Diachronic Change in English Pronoun Usage

Ramses Mulder s5363144

Abstract

This project investigates diachronic change in English pronoun usage, focusing on the relative frequency of singular *they* compared to the gendered pronouns *he* and *she*. Although the use of singular *they* has been present since the fourteenth century, its usage has increased substantially in recent decades. This study compares two time periods, 1990 and 2019, to examine how social changes related to gender and inclusivity are reflected in language use. By quantitatively analyzing pronoun frequencies across these periods, the project aims to contribute to a broader understanding of linguistic change and its social motivations.

1 Introduction

Language change is often shaped by social norms and values. One case in point is the increasing prominence of *they* as a singular pronoun in English. Traditionally excluded in prescriptive grammar for its association with plural reference, singular *they* has nonetheless persisted for centuries and become more socially visible through discourse on gender identity. Understanding how real usage patterns have shifted over time requires systematic corpus analysis.

The central research question is:

How has the relative frequency of singular *they* changed compared to that of *he* and *she* between 1990 and 2019?

The following hypotheses are tested:

- H1: Relative frequency of singular *they* is higher in 2019 than in 1990.

- H2: Combined relative frequency of *he* and *she* is higher in 1990 than in 2019.

2 Related Work

Singular *they* has been widely studied in linguistics as both a historical and contemporary phenomenon. The Oxford English Dictionary documents its usage back to the 14th century, indicating that modern debates over grammaticality contrast with its historical presence.

LaScotte (2016)'s empirical study finds that singular *they* is commonly used as a gender-neutral alternative to *he* or *she*, with most speakers preferring inclusive forms when gender is unspecified, showing that usage patterns reflect social attitudes toward gender.

Foertsch and Gernsbacher (1997)'s experimental research demonstrates that singular *they* functions efficiently in comprehension tasks, particularly when antecedents are non-referential, which supports claims that it is not only socially motivated but also cognitively grounded.

Corpus-based research by Stormbom (2020) reports that non-gendered singular *they* appears as the most frequent choice among epicene pronouns in academic articles, outpacing *he* or *she* combinations, suggesting a shift in written registers toward gender neutrality over recent years.

Acceptability studies show that singular *they* is judged as grammatical by many speakers, especially in generic contexts. However, specific uses, for example, referring to known individuals, still vary by speaker attitudes, reflecting ongoing norm change.

Moulton et al., 2020's syntactic research further shows that the acceptance and use of singular *they* are influenced by structural context and antecedent presence, reinforcing that linguistic patterns intersect with pragmatic conventions.

Together, these studies provide a robust foundation for expecting increased singular *they* usage over time and contextualize pronoun choice as

both a linguistic and sociocultural process.

3 Data

This study uses the Corpus of Contemporary American English (COCA), which contains multi-genre English totaling over one billion words. Subcorpora for 1990 and 2019 were selected to represent English usage in those years.

For each period:

- Occurrences of *they* were extracted and manually filtered to include only singular uses (excluding clearly plural contexts).

- Counts of *he* and *she* were combined.

- Relative frequency per million words was calculated to allow meaningful comparison across time periods of potentially differing sizes.

Table 1 provides a summary of the data that will be used in this study.

Time Period	Pronouns Counted
1990	singular *they, he, she*
2019	singular *they, he, she*

Table 1: Overview of the study dataset, showing the two COCA time periods (1990 and 2019) and the pronoun categories counted (singular *they*, *he*, and *she*).

4 Predicted Results

Based on prior corpus and experimental research, we predict an increase in the relative frequency of singular *they* from 1990 to 2019 and a decrease in the combined relative frequency of *he* and *she*. This expectation aligns with findings that singular *they* has become a widely used epicene strategy in written English and that speakers readily produce it in generic contexts Baranowski, 2002; LaScotte, 2016. If the COCA results follow this pattern, *they* would provide additional evidence that pronoun choice is participating in a broader shift toward gender-fair language.

