11.01)	6 (15.38%)	33 (84.62%)	39	CAREGIVER JOEL	0 (0%)	168 (100%)	168
LIZ (1;11.09)	0 (0%)	24 (100%)	24	CAREGIVER LIZ	4 (2.78%)	140 (97.22%)	144
WARREN (1;10.06)	0 (0%)	43 (100%)	43	CAREGIVER WARREN	1 (0.47%)	211 (81.40%)	212
	7 (3.93%)	171 (96.07%)	178		7 (1.02%)	681 (98.98%)	688

In the CHILDES data both the children and the caregivers produced equal numbers of tokens of singular *they*, although the overall number of pronouns in the dataset was four times more for the adults than the children. In percentage terms, the occurrence of singular *they* had values of 3.93% of the children’s pronominal output and only 1.02% of their input. Interestingly the percentage of occurrence of singular *they* across all the data (1.58%) was five times higher than the corresponding figure for plural *you* (0.3%). This finding is reinforced by the fact that, even though the total number of tokens of *they* was smaller than the totals for *you*, there was one more raw token of singular *they* in the data than plural *you*. If the argument holds that the limited occurrence of plural *you* in input (above), is enough to support a child’s acquisition of two forms of the second person pronoun then, in

relation to research question four, there is no reason to argue that pronoun acquisition cannot facilitate the inclusion of two syntactically distinct forms of *they* in the mental lexicon. This position is reinforced by the fact that the adult (input) value for singular *they* (1.02%) was higher than its counterpart for plural *you* (0.08%).

Interestingly it is notable that three of the caregivers use singular *they*, whereas the result was the opposite for plural *you*, indicating that they are one of the sources providing the children with L1 input including singular *they* (even though the raw values are relatively low). The only child not to receive singular *they* as input in the CHILDES transcripts is Joel, who has a 15.38% occurrence rate for singular *they*. These figures again indicate that, although certain constructions do not occur in input in the CHILDES transcripts, there is no evidence to suggest that the children never receive the target forms from their primary caregiver (or other sources of linguistic input) outside of the experimental recordings.

The data in Table 6.4 and Table 6.5 shows that there are parallels between the acquisition data for singular and plural *you* and singular and plural *they*. In both cases, the children received limited input of one version of the pronoun from their primary caregivers, yet they still produced both forms. In terms of percentages and raw tokens the children actually produced more instances of plural *you* and singular *they* than their caregivers. Thus, in respect of the issues I raised in research question four, the numerical data does appear to support the argument that there are parallels in the pronominal acquisition of *they* and *you* (although the limitations of such small numbers of tokens must be acknowledged). Furthermore, the data on singular *they* supports the fourth hypothesis set out in section 3.2.3 that children within the established age range of personal pronoun acquisition will receive singular *they* as L1 input.

6.2.3: MASCULINE AND GENERIC *HE*

The third part of my analysis of pronoun acquisition focuses on collecting data that is comparable with the BE06 sub-corpora discussed in section 4.1 in order to evaluate the epicene input given to language acquiring children. To this end, I also catalogued all instances of *he* and *him*

in the CHILDES corpus based on whether they were instances of masculine or generic *he*, using the same classification criteria I used for the BE06 data (section 4.1.2). The results and raw figures are shown in Table 6.6, and it initially appears that, when compared with the figures for singular *they*, discussed above, generic *he* was much more frequent in the data both in terms of input and child output. However, the analyses in section 6.3 show that the numerical data oversimplifies the use of this form, indicating that the dominance of generic *he* is not as clear-cut as it initially appears.

TABLE 6.6: DISTRIBUTION OF MASCULINE AND GENERIC *HE*

	PRODUCTION				INPUT		
	MASCULINE	GENERIC	TOTAL		MASCULINE	GENERIC	TOTAL
BECKY (2;0.07)	63 (70.00%)	27 (30.00%)	90	CAREGIVER BECKY	232 (93.17%)	17 (6.83%)	249
JOEL (1;11.01)	87 (100%)	0 (0%)	87	CAREGIVER JOEL	178 (97.80%)	4 (2.20%)	182
LIZ (1;11.09)	22 (88.00%)	3 (12.00%)	25	CAREGIVER LIZ	141 (97.92%)	3 (2.08%)	144
WARREN (1;10.06)	59 (93.65%)	4 (6.35%)	63	CAREGIVER WARREN	171 (91.94%)	15 (8.06%)	186
	231 (87.17%)	34 (12.83%)	265		722 (94.88%)	39 (5.12%)	761

Again, as with the results for *they* and *you*, some tokens in the original 1193 (see Table 6.2) were eliminated due to their antecedents not being recoverable. The final number of was 1026 (a drop of 167 tokens). Interestingly, the overall ratio of tokens of *he*-to-*they*, 1026:866, is similar to the overall distribution of tokens of the pronouns in the BE06 sub-corpora in section 4.2 (which have corresponding values of 4845:4373, simplified to 1.18:1 and 1.11:1 respectively). Thus, the higher number of tokens of *he* in the data is not problematic,¹⁴³ and its comparability with the BE06 data is evidence for the position that the acquisition data follows a regular distribution, representative of wider language use.

However, what is significant is that there were 73 tokens of generic *he* across the CHILDES corpus, which represented 7.12% of the tokens of *he*. This is much higher than the corresponding value for singular *they* in the CHILDES data, and is more than double the percentage value for generic *he* in the BE06 data where the figure was 3.22%. Also, there was little

¹⁴³ This higher figure can be accounted for by the fact that two of the children had pets (a dog and a cat) which were biologically male, which were referred to on multiple occasions by both conversational participants.

difference between the number of tokens of generic *he* produced by the children and the adults - figures of 34 and 39 respectively - although in terms of percentages, the value was higher for the children (12.83%) as they produced less tokens of *he* overall (compared with an adult value of 5.12%).

However, a closer look at Table 6.6 shows that the data for generic *he* is not uniform. The overall children's total is inflated by Becky's large number of raw tokens, whilst Joel does not use the form at all, and the others use generic *he* much less than Becky. Thus, the usage is not as consistent as for the other tested pronouns, and indeed Becky's use of the form is problematic, as I show in section 6.3.2. Nevertheless, the data follows the pattern, seen throughout this analysis, whereby the children produce proportionally more instances of the target pronoun than their caregivers, yet based on the analysis of BE06 in section 4.1, the overwhelming dominance of generic *he* in the CHILDES corpus is surprising. Comparing Table 6.6 with Table 6.5 shows that proportionally the children are 8.9 percentage points more likely to produce generic *he* than singular *they*, and similarly for the adults, this value is 4.1 percentage points¹⁴⁴. Fundamentally, the use of generic *he* in the CHILDES data does not follow the patterns of pronominal use in written British standard English found in section 4.2.1.

In this section I have evaluated the data from my quantitative analysis of the second- and third-person pronouns used by children, and provided to them as L1 input, whilst they are young enough to be acquiring their closed-class paradigms. Beginning with my discussion of the second-person forms, the values for plural *you* set a benchmark by which it was assumed that the children would receive enough input to facilitate the acquisition of both a plural and a singular form of the pronoun. Comparing this with the data for *they*, it is significant that proportionally, and in terms of raw figures, singular *they* occurs more frequently in comparison with plural *they*, than plural *you* occurs in relation to singular *you*. In addition, more of the caregivers used singular *they* than plural *you*.

¹⁴⁴ These figures are calculated by subtracting the percentage value for occurrences of singular *they* from the percentage value for occurrences of generic *he*.

However, when analysing the corresponding data for *he*, in order to facilitate a comparison between epicene forms, and to measure whether occurrences of generic *he* reached the benchmark level set based on the values for plural *you*, the results were surprising. Based on the numerical figures only, it appears that generic *he* is the epicene of choice in the CHILDES corpus, going against the grain of previous research, which has shown that the form does not occur in L1 acquisition (section 2.3.1). In addition, generic *he* occurred proportionally more in the CHILDES corpus than it did in the BE06 data. This finding is anomalous to the results of the previous two chapters, and so, to investigate the tokens of generic *he* in more detail I take a closer look at the antecedents coindexed with the form in the following section.

6.3: ANTECEDENTS IN ACQUISITION

In order to investigate why the occurrence of generic *he* was so high in the CHILDES corpus when compared with the BE06 data, I catalogued the antecedents of generic *he* and singular *they*, both syntactically as in chapter four and chapter five (section 6.3.1), and qualitatively (see section 6.3.2). This procedure raised an important issue, as it immediately became evident that the distinction of animate/inanimate, which had been included in the BE06 analysis, was blurred in the CHILDES corpus. This is because the children would often personify inanimate objects (the majority of which were toys, including teddy bears, figurines, and shapes made from plasticine) thus imbuing them with animate qualities. As such, it was impossible to separate all of the antecedents of *he* and *they* into binary categories of animate and inanimate without a high level of subjectivity, and thus to avoid any errors, this distinction was not enforced¹⁴⁵. The analyses in the following two sections highlight the problems with the tokens of generic *he*.

¹⁴⁵ As I show in section 6.3.2, there are only limited occurrences of inanimate antecedents such as *my blouse* or *telescope*, all of which came from the children. Interestingly, all such antecedents were coindexed with singular *they*, arguably suggesting that the children were assuming animacy but not gender.

6.3.1: SYNTACTIC CLASSIFICATION OF ANTECEDENTS

Following the same method of antecedent classification as I have used in the previous two chapters, I separated the antecedents of singular *they* and generic *he* on syntactic grounds, classifying them as indefinite pronouns, NPs with quantifiers, indefinite NPs, or definite NPs. As the idiosyncratic usage of individuals was not the primary focus here (c.f. criterion b) in section 6.1.1) I pooled the antecedents from the children’s outputs, and the antecedents used by their caregivers, into two groups. However, as my discussion below will show, it was sometimes important to look at an individual’s usage to highlight any anomalies in the data.

Table 6.7 shows the distribution of antecedent types for both singular *they* and generic *he*, in terms of production (child data) and input (caregiver data). The children and the adults show similar distribution patterns for singular *they*, with most of the tokens (78.57%) coindexing with indefinite or definite NPs. Only the adults use singular *they* with indefinite pronouns, and even then the tokens are limited to only three occurrences produced by two of the caregivers in the corpus. This distribution of antecedent types does not match the usage data presented in section 4.2.1 and 4.2.2, as in the CHILDES data the tokens of singular *they* do not decrease as the definiteness of referents increases. However, I do not wish to overstate the significance of these results, as there are only fourteen tokens of singular *they* in the whole corpus, a much lower figure than the corresponding result for the BE06 data.

TABLE 6.7: SYNTACTIC DISTRIBUTION OF ANTECEDENTS IN THE CHILDES CORPUS

ANTECEDENTS OF SINGULAR <i>THEY</i>					
	INDEFINITE PRO	NP WITH QUANTIFIER	INDEFINITE NP	DEFINITE NP	TOTALS
PRODUCTION	0	0	2	5	7
INPUT	3	0	1	3	7

ANTECEDENTS OF GENERIC <i>HE</i>					
	INDEFINITE PRO	NP WITH QUANTIFIER	INDEFINITE NP	DEFINITE NP	TOTALS
PRODUCTION	0	1	27	6	34
INPUT	0	0	17	22	39

In terms of generic *he*, results for which are shown in the second table, the distribution pattern is similar to the results for singular *they*, with the vast majority of occurrences referring to either indefinite or definite NPs. What is interesting is that the adults do tend towards the trend of increasing

uses of generic *he* with increased definiteness (as was the pattern for generic *he* in the BE06 data). However, as is shown in Table 6.8 below, this result is not as pronounced when multiple references to a single antecedent are condensed to just one token, and thus its significance is debatable.

As discussed in section 4.1.1 in relation to the BE06 data, analysis of the total number of tokens of the target pronouns overshadows the fact that, in many cases, pronouns are used more than once to refer to the same antecedent. The same effect also occurs in the CHILDES corpus. For example, in this data Becky used *he* fourteen times to refer to (a) *bunny* that she had invented whilst telling her mother a story¹⁴⁶ (see section 6.3.2 for further analysis of this antecedent). Therefore, following the same condensation procedure described in chapter four I eliminated all duplicate references to a single antecedent in order to show more clearly the distribution of syntactic forms coindexed with generic *he* or singular *they* across the corpus. The revised figures are shown in Table 6.8.

TABLE 6.8: SYNTACTIC DISTRIBUTION OF ANTECEDENTS IN CHILDES CORPUS **AFTER** CONDENSATION

	ANTECEDENTS OF SINGULAR <i>THEY</i> (AFTER CONDENSATION)				TOTALS
	INDEFINITE PRO	NP WITH QUANTIFIER	INDEFINITE NP	DEFINITE NP	
PRODUCTION	0	0	2	4	6
INPUT	3	0	1	2	6

	ANTECEDENTS OF GENERIC <i>HE</i> (AFTER CONDENSATION)				TOTALS
	INDEFINITE PRO	NP WITH QUANTIFIER	INDEFINITE NP	DEFINITE NP	
PRODUCTION	0	1	7	6	14
INPUT	0	0	8	11	19

For singular *they* there is not much change between Table 6.7 and Table 6.8 as only two antecedents were referred to using more than one pronoun. The effects of the condensation of antecedents are much more pronounced for generic *he*. The difference between the figures in Table 6.7 and Table 6.8 show an overall decrease from 73 tokens of generic *he* to 33 individual antecedents (a drop of 45.21%). In Table 6.8 the distribution of the children's pronouns is much closer to the adult data, and thus the differences between child and caregiver usage in Table 6.7 are neutralised. However, again the numerical data can only take the analysis so far, and

¹⁴⁶ It is possible that Becky intended that the rabbit in the story was male, but categorising such occurrences of *he* as overtly masculine without closer analysis would have been a subjective judgement. In any case, antecedents involving animals and toys are dealt with in more detail below.

thus, in the following section, I take a closer look at the antecedents in the CHILDES data.

6.3.2: PROBLEMS WITH PERSONIFICATION

When the antecedents in the CHILDES database are analysed from a qualitative perspective the number of actual tokens of singular *they* and generic *he* decreases even more¹⁴⁷. Looking at Table 6.9 which shows the antecedents of singular *they*, the majority of antecedents produced by the children refer to inanimate objects (as noted above in the introduction to this section)¹⁴⁸. For the adults on the other hand, the input tokens of singular *they* all have possibly animate antecedents, although *mouse* and *my baby* most likely refer to toys or pictures in books (based on the limited context provided by the queries). Based on this data, I therefore posit that the children in the study have not yet fully acquired the [ANIMATE] phi-feature for the pronoun *they*.

TABLE 6.9: ANTECEDENTS OF SINGULAR *THEY* IN THE CHILDES CORPUS

ANTECEDENTS OF SINGULAR <i>THEY</i>			
INDEFINITE PRONOUNS	NPs WITH QUANTIFIER	INDEFINITE NPs	DEFINITE NPs
PRODUCTION			
-	-	HAND TELESCOPE	MY BLOUSE MY TELESCOPE THE SHOP YOUR HAIR
INPUT			
EVERYBODY X 2 NOBODY	-	MOUSE	GOOD DOG MY BABY

Whilst it is impossible to directly test the animacy argument using corpus analysis, there is some useful pre-existing scholarship available. For example, Baauw (2002:8) notes that “children’s non-adult like interpretation of definite articles and pronouns is due to the underspecification of functional categories”, stating specifically that this applies to “the

147 Due to the very small number of animate antecedents (which do not refer to personified toys) it was not possible to test the CHILDES data for interaction with gender stereotypes, as was done for BE06 (section 4.2.3) and the example sentences in the grammar corpus (section 5.2.3). None of the antecedents of singular *they* or generic *he* in the CHILDES corpus were tested by Kennison and Trofe (2003) or Carreiras et al. (1996).

148 Possibly *the shop* (if taken as a metonym) could also be an example of institutional *they* as opposed to singular *they* (Sanford et al. 2007).

incomplete acquisition of the feature content of pronouns”. Although my interpretation of the children’s use of singular *they* with inanimate antecedents can only be speculative, it seems reasonable to argue that, as the children are still within the age boundaries of third-person pronoun acquisition, their pronoun specifications are not yet fully formed. A lack of specification for the [ANIMATE] phi-feature can neatly explain the children’s usage.

