Facultad de Lenguas y Estudios Extranjeros
Maestría en Lengua Inglesa

Epicene pronominal forms in written English: variation across genres

N° 289                                      Alejandro Parini

Departamento de Investigaciones
Agosto 2013
The following paper summarises the general results of a study that focuses on the complexities surrounding the use of gender-neutral pronouns in contemporary English. The corpus under scrutiny covers texts coming from journalistic and academic genres. The data were collected over a period of two years from 2008 to 2010.

I would like to thank Mónia Benito, Natalia Gerrero, Marta Moitre and Gloria Sánchez for their participation in the gathering and analysis of the data presented here.

Epicene pronominal forms in written English: variation across genres.
Abstract

The fact that the English language lacks a gender-neutral pronoun has prompted the use of forms already existing in the language such as the pronominal alternatives he, he or she and they. The purpose of this study is to analyse the choice of the epicene pronoun in two different corpora: a) a corpus of textbooks on the social sciences covering the period 1995-2005 and b) a corpus of online editions of UK and US newspapers consulted in the first half of 2007. The analysis has been carried out taking into account the occurrence of the three pronominal variants with three different types of antecedents: definite, indefinite and quantifier. We argue that pronominal choice is sensitive to the linguistic context and to the type of genre.

The results show that there is a direct relationship between the pronominal form used and the type of antecedent and an indirect relationship between the use of the epicene variants and the type of genre.

Key words: epicene pronoun, linguistic context, genre, academic writing, newspapers

1. Introduction

Many linguists have speculated about the co-relation between language and the ideas, beliefs, prejudices and stereotypes in any given society. The co-relation between language and gender has been a fertile area of linguistic inquiry since the 1970s. A number of studies within the paradigms of discourse analysis, pragmatics and sociolinguistics have provided valuable information about the linguistic behaviour of men and women. In some cases this behaviour has been somewhat stereotyped and in others it has been explained in relation to social inequalities in the modern world. According to traditional feminist thinking, these inequalities have led society to equate men with power and women with weakness. The feminist stance has been highly critical of contemporary society. Feminists, in particular, state that society is organised in terms of a patriarchal order which is based on the belief that the male is the superior sex and therefore women's interests are subordinated to those of men. Much of the meaning conveyed about what is to be female or male is conveyed through language and mainly through the words used to talk about women and men. Sexist language, feminists claim, has the effect of underrating the role of women in society in general, hence leading to their discrimination and exclusion. Demystifying the ways in which language is used to support the patriarchal order has become the major concern of feminist work (Cameron 1992, Martyna 1983). As a consequence, feminist activity over the past decades has made language users aware of the non-neutral nature of language.

One of the linguistic features in the English language that has perhaps been most vulnerable to this kind of scrutiny is the third person neutral-gender pronoun.

The controversial issue of the third person neutral-gender pronoun by means of forms such as they and he or she alongside he dates back to the Middle Ages. The pronoun they, although considered improper by early grammarians (e.g. Murray 1795) as it was regarded as purely plural and therefore cannot agree with its antecedent in number and gender, was widely used by well-known writers such as Jane Austen. The form he or she was dismissed as pedantic and cumbersome and so it was agreed that he would be chosen because it was alleged to include both men and women. Bodine (1975) claims that this choice was based on the androcentric world-view of the 18th century grammarians.

However, in the early 1970s the issue of the equality of the sexes and sexism began to play a major role in the controversy of the English epicene pronoun. Since then, the use of the pronoun he has been considered ambiguous and has been accused of excluding women and equating maleness with humanness.

In an attempt to shed more light on the topic, this paper will try to explore how this awareness has led to significant variation in the use of the epicene pronoun in written English.

We argue that this variation is sensitive to the linguistic context and to the type of genre.

The term epicene, taken from Baron (1986), is borrowed from Greek and means common gender, although, in the study, it is used to describe pronouns for which the gender is inclusive of both men and women or uncertain.
2. Epicene pronouns in the grammar of English

The rationalisation that “man embraces woman” was virtually unknown in the fifteenth century (Spender, 1998). There were three forms in English for a sex-indefinite referent: \textit{he or she}, \textit{they} and \textit{he}. Baron (1986:193) points out that singular \textit{they} has a long history in Modern English, stretching back to the mid sixteenth century, and a distinguished one as it occurs in the works of Addison, Austen, Fielding, Chesterfield, Ruskin and Scott. But by the end of the 17th century, prescriptive grammarians began to use \textit{he} as a generic pronoun in an attempt to avoid the use of singular \textit{they} which was considered ungrammatical. These grammarians accepted the definition of \textit{they} as exclusively plural, which implied that they failed to agree with a singular sex-indefinite antecedent and encouraged the use of \textit{he}, which lacks agreement of gender.

