

Anaphora Resolution in English Epicene Pronominalization: Language Reform and Linguistic Theory

Andrew M. Kato
University of California, Santa Cruz
December 10, 2022¹

Abstract

In recent years, increasing consensus among English speakers for an inclusive pronoun regardless of gender has materialized in mainstream usage of a singular neutral dimension of the word *they*. Especially among situations in which the gender identity of an interlocutor is unknown, opting for *they*, called an epicene pronoun, seems to have increased in both written and spoken English. How does this contemporary development interface with syntax, semantics, and contextual dialogue? The goal of this work is to serve, then, as a meaningful exploratory analysis of epicene pronominalization in English, focusing on disambiguation in discourse and arguing that it falls within a larger context of bottom-up language reform in the United States. The Corpus of Contemporary American English (COCA) will be referred to for real-world data of epicene pronoun usage. For accessibility to audiences beyond fields of modern theoretical linguistics, this paper will attempt to contextualize the topic at hand by introducing the study of language as a science, as a facet of philosophy, and as an extension of symbolic systems. For a holistic, historical understanding relevant to the epicene pronoun's development in English, this discussion will also briefly include the cultural debate and movements surrounding its prevalence up to today.

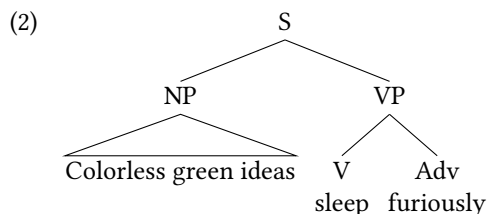
Keywords: Epicene pronouns, pronominalization, language reform, anaphora resolution, descriptivism, inclusive language.

1 Syntactic Introduction: Surveying Relevant Foundations

Modern approaches to theoretical linguistics as we know it are incredibly recent, especially in comparison to other sciences and disciplines in human history. While, for instance, relevant mathematical theorems can stand for thousands of years, the status of modern linguistics as a scientific endeavor rooted in cognition and formal reasoning dates back to as recently as the mid-20th century (Stokhof & van Lambalgen, 2011). Perhaps most notably, the theory of generative grammar, i.e. that syntax presupposes finitely many cognitive rule-based principles, spurred intense exploration of linguistic patterns and phenomena (Chomsky, 1957). This includes notions of what grammaticality might truly mean, as seen in a famous pair of constructions, which can be seen below.

- (1) a. Colorless green ideas sleep furiously. (Chomsky, 1957, p. 3)
- b. *Furiously sleep ideas green colorless. (Chomsky, 1957, p. 3)²

While the semantic value of (1a) carries no real-life accuracy, its grammaticality remains. There is no reality where an idea that is simultaneously green and colorless possesses the capacity to sleep. The sentence is nonsensical. By contrast, the structure of (1b) is both semantically and syntactically illogical. This example, though simple, helps characterize thorough theoretical investigation into the cognitive principles that might explain such discrepancies in syntactic structure. With mathematical abstractions of syntax into formal phrase-structure rules (Chomsky, 1956; Chomsky & Schützenberger, 1963; Pullum & Gazdar, 1982), sentences (1a-b) could be analyzed systematically as seen in (2)³, which is a constituency-based tree. (1a) could be considered a realization of symbolic rules in (4).



¹The format of this paper is from Pomona College's L^AT_EX template by linguists Michael Diercks and Franny Brogan, which can be found at <https://www.overleaf.com/latex/templates/pomona-linguistics-latex-template/bvdxdtwysnd>.

²This paper will utilize the convention of placing an asterisk symbol (*) at the beginning of constructions ungrammatical to native speakers.

³Note the triangle notation, which is typically meant to abstract away details that are irrelevant to the discussion at hand.

(3) [S [NP Colorless green ideas] [VP sleep furiously.]]

- (4) a. $S \rightarrow NP VP$
 b. $NP \rightarrow (A)^* N$
 c. $VP \rightarrow V (Adv)^4$

Additionally, the emergence of a theory of innate language capacity (Putnam, 1967; Wasow, 1973)⁵ developed a view of all language acquisition as the consequence of mapping stimuli to latent faculty universally present in humans. The degree to which this theory holds is debated in literature past and present. Considerations such as these nonetheless add to theoretical exploration that will become useful for analyzing pronominal usage here. Since the simplistic representations seen in (2), further theoretical developments continued to formalize and broaden the scope of what operations comprise sentence structure. This includes aspects of movement, which serve to codify and account for discrepancies between natural language and plain phrase-structure rules, e.g. (5a-c).

