

*He_i, should do his or her_i, homework
if they_i, want a good grade:*

Variable L2 epicene pronoun usage

Jack Rechsteiner (they/them) and Matthew Kanwit (he/him)
University of Pittsburgh

New Ways of Analyzing Variation 53 | November 5th, 2025

Overview



University of
Pittsburgh®

Goals of the Project

- Investigate linguistic and extralinguistic factors that may affect development of sociolinguistic competences in L2 English (Canale & Swain, 1980; Celce-Murcia, 2008; Kanwit & Solon, 2023; Knisely, 2022)
- Examine epicene (gender-neutral) pronoun usage by L2 English learners in writing assignments (Abudalbuh, 2012; Sibanda & Begede, 2015; Zhang & Yang, 2021)
- Contribute to research on the acquisition of epicene pronoun variation (Stormbom, 2018, 2019, 2023)
- Better understand L2 English learners' usage of gendered vs. nongendered language (Zhang & Yang, 2024)

Introduction

- Contexts can arise where speakers need to refer to a single referent without specifying their gender (Newman, 1998; Stormbom 2023).
- However, English lacks a distinct and standard gender-neutral (i.e., epicene) third-person singular pronoun.
- English resolves this via three primary epicene variants:
 - singular *they*
 - *he or she*
 - generic *he*

Background

- First-language (L1) and second-language (L2) English speakers have exhibited variation with the three epicene variants.
- Prior work has primarily focused on homogenous groups of L2 English learners in their countries of origin (e.g., Abudalbuh, 2012; Sibanda & Begede, 2015).
- Factors that have been seen to influence epicene variant usage include:
 - antecedent type (Zhang & Yang, 2021)
 - grammatical case (Stormbom, 2021)
 - language background (Baranowski, 2002; Stormbom, 2018)

The Gap in the Literature

- There has not been a focused examination on the roles of cultural immersion and possible L1 effects in comparison to other L2 English learners in similar contexts.
 - Immersion has been shown to facilitate the acquisition of variable structures (Geeslin & Long, 2014; Gudmestad & Edmonds, 2022), but this has not yet been studied with epicene pronouns.
- Investigating these factors in L2 epicene pronoun variation would provide insight into the development of sociocultural competence (Celce-Murcia et al., 1995; Celce-Murcia, 2008; Kanwit & Solon, 2023; Knisely, 2022; Poehner, 2023) by L2 English learners.

The Present Study

Using texts written by 83 learners from 12 L1s in the University of Pittsburgh English Language Institute Corpus (PELIC) (Juffs et al., 2020), this study asks:

1. What are the overall rates of L2 epicene pronoun variant usage in the data?
2. To what extent do case and antecedent type influence L2 epicene pronoun variation?
3. To what extent does learners' L1 influence epicene pronoun variation?
4. To what degree do the data show effects for other sociodemographic factors (i.e., age, gender, number of years spent studying English)?

Method



University of
Pittsburgh®

Data

- Texts were produced by L2 English learners enrolled in writing classes.
 - PELIC also includes texts from grammar and reading classes, but these were excluded to control for factors of formality and task effect.
- Texts containing at least two of the aforementioned three epicene pronoun variants were extracted.
- 523 total tokens of epicene pronoun variants from 83 learners.
 - After removing instances of specific *he* and plural *they*, as well as tokens from learners who were missing demographic information.

Learner Demographics ($N = 83$)

Gender	<i>n</i>
Men	49
Women	34

Years of English learning	<i>n</i>
less than 1 year	11
1-2 years	20
3-5 years	11
more than 5 years	41

L1	<i>n</i>
Arabic	35
Chinese/Taiwanese/Turkish	19
Korean	16
Japanese/Russian/Thai	8
French/Italian/Portuguese/Spanish	5

*Learners ranged in age from 18 to 45
(mean age = 26.2; std. dev. = 5.9)*

*Due to low token counts, some L1s were combined
based on similar patterns of usage in the data.*