Bodine (1975) states that early grammarians, who felt motivated by an interest in logic, accuracy and elegance, insisted that it was not only more natural to place the man before the woman but it was also proper because the male gender was the worthier gender. The same principle was repeated in the eighteenth century by Kirby (1746), who posited that, “The masculine Person answers to the general Name, which comprehends both Male and Female; as Any Person who knows what he says” (cited in Bodine, 1975).

This represented a move toward the concept of male as the universal category and therefore, the norm. This concept was based on the grammarians’ perspective of the world and dictated by an androcentric world-view that cemented men’s role as language planners and regulators (Bodine, ibid). Bodine goes on to say that, “although androcentrism was present, it had not resulted in the proscription of singular \textit{they}, which was still freely used along with \textit{he or she} and sex-indefinite \textit{he}” (p. 129).

Generally speaking, this androcentric view of the world has, Mills (2009:51) explains, contributed to the conceptualization of male terms as the unmarked forms and female terms as the marked, and this, in turn, has contributed to the invisibility of women within the language and within society as a whole.

Examining the legal meaning of man and \textit{he} the historian Charlotte Stopes (1908, quoted in Baron, 1986) demonstrated that nineteenth-century English lawyers were willing to admit the gender-neutral terms -both genders combined- when there was a penalty to be incurred but never when there was a privilege to be conferred. This, together with the attack of grammarians on singular \textit{they}, culminated in an Act of Parliament in 1850 (Acts Interpretations Act), which gave official sanction to the recently invented concept of generic \textit{he}. The text, as quoted by Baron (1986:140) reads:

\begin{quote}
Be it enacted, That in all Acts to be hereafter made Words importing the Masculine Gender shall be deemed and taken to include Females, and the Singular to include the Plural, and the Plural the Singular, unless the contrary as to Gender or Number is expressly provided.‘ [British Sessions Papers (1850) 338.I.5]
\end{quote}

The advocacy of sex-indefinite \textit{he} and the androcentric world-view underlying it were also strong in America. In 1880 an American prescriptive grammarian, White (cited in Bodine, 2002:131) stated that \textit{he} “is the representative pronoun, as mankind includes both men and women” because \textit{he or she} “seems to me very finical and pedantic”.

This view somehow prevailed up until the 1970s when Bodine (1975) examined thirty-three of the school grammars used in American junior and senior schools and found that twenty-eight of these books condemned both \textit{he or she} and singular \textit{they} because they were clumsy and inaccurate. She concluded that the pupils were taught to achieve both elegance of expression and accuracy by referring to women as \textit{he}. Bodine went on to say that in the light of the fact that children are exposed to the pattern of everyday conversational usage from the day they are born, and later on, they meet the overwhelming use of the masculine reference in books at school, it comes as no surprise that they cannot escape internalising the notion of generic \textit{he}, which renders women invisible.

With the growth of feminism there has been ample opportunity for feminist linguists to demystify the ways in which language is used to support this patriarchal order, but research into language and sex did not really begin in a systematic or serious way until the early 1970s. With the emergence of the feminist movement in the 20th century the use of generic \textit{he} started to be questioned due to an increasing number of experimental investigations of the generic masculine that confirmed that most people understand \textit{he} to refer to men only (MacKay, 1983).
Although an attempt to change pronominal usage was considered futile (Lakoff, 2004), there were many attempts to create sex-neutral pronouns such as *ley*, *thon*, *na*, *per*, *po*, and *person* (Crystal 1997:46) but they had no support. Nevertheless, current research (Baranowski, 2002; Parini 2004) has confirmed that *he* seems to be declining to singular *they* and to *he or she* with the use of semantically neutral expressions in contemporary written English.

Taking a diachronic perspective, Laitinen (2008) studied the anaphoric use of the epicene pronouns *he* and *they* and the cataphoric use of *he*, *they* and *those* in formal and informal letters during the periods of Early and Late Modern English. He also estimated the impact of gender, register and geographical area on the writer’s choice.

In his gender analysis, Laitinen concluded that in anaphora, women preferred *they* in over 80% of the tokens throughout the three centuries, whereas men were more equally divided between the two variants. Variation in terms of gender largely exceeded the influence of register and geographical area. In cataphora, the development towards *those* seemed to have been unconscious and independent of the social variables analysed.