- (5) a. Extraposition: The party guest _ danced outside **without sunglasses**.⁶ .
 b. Wh-movement: **Whom** do you like eating at the diner across the street each weekend with _?⁷
 c. Passivization: \emptyset **She**_{NOM} is __{ACC} encouraged to finish the race.

Take (5c), for instance, which depicts a traditional example of a passive sentence. The average English speaker can generally identify that there exists a relationship between active and passive voice, in essence as two sides of the same coin. (6) represents a valid active version, with relevant morphological glossing.

- (6) Somebody encourage-s her to finish the race. **English**
 Somebody encourage-3.SG.PRES PRO.3.SG.ACC to finish the race.
 ‘Somebody encourages her to finish the race.’

1.1 Transformational Grammar and Standard Theory

Phrase-structure rules alone will not be able to generalize and account for the syntactic relationship between active and passive voice in (5c) and (6). The constructions don’t necessarily exist in isolation, and their connection is evident among native speakers of the language. For every sentence of an active voice, then, there also exists a passively-voiced version, sharing identical meaning. This repetition is unavoidable with phrase-structure rules alone, since there is nothing to indicate the relationship between the two. Chomsky identifies this limitation (7). A possible solution to this is to introduce the pair via another rule, but Chomsky, too, says this ignores that active and passive voice are structurally deterministic of one another.

- (7) “This inelegant duplication, as well as the special restrictions involving the element $b+en$, can be avoided only if we deliberately exclude passives from the grammar of phrase structure, and reintroduce them by a rule.” (Chomsky, 1957, p. 43)

As a result of issues such as these, proposals of transformations arose over the years and decades — which necessitated a richer dimensionality to how structure arises. Rather than taking sentences for face value, syntacticians began considering (and debating over) sentences as end results of abstract representations stepping through processes

⁴The production rules depicted here are simplified, representing only the elementary details shown in (2) and (3). Parentheses indicate optionality, and the symbol * in this scenario is the Kleene star, which indicates that a particular element can appear any number of times consecutively — for instance: *The tall cloudy unstable inconsiderate thoughtless overbearing [...]*. In early conceptions of grammar, reflected in these rules, a sentence *S* is composed of an obligatory noun phrase *NP* and an obligatory verb phrase *VP*, in that order. These child constituents, *NP* and *VP*, have their own rules as well, in similar fashion, as do prepositional phrases (*PP*), adjective phrases (*AP*), and others.

⁵This viewpoint is often termed the *Innateness Hypothesis* (IH). The ideas themselves are attributed to theoretical advancements by Chomsky.

⁶Extraposition is the operation of positioning (hence the name) a constituent outside of its canonical, or standard, place in a relative clause, often no matter the distance. The prepositional phrase (*PP*) *without sunglasses*, although directly modifying the subject, is significantly distanced. Semantically, this effects how the *PP* applies to the subject, i.e. that the party guest is without sunglasses perhaps only while outside, which may be unusual

⁷The long-distance movement of *who*, *what*, *when*, *where*, *why*, and *how* in English interrogative constructions is a closely studied phenomenon (Chomsky, 1977; Kiss, 1993). Note the extended distance between *whom* and the location it would otherwise be sitting in.