In terms of generic *he*, a qualitative analysis of the antecedents (displayed in Table 6.10) indicates why there were so many tokens of this particular pronoun in the CHILDES corpus. All of the antecedents are examples of creatures that may be animate (even if the physical referent is a toy or picture), whether they represent real animals (such as *a fly*) or mythical ones (*a big bad baby monster*), or are inanimate objects personified (*my train*).

TABLE 6.10: ANTECEDENTS OF GENERIC *HE* IN THE CHILDES CORPUS

ANTECEDENTS OF GENERIC <i>HE</i>			
INDEFINITE PRONOUNS	NPs WITH QUANTIFIER	INDEFINITE NPs	DEFINITE NPs
PRODUCTION			
-	ANOTHER BUTTERFLY	(A) BUNNY A BIG BAD BABY MONSTER A FLY A PARROT A TEDDY IN A BOX HORSIE RABBIT	MY FROGGY MY TEDDY MY TRAIN THAT TEDDY THIS PUPPY THIS RABBIT
INPUT			
-	-	[FLY] A BIG BIG ELEPHANT A BLUE TIT A CHICKEN A DOG A FLY A HORSIE MONSTER	BAA BAA SHEEP THAT ALLIGATOR THAT RHINO THE BIRD THE DRIVER THE ELEPHANT THE GORILLA THE MONKEY THE MOUSE THE PARROT THE ZEBRAS

Therefore, following on from the discussion of animacy above, an initial analysis of the data suggests that the preferred epicene for potentially animate referents is generic *he*. However, a closer look at some of the context provided in the queries suggests that the case may not be this clear cut. There is no way to tell whether the instances of *he* coindexing with antecedents such as *the bird*, *a horsie*, *a parrot*, etc. are examples of

masculine or generic reference. In each case it could be that the child, or the caregiver has already assigned a gender identity to a particular toy before the recording process which led to the transcripts in the CHILDES database began. I base this argument on the assumption that children create individual characters and identities for their toys and stuffed animals, as it is highly unlikely that an inanimate object personified as *my froggy* does not have some form of gender associated with it. Indeed, Wales (1996:141-144) argues that pronoun choice and personification is not straightforward, discussing the different factors which can affect the selection of pronouns for animals¹⁴⁹.

The process by which a child creates a character and assigns it a gender (or rather, biological sex) is illustrated by Becky's use of *a bunny* with generic *he*. During one recording she makes up a story about a rabbit that she tells to her mother, from the start of the story she refers to the creature as *he*, and this is reinforced by her caregiver¹⁵⁰ who repeats the pronoun in relation to the rabbit. I argue that it is highly unlikely that Becky's story involves a generic rabbit which could be either male or female, especially as the masculine pronoun is used in reference to the creature a total of 14 times. It is much more likely to be the case that Becky has created a male character for her story and is thus using the appropriately gender-marked pronoun.

Although I am aware that this is only one interpretation of Becky's actions, and the same process may not hold across all of the antecedents of generic *he* in Table 6.10, the explanation nevertheless problematises the tokens of generic *he*. It is thus unclear from the data whether or not these particular instances of *he*, which refer back to indefinite or inanimate antecedents, are actually generic, or whether they represent uses of the masculine form to aid in the personification of the children's toys. Eliminating all the problematic antecedents would lead to extremely small totals for both singular *they* and generic *he*, and therefore much more data would need to be collected in order to see which epicene pronoun children are predominantly exposed to.

149 Wales also looks at personification and sexism (1996:146-152).

150 The reason *bunny* is not an antecedent of generic *he* in the input section of Table 10 is because the caregiver is merely replicating the pronoun use of the child.

When the problematic antecedents are removed, there are effectively no instances of generic *he* in the CHILDES corpus. Yet there are seven instances of singular *they*: three indefinite pronouns, one indefinite NP (*mouse*), and three definite NPs (*the policeman*, *good dog*, and *my baby*). These final figures actually suggest that singular *they* is the favoured epicene pronoun in the CHILDES corpus, insofar as the instances of generic *he* are not actually generic. Significantly, it is important to note that all of these antecedents are tokens of caregiver speech, suggesting that the children in the corpus receive singular *they* as L1 input. Thus, the use of singular *they* as the (unproblematic) epicene pronoun of choice mirrors the BE06 data presented in chapter four, where singular *they* was the dominant choice of epicene in written British English. However, due to the limitations of the data, this area clearly needs more research.

In this final section I took a closer look at the antecedents used with singular *they* and generic *he* in the CHILDES corpus, problematising the quantitative results by illustrating that the distinction between animate and inanimate antecedent was blurred, and showing that issues surrounding personification interfere with the data for generic *he*. When the problematic antecedents of singular *they* and generic *he* were eliminated, only tokens of singular *they* were unaffected, thus bringing the CHILDES data back into line with the results from the BE06 data and indicating a preference for *they* in epicene contexts. However, the data on epicene usage from the CHILDES corpus is limited, and as I have noted above, more research will have to be done in order to make any firm conclusions about the epicene pronoun of choice in child language acquisition.

CONCLUSIONS

In relation to research question four I have shown that there are numerical parallels between the input given to children and the output they produce whilst they are acquiring the second person pronoun *you* and the third person pronoun *they*. Although the data does have its limitations, I argue that the sample I chose for the CHILDES corpus is representative of the

rest of the dataset and arguably of child language acquisition in British English in a wider sense, due to the suitability of the data under the criteria (a) to (f) which I justified in section 6.1.2. Working with the assumption that the child's primary caregiver is (one of) their main sources of linguistic input (especially in terms of dialogue) the analysis of this data suggests the following:

- Children receive both singular and plural *you* as input (Table 6.4).
- The limited uses of plural *you* (0.08%) is arguably enough to facilitate the child's acquisition of singular and plural *you*, which are part of the standard English personal pronoun paradigm¹⁵¹.
- The children in the study produce proportionally more tokens of plural *you* (2.21%) than the adults providing the input, suggesting that they could expand upon the input received.
- The children also receive both singular and plural *they* as input, with the vast majority of instances being plural (98.98%).
- The children in the study produce proportionally more tokens of singular *they* (3.93%) than they received as input.
- In the same vein as with the second person, the limited adult uses of singular *they* (1.02%) is therefore enough to allow the child to acquire the two forms of the pronoun, even though the singular form is not endorsed as part of the standard English personal pronoun paradigm (c.f. chapter two).

Thus in relation to research question four, it appears that there are clear parallels in the input given to children during their acquisition of second and third person pronouns. Similarly, there also appear to be parallels in pronominal production, as the children in the study expand upon the pronominal input they are given in much the same way for *you* and *they*. In this chapter I have focused on spoken epicleses only, and I am aware that there are of course other forms of linguistic input given to children, such as books, television, etc., which may or may not influence their pronominal acquisition, but they are not included within this study. Whilst further research on parallels in pronoun acquisition is therefore necessary, the data

151 Whilst the children's primary caregivers are not their only source of L1 input, I argue that the transcripts in the CHILDES corpus are representative of the vast majority of a child's daily interactions, and thus the importance of the pronoun use of the primary caregivers must not be underestimated.

provided here is both in keeping with the data used in the rest of this thesis, insofar as it is corpus-based, and sufficient to address the research questions set out herein.

Of course there are many other methods which could be used to test hypotheses concerned with pronoun acquisition, such as experiments that test how children conceptualise the referents of the forms *they* and *you*. However, further research on this topic is beyond the focus of this thesis, as it would stray from the corpus-based approach I have used in the previous two chapters. In addition, Chiat (1986:345) noted that in other forms of pronoun-based experiments “it is difficult to obtain reliable responses which are not experimental artefacts from children at the crucial stages of development”. She also claims that “By the time the child responds reliably, the acquisition process may be well advanced, and the experiment may fail to tap it” (1986:345). Therefore, by using corpus analysis I am able to avoid the problems with elicitation experiments etc. and analyse the actual language used by the pronoun-acquiring child.

To link back to the primary aim of my research, the analysis of epicene pronouns, and my hypothesis that children would receive singular *they* as L1 input, I compared the relative frequencies of singular *they* and generic *he* in the CHILDES corpus. Initially the quantitative analysis showed that generic *he* occurred more frequently than singular *they* in the data. These results stand in opposition to the BE06 data presented in section 4.1, as well as other similar corpus-based studies of epicenes such as Newman 1992, Stringer and Hopper 1998, Baranowski 2002, and Paterson 2011. However, on closer inspection from a qualitative perspective it became apparent that some of the antecedents of both singular *they* and generic *he* were problematic. The children (and adults) used personal pronouns in relation to inanimate objects, potentially as an aid to personification. Due to this usage it was extremely difficult to isolate instances where *he* was used generically and where it was used to imbue an inanimate object, such as a toy figurine of an animal, with masculine gender. The removal of the problematic antecedents meant the elimination of all the children’s tokens of generic *he*, suggesting that epicene usage was restricted to the adults’ use of singular *they*, a finding which appears to support the dominance of this particular epicene in the BE06 data of chapter four.

CHAPTER 7: FURTHERING EPICENE RESEARCH

7.1: THE ENGLISH EPICENE	226
7.2: EPICENES AND LANGUAGE-EXTERNAL FORCES	232
7.3: THE ACQUISITION OF THE EPICENE PRONOUN	238
7.4: EXPLAINING EPICENE REFERENCE: THE HOMONYMY OF <i>THEY</i>	243
CONCLUSIONS	247

The studies presented across the previous three chapters are not isolates, both in terms of this thesis and in terms of the large body of literature on English epicene pronouns. My goal in this chapter is to synthesise the results of my corpus analyses with the existing scholarship I evaluated in chapters one and two. Furthermore, I also reiterate how my results have informed my responses to my research questions, and my related hypotheses. In the sections that follow, I show how my research fulfils the aims of my thesis and contributes to debates in wider literature.

One of the key aims of this thesis, as set out in the introduction, was to bring epicene research up to date, having identified in chapters one and two that previous work on epicene pronouns could not necessarily account for modern usage, as existing studies were generally over fifteen years old. I hypothesised that the investigation of current usage would indicate that singular *they* is the epicene of choice in speech and writing in British English today. Furthermore, I also hypothesised that modern grammar texts would continue the trend evident in the literature and reject generic *he* in favour of prescribing epicene avoidance tactics, and that modern grammar prescriptions would be influenced by language-external factors. In order to address these hypotheses I used corpora to compare language use with language prescriptions, documenting current usage and investigating whether the data from previous studies, showing the continuing condemnation of singular *they*, still held post-2000.

Additionally I noted the lack of discussion about the role that acquisition played in the epicene debate, hypothesising that, as there is support for the position that singular *they* is the spoken epicene of choice,

children who are acquiring their pronoun paradigms will receive it as L1 input. I not only proposed that acquisition could account for the apparent suitability of singular *they* as an epicene pronoun, I analysed children's usage to investigate whether they received singular *they* when they were acquiring their personal pronoun paradigms. Finally, I noted that current theories of epicene reference were insufficient to account for the use of singular *they* with formally singular definite NPs and in response proposed my adaptation of Homonymy Theory.

The analyses I carried out on the BE06 sub-corpora, the grammar corpus, and the CHILDES corpus all indicate that the two major epicene candidates in British English are singular *they* and generic *he*, whilst other possible epicene candidates, such as generic *she*, neologisms, and combined pronouns are extremely rare.¹⁵² There were no evident cases of neologisms in any of the corpora in this research and there were only limited occurrences of combined pronouns in the BE06 sub-corpora (section 4.1.2), despite the fact that the promotion of such combined forms for written use is the primary avoidance tactic promoted by authors in the grammar corpus. These results confirm my assertion, made at the beginning of this thesis, that singular *they* and generic *he* are the two main candidates for epicene status in British English.

Further justification for this claim is that these two pronouns are in direct opposition to one another throughout the wider literature on epicene reference. Generic *he* is the formally endorsed epicene of traditional grammatical prescriptivism (which I discussed in detail in chapter two), whilst singular *they* is the pronoun in common use that has been proscribed by grammarians using arguments concerning number concord. The results I have presented in the previous three chapters reinforce the relative positions of singular *they* and generic *he* as prime contenders in the epicene debate.

Having established that generic *he* and singular *they* appear to be in direct competition for epicene status, throughout the rest of this chapter I

¹⁵² Although I did not focus specifically on generic *she* in this research, I previously tested a sample of two hundred concordance lines (out of 963 tokens) from the BNC and found that generic *she* only occurred with the heavily stereotyped *the/your receptionist* (see Paterson *forthcoming a*). Using stereotype data from Carreiras et al. (1996), I found that *receptionist* was ranked as highly female-stereotyped with a score of 9.57 out of 11.0, meaning that the generic nature of this use of the pronoun can be called into question.

explore how the results of my analyses of these two forms have impacted upon the epicene debate. In section 7.1 I document the dominance of singular *they* which has been evident throughout this thesis. I focus mainly on the data from the BE06 sub-corpora and show how my research has shown singular *they* to be a versatile pronoun, which can coindex with all types of formally syntactically singular antecedents. However, I also show how, despite the impact of language-external forces, such as non-sexist language reform, generic *he* still occurs in my data and thus has not been eliminated from usage.

Section 7.2 includes my evaluation of the results from the grammar corpus in chapter five and shows how they contradict previous literature on grammar handbooks. My results suggest that not only are the modern grammar authors aware of language-external pressures on epicene pronoun choice, they have reacted to such pressures and moved away from the endorsement of generic *he*. Yet, although there is some limited endorsement of singular *they* epicene avoidance tactics predominate as the grammarians' choice for tackling epicene reference.

In section 7.3 I show how my research contributes to the understanding of the position of singular *they* in the current standard pronoun paradigm. I rely on the data from the CHILDES corpus to suggest that the form is available as L1 input during acquisition and can thus be integrated into people's mental grammars. Conversely, I show that the children do not use generic *he* with any consistency as an epicene pronoun, yet it does appear to be their default choice of pronoun for personification of inanimate objects. Using my data from the CHILDS corpus, I also highlight the parallels that singular and plural *they* have with the second-person pronouns, suggesting that regularity in the rote-learned system is an important aid to pronoun acquisition.

Finally, in section 7.4, I show that my research provides support for my adaptation of Whitley's (1978) Homonymy Theory. Casting aside the common consensus that singular *they* is a manipulation of the plural, I show that my results can support the theory that there are two syntactically distinct forms of *they* in the mental lexicon, which represent a pronoun syncretism. However, I do not wish to overstate my case and note that the proposal of Homonymy Theory is still in its infancy and more research needs to be done to test the theory directly.

7.1: THE ENGLISH EPICENE

The dominance of singular *they* as an epicene pronoun is evident throughout this thesis. I began my research with a comprehensive review of the literature on epicene pronouns, focusing in chapter one on the many studies that have tested the production and comprehension of singular *they* and generic *he*. Documentation of the results of such studies, which included Gastil (1990), Foertsch and Gernsbacher (1997), and Sanford and Filik (2006), illustrated that singular *they* appears to be an unproblematic epicene, causing little-to-no processing difficulties¹⁵³ and causing no gender-mismatches with its antecedents. My initial review of previous studies also showed that there is very little empirical data in favour of generic *he* as an epicene pronoun.

The vast majority of studies in chapter one indicated that *he* is perceived as masculine; it also creates mental images of referents which are biologically male, and causes a gender-mismatch with feminine-stereotyped antecedents which initiates the P600 brainwave associated with syntactic anomalies (c.f. Osterhout, Bersick, and McLaughlin 1997). Based on such findings I conclude that there is a large body of literature to support the position that *he* has a default masculine interpretation and thus it cannot be unproblematically processed as a generic, genderless pronoun.

