

WHAT CAN PRONOUNS TELL US?
A CASE STUDY OF ENGLISH EPICENES

MICHAEL NEWMAN
Queens College, City University of New York

ABSTRACT

The article reports an oral corpus-based study of epicene pronominal constructions (i.e., pronouns coreferent with singular antecedents and referring to referents of indeterminate sex) in English. *They* is used in 60% of the tokens, *he* in 25%, and other forms were used minimally. That variation corresponds to three semantic factors: perceived sex stereotypes associated with the referent, notional number, and, surprisingly, degree of individuation. These findings support accounts of the importance of agreement as a discourse-level phenomenon (e.g., Barlow's Discourse-Linking Theory), and of pronouns as elements whose informational content goes beyond mere denotation.

1. Introduction

1.1 *Pronoun theory*

Intuitively, a pronoun coreferent with an antecedent can be described as being in a relation expressed as a triangle such as that shown in Figure 1. The pronoun is in the right top corner, the antecedent in the left top corner, and their common referent at the bottom. The top side of the triangle represents a direct anaphoric relation, a semantic or syntactic link between the pronoun and antecedent. The right side indicates a referential relation linking the pronoun to the referent, and the left side denotes a similar link connecting the antecedent to the same referent.

As an intuitive description, this diagram makes few claims about the particular relations and entities shown respectively as lines and vertices.

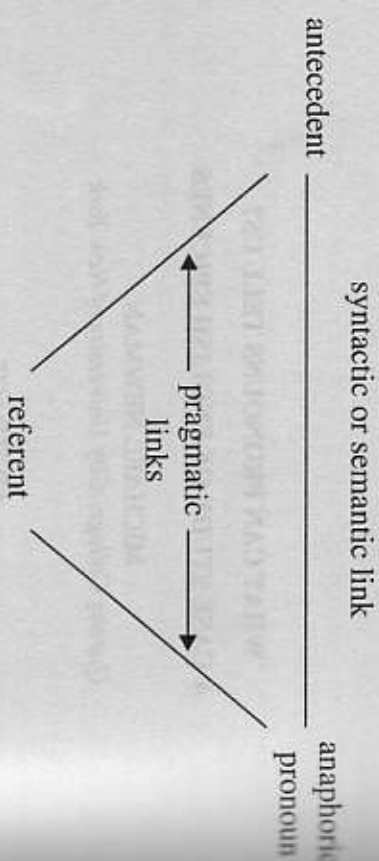


Figure 1. *The potential relations between pronoun-antecedent and referent*

Nevertheless, even a brief examination highlights an important distinction between two types of relations: The top line marks a link between two linguistic elements whereas the laterals indicate a mapping or encoding of some form of meaning of the referent. Therefore, any full account of anaphoric pronouns will necessarily be complex because it requires the integration of at least two different treatments. The approach needed to elucidate the anaphoric relation will be crucially unlike that needed to understand the anchoring — to use Barwise and Perry's (1983) and Pollard and Sag's (1994) term — of the pronoun and antecedent to the referent.

Given these considerations it can only be considered unfortunate that the bulk of recent research has concentrated exclusively on the anaphoric relation represented by the top line. For example, the following five issues, which have dominated pronoun theorists' research agendas, are all primarily or exclusively efforts to explore the relation between antecedent and pronoun: [i] the circumstances under which the anaphoric link can be described as a one of semantic dependence of pronoun upon antecedent (e.g. Lasnik 1976; Bolinger 1979; Evans 1980; Bosch 1983; Reinhart 1983; Rigau 1986; Van Hoek 1992; Fiengo & May 1994); [ii] the nature of semantic dependence, and how can it be modeled formally (e.g. Lasnik 1976; Evans 1979; Bosch 1983; Fiengo & May 1994); [iii] the nature of purely discursive anaphoric relations (e.g. Donnellan 1978; Bosch 1983; Cornish 1987; Fox 1987; Roberts 1989, 1995), [iv] how the anaphoric relation is resolved by linguistic agents, human or machine (e.g. Webber 1980; Sidner 1983;

Dahlback 1991; Grodzinsky, Wexler, Chien, Marakovitz 1993), and [v] the acquisition of anaphoric relations between pronouns and antecedents by children (e.g. Chien & Wexler 1990, 1991; Deutch, Koster, & Koster 1986; Fowler-Cohen 1994; McKee 1992; Mazuka & Lust 1994).

By contrast, the right line connecting the pronoun to the referent has not been given much consideration. As a result of this imbalance in research we are left with an incomplete understanding of how pronouns function as referring expressions. Instead, they are typically assumed to be simple designators; their presence is presumed to obey only principles of economy, and their form is considered to be entirely dependent on factors of antecedent morphology and syntactic structure.