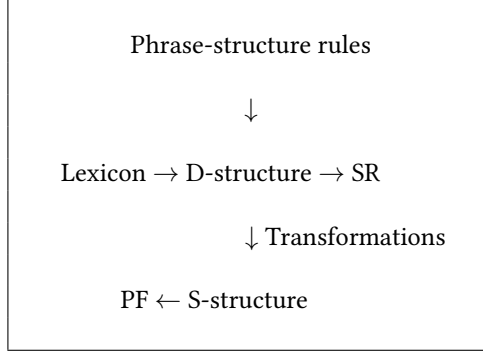


Figure 1: Standard Theory

of operations given particular prerequisites. The starting point of these processes for a given sentence is termed a deep-structure (D-structure), while the final output is the surface-structure (S-structure) (Chomsky, 1957, 1965). This in turn heavily expands the strength of generative grammar to represent natural language by extending structure beyond the constraints of the original context-free approach. Out of this comes the Standard Theory of syntax (Chomsky, 1965), the foundation of Chomskyan generative syntax and the typical model relied upon in theoretical linguistics for the time. Instead of phrase-structure rules (PSRs going forward), alone, dictating sentence realization, multiple steps become codified in this process — after PSRs and lexical items (e.g. words, phrases) form the core representation in D-structure, which also forms the basis of semantic representation (SR). In this way, meaning is determined not by S-structure, but is rather by D-structure. Notably, the transformational operations that follow, then, do not alter meaning and simultaneously conform to PSRs with each step *structure-preserving* (Emonds, 1970). After this series of operations, like those that form examples (5a-c), it is shipped off to phonetic form (PF) — which determines features such as prosody and stress, phoneme insertion and deletion, and sound assimilation⁸. This proposed cognitive process is depicted in Figure 1, with movement from D- to S-structure, known as derivation, being the emphasis of this process. With as many special cases and scenarios seen across human languages, though, Standard Theory can introduce significant complications over transformation ordering along with operations done as technicalities to properly generate the S-structures seen in natural language.

1.2 The Role of Semantics

From Standard Theory came considerable debate over how D-structure should be interpreted, and how it fits into Standard Theory. Multiple phases of revision of it have taken place since. The growth of generative semantics by George Lakoff and Paul Postal, among others, furthered an approach that placed meaning as the basis for syntactic structure — the opposite of Chomskyan generative linguistics at that time (Lakoff, 1971, 1976). Instead of PSRs constructing D-structure, derivations in this theory are from meaning transformationally building the appearance seen in S-structure. Generative semantics, itself, has since fallen out of the linguistics scene, although the fierce dialogue over how semantics fits into grammar has influenced the meaning- and thought-driven approach of Cognitive Linguistics (see Harris, 2021; Lakoff, 1976; Langacker, 1986). These semantic developments provide insight into the history of meaning relevant for discussion of ambiguity in future.

1.3 Minimalist Operations and Head-Patterns

In present-day investigations, though, syntacticians generally opt for a simpler framework, which will also be assumed going forward in this paper. This includes more recent developments such as X'-theory for generalizing a constituent-head pattern and refined takes on movement operations, including those of the Minimalist Program (MP)

⁸Phonology is a branch of theoretical linguistics concerned with sound systems and patterns, and why those sound units, called phonemes, adjust based on environment. One of the more basic examples is, for instance, why *dogs* and *cats*, despite sharing the same plural morpheme [-s], have their final consonant pronounced differently among most English speakers — a voiced alveolar fricative in ['dɒgz] versus a voiceless version in ['kʰæts]. In this case, they are allophones of the same phoneme. The related study of sound production, which is more so involved with phonemes in real-world time and space, i.e. the *how*, is known as phonetics (for past and present inquiry into theoretical and experimental phonological systems of language, see Chomsky, 1967; Jescheniak & Levelt, 1994; Cohn et al., 2012)

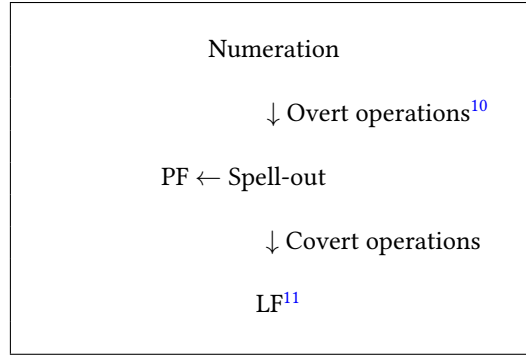


Figure 2: The Minimalist Program (Chomsky, 1995; Ko, 2014)

(Chomsky, 1995). MP defines a model of syntax as that seen in Figure 2 (Ko, 2014, p. 2), emphasizing any operations occurring only as constraints of economy allow it to. This minimalism strips down the process of derivation, removes the notion of PSRs in favor of two basic operations MERGE (of generalized elements α and β) and MOVE, and supports a bottom-up approach to tree-building (compared to top-down in Standard Theory) in phases (where each phase represents the domain that a constituent-phrase would otherwise be). Movement is seen as the result of *escape-hatches* of *strong phases* CP (complementizer-phrase) and VP (verb-phrase) – and all other phases are subject to the *Phase-Impenetrability Condition* (PIC) that dictates heavy constraint on item movement. The discussion at hand will not adopt quite as radical a framework as MP, but will survey its usability in future as a means of explaining pronominal movement⁹.