However, my review of the existing literature indicated that there have been very few empirical studies on epicene pronouns in the last ten-to-fifteen years. Therefore, following on from my evaluation of previous works on epicene pronouns, I set out to update the epicene literature. I posed sets of research questions at the end of chapters one and two in order to focus my research, and address the issues of epicene reference, prescription, usage, and acquisition that came to the fore through my synthesis of the current literature. Significantly, what was evident throughout my review of existing research was that, because of the small number of modern studies, researchers were still making conclusions based on older data, and thus one

153 The results from Sanford and Filik (2006) did indicate that there was an overall processing lag when singular *they* was used. However, I suggest that this impact could be due to their participants' awareness of traditional prescriptions of generic *he*, although such claims can only be speculative as Sanford and Filik did not control for this variable.

of the primary aims of my research questions was to update the state of knowledge about how epicenes are used in British English.

My analysis of modern standard English in the BE06 sub-corpora in chapter four showed that on all measures I tested singular *they* was the most-used epicene, and thus the top candidate for epicene status in standard written British English. These results illustrate a continuing trend for singular *they* to be the dominant epicene in corpus data, as my results reinforce the findings of Newman (1998), Baranowski (2002), and Pauwels (2001), all of whom took a corpus-based approach to epicene research. Significantly, I showed that singular *they* was used with all antecedent types in both `pronouns_press` and `pronouns_general_prose` - from indefinite pronouns, where it is most popular, to definite NPs (where its use was still more frequent than generic *he*). This finding contradicts the notion that singular *they* can only coindex with indefinite pronouns, and other limited classes of antecedents, as discussed by Kolln (1986), and Foertsch and Gernsbacher (1997).

My close analysis of the antecedents of singular *they* in section 4.3.2 also shone light on the specifics of singular *they*, showing, for example, that the pronoun was used with definite NPs where the writer presumably knew the biological sex of the referent, but chose not to make it salient. For example, in text F26 in `pronouns_general_prose` a *protester* is referred to as *him*, before the writer refers to the same referent with *they* in the following sentence:

- (1) [The protester]_i wouldn't and after I had spoken with [them]_i asking [them]_i not to disrupt the meeting - which [they]_i refused - [they]_i were escorted out. (Text F26)

The use of *he* with any such antecedent could not be generic, as it would have to match the biological sex of the referent, thus blocking any secondary generic interpretation of the pronoun. This use of *they* with definite NPs, which I observed in the BE06 sub-corpora, is yet more evidence that previous arguments (c.f. Kolln 1986; Foertsch and Gernsbacher 1997) limiting the use of singular *they* to indefinite pronouns do not hold in practice. Therefore my results have updated the state of knowledge about what types of antecedent singular *they* is coindexed with in attested British English. This result also provides support for my argument that the grammar authors in chapter five simplified their discussions of epicene reference by

choosing unproblematic antecedents for their example sentences and over representing the use of singular *they* with indefinite pronouns (see section 7.2). My results clearly show that the limits imposed on singular *they* in the grammar corpus of chapter five do not reflect language use evident in chapter four (although as I discuss in section 7.2 the grammar corpus data does indicate that such texts are more in line with language use than previous studies have suggested).

Furthermore, there were limited occurrences of singular *they* in the BE06 data with antecedents that were lexically specified for sex, such as *one in nine women* and *a man* (see section 4.3.2). These results expand on similar data from previous studies, such as Green (1977) who found *they* coindexed with *actress*, Baranowski (2002) whose *an average PR woman* example is discussed in section 1.3.2, and Paterson (2011:179) where I found *a man*, *any girlfriend* and *any woman* used as antecedents of singular *they* in a corpus of newspaper texts taken from 2007-2008. Such uses of singular *they* provide more evidence that the form does not cause a gender mismatch between pronoun and antecedent. The attested uses in my data suggest that, in line with the results for *actress* etc., singular *they* can be used as a suitable pronoun for coreference with antecedents that are lexically specified for grammatical gender (or biological sex).

In addition, my analysis of written British English also showed that singular *they* is used with both masculine and feminine gender-stereotyped antecedents, demonstrating that the form is an extremely versatile epicene pronoun. I am aware that I did not directly test the processing difficulties that may or may not have resulted from using singular *they* with gender-stereotypes, as this was not possible using a corpus methodology. However, the use of singular *they* in British English with such stereotypes, suggests that it can be processed unproblematically with antecedents such as the masculine-stereotyped *officer* and the feminine-stereotyped *victim* (taken from the BE06 data). Although I did not test mental processing directly, existing research (c.f. Foertsch and Gernsbacher 1997) has shown that singular *they* can be processed with relative ease with gender-stereotyped antecedents, and my BE06 analysis provided further examples of the types of construction tested in such studies (Foertsch and Gernsbacher 1997; Kennison and Trofe 2003). I thus argue that the occurrence of singular *they* with gender-stereotypes in the BE06 data is evidence of its acceptability

from a processing point of view, especially considering that the written texts in the BE06 sub-corpora had presumably been through some form of editing which would have eliminated the majority of problematic constructions.

However, although the dominance of singular *they* is clear from my results, my research also showed that generic *he* has not been eliminated from current epicene usage. My analyses therefore update the state of knowledge on the use of generic *he* showing that it is still in use as an epicene pronoun. The form occurs both in the BE06 sub-corpora and in the CHILDES data (albeit in a limited capacity, see section 7.3). Whilst the occurrence of generic *he* can be seen as evidence for the impact of traditional grammatical prescriptivism, I propose that this is only one possible explanation. Alternatively, I argue that the continued use of generic *he* is evidence in favour of McConnell-Ginet's (1979) argument that it is difficult to imagine a referent without biological sex (discussed in section 1.2.3). Indeed, I showed in chapter four that as the definiteness of the antecedents in the BE06 data increased, so did the use of generic *he*. I propose that this increase in the use of generic *he* relates to the ease with which the referent can be imagined or visualised. It is arguably easier to produce a mental image for a definite NP, such as *the guitarist*, than it is to imagine a single figure for the indefinite pronoun *someone* (most likely because of the semantic value of the definite NP).

Following McConnell-Ginet's (1979) argument, I suggest that a person's pronoun selection can be swayed by the mental image created by an imagined antecedent, and thus *he* may have been used pseudo-generically in the BE06 data with definite NPs because it matched the biological sex of the imagined referent. Incidentally, this explanation could also be used for the children's use of *he* for toys and imagined characters in the CHILDES data (see section 7.3). However, I argue that in such cases grammatical gender is used to reflect the language-user's imagined biological sex for the referent, and therefore *he* cannot be truly generic, as the selection of the masculine pronoun imbues the imagined referent with male sex, thus excluding biologically female referents. My results thus support the claims of Martyna (1980), Hellinger and Schräpel (1983), Crawford and English (1984), Crawford and Chaffin (1986), and others, that the use of the masculine pronoun for generic reference precludes a gender-inclusive interpretation,

and thus excludes women as potential referents of the antecedent/pronoun combination.

Therefore, the selection of generic *he* with antecedents whose referents are more easily visualised is not necessarily a matter of epicene reference, as the definiteness of the antecedent may influence pronoun choice (and the imagined referent may be assigned a biological sex). Also, the stereotypical value of the antecedent may come into play, insofar as it is less likely for feminine-stereotypes such as *nurse* to be visualised as masculine and coindex with *he*. My results from the stereotype analysis of the BE06 antecedents indicate that, in line with the argument that *he* has a default masculine interpretation (see section 1.2.2), generic *he* has a stronger tendency to occur with masculine-stereotypes than singular *they* (although this result is marginal). I argue that the use of generic *he* with masculine-stereotypes means that interpretation of the pronoun will default to the masculine, and thus referents will be visualised as male-only. Thus, the generic nature of such usage is again called into question.

My results from the stereotype analysis of the BE06 antecedents, and the correlation between generic *he* and definiteness of referent, suggests that it is incorrect for Madson and Hessling (1999:569) to claim that generic *he* does not “necessarily provide information about the characteristics of the referent”. Even children appear to process the form as a male-only pronoun. In chapter one I documented Bennett-Kastor’s (1996:290) research on the epicene comprehension of nine to twelve-year-old children which showed that *he* and *she* were only perceived gender-neutrally 4% and 2% of the time respectively, whilst their gender-specific interpretation reached 53% and 61%. However, I also noted Bennett-Kastor’s (1996:287) acknowledgement of Fisk’s (1985) study which indicated that not only was generic *he* interpreted as a masculine pronoun in children, singular *they* also had this interpretation “prior to first grade in boys and prior to kindergarten in girls”. Thus, Fisk’s results suggest that children as old as seven may interpret *they* as a masculine form.

Yet, Fisk’s results are not repeated anywhere else in the literature, either due to the lack of research on children’s pronouns, or because they are anomalous and the results could have been a feature of the experiment. Such results were certainly not repeated in my data, but warrant further investigation. In relation to Bennet-Kastor’s figures however, my results

from the CHILDES corpus suggest that for younger children, still acquiring the pronoun paradigm, the figures for epicene interpretation of *he* are even smaller than 4%, as when the problematic examples of generic *he* were removed from the sample generic *he* did not occur in the children's output. Thus, my research has expanded on the findings of Bennett-Kastor's work to show that the likelihood of a generic interpretation of *he* is even smaller than her results indicate.

However, one downfall of the corpus methodology chosen for this research is that, there is no way to test whether the gender-stereotype value of an antecedent was definitely a factor in pronoun choice. There is no way to question the authors of the texts in the BE06 sub-corpora about any subconscious mental images they may or may not have had when writing sentences which included generic *he*. Nor is it possible to test the language-acquiring children in the CHILDES corpus on their visualisation of pronoun referents, as children of that age do not possess the relevant meta-language to provide such information (and the children in the Manchester Dataset will now be teenagers). Therefore such arguments may need further research, although the case in the previous literature for the default masculine interpretation of generic *he* is rather solid (section 1.2.2).

Nevertheless, despite the limited occurrence of tokens of generic *he* that my analyses have highlighted - whether or not they are truly generic - it is clear that the results from the BE06 sub-corpora, combined with the (limited) acceptance of singular *they* in grammar books (see section 7.2), and its use in the CHILDES corpus (section 7.3), show that singular *they* is the epicene of British English. Based on the BE06 data I echo Green (1977:152) who argued thirty years ago that singular *they* represents "normal usage", a claim which still holds across the data in chapter four and chapter six.

In response to research question one, and my hypothesis that singular *they* is the epicene of choice in British English, my data has shown singular *they* to be a versatile epicene, occurring with all antecedent types, including definite NPs, and co-occurring with antecedents that are either lexically specified or stereotyped for gender. These results provide strong support for the argument that *they* is a viable epicene pronoun (c.f. section 1.2.2) which can be processed unproblematically with different antecedents,

without causing a gender mismatch effect, or imbuing imagined referents with a biological sex.

The finding that singular *they* is the epicene of choice across the BE06 data can be contextualised within the history of the pronoun paradigm that I presented in section 1.1.1. In my discussion, I showed that the use of *they* as a singular epicene form has a long and relatively well documented history. For example, Nevalainen (2006) argued that singular *they* can be found alongside generic *he* in documents from Middle and Early Modern English, and data from the *OED* shows that the first citation of singular *them* was in 1389. Thus, the use of *they* as a singular form evident in the modern data I have presented is not a new phenomenon, and thus my research emphasizes the continuing use of the pronoun, despite language-external factors, such as its rejection by traditional grammatical prescriptivists.

7.2: EPICENES AND LANGUAGE-EXTERNAL FORCES

As part of my analysis of modern grammar texts in chapter five I showed that the authors of newer grammar textbooks appear to be aware of epicene history and the promotion of generic *he* within traditional grammatical prescriptivism. Not only was generic *he* classified as old-fashioned usage, my keyword searches in relation to research question three indicated that four of the grammars specifically used the term *traditional* to label generic *he*. The occurrence of this keyword in the grammar corpus reinforces the notion that generic *he* was favoured as the “correct” form in traditional grammars, and thus adds to the large body of research, discussed in chapter two, which supports the conclusion that generic *he* was the promoted epicene in traditional grammatical prescriptivism (c.f. Bodine 1975; Stanley 1978; Baron 1986). Such results support my hypothesis that there would be explicit reference to language-external factors in the modern grammar handbooks I analysed.

Furthermore, as with my data on epicene usage, discussed above, my grammar corpus analysis brings data on epicene prescriptions up to date, as before this research the most recent published data on epicenes in grammar texts was Zuber and Reed’s eighteen-year-old study on American college

handbooks. Thus, the work I have presented in chapter five not only modernises research on the treatment of epicenes in educational materials, it also presents new data on grammar guidelines for British English. In relation to my hypothesis that modern grammars would continue the trend evident in the previous literature and reject generic *he* in favour of epicene avoidance tactics, my results do seem to support this position. However, my results also go beyond this hypothesis, as the increased reference to singular *they* in grammatical prescriptions could not be hypothesised based on previous research. Therefore, the results in chapter five update the state of knowledge about epicene prescriptions in grammar handbooks; no longer is there the out-and-out rejection of singular *they*, or the endorsement of generic *he* that was found in older studies.

However, it is important to note that there was not uniform treatment of epicenes in the grammar corpus as different authors endorsed different candidates to different extents (see section 5.1.1). Nevertheless, whilst my analysis of modern grammars illustrated that the authors did not all take up the same position in relation to epicene pronouns, the consistency of mention of epicene pronouns I found in chapter five highlights the position of epicene pronouns as a debating point of English grammar. I showed that thirteen out of twenty grammars included a discussion of epicene reference, thus providing supporting evidence for claims such as that made by Gershuny (1989:98), who argued that grammars consistently “include short discussions” of epicene pronouns. If Gershuny’s statement is correct, then my results, combined with previous works on epicenes in grammars (section 2.3.1), indicate that the problem of epicene reference has been, and still is, a regular feature of English grammars.

I also argue that the conflict over epicenes, still evident in the grammar corpus, indicates that traditional grammatical prescriptivism, which promoted generic *he*, has influenced, and indeed still influences epicene use. The findings of my literature review in chapter two illustrated that there was no linguistic reason to privilege number concord above gender concord and promote generic *he*. Yet the limited use of generic *he* found in my research in chapters four and six illustrates that the traditional grammatical prescription of the form has impacted upon, and still influences language use, despite its foundation in non-linguistic arguments. Thus, although studies such as Pauwels (2001) have shown that non-sexist language

reform can influence language use and lead to decreasing numbers of tokens of generic *he*, my results indicate that the form has not been eliminated.

In contrast with previous corpus studies on grammar handbooks, such as Bodine (1975), Sunderland (1986), and Zuber and Reed (1993), the results from my analysis of modern grammars showed that the outright rejection of singular *they* is highly uncommon, with the vast majority of texts in the grammar corpus noting its usage. Indeed, the form is endorsed without caveat in one of the texts. My results indicate that whilst previous studies on corpora of grammar books may depict the state of prescriptions in the 1970s the modern data is different, and thus conclusions about epicene prescriptions should no longer be made based on the older material, as my analysis shows that earlier studies do not reflect current attitudes to epicenes in grammar textbooks.

I argue that, based on my grammar corpus results, there has been an observable impact of non-sexist language campaigns on the treatment of generic *he*, with the explicit reference to language-external factors I hypothesised at the end of chapter two clearly evident in my data. The data on agency in the grammar corpus, which I presented in section 5.2.2, illustrates that the grammar authors were aware that language-external factors may affect a person's choice of epicene (c.f. Harrigan and Lucic 1988; section 2.2.2). Although the grammar authors did not necessarily align themselves with groups associated with politically motivated language change (i.e. feminists), they could not deny the impact that non-sexist language campaigns have had. Their knowledge of such campaigns is evident in the occurrence of keywords such as *non-sexist language* and *feminism* (section 5.2.2).

The results of my analysis of the grammar corpus corroborate previous works, such as Green (1977:152) who attributed the rejection of generic *he* to the impact of sexual politics, and Jacobson and Insko (1985), Harrigan and Lucic (1988) and Rubin and Greene (1991) who showed that there is interaction between a person's political allegiance and their opinions/choice of epicene (section 2.2.2). The prescriptions in the grammar corpus suggest that, in the face of non-sexist language reform there has been a move away from the advocacy of generic *he*, and therefore the endorsement of the pronoun in such texts has died down since previous studies such as Bodine (1975), Sunderland (1986) and Zuber and Reed (1993).