Data Coding

- The 523 epicene pronoun tokens were manually coded for:
 - Pronoun form (dependent variable)
 - generic *he*, *he or she*, or singular *they* (three variants)
 - Three linguistic predictors:
 - Case
 - subject, object, possessive, or reflexive
 - Antecedent
 - definite NP (*the noun*), indefinite NP (*a noun*), qualified NP (*every noun*), indefinite pronoun (*someone*), epicene pronoun, or no antecedent
 - Location of pronoun in relation to antecedent
 - same clause, different clause, or no antecedent
 - Four aforementioned sociodemographic predictors:
 - Learner gender, years of English learning, L1, age

Sample Coding

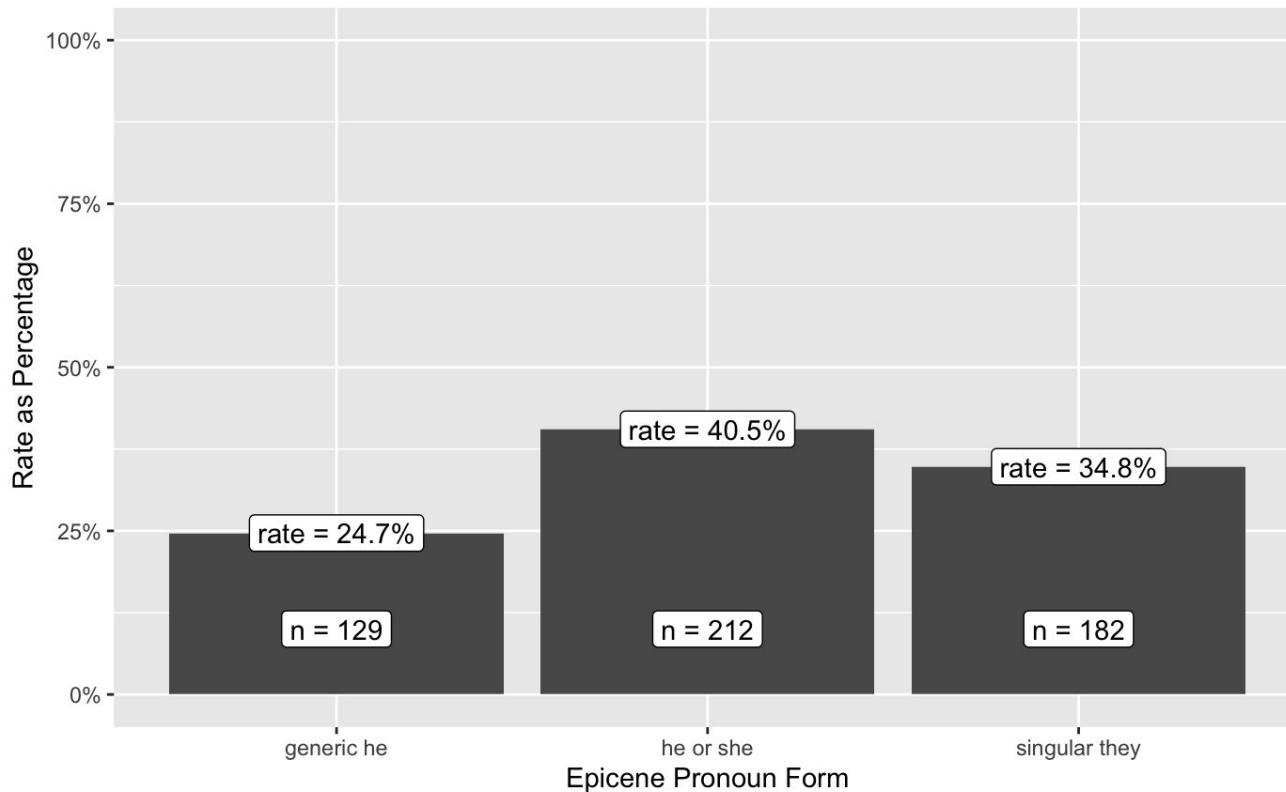
“You should communicate with a person for a long time, you might be able to know who **he or she** is.”