2 Reference and Pronouns

2.1 Nominalization in Brief

The ability to nominalize, or employ as a noun, various parts of speech emerges across languages, including English.

- (8) a. Grocery shoppers often [_{VP} buy [_{NP} lemons]].
 b. Grocery shoppers [_{VP} love [_{NP} **buying lemons**]].

In this case, the act specified in (8a) is represented as a NP in (8b). The word-level change, *buy* → *buy+[-ing]*, represents a form of derivational morphology (Lieber, 2017). With English in particular, this represents one of multiple forms of nominalization (Chomsky, 1970; Siegel, 1998).

- (9) a. Gerundive nominalization: [_{NP} **Cleaning electronic equipment**] proves difficult.
 b. Non-gerundive nominalization: [_{NP} **The destruction of the memo**] took an hour. (Siegel, 1998, p.4)
 c. Adjectival nominalization: [_{NP} **The strong**] overcome [_{NP} **the weak**].

Notice the distinction between sentences such as (9a) versus those such as (9b) – non-gerundive nominalizations lack the morphological addition of the suffix *[-ing]*. Both involve some action, the (9a) being atelic and (9b)

⁹Even with more generalized frameworks such as these in generative syntax, it's worth noting that the ordering of transformations across languages is still debated. A notable example is the timing of derivational ellipsis, which is thought by some to occur prior to pronounced expression in PF (Güneş & Lipták, 2011). Ellipsis is constituent omission generally due to the desire to avoid repetition or based on presupposed knowledge, i.e. (1a-c).

- (1) a. I drink coffee but she does not ~~drink coffee~~.
 b. On **THIS** building there used to fly a Confederate flag, and on **THAT** building there did ~~used to fly a Confederate flag~~, too. (Güneş & Lipták, 2011, p. 107)
 c. Gwendolyn smokes marijuana, but ~~Gwendolyn~~ seldom ~~smokes marijuana~~ in her own apartment. (Hankamer & Sag, 1976, p. 409)

being durative in telicity¹². The base verb *destruct* changes in part-of-speech category, verb to noun, which is a form of derivational morphology (as opposed to inflectional, in which the given word's category remains constant), and also proves to be the case in (9a).

2.2 Pronominalization

The usage of pronouns in natural language primarily serve the well-documented purpose of simplifying communication and discussion (Postal, 1966; Kayne, 2002), as seen in the following examples. This typically involves avoiding repetition and presuming a level of knowledge specific to a given sentence structure or context (pragmatics). Along the same lines of the previous discussion, pronominalization involves the process of representing some constituent as a pronoun.

- (10) Tisha saw the loaves of bread but didn't want any of **them**.

In this sense, the pronoun *them* refers back to the NP *the loaves of bread*. This is only a surface-level description, though, and belies the complexity of reference and relationships between constituent meaning. Below demonstrates an example of how the act of nominalization can become salient in a pronominal construction. (12a-b) demonstrate the necessity for the 2nd-person pronoun *you* in order to directly converse. In most any situation, a pronoun can be used to replace a given NP, whether or not that NP is the result of nominalization.

- (11) Pronominalizing (9a): [_{NP} Cleaning electronic equipment] proves difficult, especially when [_{NP} it] involves small devices and older machinery.
- (12) a. Person A: I_k had to go to my class last night in the rain!
b. Person B: Did you_k at least have an umbrella?
c. Person B: #Did Person A_k at bring a raincoat?

(12c) allows for some light to be shed on the necessity of pronominalization when referring to people, both in direct conversation and in passing reference. This draws upon theoretical representations of how meaning can be carried across multiple degrees of constituent separation.

2.3 Anaphora and Pragmatics

The relationship between pronouns and anaphoric relationships will be predicated going forward on the following definitions from Huang (2013) (emphasis added in bold), developed from MP and Chomsky (1995).

- (13) a. 'Anaphora can be described as **a relation between two linguistic elements**, in which the interpretation of one (called an anaphoric expression) is in some way determined by the interpretation of the other (called an antecedent)' (Huang, 2013, p. 73).
b. 'An anaphor is **a feature representation of an NP which must be referentially dependent** and which must be bound within an appropriately defined minimal syntactic domain' (Huang, 2013, p. 74)¹³.

From this analysis comes distinctions between NPs, namely anaphors (bound to a local domain), pronominals (free in a local domain), and referential expressions (r-expressions going forward; free regardless of domains) (Huang, 2013). The locality of reference distinguishes them from one another. Reflexives within a space enclosed by a shared head can be considered anaphora in this context.