Another key finding in my grammar corpus analysis is that, also in contrast with the previous works noted above, the prescriptions in the modern grammars are more in-line with the usage data from the BE06 sub-corpora (although the promotion of avoidance tactics still predominates, see below). Thus, it appears that, rather than rejecting language change, which is a criticism usually levelled at prescriptive grammar authors (section 2.1.1), the modern prescriptions actually more accurately reflect language. The results from the grammar corpus indicate that the generic masculine is now reviewed negatively in modern grammar texts and thus I must reject Silveira's (1980:174) claim that students are "rewarded" for using generic *he* and "punished for not doing so". My results also reinforce Pauwels and Winter's (2006) finding that 60% the teachers they polled (see section 2.3.3) would correct a pupil's use of generic *he*.

However, my close analysis of the treatment of epicenes in the grammar corpus also highlighted the fact that limitations are still placed on singular *they* due to its classification as a purely plural pronoun. None of the texts in my grammar corpus analysis in chapter five depicted a paradigm including singular *they*. Thus my results can be used to support Madson and Hessling's claim that singular *they* cannot be separated from the prescription that "it is grammatically incorrect to use plural pronouns to refer to a singular antecedent"(1999:561), as in the grammar corpus the form *they* is still classified as a purely plural form. Therefore, my research has shown that *they* has not been considered as a formally singular pronoun in the grammars analysed in chapter five (or in the wider literature on the pronoun paradigm, see section 7.4).

Moving on to my consideration of the example sentences used in the grammar corpus, an area of research which, as yet, has not received much space in the wider literature, I found that analysing the sentences given in the grammars illuminated the simplification of the epicene debate by grammar authors. The results from the analysis of the antecedents of the example sentences indicate that, rather than engage with the more problematic aspects of the epicene debate, such as the coindexation of singular *they* and a singular definite NP (see section 7.4), the grammar authors chose unproblematic antecedents to illustrate their positions. The overwhelming predominance of indefinite pronouns present in the example sentences of the grammar corpus illustrates that the texts do not reflect the

proportional use of singular *they* with different types of syntactically singular antecedent found in the BE06 sub-corpora (section 4.2.1). Although it is not clear whether the grammar authors in the corpora used in previous studies, such as Zuber and Reed (1993), also used this tactic, my findings indicate that in the post-2000 texts of chapter five, the grammar authors present only a simplified view of epicene reference.

In any case, the predominance of epicene avoidance tactics, which is evident in the results from the grammar corpus, suggests that some of the modern grammar authors avoid endorsing epicene reference altogether. They choose not to promote singular *they* or generic *he*, instead opting for alternatives such as recasting or combined pronouns. My analysis of the modern grammars has shown that the avoidance tactics noted by Treichler and Frank (1989) which I summarised in section 2.3.2 still occur in discussions of epicene pronouns. Although not all of Treichler and Frank's noted tactics are explicitly evident in the modern data, the parallels between their data and mine highlight the fact that the choice of proposed alternatives for epicene pronouns has been relatively static in grammars since their work twenty years ago. This finding compounds my argument that the grammar authors in chapter five simplified their discussions of epicene reference, using well established alternatives to endorsing a particular epicene pronoun.

However, whilst my results provide some support for Pauwels' (1998:181) argument that singular *they* may still be criticised "on the grounds that it is linguistically incorrect", its position in the grammar corpus suggests that it is currently receiving a more favourable review than it has done in previous studies with similar methodologies. Far from grammarians endorsing generic *he* in an "androcentric tradition" (Bodine 1975:139) and dismissing singular *they*, there is much more support for the latter form in the modern texts. Based on my results I reject the claim of Crawford and Fox (2007:483) that the proposal of an alternative epicene to replace the traditionally endorse generic *he* has "failed spectacularly". The data from the grammar corpus shows increased endorsement of singular *they* in comparison with previous research. In addition, the data from the BE06 sub-corpora, which indicates that singular *they* is a viable epicene that is used in British written standard English, also provides evidence for my claim that Crawford and Fox are incorrect in their assertion. Whilst the limits imposed

on the use of singular *they* in the grammar corpus suggest that the traditional grammatical prescription on number concord still holds to some extent, the use of singular *they* in the BE06 sub-corpora indicates that it is the majority epicene.

Furthermore, the results from the grammar corpus which showed overwhelming rejection of generic *he* thus goes against Silveira's (1980:174) argument (discussed in section 2.3.1) that when using grammar handbooks pupils can "develop a people=male bias", as such guidelines no longer predominate in modern grammar texts. The results from the grammar corpus suggest that Bodine's (1975:1339) assertion that grammarians are "the docile heirs to the androcentric tradition of the prescriptive grammarians" no longer holds. The rejection of generic *he*, and the overt mention of feminism and non-sexist language in the grammar corpus, suggests that feminism-based language reform has had an impact on epicene prescriptions. No longer is singular *they* the target of rejection on the grounds that it violates number concord, in fact it receives some limited support in modern grammar texts.

In contrast with the argument that grammars generally reject language change (c.f. Drake 1977; Aitchison 2001), my results show that the prescriptions in modern grammars more accurately reflect the language use in my corpus analysis of written British English, with the increasing endorsement of singular *they*. Although there are limitations placed on its usage, and epicene tactics predominate in the corpus, the (limited) positive evaluation of singular *they* evident in my data, was not found in the previous research I discussed in section 2.3. I also showed that the authors of the modern grammars simplified their example sentences with unproblematic antecedents, such as indefinite pronouns, or NPs that were not gender-stereotyped, thus avoiding the more controversial aspects of the epicene debate. This simplification of the issue corresponds well with (more general) claims about how grammar texts do not accurately reflect language use (see Zuber and Reed 1993 in section 2.3.3).

By undertaking an analysis of modern grammars I have updated the literature on epicene prescriptions, providing new data on the treatment of epicene pronouns in British English. In addressing research questions two and three, I have shown that epicene pronouns are still a point of contention in English grammars, yet in contrast with previous studies (Bodine 1975; Zuber

and Reed 1993), my results show that generic *he* is now rejected in modern texts. Therefore, previous studies on corpora of grammars no longer reflect the state of epicene prescriptions.

7.3: THE ACQUISITION OF THE EPICENE PRONOUN

Whilst the previous two sections have dealt with epicene use and prescription, both topics that have been researched relatively widely before, my literature review in chapters one and two indicated that there has been little-to-no research undertaken on the acquisition of epicene pronouns. In chapter one I questioned the role played by acquisition in the construction of the personal pronoun paradigm, and the potential inclusion of an epicene pronoun, but noted that this area of interest is not well developed in current scholarship. My analysis of the CHILDES transcripts in chapter six thus begins to provide data on children's experiences of epicene pronouns, highlighting the types of epicenes they use and receive as L1 input. Although there has been some limited research on children's epicenes, notably Bennet-Kastor's (1996) study discussed in section 1.2.2, there are no previous studies on the epicene usage of children still in the stages of pronominal acquisition.

I hypothesised, based on the evidence of previous studies on spoken epicenes (c.f. Stringer and Hopper 1998), that the children in my corpus would receive singular *they* as L1 input. Positively, my analysis of the CHILDES data indicates that this hypothesis holds, and furthermore, despite limited input from their primary caregivers, three out of the four children I considered actually used singular *they* as an epicene pronoun (although there were issues concerning animacy). In addition, none of the previous works in the literature on spoken epicene pronouns, such as Newman (1992), Stringer and Hopper (1998), and Pauwels (2001), are primarily concerned with British English. Therefore, although my sample was relatively small, the CHILDES corpus also provides new data on the use of spoken epicenes by British children and British adults.

However, one unexpected finding in my analysis was the prevalence of generic *he* evident in the CHILDES corpus. The occurrence of generic *he* could be used as evidence against my hypothesis about epicenes in L1 input,

as whilst I never stated explicitly that generic *he* would not occur in the data, its non-occurrence is potentially inferred from the hypothesis that singular *they* would occur. Yet, a closer analysis of the CHILDES corpus showed that the tokens of generic *he* used by the children were entangled with issues of animacy and personification. Furthermore the children did not receive unproblematic tokens of generic *he* as L1 input from their primary caregivers. Therefore, it appears that my hypothesis is somewhat unaffected by the anomalous occurrences of (pseudo)generic *he*.

A closer analysis of the problematic tokens of *he* yielded an unexpected, yet highly significant finding in the CHILDES data. Although my results indicate that children do not use generic *he* for animate singular (human) antecedents, the use of indefinite forms of *he* is relatively common when compared to the data for *they* and *you*. My analysis shows that the children in the CHILDES corpus use *he* in a very interesting way, as it appears to be the default pronoun for the personification of inanimate objects, such as toys, and it is also the preferred pronoun for animals or imagined referents (such as Becky's *bunny* in section 6.3.2). I argue that this use of *he* reinforces the fact that the pronoun specifies gender and cannot be used generically, as to make a toy or character "real" to a child they assign the toy a gender identity. Such uses of *he* are not related to epicene reference, as I argue that a toy is not intended to be a genderless generic, rather it is meant to be a particular entity, complete with a gender specification¹⁵⁴.

I am aware that this is a strong claim, and currently, there appears to be no conclusive evidence on this topic in the language acquisition literature. However, Nilsen (1977) did test for something similar when she asked fifty pairs of children aged between four and twelve to describe animals to each other. Nilsen noted that out of a total of 3020¹⁵⁵ pronouns used, only seventeen were feminine, despite the fact that the images the children saw were not marked for gender in any way. Her findings led her to conclude that "unless we have evidence to prove than an animal is female, we use the masculine pronoun, whether or not we make a conscious decision that the

154 Although not tested specifically in this thesis, potential further research based on the finding that *he* is used as a default pronoun for personification could include investigating whether a similar phenomenon occurs with the pronoun *she*.

155 Out of the total number of tokens, 1461 pronouns were neuter, in this case *it*, and 1524 were masculine, showing a clear trend towards the children characterising the animals as male.

animal is male” (1977:170). Therefore, using Nilsen’s results as a foundation for interpreting the CHILDES data, it appears that children use the pronoun *he* as a default, but there is no indication that their use of the pronoun is meant to signify that the referent can be either male or female. Similarly, in her analysis of children’s conversations Weatherall (2002:771) notes that “there were no cases in which the children spontaneously ‘noticed’ or repaired the default assumption of a male referent in relation to a pseudo-generic”. That is, when *he* was used to refer to an unknown referent the children did not challenge the use of the pronoun; even though the referent could have been female, they accepted the default masculine that *a fire fighter* or *the dog* was presumed to be male.

The finding that *he* is used in this complex manner is not evident in previous epicene (or language acquisition) research, and thus my research has made a welcome, but unexpected contribution to wider debates on pronoun usage. For example, such data could be used in sociolinguistic debates on whether the masculine is (still) the unmarked form in English. Further research is needed on this topic, but it is sufficient for my arguments, and my hypothesis, to show that children do not use *he* as a generic pronoun, and to show that they do receive examples of singular *they*, without the complications associated with personification and anthropomorphism.

In the CHILDES data three of the primary caregivers used singular *they* in speech to their children at a comparable ratio to their use of plural *you*, which children undisputedly acquire¹⁵⁶. Significantly, the CHILDES results also show that the children in the study use singular *they* proportionally more than their primary caregivers (although I concede that some of these tokens may be problematic due to under specification of the phi-feature [ANIMATE] in the children’s still-developing mental pronoun paradigms). Nevertheless, this result is directly comparable with Meyers’ (1990) short note of an unreferenced study where Minnesota children used singular *they* more frequently than their parents. Such data can be accounted for by the argument that the children had developed different pronoun paradigms from

156 That is, there is no current evidence that children do not acquire a plural and singular form of *you*, and the second-person forms are generally considered to be bona fide members of the closed-class pronoun paradigm. The use of *you* for both plural and singular reference is not disputed in the current literature.

their parents, and had incorporated singular *they* into their mental grammars, thus reinforcing the arguments of Yang (2000) and Egerland (2005), which I discussed in section 1.1.3.

My data can be used in support of the argument that singular *they* can be acquired alongside the plural form of the pronoun (see section 7.4). The results of my analyses of child language support previous research, such as Bennett-Kastor's (1996) study which shows that children can process singular *they* unproblematically in both production and comprehension. The CHILDES data provides more evidence in support of my claim that singular *they* is the epicene pronoun of choice in British English, supporting the results from the BE06 sub-corpora. My results also reinforce the idea, first noted by Graham (1973) in her study on pronouns in educational materials, that children are not exposed to generic *he* (in any great quantity) in L1 input, and thus provide more evidence that generic *he* is not part of L1 acquisition. Consequently, I argue that if children do not receive generic *he* as L1 input, then the form cannot become part of the pronoun paradigm, as lack of exposure to the form means that it cannot be incorporated into the closed-class paradigm as it is being built in a child's mental grammar. My results present further evidence for the argument that, if generic *he* is ever learned, this process takes place after the pronoun paradigm is already fully formed (c.f. Graham 1973; Nilsen 1977; Bennett-Kastor 1996).

As part of my argument surrounding the acquisition of singular *they* (see section 1.1.3) I noted that it is unproblematic to have a singular and a plural pronoun which are morphologically identical, yet syntactically distinct, as the current English personal pronoun paradigm has such a set of pronouns in the second-person. It should not therefore be a problem if the same syncretism that accounts for both forms of *you* (see Harley 2008 discussed in section 1.1.3) occurred in the third-person with the two forms of *they*. Not only would such a pattern create regularity in the pronoun paradigm, it would also be in line with Baron's (1986:193) observation that it is a common pattern to use plural pronouns as singular forms in English (e.g. singular *you*, the royal *we*).

The results from the CHILDES data showed that both singular and plural *you* were provided in input (and produced by the children) with similar relative frequencies for singular and plural *they*, which leads me to argue that, if limited occurrence of plural *you* in input is sufficient for

children to acquire both forms of the pronoun, then there is no reason why a similar proportion of occurrences of singular *they* cannot facilitate the acquisition of both a plural and singular form of the pronoun. By identifying the parallels between *you* and *they*, in terms of their availability in L1 input during the timeframe for pronoun acquisition, I have provided evidence for my argument that there is regularity in the pronoun paradigm. The implications of the identification of this pattern is that arguably, Homonymy Theory (which is discussed in section 7.4) can account for the syncretism in both the second and third-person forms, insofar as the acquisition parallels can facilitate the inclusion of two entries for *you* and two for *they* in the mental lexicon.

The pattern in acquisition data between *you* and *they* supported by the CHILDES data cannot be overlooked, as any parallels in a rote-learned system are significant. I see no problem in adapting the mechanism children use for acquiring both the singular and plural forms of *you* to the acquisition of two forms of *they* (although this thesis does not focus explicitly on what this mechanism may be). This process of acquisition would actually represent a pattern in the usually irregular pronouns. The rigidity of the pronoun paradigm is well established in wider research on pronouns and on closed-classes in general. Chafetz (1994:275) claimed that even children are aware that closed-classes are “not likely to admit new members”. However, there have been some studies (e.g. Gerken 1987) which have shown that children can adapt to new pronouns whilst they are still acquiring the forms (see section 1.1.3), and thus, despite Chafetz’s assertion, it is still possible for a closed-class to change through the acquisition of different grammars in each generation (c.f. Yang 2000). My CHILDES data showed that, in comparison with the data for plural *you*, singular *they* occurs proportionally enough in input to arguably facilitate its acquisition.

Thus in relation to research question four, my evaluation of the data from the CHILDES corpus indicates that the conditions for the acquisition of singular *they* are met by the caregiver, insofar as the proportional values of singular and plural *you* in the input data were comparable to the values for singular and plural *they*. Having shown that that singular *they* is the most-used epicene in written British English, based on the strong evidence from the BE06 sub-corpora, in this section I have provided evidence which

supports the position that it is possible to acquire singular *they* as a bona fide member of the pronoun paradigm in a person's mental grammar.