<i>Form of epicene pronoun</i>	he or she
<i>Grammatical case</i>	Subject
<i>Antecedent type</i>	Indefinite noun phrase (“ <i>a person</i> ”)
<i>Clause location</i>	Different clause
<i>L1</i>	Chinese
<i>Age</i>	28
<i>Gender</i>	Woman
<i>Years spent learning English</i>	> 5 years



Results

RQ 1: Counts and rates of epicene pronoun variants



Statistical Modeling for RQs 2, 3, and 4

A mixed-effects multinomial regression was performed using Jamovi for R to compare the three epicene pronoun variants according to the following predictors:

- Linguistic predictors:
 - grammatical case, antecedent type, and clause location
- Extralinguistic predictors:
 - gender, years of English study, age, and L1

Summary of Mixed-Effects Multinomial Regression

Linguistic Variables		
Antecedent	Case	Clause Location
.020*	.705	.428

Extralinguistic Variables			
Age	Gender	L1	Years of English Learning
.417	.085	<.001***	.102

* = $p < .05$, *** = $p < .001$

Coefficients of Significant Predictors

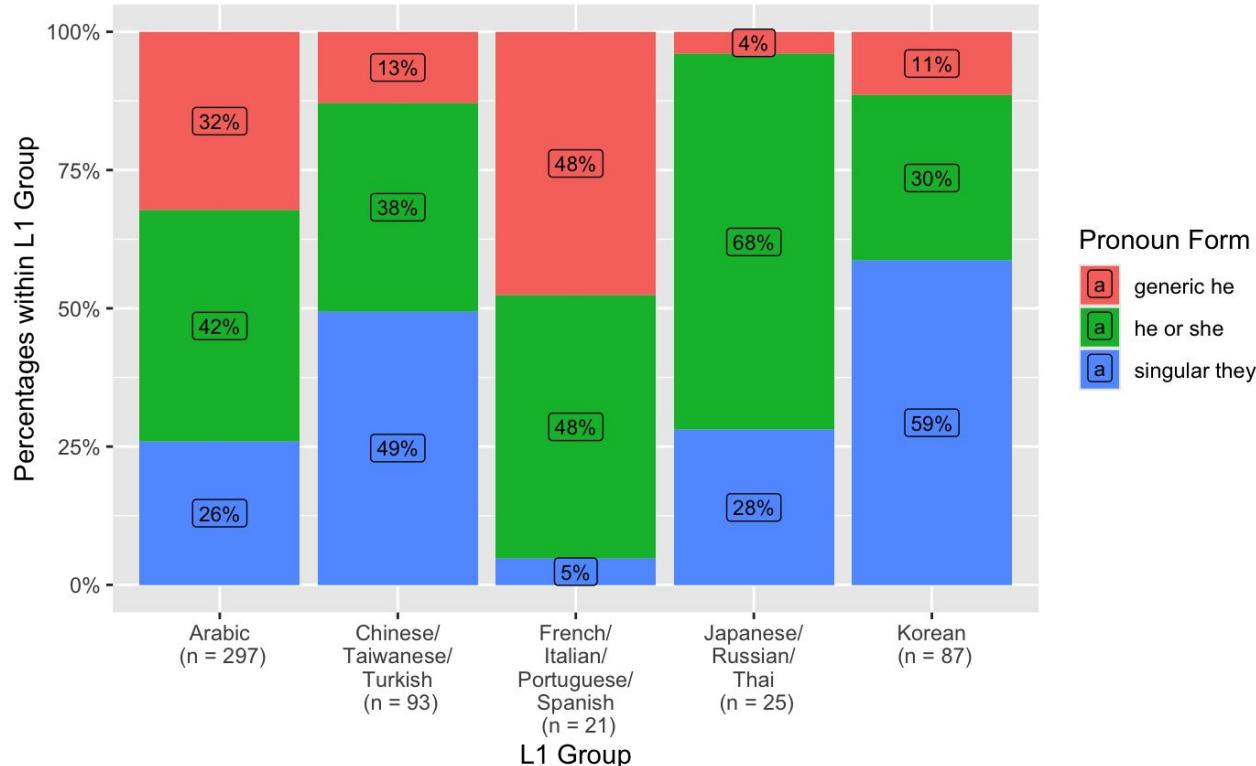
Variable	generic <i>he</i> compared to <i>he or she</i>					singular <i>they</i> compared to <i>he or she</i>				
	Est.	95% CI	Std. Error	z-score	p-value	Est.	95% CI	Std. Error	z-score	p-value
(Intercept)	-1.37	[-2.05, -.70]	0.34	-4.02	<0.001	-0.68	[-1.27, -.09]	0.30	-2.26	0.024
L1 (<i>Ref. level = Arabic</i>)										
Chinese/Taiwanese/Turkish	-1.61	[-2.95, -.26]	0.69	-2.24	0.020*	1.01	[0.09, 1.94]	0.47	2.15	0.032*
French/Italian/Portuguese/Spanish	0.34	[-1.54, 2.22]	0.95	0.36	0.721	-1.21	[-3.67, 1.23]	1.25	-0.97	0.331
Japanese/Russian/Thai	-2.77	[-5.28, -.27]	1.27	-2.18	0.030*	-0.09	[-1.46, 1.28]	0.70	-0.13	0.900
Korean	-2.27	[-4.03, -.52]	0.89	-2.54	0.011*	1.33	[.36, 2.31]	0.50	2.67	0.008**
Antecedent (<i>Ref. level = Definite NP</i>)										
Epicene pronoun	0.34	[-.41, 1.09]	0.38	0.89	0.374	0.99	[.24, 1.73]	0.38	2.59	0.010*
Indefinite NP	-0.14	[-1.05, .76]	0.46	-0.31	0.760	0.32	[-.52, 1.16]	0.43	0.75	0.455
Indefinite pronoun & Qualified NP	-1.12	[-2.09, -.15]	0.50	-2.27	0.023*	-0.03	[-.88, .82]	0.44	-0.07	0.940
No antecedent	0.08	[-0.88, 1.04]	0.49	0.16	0.872	0.41	[-.43, 1.25]	0.43	0.95	0.341