Such an approach is problematic because it implies pronouns are static elements, a view soundly criticized by Bolinger (1979), Bosch (1983), Cornish (1986, 1987), and Wiese (1983), and Fiengo & May (1994), among others. For example, it leaves no way to account for the pronominal variation evident, just in the case of English, in reference to collectives, higher animals, and epicenes. Similarly, pronominal variation arising from contrasts between natural and formal gender in many languages is left without explanation, as is the indexing of sociolinguistic factors as in TV and honorific systems. Finally, there are indications (Ulakis 1995, 1997) that referential characteristics of pronouns, such as definiteness, play a significant role in the acquisition process.

To be fair, it should be pointed out that this neglect of pronoun meaning has not been complete; there has long been some speculation in the literature about what the relation between pronoun and referent might look like. Lasnik (1976:9), for example, tentatively offers what might be called a *lexicalist* account of pronoun reference, that is one that assimilates the right to the left side of triangle:

One interpretation [...] although certainly not the only possible one, is that *he*, for example, means 'male human being,' and consequently, that that pronoun can be used to refer to any member of that class.

Later, another generative linguist, Rigau (1986:146), proposed an alternative, though related, picture in which full pronouns lie between more typically grammatical elements, such as clitics, and lexical items in their semantic properties:

Pronouns are not descriptions. Deictic pronouns designate but do not describe. In spite of that, the overt pronouns are closer to definite descrip-

tions — terms according to Hornstein — than are clitic or empty pronouns. Because of their lexical characteristics and the logical feature of uniqueness — i.e., because of their inherent semantic content — strong pronouns may be dealt with as terms.

Whatever the difference in the respective theorists' overall linguistic frameworks, these accounts resemble Bolinger's (1979) suggestion that pronouns are similar to lexical items, only semantically leaner.

Moreover, a few linguists have proposed more developed models of pronominal meaning. Boscch (1983, 1987) and Cornish (1986, 1987), for example, offer related though somewhat differing psycholinguistic theories to account for facts of pronoun variation in a number of languages. In both models pronoun form is determined by the presence of psycholinguistic entities in mental discourse models. Cornish goes so far as to see these 'controllers' as the true antecedents of the pronoun.¹

Rather than conceiving of pronoun form as a linguistic reflex of the properties or features of discourse entities, an intriguingly different view follows from the applications of Situation Theory (Barwise & Perry 1983, Devlin 1991). In a situated model the flow of information is the opposite of that assumed by Boscch and Cornish: pronouns are seen as sources and targets, supplying information regarding the referents they are anchored to. Speakers use that information in the construction and decoding of entities in discourse models.

The first elaborated situated view of pronoun form can be found in Barlow's (1992) Discourse-Linking Theory (DLT) of agreement. This theory provides a specific model of how pronoun selection may influence referential perspective, or the accumulation of properties ascribed to a referent beyond that of simple identity. Specifically, according to DLT, all agreement exponents — be they nouns, inflections, or pronouns — contain partial information that is added to representations of referents in discourse.

The data to be examined here provide an opportunity to confirm or deny certain predictions made by DLT. Specifically, the study examines the use of epicene and related pronouns in a corpus of extemporaneous English.² The domain of pronoun use provides a good case study for this model because pronoun variation in epicene contexts is common, making it possible to explore the different information potentially contributed by the various forms.

The layout of the paper is as follows: Section 2 consists of a description of the problems to be examined and the corpus used. Section 3 gives the

results of the study. These include the relative proportions of pronouns used and their correlations with various features of antecedents and properties of addressees. They also include qualitative analyses of individual tokens, which provide data not discernible in the numbers. Section 4 discusses the theoretical implications. Specifically, pronoun choice was found to express specific intended sets of semantic properties. These can be combined with the properties of other coreferent elements (nominal or pronominal) giving rise to various referential perspectives, precisely some of those found by Barlow (1992) in other languages. In Section 5, the conclusion, the notion that pronouns contribute information is further elaborated and suggestions are made for future research.

The case study of epicene pronouns

1.1 *Epicenes*

English epicenes are most frequently discussed as a prescriptive issue that arose from the problematic proscription of singular *they* and, relatedly, through complaints by feminist language critics about the sexism of the traditionally prescribed use of *he*. The term *epicene* is taken from Baron (1990), although as Corbett (p.c.) points out, this usage is shifted from its original Greek sense of common gender. Nevertheless, I follow Baron because the more common term used for this phenomenon, *generic*, has a different basic meaning of reference to a type or class, and no other term is readily available. Note that I use *generic* in its basic type-class sense in this study.

These terminological rough edges are perhaps symptoms of the fact that concern with this phenomenon has been more focused on critiquing and addressing usage than studying the linguistic facts behind it. If so, a more serious consequence is likely to be the frequent lack of any coherent explicit definition of the phenomenon, whatever it is called. Yet it seems that when we discuss what I am calling 'epicenes,' they consistently refer to a pronoun used in one specific set of circumstances in which no canonical form is consistently found. Specifically, *an epicene pronoun is one that is consistent with a singular antecedent and that refers to a referent of unknown or indeterminate sex*.

Despite the implicit consensus in the literature on these characteristics,