The data I have provided in the CHILDES analysis is, to my knowledge, the only current analysis of the epicene usage of children whilst they are still acquiring their closed-class forms. Previous research, such as Bennet-Kastor's (1996) study involved children who were well beyond the age of pronoun acquisition (using figures from Rispoli 1994, and Owens Jr 2007). Mine is also the only research undertaken to document the direct input children are given during this period. Therefore, by completing this study I have added to the literature on epicene pronouns, and the literature on how linguistic input from primary caregivers may impact upon which forms children incorporate into closed-classes in their mental grammars. Furthermore, I have provided evidence for my hypothesis that singular *they* is present in L1 input in British English.

7.4: EXPLAINING EPICENE REFERENCE: THE HOMONYMY OF THEY

The final issue that I intended to tackle in this thesis is the disagreement in the literature over how to account for syntactic coreference between *they* and a formally singular antecedent. In my discussion of the development of the pronoun paradigm in section 1.1.1 and my consideration of grammatical prescriptivism in chapter two, I showed that the pronoun has historically been classified as a purely plural form. However, in section 1.3.3 I began to challenge this notion, arguing that it is possible that there are two forms of *they* in the mental lexicon, one of which is marked [-PLURAL] and can formally agree in number with singular antecedents. In this section I explore how my results feed into this theory.

Hypothesising that Homonymy Theory can account for number concord between *they* and a singular antecedent is a novel approach to epicenes which has not been considered in the wider literature. As I noted in chapters one and two the vast majority of research on singular *they* treats the pronoun as a mere manipulation of the plural form, whether or not the researchers state this explicitly. For example, Weidmann (1984:68) argues for the versatility of *they* but claims that its plurality is its "only drawback",

whilst Baranowski (2002:378) notes that the grammarians who proscribed singular *they* classified it as “purely plural” noting that using the pronoun as a singular “disagreed in number” with its antecedent. The presupposition behind such arguments is arguably that the singular form is in some way deficient, or not a true member of the pronoun paradigm. My results from the grammar corpus indicate a continuation of this presupposition, as none of the texts analysed in chapter five consider the idea that singular *they* is anything other than an alternative use or manipulation of the plural form, nor do they draw parallels between the second and third-person pronouns.

However, the results of my analyses can be used to contradict this argument, insofar as singular *they* is used consistently in my BE06 data as a workable pronoun of written British English, and thus there is no evidence to suggest that singular *they* is anything other than a bona fide personal pronoun. In contrast with previous assumptions about singular *they* I argue that it is not a reinterpretation of the plural pronoun; it is instead its own form in the mental lexicon. The singular form is morphologically identical to plural *they* but the two forms are differentiated on syntactic grounds as they are differently marked for number in their phi-feature specifications. I argue that the evidence from the CHILDES corpus illustrates that singular *they* is both received as L1 input (in a relative proportion to plural *you*), and is produced by children acquiring the personal pronoun. Such evidence supports the argument that singular *they* can be acquired as its own form¹⁵⁷. Therefore I reject the claims in the literature that *they* is simply “very tolerant” of different antecedents as Sanford et al. (2007:373) claimed, or just “accommodated as an acceptable deviation” from the plural (Sanford and Filik 2006:373).

A key piece of research in my proposition of Homonymy Theory is that of Harley (2008:275; section 1.3.3), who notes that there are three English metasyncretisms in the pronouns; the first is that “gender is not marked in the personal pronouns (first and second-person)”, the second is that “gender

157 As an additional note, Storkel and Young (2004) showed how a group of twenty-eight children of circa four years’ old could acquire homonyms for words they already knew faster than a set of novel words. Their results showed that “known phonological forms appeared to facilitate learning” (2004:482). Furthermore, Storkel and Young argued that their results supported the hypothesis that “children collect homonyms to promote rapid expansion of the lexicon” and therefore contradicted “claims that children avoid homonymy in word learning or fail to recognise form similarity” (2004:582). Applied to the personal pronouns, this research can also be used in support of Homonymy Theory.

is not marked in the plural pronouns” and finally, “number is not marked in the second-person”. I propose that the final metasyncretism can also include the criterion that number is not marked in the third-person when singular *they* is used for generic reference. Alternatively, and perhaps more neatly, it could be argued that the metasyncretism extends over the whole personal pronoun paradigm, as number is not marked on any pronoun, except where gender is also explicitly marked. Such a metasyncretism would fit in with Noyer’s (1992) feature hierarchy, which shows that that gender marking does not occur in a language unless number marking is already present.

My CHILDES data can also account for why Homonymy Theory cannot apply to the acquisition of generic *he*. I showed in section 6.3.2 that once the problematic tokens of generic *he* were eliminated from my data, the form did not occur in the caregivers’ speech in sufficient numbers to facilitate its acquisition, insofar as the percentage of tokens was well below the threshold of the tokens for both forms of *you*. Therefore I argue that the children in my data were not exposed to generic *he* often enough to incorporate it into their pronoun paradigms (potentially adapting the [+MASCULINE] [+FEMININE] phi-feature hypothesised in section 1.3.3). Based on the CHILDES data I argue that the extremely small proportion of tokens of generic *he* in L1 input (which are still problematic), means that children will not acquire a generic version of *he*, yet they receive enough tokens of the masculine use of *he* to acquire it as a member of their pronoun paradigms. The acquisition data therefore reinforces the results of the many studies I noted in chapter one which showed that generic *he* is almost always interpreted as a masculine pronoun, as well as indicating why Homonymy Theory cannot apply to both generic *he* and singular *they*.

As part of my argument that singular *they* is not a manipulation of the plural form I also contest Frank and Treichler’s (1989) claim that singular *they* may be “inherently prone to ambiguity of reference because it increases the number of potential antecedents” (cited in Newman 1992:456). Throughout the data from the BE06 sub-corpora it is clear that singular *they* only selects one real-world referent when it occurs with NPs with quantifiers, indefinite NPs, or definite NPs. It is not the case that a singular NP coindexed with singular *they* refers to multiple, or plural referents. I concede that the coreference of singular *they* with indefinite pronouns may be slightly more problematic, insofar as it is more difficult to imagine a

single specific referent (see section 7.1), but I argue that this is due to the semantic value of the indefinite pronoun, not the syntactic number marking on singular *they*. I have already established in section 1.3.1 that indefinite pronouns are formally singular NPs, and thus to avoid processing problems, singular *they* must match the pronoun for number in the deep structure. Using Homonymy Theory this number concord does occur.

Homonymy Theory can also account for why the use of singular *they* is not just a matter of variable reference, as the focus here is not on the semantic interpretation of pronouns, but on their syntactic processing. Using Homonymy Theory, singular *they* is syntactically singular in the deep structure of sentence processing in the brain, independent of whether the use of the pronoun refers to a specific antecedent or not. Therefore, arguments from traditional prescriptive grammar that singular *they* represents a violation of number concord (section 2.1.2) are unfounded.

One pitfall in my proposition of Homonymy Theory is that such theoretical concepts are difficult to test outright. It is not possible to ask language acquiring children how they are processing certain forms, as they have not yet mastered the basics of language, let alone gained the terms to express their metalinguistic knowledge. Therefore, on this point the CHILDES data is limited. Equally it is difficult to test adults on how they process plural and singular *they* without interference from language-external factors such as non-sexist language reform and grammatical prescriptions. For example, in Pauwels and Winter's (2006) study on the epicene use of teachers in Australia, at least one participant explicitly stated that they knew singular *they* was not prescriptive convention, yet they chose to use it anyway (section 2.3.3). Therefore, it is not unscientific to argue that an awareness of such social norms may influence participants in an experimental setting. Indeed, in section 7.1 I have argued that the continuing use of generic *he* in the BE06 data can be attributed to the influence of language-external factors, such as the prescription of the form in traditional grammar.

I have argued that the data I presented in chapters four, five, and six, largely supports my adaptation of Whitley's (1978) Homonymy Theory in which I propose that there are two morphologically identical, yet syntactically distinct forms of *they* in the mental lexicon. I showed how the CHILDES data is sufficient to support the argument that singular *they* is

received in L1 input, and therefore can be acquired as an entry in a child's mental grammar. The parallels between *they* and *you* evident in the acquisition data I presented in chapter six allowed me to argue that the metasyncretism which covers number marking in the second-person pronouns can be extended to singular and plural *they*. In addition, my CHILDES data indicated that the children in the data received insufficient tokens of generic *he* to acquire it as anything other than a masculine-only pronoun, thus providing evidence for my claim that Homonymy Theory cannot apply to generic and masculine *he*.

Whilst I acknowledge that I have not been able to test the theory directly within the bounds of this thesis, my corpus-based data is still useful in supporting my arguments. I have shown that, even though there has not been any consideration in the wider literature, nor in the grammar corpus, of the notion that singular *they* is anything other than a manipulation of the plural pronoun, the results from my BE06 analysis show that *they* is consistently used as a singular pronoun with different types of syntactically singular NPs, and selects only singular referents.

CONCLUSIONS

Based on the studies I presented in chapters four through six, and the previous research that I discussed in the first two chapters of this thesis, I argue that it is quite clear that the epicene of choice in written British standard English is singular *they*. My work has expanded upon the results of previous corpus-based studies, such as Newman (1998) and Baranowski (2002), to show a continuing trend for the predominance of singular *they* in epicene reference. Also, my results reinforce the conclusions of studies on spoken epicenes, such as Newman (1992), Stringer and Hopper (1998) and Pauwels (2001) which showed that singular *they* was the majority epicene in different varieties of spoken English. My CHILDES data shows that most of the primary caregivers analysed in chapter six used singular *they* in relative proportion to plural *you*. These results are evidence in support of my arguments about how singular *they* can be acquired as its own form, as children receive it as input (section 6.2.2).

However, despite the positive evidence from the CHILDES corpus, the usage data I provided from my analysis of the BE06 sub-corpora showed that generic *he* is still used alongside singular *they* and thus there has not been a blanket acceptance of one particular epicene. This parallel usage of singular *they* and generic *he* can be accounted for in terms of the conflicting language-external forces influencing pronoun choice. The concordance data from the grammar corpus in chapter five indicates that the prescription of generic *he* in traditional grammars, and the formal tuition of grammar “rules” during the twentieth century (see Paterson 2010), is in direct opposition to non-sexist language reform, and the promotion of singular *they* as the English epicene. I argue that the conflict of these two language-external positions on epicene usage is the primary factor for the predominance of avoidance tactics that was evident in the grammar corpus data.

Thus the limited use of generic *he* compared with singular *they* in my corpus studies is reflected in the changes in epicene guidelines evident when my data is compared to previous scholarship on grammar books. Whilst the continued use of generic *he* can be explained by strict adherence to traditional prescriptive norms or formal guidelines (in written British English at least), my documentation of changes in the treatment of generic *he* can be used to account (in part) for the decrease in usage of generic *he* when data from the BE06 sub-corpora is compared with older research. For example, my analysis of a sample of the 1961 LOB corpus in Paterson (2011) showed that generic *he* was the epicene of choice, whilst a newer corpus of 2007-2008 newspaper texts was dominated by the use of singular *they*. The BE06 analysis reinforces the results of my previous study and reasserts the dominance of singular *they* as the epicene pronoun of British English (despite lingering limited uses of generic *he*).

On the other hand, given the long standing history of use of singular *they* - it has been found in documents from (at least) Middle English onwards (see Nevalainen 2006) - I argue that the form has (potentially) been part of people’s mental grammars for a long time, although unfortunately the lack of data in the literature on the acquisition of pronouns means that there can be no direct evidence for this claim. Nevertheless, there is some limited evidence available. For example, the processing data from studies such as Foertsch and Gernsbacher (1997) and Sanford and Filik (2006) indicates that singular *they* can be processed with singular antecedents by participants

who acquired their pronoun paradigms before any impact of non-sexist language reform. However, such an argument would need further research, and at present is beyond the scope of this thesis.

What I have shown is that singular *they* is present in the linguistic environment of language-acquiring children. Combining my CHILDES results with the results of studies such as Stringer and Hopper (1998), and Pauwels (2001) which show that singular *they* is the most common spoken epicene, and with analyses indicating that generic *he* occurs very rarely in child-directed materials (e.g. Graham, 1973), suggests that singular *they* will most likely be the epicene form that children are exposed to during their acquisition period. Assuming that a child possesses an innate predisposition to acquire pronouns as a closed-class which will become fixed in their mental grammars (thus following Chafetz 1994, see section 1.1.3), the acquisition of the paradigm must be influenced by the pronouns that children hear. As the child is exposed to singular *they* in speech, it is more likely to become part of their pronoun paradigm, than any other potential epicene form.

Finally, whilst my adaptation of Homonymy Theory is still in its infancy, the preliminary data presented here looks positive. For the theory to be supported, there would have to be evidence that, a) the singular form of the pronoun *they* is used in British English, b) that the form was available in L1 input during pronoun acquisition, and c) that singular *they* did not cause processing difficulties in the brain. I argue that throughout this thesis I have provided positive evidence for all of these three criteria, although there is always room for more research on this topic. Nevertheless, not only have I updated research on epicene pronouns, reaffirming the position of *they* as a commonly used gender-neutral singular pronoun, and highlighted the fact that there have been changes in grammatical prescriptions based on the impact of non-sexist language campaigns, I have also proposed a theory which can account for epicene reference with all singular antecedents.

CONCLUSION

My primary goal in this thesis was to analyse epicene pronouns in British English, focusing on two primary candidates: singular *they* and generic *he*. My review of the existing literature in the first two chapters led me to highlight four areas of interest which could not be fully explained by current scholarship. I thus set out to update the epicene debate by providing new data on the following:

- The use of epicene pronoun in written (and spoken) British English,
- The acquisition of epicene pronouns in British English,
- The prescription of epicene pronouns in British English, and
- Whether there was overt evidence of the impact of language-external factors on epicene use and prescription.

I hypothesised that my results would indicate that singular *they* is the epicene of choice in speech and writing in British English, and furthermore, that children would receive singular *they* as L1 input. In relation to the language-external elements of the epicene debate I hypothesised that modern grammar prescriptions would continue the trend evident in the literature and reject generic *he*, and that such prescriptions would be overtly influenced by language-external factors.

As my discussion in chapter seven has shown my hypothesis about the use of singular *they* in written British English does hold across my BE06 data, and my hypothesis about L1 input is supported by my analysis of the CHILDES corpus. In terms of prescriptions, my hypothesis about modern grammar books is upheld by my analysis of the grammar corpus, as my findings indicate that the trend to reject generic *he* in favour of epicene avoidance tactics does occur. In addition, my analysis of the keywords I set out in research question three indicates that there is some overt acknowledgement of language-external factors in modern grammar prescriptions (which are of course also examples of language-external forces on epicene usage), and therefore my final hypothesis about the impact of language-external factors on the epicene debate is also supported.

Using my chosen corpus-based methodology, the modern data I have provided highlights changes in epicene use and prescription that have not

been evident in previous research. As there has only been a limited amount of research on epicene pronouns in the last ten to twenty years, with most works drawing upon studies that are even older than this, I argue that the topic had become stale. One potential reason for this state of affairs is that the many studies on epicenes undertaken towards the end of the twentieth century gave complementary results (as illustrated in section 1.2). However, because researchers were still working with, for example, Bodine's figures on grammar prescriptions from 1975, the language change of the past thirty years was hidden. My data has shown that, contrary to previous works, the data for modern epicene usage and epicene prescriptions does not necessarily follow established patterns. Indeed, in many cases my data shows that the rejection of generic *he*, which was called for by those such as Bodine and others, has occurred, and whilst the corresponding formal endorsement of singular *they* is not yet the norm, its acknowledgement as a viable epicene candidate in the grammar corpus is arguably a step towards this.