* = $p < .05$, ** = $p < .01$

Relative Rates by L1

Significant relations:

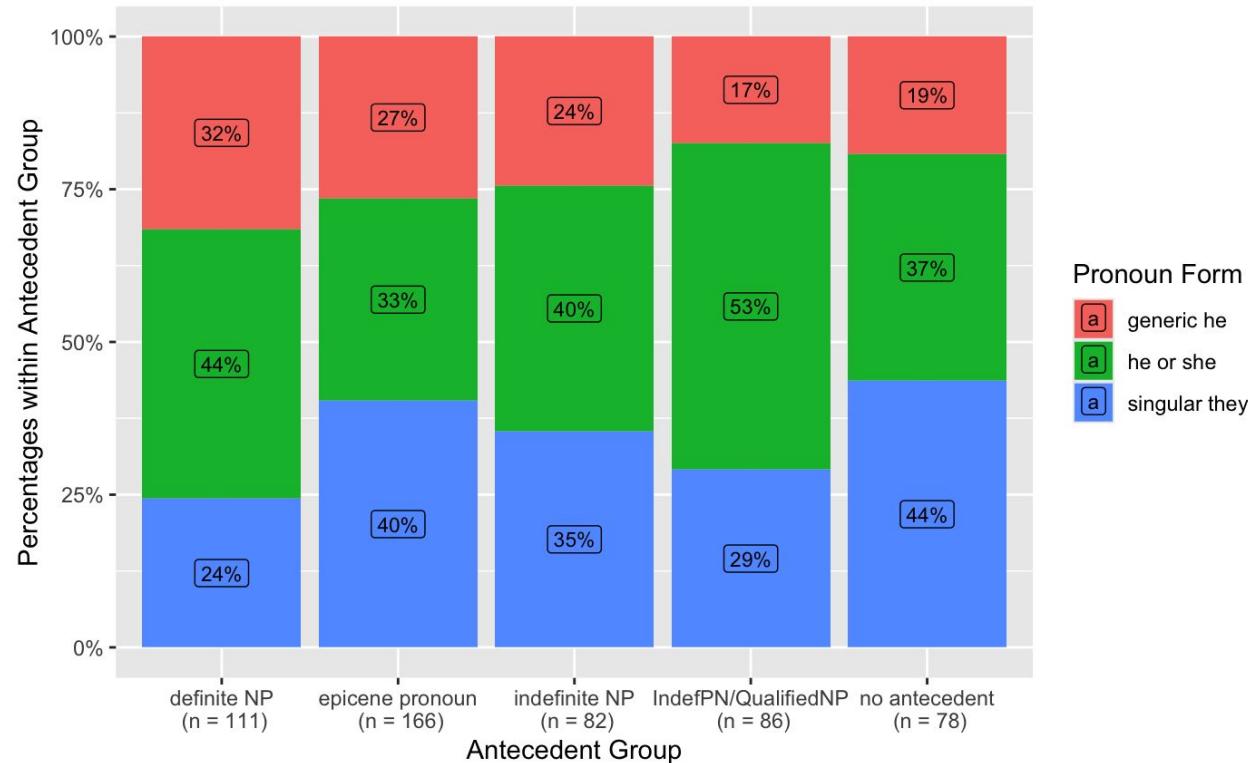
- Reference level: Arabic
- Chinese/Taiwanese/Turkish and Korean using less *he* and more *they*
- Japanese/Russian/Thai using less *he* and more *he or she*