Looking at epicene pronominal use in an academic setting, Bate (1978) studied language preferences of 20 faculty members at the University of Oregon. The part concerning the use of epicene pronouns revealed that half of the informants were not fully comfortable with *he*, whereas the majority accepted *he or she* without hesitation. Most participants expressed disapproval of singular *they* as non-standard. However, in following interviews half of the informants condemning *they* used it themselves in reference to expressions such as anyone and a faculty member. Bate concluded that singular *they* might eventually come into acceptance in written English.

More recently, in her study of academic discourse in English, Zapletalová (2009) concludes that there is a strong relation between the type of antecedent and the epicene pronominal forms that occur in the two academic journals she examined. Although her corpus for analysis is substantially small, she also reports on the trend towards the use of *they* as the preferred pronominal choice.

Turning to journalistic writing, a study carried out by Baranowski (2002) has shown that the use of singular *they* is becoming increasingly common in written English. After analysing a body of data obtained from The Independent of London and the San Francisco Chronicle newspapers, Baranowski concludes that *he* is no longer the preferred singular epicene pronoun in English, and that, although the three forms of the epicene pronoun exist alongside each other, singular *they* is now the predominant form.

Also, Anne Pauwels (2001) in her study of the use of gender-neutral pronouns in public speech as exemplified in radio programmes broadcast in Australia concludes that “...singular *they* is or has become the unmarked generic pronoun in semi-formal, non-scripted speech among professionals and its use far exceeds that of any other alternative, including normative *he*”.

The use of epicene pronominal choice in spoken English has also been studied by Romaine (2001:161) who, in her study of American television interviews and talk shows, reports the speakers’ preference for the use of *they* with singular antecedents of indeterminate gender like *person, everyone, anyone*, etc.

In consonance with this extensive body of research, our exploratory study will examine epicene pronominal choice in relation to the linguistic context and the type of genre taken as independent variables.

### 3. Data and Method

This study aims at showing how third person singular pronominal reference (*he, he or she, they* and their inflected forms *him, her, his, hers, herself*, etc which have the same semantic ranges) is reflected in newspapers and in academic writing. More specifically, it focuses on the relation between a dependent variable (i.e. a form of the epicene pronoun) and two independent variables: a) the linguistic context (i.e. type of antecedent) and b) the type of genre (journalistic and academic). The pronominal forms taken into account in the analysis are *he, he or she* and *they*, and their inflected forms. Combinations of *he* and *she*
Epicene pronominal forms in written English: variation across genres

(for example, she or he, s/he, etc.) were all counted as instances of he and she. The pronouns she, it and impersonal they have not been taken into account as their frequency of occurrence in the corpus is very low.

The analysis is based on two different corpora of contemporary English language writing: a corpus of academic writing and a corpus of journalistic writing.

The collection of academic texts comes from textbooks from the social sciences covering the period 1995 – 2005. The texts analysed make up a corpus of 2,500,000 words and comprise the following disciplines: Business Administration, Economics, History, Law, Linguistics, Media Studies, Psychology, Anthropology, Political Sciences and Sociology.


In all newspapers, advertisements were disregarded and quotations of direct speech were not considered in the search for epicene pronouns.

Content analysis (Krippendorff, 1980) has been employed to explore and observe how epicene pronominal forms are used in these publications and how these forms can be correlated with the type of linguistic context and with the type of genre. The categorisation of the items collected for the analysis has been expressed in terms of manifest content, that is, those items - the different epicene variants - which are physically present in the texts and which can be regarded as low-inference items.

The analysis has been carried out taking into account the occurrence of the three pronominal variants with three different types of antecedents:

a) indefinite, when the noun phrase is preceded by the indefinite article a/an:

  e.g. “An individual is entitled to have access to his records.”
  (Social Science textbook, March 2004)

b) definite, when the noun phrase is preceded by the definite article the or a possessive adjective:

  e.g. “….it is as if the reader himself is being warned against making judgments.”
  (The Independent, May 2007)

c) quantifiers, forms such as nobody/ no one, each, any, every, etc.:

  e.g. “No one wants to take big risks with their money.”
  (The Washington Post, July 2007)

4. Findings

For the sake of clarity and organisation, the results for each corpus are presented separately and are followed by a discussion section that includes general interpretative comments on the findings for both corpora.

a) Social Sciences

From a total count of 996 tokens, the corpus analysed shows that he is the preferred variant (37% - 370 tokens), followed by he or she (33% - 328 tokens) and they (30% - 298 tokens). As regards the frequency of occurrence of the three different types of antecedents, the analysis reveals that definite antecedents are the most frequently used (50%) followed by indefinite antecedents (34%) and quantifiers (16%).