In addition, my data on spoken epicenes taken from my analysis of the CHILDES corpus represents the investigation of a previously under-researched area of the epicene debate. Although generally, my results correspond with previous studies on spoken epicenes, such as Stringer and Hopper (1998) and Pauwels (2001), significantly my analyses show that generic *he* does not occur in child-directed speech. Conversely, I showed that, alongside both plural and singular tokens of *you*, singular and plural *they* were used by the caregivers with similar relative percentages. My data supports the argument that children receive singular *they* as L1 input. Thus, based on the proposals of Yang (2000) and Egerland (2005), as discussed in section 1.1.3, my evidence would suggest that children can incorporate singular *they* into their pronoun paradigms during language acquisition.

Furthermore, the phi-feature specifications for the two forms of *they*, which I proposed under my adaptation of Homonymy Theory (section 1.3.3), are easily adapted to fit the second-person forms, by changing person reference only, thus supporting the position that there are similarities between *they* and *you*. As the pronoun paradigm is rote learned, due to its irregularities (following the arguments noted by Rispoli 1994), I argue that any parallels between pronominal forms are significant, and thus posit that the similarities between the use of *they* and *you*, such as the similarities I

found in the CHILDES data, may aid the acquisition of the usually-irregular personal pronouns.

In addition, the unexpected finding of the CHILDES data analysis, which showed that children used masculine *he* as the default pronoun for personification and anthropomorphism, combined with the finding that children do not receive generic *he* as input, is evidence for the position that the epicene form of *he* is never acquired (c.f. Nilsen 1977; Martyna 1980; Gibbon 1999). Combining these results with the well-established argument that *he* carries its own semantic value, I am therefore in agreement with Martyna's statement, noted in chapter two, that epicene reference and the use of generic *he* as a gender-neutral pronoun "involves much more than what the Harvard linguists once termed 'pronoun envy'" (1980:70). This study therefore presents a potential starting point for further research on pronoun acquisition and the use of generic *he*.

Similarly, my adaptation of Homonymy Theory is also in its infancy and warrants further research. My argument that there are two forms of *they* (and two forms of *you*) in the mental lexicon, differentiated by the phi-feature [PLURAL] can account for why singular *they* is found in coreference with formally singular antecedents, not only in the corpus-data I have provided here, but also in the wider literature (c.f. Baranowski 2002). Homonymy theory can also account for the syntactic agreement of anomalous uses of singular *they* which have been found in other research, such as Green's (1977) observation of singular *they* with the lexically specified antecedent *actress*, and Weidmann's (1984) *Barbara Wassman* example (discussed in detail in section 1.3.1¹⁵⁸). Nevertheless, Homonymy Theory needs further testing, as due to the limitations of my research, I was unable to test my adaptation of Whitley's theory directly.

In the introduction I noted that it was possible to theorise a space in the personal pronoun paradigm for a gender-neutral third-person singular pronoun, as none currently exists in the standard English written paradigm.

158 In this thesis, I focused on accounting for syntactic reference only, and as such I do not consider the semantic implications of specific uses of singular *they* with definite referents, such as *Barbara Wassman*. However, the semantic value of the antecedent does not impact upon the applicability of Homonymy theory. Even if the speaker in Weidmann's example wished to express "uncertainly about the presence" of the referent (1984:65), there is still grammatical agreement between the pronoun and its antecedent whether the referent exists or not.

Based on the standard paradigm there is currently no pronoun which can mark singular gender-less reference with a non-gender-specific antecedent. My research has added to the body of literature that has shown singular *they* to fulfil this function, whether it is formally sanctioned by language institutions such as dictionaries and grammars or not. The synthesis of my findings with the wider literature on epicene pronouns leads me to the conclusion that, despite any such restrictions on the written pronoun paradigm singular *they* is the epicene pronoun of British English.

Furthermore, I argue that it seems unlikely that there would be a pronoun in common use (especially in language directed at children) that was difficult to process and which caused ambiguity. Singular *they* appears to be the path of least resistance in terms of sentence processing and gender-neutrality. Indeed, my data from the BE06 sub-corpora analysis indicates that *they* is an extremely versatile epicene pronoun, coindexing with a wide variety of syntactic forms of antecedent (and across a wide range of semantic fields). Therefore, based on my data I must conclude that singular *they* is the only epicene form currently and consistently used in British English.

APPENDIX A: THE GRAMMAR CORPUS

Below is a list of the texts included in the grammar corpus. The references highlighted in grey signify those texts which covered the topic of pronouns but did not discuss epicene reference.

Text 1	Klammer, Thomas P., Muriel R. Schulz, and Angela Della Volpe. 2000. <i>Analyzing English Grammar</i> (3 rd edition). Boston, Massachusetts; London: Allyn and Bacon.
Text 2	Leech, Geoffrey N., Benita Cruickshank, and Roz Ivanič. 2001. <i>An A-Z of English Grammar and Usage</i> (2 nd edition). Harlow: Longman.
Text 3	Biber, Douglas, Susan Conrad, and Geoffrey Leech. 2002 [2003]. <i>Longman Student Grammar of Spoken and Written English</i> . Harlow: Longman.
Text 4	Law, Jonathan (Ed.). 2002. <i>The Language Toolkit: Practical Advice on English Grammar and Usage</i> . Oxford: Oxford University Press.
Text 5	Azar, Betty Schramper. 2003. <i>Fundamentals of English Grammar</i> (3 rd edition). New York: Longman.
Text 6	Azar, Betty Schramper. 2003. <i>Fundamentals of English Grammar: Student Book</i> (3 rd edition). White Plains, New York; Harlow: Longman.
Text 7	Collins Cobuild. 2003. <i>Collins COBUILD English Grammar</i> . London: Collins.
Text 8	Humphries, Carolyn. 2003. <i>Really Simple English Grammar</i> . London: Foulsham.
Text 9	Jones, Eileen. 2005 [2008]. <i>Key Grammar. Book 1</i> . Harlow: Ginn.
Text 10	Carter, Ronald and Michael McCarthy. 2006. <i>Cambridge Grammar of English: A Comprehensive Guide: Spoken and Written English Grammar and Usage</i> . Cambridge: Cambridge University Press.
Text 11	Jarvie, Gordon. 2007. <i>Bloomsbury Grammar Guide</i> . London: A. and C. Black.
Text 12	Seely, John. 2007. <i>Oxford A-Z of Grammar and Punctuation</i> . Oxford: Oxford University Press.
Text 13	Sinclair, Christine. 2007. <i>Grammar: A Friendly Approach</i> . Maidenhead: McGraw-Hill/Open University Press.
Text 14	Stobbe, Gabriele. 2008. <i>Just Enough English Grammar Illustrated</i> . New York; London: McGraw-Hill.
Text 15	Taggart, Caroline and J.A. Wines. 2008. <i>My Grammar and I (or should that be 'me?'): Old-school Ways to Sharpen Your English</i> . London: Michael O'Mara Books.

Text 16 Collins, Peter and Carmella Hollo. 2009. *English Grammar: An Introduction* (2nd edition). Basingstoke: Palgrave Macmillan.

Text 17 King, Graham. 2009. *Collins Improve Your English*. London: Collins.

Text 18 Altenberg, Evelyn P. and Robert M. Vago. 2010. *English Grammar: Understanding The Basics*. Cambridge: Cambridge University Press.

Text 19 O'Sullivan, Nuala, and Geraldine Woods. 2010. *English Grammar Workbook for Dummies*. London: John Wiley and Sons.

Text 20 Pyrcz, Heather Elizabeth. 2010. *Writing with Style: Grammar in Context*. London: Oxford University Press.

REFERENCES

- Aarts, Bas. 2000. Corpus linguistics, Chomsky and fuzzy tree fragments. In Christian Mair and Marianne Hurst (Eds.). 2000. *Corpus Linguistics and Linguistic Theory: Papers from the Twentieth International Conference on English Language Research and Computerized Corpora*. Amsterdam: Rodopi. pp 5-13.
- Abney, S. P. 1987. The English Noun Phrase in its Sentential Aspect, Doctoral dissertation, MIT. Cambridge, Mass.
- Adamsky, Cathryn. 1981. Changes in pronominal usage in a classroom situation. *Psychology of Women Quarterly* 5(5): 773-79.
- Aitchison, Jean. 2001. *Language Change: Progress or Decay?* (3rd edition). Cambridge: Cambridge University Press.
- Amare, Nicole. 2007. Where is she? Gender occurrences in online grammar guides. *Research in the Teaching of English* 42(2): 165-189.
- Arnold, Jennifer E., Janet G. Eisenband, Sarah Brown-Schmidt and John Trueswell. 2000. The rapid use of gender information: evidence of the time course of pronoun resolution from eyetracking. *Cognition* 76(1):B13-B26.
- Baauw, Sergio. 2002. *Grammatical Features and the Acquisition of Reference: A Comparative Study of Dutch and Spanish*. New York: Routledge.
- Baine, Alexander. 1879. *A Higher English Grammar* (new edition). London: Longmans.
- Baker, Paul and Sibonile Ellece. 2011. *Key Words in Discourse Analysis*. London: Continuum.
- Baker, Paul. 2006. *Using Corpora in Discourse Analysis*. London: Continuum.
- Baker, Paul. 2009. The BE06 corpus of British English and recent language change. *International Journal of Corpus Linguistics* 14(3): 312-337.
- Balhorn, Mark. 2004. The rise of epicene they. *Journal of English Linguistics* 32(2): 79-104.
- Balhorn, Mark. 2009. The epicene pronoun in contemporary newspaper prose. *American Speech* 84(4): 391-413.
- Baranowski, Maciei. 2002. Current usage of the epicene pronoun in written English. *Journal of Sociolinguistics* 6(3): 378-397.
- Baron, Dennis E. 1986. *Grammar and Gender*. London: Yale University Press.
- Baron, Naomi S. 2001. *Alphabet to Email*. London: Routledge.
- Bate, Barbara. 1978. Nonsexist language use in transition. *Journal of Communication* 28(1): 139-149.
- Bennett-Kastor, Tina L. 1996. Nonanaphora and the generic use of pronouns by children. *American Speech* 71(3): 285-301.
- Bhat, Darbhe Narayana Shankara. 2004. *Pronouns*. Oxford: Oxford University Press.
- Biber, Douglas, S. Conrad, and R. Reppen. 1994. Corpus-based approaches to issues in applied linguistics. *Applied Linguistics* 15(2): 169-189.
- Black, Maria and Rosalind Coward. 1988. Linguistics, social and sexual relations: A review of Dale Spender's *Man Made Language*. In Deborah Cameron (Ed.). 1998. *The Feminist Critique of Language* (2nd edition). London: Routledge. pp 100-118.
- Bock, Kathryn, Anne Cutler, Kathleen M. Eberhard, Sally Butterfield, J. Cooper Cutting and Karin R. Humphreys. 2006. Number agreement in British and American English: Disagreeing to agree collectively. *Language* 82(1):64-113.
- Bock, Kathryn, Kathleen. M. Eberhard and J. Cooper Cutting. 2004. Producing number agreement: How pronouns equal verbs. *Journal of Memory and Language* 51(2): 251-278.

- Bodine, A. 1975. Androcentrism is prescriptive grammar: Singular “they”, sex-indefinite “he”, and “he or she”. *Language in Society* 4: 129-146.
- Boseley, Sarah. 2008. Many GPs fail to prioritise high-risk breast cancer groups. *The Guardian*. 1st January 2008.
- Bradley, D. C. and M. F. Garrett. 1983. Hemisphere differences in the recognition of closed and open class words. *Neuropsychologia* 21: 155-160.
- Brown, Goold. 1800 [1851]. *Grammar of English Grammars* (10th edition). New York: (no imprint).
- Brown, Goold. 1825. *A Key to the Exercises for Writing, Contained in the Institutes of English Grammar*. New York: (no imprint).
- Brown, Goold. 1832. Critical observations of Murray’s grammars. *American Annals of Education* 1(15): 558-559.
- Brown, Roger and Albert Gilman. 1989. Politeness theory and Shakespeare’s four major tragedies. *Language in Society* 18: 159-213.
- Burke, Nancy. 1998. *Gender and Envy*. London: Routledge.
- Burridge, Kate. 2010. Linguistic cleanliness is next to godliness: Taboo and purism. *English Today* 102 26(2): 3-13.
- Cameron, Deborah. 1992. *Feminism and Linguistic Theory* (2nd edition). New York: St Martin’s Press.
- Cameron, Deborah. 1995. *Verbal Hygiene*. London: Routledge.
- Cameron, Deborah. 1998. Feminist linguistic theories. In Stevi Jackson and Jackie Jones (Eds.). 1998. *Contemporary Feminist Theories*. Edinburgh: Edinburgh University Press. pp 147-161.
- Cameron, Deborah. 2006. *On Language and Sexual Politics*. London: Routledge.
- Cardell, William S. 1825. *Essays on Language*. New York: (No imprint). pp 29-30.
- Carreiras, Manuel, Alan Garnham, Jake Oakhill, and Kate Cain. 1996. The use of stereotypical gender information in constructing a mental model: Evidence from English and Spanish. *The Quarterly Journal of Experimental Psychology - Section A* 49(3): 639-663.
- Chafetz, J. 1994. The closed-class vocabulary as a closed set. *Applied Psycholinguistics* 15(3): 273:287.
- Chiat, Shulamuth. 1986. Personal pronouns. In P. Fletcher and M. Garman (Eds.). 1986. *Language Acquisition: Studies in First Language Development*. Cambridge: Cambridge University Press. pp 339-355.
- Choy, Jung Won Janet, and Cynthia K. Thompson. 2005. Online comprehension of anaphora and pronoun constructions in Broca’s aphasia: Evidence from eyetracking. *Brain and Language* 95(1): 119-120.
- Clark, Urzula. 2001. *War Words: Language, History and the Disciplining of English*. Bingley: Emerald Group Publishing.
- Cobbett, William. 1818. *A Grammar of the English Language in a Series of Letters Intended for the Use of Schools and of Young Persons in General, but More Especially for the Use of Soldiers, Sailors, Apprentices, and Plough-boys*. New York: (no imprint).
- Cook, Stanley J. and Richard W. Suter. 1980. *The Scope of Grammar: A Study of Modern English*. New York: McGraw.
- Cooper, Robert Leon. 1989. *Language Planning and Social Change*. Cambridge: Cambridge University Press.
- Corbett, G. 1991. *Gender*. London: Cambridge University Press.
- Crawford, M. and R. Chaffin. 1986. The reader’s construction of meaning: Cognitive research of gender comprehension. In E. Flynn and P. Schweikart (Eds.). 1986. *Gender and Reading: Essays on Readers, Texts and Contexts*. Baltimore: The Johns Hopkins University Press. pp 3-30.