Relative Rates by Antecedent

Significant relations:

- Reference level: Definite NP
- Indefinite pronoun/Qualified NP using less *he* and more *he or she*
- Epicene pronoun using less *he or she* and more *they*



Discussion



University of
Pittsburgh®

Discussion: Answers to Research Questions

1. L2 English learners in PELIC used *he or she* most frequently (40.5%) and generic *he* least frequently (24.7%).
2. Antecedent had a significant effect on epicene pronoun variation, but case did not.
 - Indefinite pronoun & qualified NP antecedents disfavored generic *he* in comparison to *he or she*.
 - Epicene pronoun antecedents favored singular *they* in comparison to *he or she*.
3. L1s without grammatical gender tended to disfavor generic *he* and favor singular *they* in comparison to *he or she*.
 - Russian was the exception, but there were only 3 tokens from a Russian speaker.
4. No significant effects for sociodemographic factors (other than L1) were found.

Discussion: Interpretation of Results

- The high rate of *he or she* may reflect that PELIC learners likely receive *he or she* forms as input in prescriptive American English sources.
 - Research outside the US has shown high rates of generic *he* (e.g., Abudalbuh, 2012; Zhang & Yang, 2021).
 - American English favors *he or she* in formal writing (Curzan, 2014).
- Indefinite pronoun and qualified NP antecedents disfavor generic *he* because they are semantically plural (cf. Baranowski, 2002).
- Epicene pronoun antecedents favor singular *they* due to their gender neutrality.
- The L1 effects observed here are consistent with findings for L2 English learners in settings outside of the US (Stormbom, 2018, 2021).



Considerations and Limitations

- Method is different from prior studies on L2 epicene pronoun variation.
 - Previous studies included learners exhibiting categorical usage.
 - To focus on speakers who demonstrated variable usage (following Tagliamonte, 2025), we excluded texts with categorical epicene pronoun usage.
- For antecedent predictor, tokens were coded as “no antecedent” if the antecedent did not occur earlier in the sentence.
 - e.g., *He or she should study before a test* was coded as “no antecedent”.
 - Learners tended to introduce and maintain reference at the sentential level (85% of tokens). Future work will investigate potential differences between sentence-level and discourse-level antecedents.

Contributions

- This study provides insight on the influence of L1 and sociocultural background on patterns of gendered speech variation in L2 English.
- The patterns seen here don't align with L1 English speaker patterns, which have shown declining usage of *he or she* (Stormbom, 2021, 2023).
 - Learners are often observed to use more formal/prescriptive language than L1 English speakers (Bayley & Tarone, 2009; Mougeon et al., 2010)
 - Learners could potentially be misinterpreted as using exclusionary language or conveying the wrong information.

Future Directions

- Future work will benefit from greater consideration of epicene pronoun forms in L2 English textbooks in order to identify areas for improvement (e.g., Paiz, 2015).
 - High rates of *he* or *she* may be due to outdated language teaching materials.
- Understanding the usage and pedagogical sources of gendered and non-gendered language helps contribute to gender-just language teaching (e.g., Knisely, 2022; Zhang & Yang, 2024).

Thank you!