- (14) Joshua_i smiled at himself_j in the mirror.

¹²Telicity is a verbal property defining the completeness of an action. (9b) is a durative telic construction, also known as describing an accomplishment, in that the action (destroying the memo) occurs over some non-instant period of time (an hour). (9a), by contrast, specifies an action without a specified endpoint.

¹³Think of a syntactic domain as some space in a given construction. This could be a single NP, like *an umbrella*, or something that dominates a larger tree, like the VP *running faster than the speed of light*

For specificity here, the subscript j represents coindexing between the antecedent *Joshua* and the anaphor *himself*. Using indices in sentences such as (12a-c) and (14) help identify sources of meaning, and when a NP refers to another. Syntactically, r-expressions are domain-unconstrained, meaning that, no matter the conversation or situation, they can be inserted (given grammatical structure conformity). The constraints on relationships of anaphora are consequently termed binding conditions based on the formal syntactic approach of Chomsky (1995).

3 Where Language Reform Meets Linguistic Theory

3.1 Language Reform and History

Within the past decade, and within the past few years in particular, awareness has increased over various aspects of the English language seen as exclusionary. As small steps by a growing range of English speakers to modify the language, changes such as the following are being seen more often in natural use.

- (15) a. Orthographic character-switching: Woman → womxn; latino → latinx; folks → folx
- b. Characteristic distancing: Disabled person → person with a disability; colored people → people of color
- c. Gendered opt-out: Waitress → server; actress → actor

Common examples, seen above, involve mitigating perceived prevalence of oppression in natural language. This includes modifying markers or reminders of male gender, e.g. *man*, commonly with the letter *x* (Ashlee et al., 2017). The *folks*-to-*folx* change tends to arise due to the inclusivity that [x] is associated with from the other use cases. With (15b), a syntactic transformation, an Adj-to-P movement that I will refer to as characteristic distancing (CD going forward), seeks to place distance between a N and a historically marginalized characteristic (Adj or Adv) of that N — such that the characteristic’s emphasis on the meaning of N is reasonably minimized. This is sometimes called *identity-first language*, as a result (Dunn & Andrews, 2015). Additionally, when there exists at least one viable ungendered variant, use cases tend toward avoiding parts-of-speech that morphologically carry markers of physiological gender or gender identity (15c). These developments represent a rare pattern in comparison to the structural and formal theories of transformational grammar among syntacticians: these movements, although overt in their changes up to S-structure, reflect a solely semantic motive. The growing desire to promote inclusivity in natural language in the US in particular is a budding example of bottom-up language reform. While language creation, e.g. Esperanto (Fiedler, 2015), and systemic language modifications, e.g. Turkish (Tachau, 1964), are well-documented examples of government- or policy-led reform, the modifications in (15) were not spurred by mandated political influence. Instead, they started first with average native speakers. From there, language choices have modified to accommodate.

Perhaps the most prominent element of this active movement in language reform is the popularity of *they* regardless of gender identity (Zimman, 2017), commonly called an epicene pronoun. Building off of the anaphora relationships discussed in Arnold et al. (2021) and Balhorn (2004), the usage of *they* as a pronoun whose semantic antecedent is a singular person is increasingly used today.

- (16) ‘The pronoun *they* can be either plural or singular, perhaps referring to an individual who identifies as non-binary’ (Arnold et al., 2021, p. 1).

A distinct feature of this development is how it is perhaps the culmination of multiple efforts and patterns of *they* usage in English language history. For instance, Baron (1981), at the outset states:

- (17) a. ‘The creation of an epicene or bisexual pronoun stands out as the one most often advocated and attempted, and the one that has most often failed’ (Baron, 1981, p. 83).
- b. ‘Use of the plural [*they*-forms] was labeled downright wrong, if not worthy of eternal damnation, and pairing of the masculine and feminine [...] was most often rejected as ugly and cumbersome’ (Baron, 1981, p. 83).

4 Agreement in Syntax and Morphology

4.1 Corpus Numbers

Citing an example as early as 1878, Baron considers past avoidance (and, at that time, current avoidance) of an ungendered epicene pronoun. For instance: between 1990 and 2019, the Corpus of Contemporary American English (COCA) (Davies, 2008) — a collection of texts from digital and physical media in the US — counts 375,141 instances of the string *they are*, across sources including television, news, and academic articles. From 1990 to 1994, 39,588 total usages of the string emerged, which is about the same as the number from 2015 - 2019. The use of the *they* itself reaches past 500,000 in total, with most usages resulting from within recent years. And, yet, over all the data available to COCA, a mere 159 total cases of the string *they is* appear. This reflects the morphosyntactic necessity of agreement Barron and Chomsky separately identify.