- Crawford, Mary and Annie Fox. 2007. From sex to gender and back again: Co-optation of a feminist language reform. *Feminism and Psychology* 17(4): 481-486.
- Crawford, Mary and L. English. 1984. Generic versus specific inclusion of women in language: Effects on recall. *Journal of Psycholinguistic Research* 13: 373-381.
- Crawford, Mary. 2004. Gender and language. In Rhoda K. Unger (Ed.). 2004. *Handbook of the Psychology of Women and Gender*. New York: John Wiley and Sons. pp 228-244.
- Cruse, Alan. 2006. *A Glossary of Semantics and Pragmatics*. Edinburgh: Edinburgh University Press.
- Cruttenden, Alan. 1978. Assimilation in child language and elsewhere. *Journal of Child Language* 5: 373-378.
- Crystal, David. 1984. *Who Cares About English Usage?* Hammondsworth: Penguin.
- Crystal, David. 2001 [1995]. *The Cambridge Encyclopaedia of the English Language*. Cambridge: Cambridge University Press.
- Crystal, David. 2002. *The English Language* (2nd edition). London: Penguin.
- Crystal, David. 2004. *The Stories of English*. London: Penguin.
- Curme, George O. 1931. *A Grammar of the English Language Volume III: Syntax*. Boston: DC Heath.
- Curme, George O. 1947. *English Grammar*. New York: Harper and Row.
- Curzan, Ann. 2003. *Gender Shifts in the History of English*. Cambridge: Cambridge University Press.
- Cutler, Anne. 1993. Phonological cues to open- and closed-class words in the processing of spoken sentences. *Journal of Psycholinguistic Research* 22(2): 109-131.
- Daily Mail*. Chelsea celebrates Harry's return at a party with student pals. 1st March 2008.
- Dale, Jennifer and Peggy Foster. 1986. *Feminists and State Welfare*. London: Routledge.
- Dawson, Mary Wood. 1956. The passing of the pronoun. *English Journal* 45(1): 34-37.
- De Vincenzi, M. 1999. Differences between the morphology of gender and number: Evidence from establishing coreferences. *Journal of Psycholinguistic Research* 28(5): 537-553.
- Déchaine, R. M. and M. Wiltschko. 2002. Decomposing pronouns. *Linguistic Enquiry* 33(3): 409-442.
- Drake, G. 1977. American linguistic prescriptivism - its decline and revival in the 19th century. *Language in Society* 6(3): 323-340.
- Eckert, Penelope and Sally McConnell-Ginet. 2003. *Language and Gender*. Cambridge: Cambridge University Press.
- Egerland, Verner. 2005. Diachronic change and pronoun status: Italian dative *loro*. *Linguistics* 43(6): 1105-1130.
- Ehrlich, K. and K. Rayner. 1983. Pronoun assignment and semantic integration during reading - Eye movements and immediacy of processing. *Journal of Verbal Learning and Verbal Behaviour* 22(1): 75-87.
- Evans, B and C. Evans. 1957. *A Dictionary of Contemporary American Usage*. New York: Random House.
- Fajardo-Acosta, Fidel. 2002. Middle English. <<http://mockingbird.creighton.edu/english/fajardo/teaching/eng520/mideng2.htm>> [accessed November 2008].
- Febvre, L. and H. J. Martin. 1958. *The Coming of the Book: The Impact of Printing 1450-1800*. London: Verso.

- Fennell, Barbara. 2001. *A History of English*. Oxford: Blackwell.
- Fisk, W. 1985. Responses to neutral pronoun presentation and the development of sex-biased responding. *Developmental Psychology* 21: 481-485.
- Fletcher, William H. 2004. Concordancing the web: promise and problems, tools and techniques. In Marianne Hundt, Nadja Nesselhauf, and Carolin Biewer (Eds.). 2004. *Corpus linguistics and the web. Language and Computers: Studies in Practical Linguistics, Symposium on Corpus Linguistics - Perspectives for the Future*. Amsterdam: Rodopi. pp 25-46.
- Foertsch, Julie and Ann Morton Gernsbacher. 1997. In search of gender neutrality: Is singular they a cognitively efficient substitute for generic He? *Psychological Science* 8(2):106-111.
- Foster-Cohen, S. H. 1994. Exploring the boundary between syntax and pragmatics - relevance and the binding of pronouns. *Journal of Child Language* 21(1): 237-255.
- Fowler, H. R. and J. E. Aaron. 1983. *The Little Brown Handbook* (3rd edition). Illinois: Scott, Foresman.
- Frank, Francine Wattman and Frank Anshen. 1983. *Language and the Sexes*. New York: SUNY Press.
- Frank, Francine Wattman and Paula Treichler. 1989. *Language, Gender and Professional Writing*. New York: The Modern Language Association of America.
- Frank, Francine Wattman. 1978. Sexism, Grammatical Gender and Social Change. *Ninth World Congress on Sociology, Sweden 1978*.
- Frank, Francine Wattman. 1989. Language planning, language reform, and language change: A review of guidelines for non-sexist usage. In Francine W. Frank and Paula A. Treichler (Eds.). 1989. *Language, Gender and Professional Writing*. New York: Modern Language Association of America. pp 205-136.
- Garnham, A., J. Oakhill and H. Cruttenden. 1992. The role of implicit causality and gender cue in the interpretation of pronouns. *Language and Cognitive Processes* 7(3-4): 231-255.
- Gastil, John. 1990. Generic pronouns and sexist language - the oxymoronic character of masculine generics. *Sex Roles* 23(11-12): 629-643.
- Gerken, L. 1987. Telegraphic Speaking Does Not Imply Telegraphic Listening. *Papers and Reports on Child Language Development, No. 26*. Stanford University, Department of Linguistics: Stanford, California.
- Gershuny, H. Lee. 1989. English handbooks 1979-85: Case studies in sexist and non-sexist usage. In Francine Frank and Paula Treichler (Eds.). *Language, Gender and Professional Writing*. 1989. New York: Modern Language Association of America. pp 95-104.
- Geurts, Bart. 2000. Explaining grammaticalization (the standard way). *Linguistics* 38(4): 781-788.
- Gibbon, Margaret. 1999. *Feminist Perspectives on Language*. London: Longman.
- Gillquin, Gaetanelle, and Stefan Th. Gries. 2009. Corpora and experimental methods: A state-of-the-art review. *Corpus Linguistics and Linguistic Theory* 5(1): 1-26.
- Graham, Alma. 1973. The making of a nonsexist dictionary. Reprinted in B. Thorne and N. M. Henley (Eds.). 1975. *Language and Sex: Difference and Dominance*. Rowley, Massachusetts: Newbury House.
- Green, William H. 1977. Singular pronouns and sexual politics. *College Composition and Communication* 28(2): 150-153.
- Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (Ed.). 1963. *Universals of Language*. Cambridge, Massachusetts: MIT Press. pp 73-113.

- Greene, S. B., G. McKoon and R. Ratcliff. Pronoun resolution and discourse models. *Journal of Experimental Psychology: Learning, Memory and Cognition* 18(2): 266-283.
- Greene, Samuel S. 1863. *Grammar of the English Language*. New York: H Copperthwait and Company.
- Halpern, Mark. 2008. *Language and Human Nature*. New Jersey: Transaction Books.
- Hardie, Andrew. *forthcoming*. CQPweb - combining power, flexibility and usability in a corpus analysis tool. <www.lancs.ac.uk/staff/hardiea/cqpweb-paper.pdf> [accessed February 2010].
- Harley, Heidi. 2008. When is a syncretism more than a syncretism? In Daniel Harbour, David Adger and Susana Béjar (Eds.). 2008. *Phi Theory: Phi-features Across Modules and Interfaces*. Oxford: Oxford University Press. pp 251-294.
- Harrigan, J. and K. Lucic. 1988. Attitudes towards gender bias in language: A reevaluation. *Sex Roles* 10: 129-140.
- Hartmann, R.R.K. and Gregory James. 1998. *Dictionary of Lexicography*. London: Routledge.
- Hellinger, Marlis and Beate Schräpel. 1983. Über de sprachliche Gleichbehandlung von Frauen und Männern. *Jahrbuch für Internationale Germanistik* 15: 40-69.
- Herron, D. T. and E. A. Bates. 1997. Sentential and acoustic factors in the recognition of open- and closed-class words. *Journal of Memory and Language* 37(2): 217-239.
- Hirst, W. and G. A. Brill. 1980. Contextual aspects of pronoun assignment. *Journal of Verbal Learning and Verbal Behaviour* 19(2): 168-175.
- Holmes, Janet. 2000. Ladies and gentlemen: Corpus analysis and linguistic sexism. In Christian Mair and Marianne Hurst (Eds.). 2000. *Corpus Linguistics and Linguistic Theory: Papers from the Twentieth International Conference on English Language Research and Computerized Corpora*. Amsterdam: Rodopi. pp 141-155.
- Howe, Stephen. 1996. *The Personal Pronouns in the Germanic Languages: A Study of Personal Morphology and Change in the Germanic Languages from First Records to Present Day*. Berlin: Walter de Gruyter and Company.
- Humphreys, Karin R. and Kathryn Bock. 2005. Notional number agreement in English. *Psychonomic Bulletin & Review* 12(4): 689-695.
- Hundt, Marianne. 2006. The Committee Has/Have Decided...: On Concord Patterns with Collective Nouns in Inner- and Outer-Circle Varieties of English. *Journal of English Linguistics* 34: 206-232.
- Jacobson, M. and W. Insko. 1985. Use of non-sexist pronouns as a function of one's feminist orientation. *Sex Roles* 13: 1-7.
- Jespersen, Otto. 1922. *Language: Its Nature, Development, and Origin*. London: George, Allen and Unwin.
- Jochowitz, George. 1982. Everybody likes pizza, doesn't he or she? *American Speech* 57(3): 198-203.
- Johnson, Edward D. 1982. *Washington Square Press Handbook of Good English*. New York: Washington Square Press.
- Johnson, Samuel. 1747 [1970 reprint]. *The Plan of a Dictionary of the English Language*. Menston: Scholar Press.
- Joseph, B.D. 2003. The editor's department - Editing, prescriptivism, and free speech. *Language* 79(1): 1-4.
- Kennedy, Graeme D. 1998. *An Introduction to Corpus Linguistics*. London: Longman.
- Kennison, Sheila M. 2003. Comprehending the pronouns *her*, *him*, and *his*: Implications for theories of referential processing. *Journal of Memory and Language* 49(3): 335-352.

- Kennison, Sheila M. and Jessie L. Trofe. 2003. Comprehending pronouns: A role for word specific gender stereotype information. *Journal of Psycholinguistic Research* 32(3): 355-378.
- Kilgarriff, Adam, Pavel Rychly, Pavel Smrz, and David Tugwell. 2004. The Sketch Engine. *EURALEX 2004, Lorient, France*. <www.sketchengine.co.uk>.
- Kingsolver, Pamela Schroder and Harold V. Cordry. 1987. Gender and the press: An update. In Lawrence B. Nadler, Marjorie Keeshan Nadler and William R. Todd-Mancillas (Eds.). 1987. *Advances in Gender and Communication Research*. Lanham, MD: University Press of America. pp 307-15.
- Kirby, J. 1764. *A New English Grammar*. Menston, England: Scholar Press Facsimile.
- Kolln, Martha. 1986. Everyone's right to their own language. *College Composition and Communication* 37(1): 100-102.
- Kreiner, Hamutal, Patrick Sturt and Simon Garrod. 2008. Processing definitional and stereotypical gender in reference resolution: Evidence from eye-movements. *Journal of Memory and Language* 58(2): 239-261.
- Labov, William. 2001. *Principles of Linguistic Change, Vol. 1: Internal Factors*. Oxford: Blackwell.
- Labov, William. 2001. *Principles of Linguistic Change, Vol. 2: External Factors*. Oxford: Blackwell.
- Laitinen, Mikko. 2007. *Agreement Patterns in English: Diachronic Corpus Studies on Common-Number Pronouns*. Helsinki: Mémoires de la Société Néophilologique de Helsinki LXXI.
- Lakoff, Robin Tolmach. 1975. *Language and Woman's Place*. New York: Harper and Row.
- Leonard, S. A. 1929. *The Doctrine of Correctness in English Usage 1700-1800*. Madison: University of Wisconsin Studies in Language and Literature.
- Lowth, Robert. 1762. *A Short Introduction to English Grammar*. Menston, England: Scholar Press Facsimile.
- MacKay, Donald G. 1980. On the goals, principles, and procedures for prescriptive grammar: Singular they. *Language in Society* 9: 349-367.
- MacWhinney, B. 2000. *The CHILDES Project: Tools for Analyzing Talk* (3rd edition), vol. 1: The format and programs. Mahwah, NJ: Lawrence Erlbaum Associates.
- MacWhinney, B. 2000. *The CHILDES Project: Tools for Analyzing Talk* (3rd edition), vol. 2: The database. Mahwah, NJ: Lawrence Erlbaum Associates.
- Madson, Laura and Robert M. Hessling. 1999. Does alternating between masculine and feminine pronouns eliminate perceived gender bias in text? *Sex Roles* 41(7-8): 559-575.
- Madson, Laura and Robert M. Hessling. 2001. Readers' perceptions of four alternatives to masculine generic pronouns. *Journal of Social Psychology* 141(1): 156-158.
- Magner, Thomas F. 2002. Sexist and non-sexist usages in the English language. *Studia Romanica et Anglicana Zagradiensia* 47-48: 271-281.
- Mair, Christian. 2004. Change and variation in present-day English: integrating the analysis of closed-corpora and web-based monitoring. In Marianne Hundt, Nadja Nesselhauf, and Carolin Biewer (Eds.). 2004. *Corpus linguistics and the web. Language and Computers: Studies in Practical Linguistics, Symposium on Corpus Linguistics - Perspectives for the Future*. Amsterdam: Rodopi. pp 233-248.
- Martyna, Wendy. 1978. What does "he" mean? *Journal of Communication* 28(1): 131-138.
- Martyna, Wendy. 1980. The psychology of the generic masculine. In Sally McConnell-Ginet, Ruth Borker and Nelly Furman (Eds.). 1980. *Women and Language in Literature and Society*. New York: Praeger. pp 69-78.

- Mautner, Gerlinde. 2007. Mining large corpora for social information: The case of elderly. *Language in Society* 36(1): 51-72.
- Maynes, Mary Jo. 2004. Gender, labor, and globalization in historical perspective - European spinsters in the international textile industry, 1750-1900. *Journal of Women's History* 15(4): 47-66.
- Maynor, Natalie. 1996. The pronoun y'all. *Discourse and Society* 3(2): 151-166.
- McCawley, J. 1974. Letter to the editor. *New York Times Magazine*. 10th November.
- McConnell-Ginet, Sally. 1979. Prototypes, pronouns and persons. In M. Mathiot (Ed.). 1979. *Ethnolinguistics: Boas, Sapir, and Whorf Revisited*. The Hague: Mouton. pp 63-83.
- McConnell-Ginet, Sally. 1989. The sexual reproduction of meaning: A discourse-based theory. In Francine Wattman Frank and Paula Treichler (Eds.). 1989. *Language, Gender and Professional Writing*. New York: Modern Language Association of America. pp 35-50.
- McEnery, Tony, and Andrew Wilson. 2001. *Corpus Linguistics: An Introduction* (2nd edition). Edinburgh: Edinburgh University Press.
- McLaren, Margaret A. 2002. *Feminism, Foucault, and Embodied Subjectivity*. New York: SUNY Press.
- Meehan, Elizabeth. 1990. British feminism from the 1960s to the 1980s. In Harold L. Smith (Ed.). 1990. *British Feminism in the Twentieth Century*. Aldershot: Edward Elgar. pp 189-204.
- Mesthrie, Rajend, Joan Swann, Andrea Deumert and William L. Leap. 2000. *Introducing Sociolinguistics*. Edinburgh: Edinburgh University Press.
- Meurers, W. D. 2006. On the use of electronic corpora for theoretical linguistics - Case studies from the syntax of German. *Lingua* 115 (11): 1619-1639.
- Meyers, Miriam Watkins. 1990. Current generic pronoun usage: An empirical study. *American Speech* 65: 228-237.
- Meyers, Miriam Watkins. 1993. Forms of they with singular noun phrase antecedents: evidence from current educated English usage. *Word* 44:181-192.
- Meyers, Walter E. 1991. Usage glossaries in current college handbooks. *American Speech* 66(4): 341-358.
- Miller, Casey and Kate Swift. 1976. *Words and Women*. New York: Anchor/Doubleday.
- Milroy, James and Lesley Milroy. 1985. *Authority in Language*. London: Routledge.
- Minsky, Rosalind. 1996. *Psychoanalysis and Gender: An Introductory Reader*. London: Routledge.
- Mogford, Kay. 1993. Language development in twins. In Dorothy Bishop and Kay Mogford (Eds.). 1993. *Language Development in Exceptional Circumstances*. Hove: Psychology Press. pp 80-95.
- Moore, Mary Evelyn. 2001. Third person pronoun errors by children with and without language impairment. *Journal of Communication Disorders* 34: 207-228.
- Moulton, Janice, George M. Robinson and Cherin Elias. 1978. Sex bias in language use: Neutral pronouns that aren't. *American Psychologist* 33: 1032-1036.
- Mucchi-Faina, A. 2005. Visible or influential? Language reforms and gender (in)equality. *Social Science Information* 44(1): 189-215.
- Mugglestone, Lynda C. 1997. Cobbett's grammar William James Paul and the politics of prescriptivism. *Review of English Studies* XLVIII: 471-488.
- Mühlhäusler, Peter and Rom Harré. 1990. *Pronouns and People: The Linguistic Construction of Social and Personal Identity*. Oxford: Basil Blackwell.
- Nagle, Stephen J., Margaret A. Fain and Sara L. Sanders. 2000. The influence of political correctness on lexical and grammatical change in late-twentieth-century English. In Dieter Kastovsky and Arthur Mettinger (Eds.). 2000. *The*