However, as can be seen in Figure 1, the graph shows a relatively even distribution of the variants.
Although the number of *he* occurrences is slightly larger, allowing for mathematical error in calculating the percentages, we can see that there does not seem to be a clear-cut preference for any of the pronominal forms.

![Figure 1. Frequencies of epicene pronouns in the Social Sciences corpus.](image1)

![Figure 2. Distribution of epicene variants by type of antecedent in the Social-Sciences corpus.](image2)

This clearly shows a preference for *they* with quantifiers, which confirms the close relation between singular *they* and expressions such as everybody, somebody, no one, etc., a relationship that is acknowledged by most usage guides suggesting that epicene *they* is the only reasonable pronominal form to be used with an antecedent specified by *every, some, any, each* and *no*. In contrast, indefinite and definite antecedents in this corpus mostly anaphorise in *he or she* and *he*. This relationship shows statistically significant values (p < .005).
The relatively strong presence of *he* as a generic form may be due to the fact that in the academic community innovation in the form of non-sexist expressions may not yet have taken root. In the academic publications analysed, the authors seem to adhere to a conservative, slow changing style typical of this type of communication whose main objective is to share and standardise knowledge (Swales 2004). Secondary aspects, such as language appropriateness in matters of gender, are thus diluted. The exception, of course, seems to be in linguistics where authors are expected to be more aware of language issues.

**b) Newspapers**

The overall results for the newspaper corpus can be seen in Figure 3 below.

![Figure 3. Frequencies of epicene pronouns in the Newspaper corpus.](image)

The data show that *they* is the preferred variant (65%) in this type of genre followed by *he or she* (21%) and *he* (14%). As regards the type of antecedent, quantifiers account for 51% of all the occurrences followed by indefinite antecedents (30%) and definite antecedents (19%).

As can been seen in Figure 4, quantifier-specified antecedents were referred to almost exclusively by *they* (82%). Although *they* is much less frequently used with indefinite and definite antecedents, it is still the preferred pronominal choice. This distribution proves to be statistically significant (*p* < .005).
Figure 4. Distribution of epicene variants by type of antecedent in the newspaper corpus.

Definite antecedents seem to anaphorise in *they* more often than in *he or she* or *he*. However, the percentages for *they* and for *he or she* are very similar (38%, and 36% respectively).

Indefinite antecedents constitute the second largest group and show a strong preference for *they* (57%), followed by *he or she* (25%) and *he* (18%).

These findings seem to agree with Newman’s findings (Newman 1992) that *they* is likely to occur frequently with indefinite antecedents. However, our data support the claim that *they* is favoured most by notionally plural expressions and less by singular ones, and that *he* is no longer the overall preferred pronominal form. This tendency is also documented by Cooper (1984) in a study of American newspapers.

Unlike quantifiers, definite and indefinite antecedents constitute a more fuzzy category, which Huddleston and Pullum (2002) regard as examples of “multiple-situation-bound” interpretations. Thus, certain noun phrases that, regardless of the type of genre in which they appear, are traditionally imputed a high degree of masculinity are likely to be interpreted in a masculine manner and therefore assigned an androcentric pronominal form as can be seen in the example below:

*A doctor should always make sure his patient is well looked after while in hospital.*

*(The Independent, March 2007)*

Situations in which women rather than men are stereotypically involved are likely to be co-referent with *she* but these are infrequent in both corpora.

5. Discussion

When comparing the results for both corpora, the pattern of choice is noticeably different. Whereas the use of the variants taken separately was relatively even in the case of the social sciences corpus (ranging between 30% and 37%), in the newspaper corpus the difference in choice is more marked (ranging between 14% and 65%), with *they* being the most favoured option. (see figure 5).
Although the supremacy of *they* in the newspaper corpus can be explained in the light of the fact that quantifiers are the most popular type of antecedents, the data reveal that there is also a preference for the use of *they* with the other two types of antecedents, which makes this variant the overall preferred choice.