Jack Rechsteiner (they/them)

jackrechsteiner.github.io

JJR156@pitt.edu

Matthew Kanwit (he/him)

linguistics.pitt.edu/people/matthew-kanwit

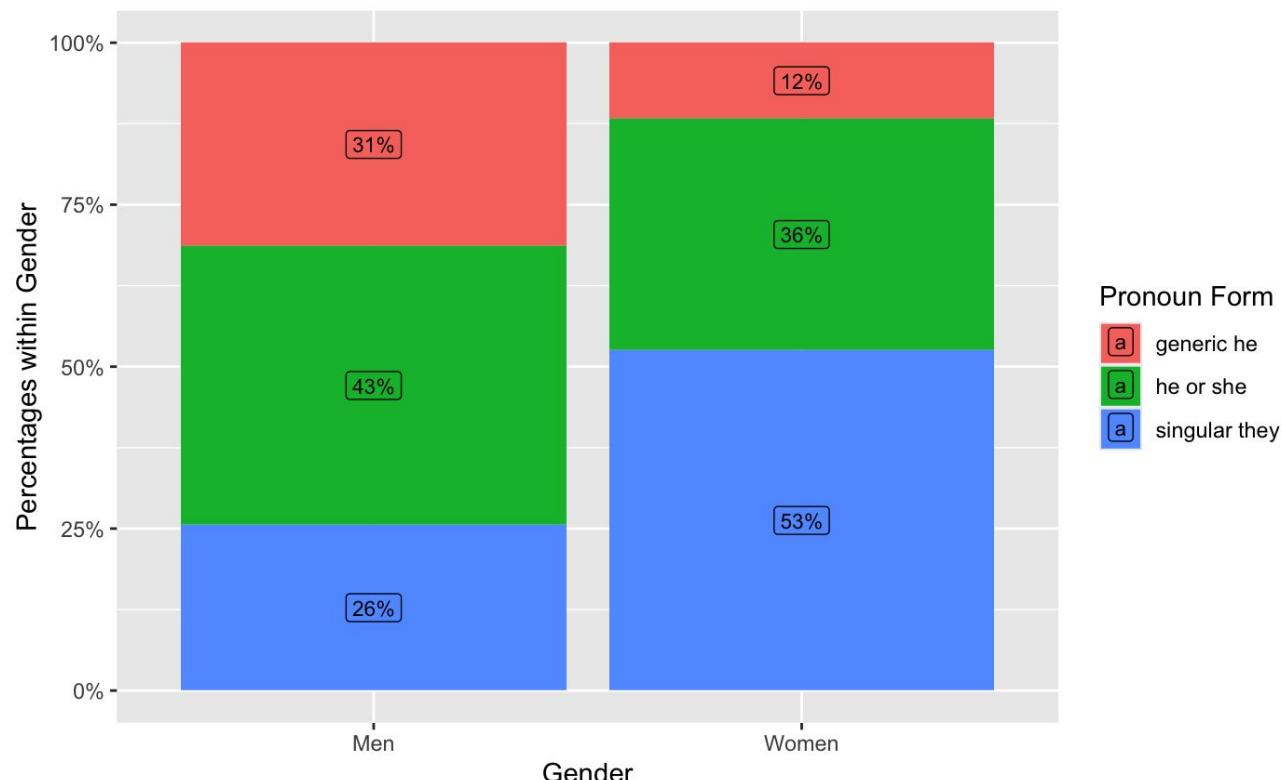
mkanwit@pitt.edu

Works Cited

- Abudalbuh, M. (2012). *Ideology, gender roles, and pronominal choice: A sociolinguistic analysis of the use of English third person generic pronouns by native speakers of Arabic* [PhD Thesis]. University of Kansas.
- Baranowski, M. (2002). Current usage of the epicene pronoun in written English. *Journal of Sociolinguistics*, 6(3), 378–397.
- Bayley, R., & Tarone, E. (2009). Variationist perspectives. In *The Routledge Handbook of Second Language Acquisition*. Routledge. <https://doi.org/10.4324/9780203808184.ch3>
- Canale, M., & Swain, M. (1980). Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing. *Applied Linguistics*, 1(1), 1–47. <https://doi.org/10.1093/applin/I.1.1>
- Celce-Murcia, M. (2007). Rethinking the Role of Communicative Competence in Language Teaching. In E. A. Soler & M. P. S. Jordà (Eds.), *Intercultural Language Use and Language Learning* (pp. 41–57). Springer Netherlands. https://doi.org/10.1007/978-1-4020-5639-0_3
- Celce-Murcia, M., Dornyei, Z., & Thurrell, S. (1995). Communicative Competence: A Pedagogically Motivated Model with Content Specifications. *Issues in Applied Linguistics*, 6(2).
- Curzan, A. (2015). *Fixing English: Prescriptivism and language history* (First paperback edition). Cambridge University Press.
- Geeslin, K. L., & Long, A. Y. (2014). *Sociolinguistics and second language acquisition: Learning to use language in context*. Routledge.
- Gudmestad, A., & Edmonds, A. (2022). Exploring crosslinguistic influence in gender marking in Spanish. *Studies in Second Language Acquisition*, 44(4), 998–1019.
- Juffs, A., Han, N.-R., & Naismith, B. (2020). *The University of Pittsburgh English Language Institute Corpus (PELIC)* (1.0) [dataset].
- Kanwit, M., & Solon, M. (2023). Introduction. In M. Kanwit & M. Solon, *Communicative Competence in a Second Language* (1st ed., pp. 1–18). Routledge. <https://doi.org/10.4324/9781003160779-1>
- Knisely, K. A. (2022). Gender-just language teaching and linguistic competence development. *Foreign Language Annals*, 55(3), 644–667.
- Mougeon, R., Nadasdi, T., & Rehner, K. (2010). *The Sociolinguistic Competence of Immersion Students*. Multilingual Matters. <https://doi.org/10.21832/9781847692405>
- Newman, M. (1998). What can Pronouns Tell Us? A Case Study of English Epicenes. *Studies in Language*, 22(2), 353–389.