- History of English in a Social Context*. Berlin: Mouton de Gruyter. pp 257-277.
- Nevalainen, Terttu. 2006. *An Introduction to Early Modern English*. Oxford: Oxford University Press.
- Neville, H., D. L. Mills, and D. L. Lawson. 1992. Fractionating language: different neural subsystems with different sensitive periods. *Cerebral Cortex* 2: 244-258.
- Newman, Michael. 1992. Pronominal disagreements: the stubborn problem of singular epicene antecedents. *Language in Society* 21:447-75.
- Newman, Michael. 1998. What can pronouns tell us? A case study of English epicenes. *Studies in Language* 22(2): 353-389.
- Ng, S. H. 1990. Androcentric coding of man and his in memory by language users. *Journal of Experimental Social Psychology* 26: 455-464.
- Nilsen, Alleen Pace. 1977. Sexism in children's books and elementary teaching materials. In Alleen Pace Nilsen (Ed.). 1977. *Sexism and Language*. Illinois: National Council of Teachers of English. pp 161-180.
- Nilsen, Alleen Pace. 1984. Winning the great he/she battle. *College English* 46: 151-157.
- Noordenbos, Greta. 2002. Women in academies of sciences: From exclusion to exception. *Women's Studies International Forum* 25(1): 127-137.
- Noyer, Robert Rolf. 1992. Features, positions and affixes in autonomous morphological structure. Cambridge, Massachusetts: MIT Dissertation.
- O'Halloran, Kieran. 2007. Critical discourse analysis and the corpus-informed interpretation of metaphor at the register level. *Applied Linguistics* 28(1): 1-24.
- O'Halloran, Kieran. 2009. Inferencing and cultural reproduction: A corpus-based critical discourse analysis. *Text and Talk* 29(1): 21-51.
- Osterhout, L. M. and P. J. Holcomb. 1992. Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language* 31: 785-806.
- Osterhout, L., M. Bersick and J. McLaughlin. 1997. Brain potentials reflect violations of gender stereotypes. *Memory and Cognition* 25(3): 273-285.
- Owens Jr, Robert E. 2007. *Language development: An introduction* (7th edition). Boston: Pearson Education.
- Parker, Frank. 1977. Typology as an inadequate explanation for language change. *Div. on Language Change, MLA Convention. Chicago, December 1977*.
- Parks, Janet B. and Mary Ann Robertson. 2005. Explaining age and gender effects on attitudes toward sexist language. *Journal of Language and Social Psychology* 24: 401-411.
- Parks, Janet B. and Mary Ann Robertson. 2008. Generation gaps in attitudes toward sexist/nonsexist language. *Journal of Language and Social Psychology* 27: 276-283.
- Paterson, Laura Louise. 2010. Grammar and the English National Curriculum. *Language and Education* 24 (6): 473-484.
- Paterson, Laura Louise. 2011. Epicene pronouns in UK national newspapers: A diachronic study. *ICAME* 36: 171-186.
- Paterson, Laura Louise. *forthcoming a*. Who is the generic masculine? *Culture, Society and Masculinities*. [at review].
- Paterson, Laura Louise. *forthcoming b*. No language variation, no change: Eighteenth century grammar in twenty-first century classrooms. *Salford Working Papers*. [accepted].

- Pauwels, Anne and Joanne Winter. 2006. Gender inclusivity or 'Grammar Rules OK'? Linguistic prescriptivism vs. linguistic discrimination in the classroom. *Language and Education* 20(2): 128-140.
- Pauwels, Anne. 1998. *Women Changing Language*. London: Longman.
- Pauwels, Anne. 2001. Non-sexist language reform and generic pronouns in Australian English. *English World-Wide* 22(1): 105-119.
- Perales-Escudero, Moisés D. 2010. To split or not to split: The split infinitive past and present. *Journal of English Linguistics* (ifirst, 20th December 2010) <<http://eng.sagepub.com/content/early/2010/12/15/0075424210380726>> [accessed January 2011]
- Perera, Katherine. 1994. Standard English: The debate. In Susan Brindley (Ed.). 1994. *Teaching English*. London: Routledge. pp 79-88.
- Pertsova, Katya. 2011. Grounding Systematic Syncretism in Learning. *Linguistic Inquiry* 42(2): 225-267.
- Pinker, Steven. 1999. The language mavens. In Rebecca S. Wheeler (Ed.). 1999. *The Workings of Language: From Prescriptions to Perspectives*. Connecticut: Praeger. pp 3-14.
- Platt, Jennifer. 2007. The Women's Movement and British Journal Articles, 1950-2004. *Sociology* 41: 961-975.
- Poole. J. 1646. *The English Accidence*. Menston, England: Scholar Press Facsimile.
- Poplack, Shana, Gerard van Herk and Dawn Harvie. 2002. Deformed in the dialects: An alternative history of non-standard English. In Richard J. Watts and Peter Trudgill (Eds.). 2002 *Alternative Histories of English* . London: Routledge. pp 87-110.
- Pullum, Geoffrey K. 2010. The Land of the Free and *The Elements of Style*. *English Today* 102 26(2): 34-44.
- Pulvermüller, Freidmann. 1999. Words in the brain's language. *Behavioural and Brain Sciences* 22: 253-336.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik. 1972. *A Grammar of Contemporary English*. London: Longman.
- Renouf, Antoinette. 2003. WebCorp: providing a renewable data source for corpus linguists. In Sylviane Granger and Stephanie Petch-Tyson (Eds.). 2003. *Extending the Scope of Corpus-based Research: New Applications and New Challenges*. Amsterdam: Rodopi. pp 39-58.
- Ricard, Marcelle, Pascale C. Girouard and Thérèse Decarie. 1999. Personal pronouns and perspective taking in toddlers. *Journal of Child Language* 26(3): 681-697.
- Richardson, Gina. 1984. Can *yall* function as a singular pronoun in southern dialect. *American Speech* 59(1): 51-59.
- Rispoli, M. 1994. Pronoun case overextensions and paradigm building. *Journal of Child Language* 21(1): 157-172.
- Rizzi, Luigi. 2004. From inventories to computations: Open/closed class items and substantive/functional heads. *Dialectica* 58(3): 437-451.
- Rowbotham, Sheila. 1989. *The Past is Before Us: Feminism in Action Since the 1960s*. London: Pandora.
- Rubin, Donald I. and Kathryn L. Greene. 1991. Effects of biological and psychological gender, age, cohort and interviewer gender on attitudes towards gender-inclusive/exclusive language. *Sex Roles* 24: 391-412.
- Sampson, Geoffrey. 2005. *The 'Language Instinct' Debate* (2nd edition). London: Continuum.

- Sanford, Anthony J. and Ruth Filik. 2006. They as a gender-unspecified singular pronoun: Eye tracking reveals processing cost. *The Quarterly Journal of Experimental Psychology* 60(2):171-178.
- Sanford, Anthony J., Ruth Filik, Catherine Emmot and Lorna Morrow. 2007. They're digging up the road again: The processing cost of Institutional They. *Quarterly Journal of Experimental Psychology* 61(3): 372-380.
- Schaffer, Deborah. 2010. Old whine online: Prescriptive grammar blogs on the internet. *English Today* 26 (4): 23-28.
- Schulz, Muriel R. 1975. How serious is sex bias in language? *College Composition and Communication* 26(2): 163-167.
- Segalowitz, S. J., and K. C. Lane. 2000. Lexical access of function versus content words. *Brain and Language* 75(3): 376-389.
- Shapiro, L. P. and L. R. Jensen. 1986. Processing open and closed class-headed nonwords - left hemisphere support for separate vocabularies. *Brain and Language* 28(2): 318-327.
- Shay, Scott. 2008. *The History of English*. San Francisco: Wardja Press.
- Shute, Sarah. 1981. Sexist Language and Sexism. In Mary Vetterling-Braggin (Ed.). 1981. *Sexist Language: A Modern Philosophical Analysis*. Maryland: Littlefield, Adams and Co. pp 23-33.
- Silveira, Jeanette. 1980. Generic masculine words and thinking. *Women Studies International Quarterly* 3: 165-178.
- Simpson, Paul. 1993. *Language, Ideology and Point of View*. London: Routledge.
- Sinclair, John M. 1987. Corpus creation. In Geoffrey Sampson and Diana McCarthy. 2004. *Corpus Linguistics: Readings in a Widening Discipline*. London: Continuum. pp 78-84.
- Skeat, Wailter William. 2009. *English Dialects from the Eighth Century to the Present Day*. Charleston, South Carolina: Bibliolife.
- Sklar, Elizabeth S. 1983. Sexist grammar revisited. *College English* 45: 348-358.
- Sklar, Elizabeth S. 1988. The tribunal of use: Agreement in indefinite constructions. *College Composition and Communication* 39: 410-422.
- Sniezek, J. and C. Jazwinski. 1986. Gender bias in English: In search of fair language. *Journal of Applied Social Psychology* 16(7): 642-662.
- Spence, J. T., and R. Helmreich, R. 1972. The Attitudes toward Women Scale: An objective instrument to measure attitudes toward the rights and roles of women in contemporary society. *JSAS Catalog of Selected Documents in Psychology* 2, 66. (Ms. No. 153).
- Stanley, Julia Penelope. 1978. Sexist grammar. *College English* 39: 800-811.
- Storkel, H. L. and J. M. Young. 2004. Homonymy in the developing mental lexicon. In Alejna Brugos, Linnea Micciulla and Christine E. Smith (Eds.). 2004. *Proceedings of the 28th Annual Boston University Conference on Language Development, vols. 1-2*. pp 577-584.
- Strabone, Jeff. 2010. Samuel Johnson: Standardizer of English Preserver of Gaelic. *English Literary History* 77 (1): 237-265.
- Strahan, Tania E. 2008. 'They' in Australian English: Non-gender-specific or specifically non-gendered. *Australian Journal of Linguistics* 28(1): 17-29.
- Streb, Judith, Frank Rösler and Erwin Hennighausen. 1999. Event-related response to pronoun and proper name anaphors in parallel and nonparallel discourse. *Brain and Language* 70: 237-286.
- Stringer, Jeffrey L. and Robert Hopper. 1998. Generic he in conversation? *Quarterly Journal of Speech* 84: 209-221.
- Strunk, W. and E. B. White. 1979. *The Elements of Style*. New York: Macmillan Publishing Co.

- Stubbs, Michael. 2001. *Words and Phrases: Corpus Studies of Lexical Semantics*. London: Blackwell.
- Sunderland, Jane. 1986. *The grammar book and the invisible woman*. Unpublished MA Thesis, Lancaster University.
- Sutton-Spence, Rachel. 1999. The influence of English on British Sign Language. *International Journal of Bilingualism* 3: 363-394.
- ter Keurs, Mariken, Colin M. Brown, Peter Hagoort, and Dick F. Stegeman. 1999. Electrophysiological manifestations of open- and closed-class words in patients with Broca's aphasia with agrammatic comprehension. An event-related brain potential study. *Brain* 122: 839-854.
- Teubert, Wolfgang, and Anne Čermáková. 2007. *Corpus Linguistics: A Short Introduction*. New York: Continuum.
- Theakston, A. L., E. V. M. Lieven, J. M. Pine, and C. F. Rowland. 2001. The role of performance limitations in the acquisition of verb-argument structure: an alternative account. *Journal of Child Language* 28: 127-152.
- Tieken-Boon van Ostade, Ingrid. 2006. English at the onset of the normative tradition. In Mugglestone, Lynda (Ed.) 2006. *The Oxford History of English*. Oxford: Oxford University Press. pp 240-273.
- Tieken-Boon van Ostade, Ingrid. 2010. The usage guide: Its birth and popularity. *English Today* 102 26(2): 14-23.
- Tieken-Boon van Ostade, Ingrid. 2011. *The Bishop's Grammar: Robert Lowth and the Rise of Prescriptivism*. Oxford: Oxford University Press.
- Tognini-Bonelli, Elena. 2001. *Corpus Linguistics at Work*. Amsterdam: John Benjamins.
- Treichler, Paula A. and Francine Wattman Frank. 1989. Common problems of sexist usage. In Francine Frank and Paula Treichler (Eds.). 1989. *Language, Gender and Professional Writing*. New York: Modern Language Association of America. pp 141-226.
- Tyrkkö, Jukka. 2010. Hyperlinks: Keywords or key words? In Marina Bondi and Mike Scott (Eds.). 2010. *Keyness in Texts*. London: John Benjamins. pp 79-92.
- Valian, Virginia. 1977. Linguistics and feminism. In M. Vetterling-Braggin, F. Ellison, J. English and M. Vetterling (Eds.). 1977. *Feminism and Philosophy*. New Jersey: Littlefield, Adams. pp 154-166.
- Vetterling-Braggin, Mary (Ed.). 1981. *Sexist Language: A Modern Philosophical Analysis*. Maryland: Littlefield, Adams and Co.
- Wales, Katie. 1996. *Personal Pronouns in Present Day English*. Cambridge: Cambridge University Press.
- Walker, Terry. 2000. The choice of second person singular pronouns in authentic and constructed dialogue in late sixteenth century English. In Christian Mair and Marianne Hundt (Eds.). 2000. *Corpus Linguistics and Linguistic Theory: Papers from the Twentieth International Conference on English Language Research on Computerised Corpora*. Amsterdam: Rodopi. pp 375-384.
- Wardhaugh, Ronald. 2006. *Introduction to Sociolinguistics* (5th edition). London: Blackwell.
- Waryas, C. L. 1973. Psycholinguistic research in language intervention programming - pronoun system. *Journal of Psycholinguistic Research* 2(3): 221-237.
- Weatherall, Ann. 2002. Towards understanding gender and talk-in-interaction. *Discourse and Society* 13(3): 767-781.
- Weidmann, U. 1984. Anaphoric they for singular expressions. In R. Watts and U. Weidmann (Eds.). 1984. *Modes of Interpretation: Essays Presented to Ernst Leisi on the Occasion of his 65th Birthday*. Tübingen: Narr. pp 59-68.

- Wells, C. G. 1981. *Learning through Interaction: The Study of Language Development*. Cambridge, UK: Cambridge University Press.
- Weseen, Maurice H. 1928. *Everyday Uses of English*. New York: Thomas Y Cromwell Co.
- Whelehan, Imelda. 1995. *Modern Feminist Thought* (2nd edition). Edinburgh: Edinburgh University Press.
- Whitley, M. Stanley. 1978. Person and number in the use of we, you, and they. *American Speech* 54(1): 18-39.
- Whorf, B. 1956. *Language, Thought, and Reality*. Cambridge, MA: MIT Press.
- Wiese, Heike and Horst J. Simon (Eds.) 2002. *Pronouns - Grammar and Representation*. Amsterdam: John Benjamins. pp 41-62.
- Wilson, Elizabeth and Sik H. Ng. 1988. Sex bias in visuals evoked by generics: A New Zealand study. *Sex Roles* 18: 159-168.
- Wilson, T. 1560 [1553]. *Arte of Rhetorique*. Gainsville: Scholars Facsimiles and Reprints.
- Wolfe, Susan J. 1989. The reconstruction of word meanings: A review of the scholarship. In Francine Frank and Paula Treichler (Eds.). 1989. *Language, Gender and Professional Writing*. New York: Modern Language Association of America. pp 80-94.
- Yang, Charles D. 2000. Internal and external forces in language change. *Language Variation and Change* 12(3): 231-250.
- Yule, George. 2006. *The Study of Language* (3rd edition). Cambridge: Cambridge University Press.
- Zuber, Sharon and Ann M. Reed. 1993. The Politics of Grammar Handbooks: Generic He and Singular They. *College English* 55(5): 515-530.