On the other hand, in the social sciences definite antecedents constitute the largest group followed by indefinite antecedents and quantifiers. Therefore, the frequencies of occurrence of the three different types of antecedents could help explain the pattern that emerges which clearly shows that there is a strong relationship between the choice of pronoun and the morphosyntactic antecedent, and that what seems to mark the difference between the two types of written genres in terms of epicene pronominal choice is the type antecedent that is favoured by one genre or the other.

Thus, while newspapers seem to prefer the use of quantifiers, academic writing prefers the use of definite antecedents, although this preference is not overwhelming. This leads to the argument that the choice of epicene pronominal form appears to be indirectly correlated with the type of genre.

Moreover, the different academic fields explored may reveal different preferences for the types of antecedents used and may even differ in the way they conceptualise their objects of study. For example, law and psychology usually deal with the individual, who has traditionally been imputed a higher degree of masculinity than femininity. Linguistics is more sensitive to language issues concerning not only correctness but also appropriateness. Sociology is bound to reflect social changes affecting society and so project a more egalitarian view of the sexes, etc.

So, breaking down the academic writing corpus by type of discipline may yield different and interesting results.

However, we may consider the possibility that being academic writing an institutionalised and formal type of genre (although by no means homogeneous), discourse conventions in terms of certain language forms are less affected in this form of communication by general social and linguistic changes in society at large. Besides, if we take into account the fact that academic writing has traditionally been a male-dominated domain, despite the fact that more and more women are joining academia, the regular occurrence of *he* should not be surprising.

Newspapers, on the other hand, are expected to be more concerned with socially sensitive issues and, consequently, their editorial polices are carefully observed. Although most newspapers have their own house-style manuals that show their linguistic idiosyncrasy, special attention is generally paid to political correctness, probably in an effort not to offend their potential readers.
5. Conclusion

This paper provides evidence that there is a link between pronominal choice and the linguistic context. Thus the variant *they* is generally preceded by quantifiers in both types of writing, whereas the pronominal forms *he* and *he or she* are mostly used with definite and indefinite antecedents.

Although the emergence of singular *they* as the predominant form and the decline of generic *he* as the normative form can be well argued for, specific linguistic contexts are likely to trigger the use of certain epicene forms and not others. For example, *he* is unlikely to occur with quantifiers and *they* is less frequently used with definite or indefinite antecedents.

Also, the occurrence of non-referential antecedents of the type *a reader or the advertiser*, which can be subject to multiple interpretations, contributes to the problematic of pronominal choice and therefore to the complexity of analytical scrutiny. Equally complex for the analysis is the treatment of those nouns or antecedents which are stereotypically interpreted as more masculine or more feminine, and those which tend to be understood as more masculine than feminine, even though they are epicene.

This is somehow reflected in the fact that there is no clear uniformity of pattern regarding the overall choice of pronoun in either of the two corpora of written language. In the newspapers analysed *they* is the most favoured option, followed by *he or she* and *he*, and it is the preferred form especially with quantifiers and with indefinite antecedents. However, as has been mentioned above, the large number of quantifier-specified antecedents found in the texts may account for the high frequency of *they* in the whole corpus. Normative *he*, while clearly dispreferred in journalistic writing, is the first option in the social sciences textbooks, although the proportions of the other variants are very similar.

Against this background, it seems logical to interpret the link between the choice of epicene pronoun and the type of genre as an indirect relation where certain types of antecedents may be more or less favoured by one genre or the other. So the preference for a particular type of antecedent is what ultimately appears to influence pronominal choice. However, a word of caution is in order here as journalistic and academic genres should not be regarded as homogeneous forms of communication but rather as text-type categories within which it is possible to find different types of sub-genres that may exploit specific rhetorical devices that can be linked to specific communicative purposes.

The tendencies revealed in our study could be explained in the light of more general patterns of linguistic and social behaviour. So, we could argue that newspapers are more readily and more likely to reflect ongoing changes in society as they are aimed at a general audience whereas academic writing is a far more formal type of communication aimed at a particular discourse community and traditionally regarded as a male enterprise.

Future research should provide a more detailed analysis of the use of avoidance strategies such as pluralisation and the repetition of noun phrases, and this analysis should include statistical procedures in order to corroborate the significance of the findings.

Moreover, other written and even spoken language styles are worth exploring with a view to examining stylistic variation in the use of epicene pronouns. Finally, given the nature of the topic, future studies should focus on gender as an independent variable in order to find out whether the choice of pronouns made by male authors differ from the choice of pronouns made by female authors. This line of inquiry may prove fruitful as genres are socially constructed in part through association with the gender of their producers.
References