4.2 Morphological and Pragmatic Constraints

- (18) a. They would rather be sleeping in the bed than on the sofa.
b. *They wants a cookie from the store.

Even across extended distances of auxiliary modifiers (18a), the epicene pronoun seeks agreement just as the plural usage of *they* does — minus a distinctive exception: reflexives as anaphora. Given, FACT: RYAN_R, A STUDENT, IS A NON-BINARY INDIVIDUAL:

- (19) a. Nominative: They_r are such a great speaker in class.
b. Accusative: The instructor gave them a great grade on their_r essay.
c. Reflexive: They_r looked at themself_r in the mirror after class.
d. Genitive (possession): The pencil is theirs_r.

Going against counterarguments of researchers such as Bickerton (1987), and even Baron (1981), who are unable to come to conclusions about the nature of reflexive anaphora, the reflexive use-case of the English epicene pronoun likely undergoes an overt operation meant to satisfy both LF and the case agreement involving singular pronouns (he, she, it). Where multiple other aspects of morphological case lend themselves to makeshift alternatives, e.g. *y'all*, *y'all's*, etc. acclimation to the epicene pronoun *they* needs not such accommodations. However, where the epicene pronoun agrees with the verbal copula in the plural form (19a), agrees with other verbs while in accusative case (19b), and satisfies the typical morphological affix denoting plural possession (19d), the reflexive anaphor in (19c) stands as an outlier. A possible proposal is that, at some point in derivation, a given sentence undergoes an operation that deterministically modifies the typical instance of the epicene reflexive to reflect a singular antecedent when used. Given how recent this contemporary development in epicene pronoun usage is, especially compared to past examples (e.g. defaulting to *he*), this should be explored in future.

5 Holistic Discussion

5.1 Concluding Remarks

The scope of this paper is certainly exploratory, briefly covering the history of modern descriptive syntactic investigation, and going into the morphosyntactic constraints surrounding a Chomskyan understanding of anaphora. In sum, the primary goal of the information at hand is to provide a unifying exploration of the most recent movement of language reform seen regarding an epicene pronoun. In surveying recent data and research-driven examples from others about the nature and usage of anaphora, the conclusion is reached that the epicene usage falls within the context of a larger language reform and exhibits an unusual morphosyntactic quality that could be the result of an overt operation in derivation.

5.2 Limitations

Some of the real-world limitations of exploratory research such as this include constraints of time and resources. The degree to which the heavily theoretical morphology of reflexive anaphora could be covered necessarily fell within the bounds of an academic quarter and the maximum number of pages allowed. More importantly, the ability to conduct a large-scale primary research experiment would have been ideal, but unrealistic given these constraints. This perhaps could provide information that might even muddle how often English speakers default to the irregular reflexive form of the epicene pronoun. Moreover, the advanced eschelons of generative linguistic theory that might more soundly codify these observations can lie beyond undergraduate curricula. Unlike syntax, semantics, and phonology, morphology is not an often taught course at UC Santa Cruz.

5.3 Implications

In future, I would recommend the unified ideas presented in this paper to be taken up by a student of graduate-level or higher, to better relate them to MP, derivational timing, and distributed morphology (the theory that syntax and morphology are in essence one and the same). Building off of the research of [Arnold et al. \(2021\)](#) would undoubtedly be a beneficial decision, especially in relating the more cognitive and sociolinguistic approach of their paper to the theoretical approach of this exploration. In future, I hope to revise and build upon this paper in my time here at UC Santa Cruz as I progress through my undergraduate career. My priorities will be fortifying a more pronounced theory of the cognitive and transformational movement that might be at work in reflexive anaphora, as well as keeping a keen eye on data about epicene pronouns and inclusive language as the US continues to evolve