- Paiz, J. M. (2015). Over the monochrome rainbow: Heteronormativity in ESL reading texts and textbooks. *Journal of Language and Sexuality*, 4(1), 77–101.
- Paterson, L. L. (2020). Non-sexist language policy and the rise (and fall?) of combined pronouns in British and American written English. *Journal of English Linguistics*, 48(3), 258–281.
- Poehner, M. E. (2023). Sociocultural considerations of communicative competence. In M. Kanwit & M. Solon, *Communicative Competence in a Second Language* (1st ed., pp. 60–75). Routledge.
- Sibanda, J., & Begede, M. P. (2015). University students' preferences for gender specific and gender neutral occupational terms and gender neutral singular pronouns: A case study. *Journal of Social Sciences*, 42(1–2), 11–21. <https://doi.org/10.1080/09718923.2015.11893389>
- Stormbom, C. (2018). Epicene pronouns in intermediate to advanced EFL writing. *International Journal of Learner Corpus Research*, 4(1), 1–22. <https://doi.org/10.1075/ijlcr.16016.sto>
- Stormbom, C. (2019). Language change in L2 academic writing: The case of epicene pronouns. *Journal of English for Academic Purposes*, 38, 95–105. <https://doi.org/10.1016/j.jeap.2019.02.001>
- Stormbom, C. (2021). *Gendered language in flux: The use of epicene pronouns in EFL and ELF writing* [PhD Thesis]. Åbo Akademi University Press.
- Stormbom, C. (2023). Epicene pronouns new and old. In L. L. Paterson (Ed.), *The Routledge Handbook of Pronouns* (1st ed., pp. 411–420). Routledge. <https://doi.org/10.4324/9781003349891-34>
- Tagliamonte, S. A. (2025). *Analysing sociolinguistic variation* (Second edition). Cambridge University Press.
- Zhang, X., & Yang, H. (2021). Gender voices in Chinese university students' English writing: A corpus study. *Linguistics and Education*, 64, 100935. <https://doi.org/10.1016/j.linged.2021.100935>
- Zhang, X., & Yang, H. (2024). Perception and usage of English epicene pronouns among L2 teachers in China - focusing on he, he or she and they. *European Journal of Education*, 59(3), e12651.

Additional Slides

Relative Rates by Gender



Different forms of *he* or *she* in the data

Subject	
Form	Count
he or she	47
he/she	25
h/she	16
she/he	7
she or he	3
s/he	1
<i>Total</i>	99

Possessive	
Form	Count
his or her	38
his/her	35
her/his	2
his/she	1
<i>Total</i>	76

Object	
Form	Count
him or her	17
him/her	14
him or she	1
<i>Total</i>	32

Reflexive	
Form	Count
her or himself	1
him or herself	1
his or herself	1
himself/herself	1
his/herself	1
<i>Total</i>	5