Table 2 summarizes the predicted results in relative frequency

Discussion If *he* and *she* decreases between 1990 and 2019, several mechanisms could be at play. One possibility is a general movement away from explicitly gendered references in contexts

Pronoun Type	1990	2019
singular *they*	lower	higher
he + she	higher	lower

Table 2: Predicted relative frequencies of singular *they* and the combined gendered pronouns *he* and *she* in COCA for 1990 and 2019.

that do not require gender specification, which would be consistent with the observed preference for epicene strategies in writing Baranowski, 2002.

An observed increase in singular *they* in 2019 relative to 1990 would support H1 and align with corpus research showing greater acceptance of gender-neutral reference. A decrease in *he* and *she* would support H2, in line with evidence that traditional gendered pronouns are increasingly being replaced in inclusive contexts LaScotte, 2016.

5 Conclusion

This study set out to examine whether English pronoun usage has changed over time by comparing COCA data from 1990 and 2019. Specifically, we asked how the relative frequency of singular *they* has changed in comparison to the combined relative frequency of *he* and *she*. Drawing on prior work on epicene pronouns and speakers' production preferences, we hypothesized (H1) that singular *they* would be more frequent in 2019 than in 1990, and (H2) that *he + she* would be more frequent in 1990 than in 2019.

If the predicted pattern is observed, it would provide quantitative support for the idea that pronoun choice is participating in a broader shift toward gender-fair language. An increase in singular *they* would be consistent with research describing *they* as an increasingly common epicene strategy, especially in generic contexts where gender is not specified. At the same time, a decrease in the combined frequency of he and *she* would be compatible with an interpretation that explicitly gendered references are being used less often in contexts where gender is irrelevant, reflecting changing norms of inclusivity.

Several limitations should be acknowledged. First, identifying singular *they* in corpus data requires contextual decisions to exclude clearly plural readings, and some tokens may remain ambiguous despite manual filtering. Second, focusing on two time points provides only a coarse pic-

ture of change; intermediate years would be necessary to determine whether the shift is gradual or concentrated in particular periods. Future work could extend the analysis to additional years and examine genre-specific patterns within COCA to test whether changes are uniform across registers or driven primarily by particular text types.

GitHub repository: https://github.com/Shaffy27/IRM

References

Baranowski, M. (2002). Current usage of the epicene pronoun in written English. *Journal of Sociolinguistics*, *6*(3), 378–397. https://doi.org/10.1111/1467-9481.00193
Rationale: Used as corpus evidence that singular they is the most frequent epicene option in written English and that generic he is no longer predominant, supporting the predicted direction of change.

Foertsch, J., & Gernsbacher, M. A. (1997). In search of gender neutrality: Is singular They a cognitively efficient substitute for generic He? *Psychological Science*, *8*(2), 106–111. https://doi.org/10.1111/j.1467-9280.1997.tb00691.x
Rationale: Used to justify the expectation that singular they can spread because it is cognitively efficient in comprehension, especially in generic/nonreferential contexts, supporting predicted increases over time.

LaScotte, D. K. (2016). Singular they: An empirical study of generic pronoun use. *American Speech*, *91*(1), 62–80. https://doi.org/10.1215/00031283-3509469
Rationale: Used to support predicted increases in singular they by showing speakers strongly prefer inclusive pronouns in generic/gender-unknown contexts.

Moulton, K., Han, C.-h., Block, T., Gendron, H., & Nederveen, S. (2020). Singular they in context. *Glossa: a journal of general linguistics*, *5*(1), 1–13. https://doi.org/10.5334/gjgl.1012
Rationale: Used to justify methodological choices for identifying singular they (context/antecedents matter) and to frame limitations around ambiguity in corpus annotation.

Stormbom, C. (2020). Gendering in open access research articles: The role of epicene pronouns. *English for Specific Purposes*, *60*, 193–204. https://doi.org/10.1016/j.esp.2020.08.001
Rationale: Used to support the expectation that singular they is increasingly common in formal written registers (research articles) and to frame epicene pronoun choice as an ongoing change shaped by gender-fair language norms.