References

- Arnold, J. E., Mayo, H. C., & Dong, L. (2021). My pronouns are they/them: Talking about pronouns changes how pronouns are understood. *Psychonomic Bulletin & Review*, 28(5), 1688–1697. <https://doi.org/10.3758/s13423-021-01905-0>
- Ashlee, A. A., Zamora, B., & Karikari, S. N. (2017). We are woke: A collaborative critical autoethnography of three “womxn” of color graduate students in higher education. *International Journal of Multicultural Education*, 19(1), 89. <https://doi.org/10.18251/ijme.v19i1.1259>
- Balhorn, M. (2004). The rise of epicene they. *Journal of English Linguistics*, 32(2), 79–104. <https://doi.org/10.1177/0075424204265824>
- 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>
- Baron, D. E. (1981). The epicene pronoun: The word that failed. *American Speech*, 56(2), 83. <https://doi.org/10.2307/455007>
- Bodine, A. (1975). Androcentrism in prescriptive grammar: Singular “they”, sex-indefinite “he”, and “he or she.” *Language in Society*, 4(2), 129–146. <http://www.jstor.org/stable/4166805>
- Bickerton, D. (1987). He himself: anaphor, pronoun, or...?. *Linguistic Inquiry*, 18(2), 345–348.
- Chomsky, N. (1956). Three models for the description of language. *IEEE Transactions on Information Theory*, 2(3), 113–124. <https://doi.org/10.1109/tit.1956.1056813>
- Chomsky, N. (1957). *Syntactic structures*. De Gruyter. <https://doi.org/10.1515/9783112316009>
- Chomsky, N. (1965). *Aspects of the theory of syntax*. The MIT Press.
- Chomsky, N. (1967). Some general properties of phonological rules. *Language*, 43(1), 102. <https://doi.org/10.2307/411387>
- Chomsky, N. (1970). Remarks on nominalization. In: R. Jacobs and P. Rosenbaum (eds.) *Reading in English Transformational Grammar*, 184–221. Waltham: Ginn.
- Chomsky, N. (1977). On wh-movement. In: P. Culicover, T. Wasow and A. Akmajian (eds.). *Formal Syntax*. New York: Academic Press, 71–132.

- Chomsky, N. (1986). *Knowledge of language: Its nature, origin and use*. Praeger.
- Chomsky, N. (1995). *The minimalist program*. MIT Press.
- Chomsky, N., & Schützenberger, M. P. (1963). The algebraic theory of context-free languages. *Computer Programming and Formal Systems*, 35(0049-237X), 118–161. [https://doi.org/10.1016/s0049-237x\(08\)72023-8](https://doi.org/10.1016/s0049-237x(08)72023-8)
- Cohn, A. C., Cécile Fougeron, & Huffman, M. K. (2012). *The Oxford handbook of laboratory phonology*. Oxford University Press.
- Davies, Mark. (2008-) The Corpus of Contemporary American English (COCA). Available online at <https://www.english-corpora.org/coca/>.
- Dunn, D. S., & Andrews, E. E. (2015). Person-first and identity-first language: Developing psychologists' cultural competence using disability language. *American Psychologist*, 70(3), 255–264. <https://doi.org/10.1037/a0038636>
- Emonds, J. E. (1970). Root and structure-preserving transformations (Doctoral dissertation, Massachusetts Institute of Technology).
- Everett, C. (2011). Gender, pronouns and thought: The ligature between epicene pronouns and a more neutral gender perception. *Gender and Language*, 5(1). <https://doi.org/10.1558/genl.v5i1.133>
- Fiedler, S. (2015). The topic of planned languages (Esperanto) in the current specialist literature. *Language Problems and Language Planning*, 39(1), 84–104. <https://doi.org/10.1075/lplp.39.1.05fie>
- Güneş, G., & Lipták, A. (2022). *The derivational timing of ellipsis*. Oxford University Press.
- Hankamer, J. (1973). Unacceptable ambiguity. *Linguistic Inquiry*, 4(1), 17–68. <http://www.jstor.org/stable/4177750>
- Hankamer, J., & Sag, I. (1976). Deep and surface anaphora. *Linguistic Inquiry*, 7(3), 391–428. <http://www.jstor.org/stable/4177933>
- Harris, R. A. (2021). *The linguistics wars: Chomsky, Lakoff, and the battle over deep structure*. Oxford University Press.
- Huang, Y. (2013). "Neo-Gricean pragmatics". In: Huang, Y. (ed.). *The Oxford handbook of pragmatics*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199697960.001.0001>
- Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(4), 824–843. <https://doi.org/10.1037/0278-7393.20.4.824>
- Kayne, R. (2002). Pronouns and their antecedents. *Derivation and explanation in the minimalist program*, 133(166).
- Kiss, K. (1993). Wh-movement and specificity. *Natural Language and Linguistic Theory*, 11(1), 85–120. <https://doi.org/10.1007/bf00993022>
- Ko, H. (2014). *Edges in syntax: Scrambling and cyclic linearization*. Oxford University Press.
- Lakoff, G. (1971). On generative semantics. In: D. D. Steinberg & L. A. Jakobovits (Eds.), *Semantics: An Interdisciplinary Reader in Philosophy, Linguistics and Psychology* (pp. 232–296). Cambridge: Cambridge University Press.
- Lakoff, G. (1976). Toward generative semantics. *Syntax and Semantics: Notes from the Linguistic Underground*, 7, 43–61. https://doi.org/10.1163/9789004368859_005
- Langacker, R. (1986). An introduction to cognitive grammar. *Cognitive Science*, 10(1), 1–40. [https://doi.org/10.1016/s0364-0213\(86\)80007-6](https://doi.org/10.1016/s0364-0213(86)80007-6)
- Lieber, R. (2017). Derivational morphology. In: *Oxford Research Encyclopedia of Linguistics* Oxford University Press. <https://doi.org/10.1093/acrefore/9780199384655.013.248>
- Meadows, M. (2022). The effect of RC type on agreement production. In: Balachandran, L., & Duff, J. (eds.). *Syntax & Semantics at Santa Cruz*, 5. Linguistics Research Center, Department of Linguistics, University of California, Santa Cruz, 59-76. <https://escholarship.org/uc/item/2s3545j1>
- Newman, M. (1992). Pronominal disagreements: The stubborn problem of singular epicene antecedents. *Language in Society*, 21(3), 447–475. <https://doi.org/10.1017/s0047404500015529>
- Postal, Paul. 1966. On so-called pronouns in English. In: Dinneen, F. (ed.). *Report of the 17th annual round table*

- meeting on languages and linguistics*, 177–206. Washington, D.C.: Georgetown University Press.
- Pullum, G. K., & Gazdar, G. (1982). Natural languages and context-free languages. *Linguistics and Philosophy*, 4(4), 471–504. <https://doi.org/10.1007/bf00360802>
- Putnam, H. (1967). The “Innateness Hypothesis” and explanatory models in linguistics. *Synthese*, 17(1), 12–22. <http://www.jstor.org/stable/20114532>
- Siegel, L. (1998). “Gerundive nominals and the role of aspect”. In: Austin, J., & Lawson, A. (eds.). *Proceedings of ESCOL 1997*.
- Stokhof, M., & van Lambalgen, M. (2011). Abstractions and idealisations: The construction of modern linguistics. *Theoretical Linguistics*, 37(1-2), 1–26. <https://doi.org/10.1515/thli.2011.001>
- Tachau, F. (1964). Language and politics: Turkish language reform. *The Review of Politics*, 26(2), 191–204. <https://doi.org/10.1017/s0034670500004733>
- Wasow, T. (1973). The innateness hypothesis and grammatical relations. *Synthese*, 26(1), 38–56. <https://doi.org/10.1007/bf00869755>
- Zimman, L. (2017). Transgender language reform. *Journal of Language and Discrimination*, 1(1), 84–105. <https://doi.org/10.1558/jld.33139>

Appendix A: Symbols

- *: In phrase-structure rules: Kleene star. In sentence constructions: Ungrammaticality.
- ∅: Deletion of a lexical item or constituent.
- #: An ill-formed, but syntactically grammatical, utterance or sentence.
- ?: Well-formedness cannot be confidently determined.

Appendix B: Morphological Glossing Abbreviations

- 1: First person agreement
- 2: Second person agreement
- 3: Third person agreement
- ACC: Accusative case
- NOM: Nominative case
- PAST: Past tense
- PL: Plural
- PRES: Present tense
- PRO: Pronoun
- PROG: Progressive
- PRT: Participle
- SG: Singular

Appendix C: Constituent and Derivational Notation

- Adj: Adjective
- AdjP: Adjective phrase
- Adv: Adverb
- AdvP: Adverbial phrase
- C: Complementizer
- CP: Complementizer phrase
- D: Determiner
- DP: Determiner phrase

LF: Logical Form
N: Noun
NP: Noun phrase
PF: Phonetic Form
PP: Prepositional phrase
Pro: Pronoun
SR: Semantic representation
T: Tense
TP: Tense phrase
V: Verb
VP: Verb phrase