How many selves for them*

Kirby Conrod, Ruth Schultz & Byron Ahn

1 Introduction

English *they*-series pronouns can be anteceded by morphosyntactically singular nominals ("singular *they*"): in (1), bracketed nominals control 3.SG *-s* on the verb and antecede *their* pronouns. This paper concerns the reflexive forms of singular *they*, focusing on two possible morphological surface realizations —*themself* and *themselves*— exemplified in (2)

(1) a. [The ideal student]$_1$ always completes their$_1$ homework on time.[1]
     b. [Each photographer]$_1$ is bringing their$_1$ own camera.

(2) a. [The ideal student]$_1$ always completes it themself$_1$/themselves$_1$.
     b. [Each photographer]$_1$ respects themself$_1$/themselves$_1$.

Using acceptability data from a large-scale survey, this study aims to: document the distributions of *themself* and *themselves* and identify which patterns are determined linguistically and/or socially. Accordingly, we formulated the following four hypotheses:

Hyp.1: Nonbinary people will rate '*themself*' higher than other people.
Hyp.2: Prescriptivists will rate '*themselves*' higher than '*themself*'.
Hyp.3: Younger speakers will rate '*themself*' higher than older speakers do.
Hyp.4: With antecedents of lower specificity, '*themselves*' will be rated higher than '*themself*'; this will be reversed for higher specificity.

To test these hypotheses, we ran an acceptability survey in which 3rd-person singular subject nominals anteceded a singular *they* reflexive (*themself*/*themselves*). Antecedents fell on a seven-point scale of continuous pragmatic specificity, shown top-to-bottom in order from least definite/specific to most definite/specific in Table 1.

---

[1]Note: For all examples with *themself*/*themselves* with singular antecedents, there is variation in acceptability across speakers – this variation is a primary focus of this research. Sentences without acceptability diacritics should be interpreted to mean that such an expression is acceptable to some speakers.)

Antecedent type	Example antecedent
quantified indefinite	*someone who is passionate about art*
quantified universal	*every child in the class*
generic kind definite	*the ideal candidate for this job*
distal definite	*the customer who wrote this review*
specific indefinite	*an employee at the movie theater*
proximate definite	*the person I talked to yesterday*
proper name	*Alex, who is quite short*

Table 1: Antecedent types arranged from least specific (top) to most (bottom)

As we will argue, this work does <u>not</u> support an analysis whereby acceptability of *themself* / *themselves* gradiently depends on its antecedent's specificity, valued on a continuous scale like the one depicted in Table 1. Instead, the ratings of different antecedents cluster into three groupings: quantified nominals, context-distal nominals, and definite specific nominals.

The paper is structured as follows. We give precedent for our hypotheses in §2, and then describe the stimuli and methodology in §3. Results are reported in §4 and discussion in §5 suggests that explanations must go beyond social variables and gradient notions of specificity. We conclude in §6 that there are underexplored grammatical constraints, laying groundwork for future work in this domain.

2 Background: Sociolinguistics and Morphosyntax

Sociolinguistic variation of singular *they* We begin by discussing sociolinguistic variation in (non-reflexive) "singular *they*" pronouns. Previous work has shown the use and acceptability of singular *they* varies sociolinguistically, especially with respect to definite specific singular *they* (henceforth DST).

Previous descriptive work on DST has analyzed it in the Language Variation and Change framework ("LVC"). LVC's primitives are sociolinguistic *variables*, with *variants* stochastically deployed, yielding both synchronic variation (within and between speakers) and diachronic variation. Within this framework, the Apparent Time Hypothesis (Weinreich et al., 1968) posits that diachronic change can be inferred using synchronic data on the basis of variation by speaker age. Conrod's 2019 LVC-based exploration of DST uncovered both "innovative" grammars (ones where a new feature is picking up steam) and "conservative" ones (which lack such an innovation).[2]

Speaker-related social categories have been found to influence the ratings of singular *they* with definite specific antecedents differently than with indefinite, quantified, and generic antecedents. Conrod (2019) showed ratings of DST were negatively correlated with age when the antecedent was a proper name, but not when the antecedent was indefinite or generic. Conrod also found that non-binary participants rated DST higher than men and women, who rated it similarly; additionally, they found that men, women, and non-binary respondents who identified themselves as trans rated DST higher than those who did not. Similar effects on acceptability ratings of DST have also been observed by Block (2019), and parallel findings have been made in a production experiment by Sheydaei (2021).

In addition to the effects of social categories, several studies have examined the ideological beliefs in relation to DST. Bradley (2020) found that endorsement of linguistic prescriptivist ideology

---

[2]Bjorkman (2017) and Konnelly & Cowper (2020) also use "*innovative*" and "*conservative*" to describe variants, but do not explicitly situate themselves in the LVC framework.

was correlated with lower ratings of DST, as was endorsement of benevolent sexist ideology. Hernandez (2020) found that while there was a correlation between prescriptivism and DST ratings in general, participants who identified themselves as heterosexual and cisgender had a stronger correlation between prescriptivist ideology and (lower) ratings of DST than other groups.

Morphosyntax of singular *they* To begin a discussion of the morphosyntax of singular *they* forms, consider the small dataset in (3), which we will reference in the paragraphs ahead. (*Grammatical versions of these sentences come from abcnews.com.*)

(3) a. ...[whoever has custody]$_1$ when they$_1$ walk/*walks into court has custody...
    b. ...[the person]$_1$ is/*are not turning themselves$_1$ in now...
    c. ...[he or she]$_1$ finds/*find themself$_1$ facing an enemy...

Beginning with non-reflexive contexts in (3a), note that *they* triggers plural verb agreement, indicating it is grammatically plural even when it has a [sg] antecedent. Previous analyses of singular *they* have proposed that in order to allow for the gender-neutral, specific definite, and non-specific definite uses, '*they*' lacks a privative singular [sg] feature (Bjorkman 2017; Konnelly & Cowper 2020; Conrod 2019). In lacking such a feature, the semantic content of '*they*' does not limit felicitous use to contexts where the antecedent is plural; instead, it is referentially compatible with both singular and plural referents.[3] Analyses of this sort (e.g., Bjorkman 2017; Wiltschko 2008) propose that gender features are also optional ('adjunct features'), and so singular *they* is radically underspecified for all ϕ features (person, number, and gender).

We expect *themselves* to allow [sg] antecedents, as in (3b), if reflexive anaphors like *themselves* morphosyntactically contain a *they* pronoun alongside a separate *self* nominal.

(4) [$_{DP}$ [$_{DP}$ them] [$_{NP}$ selves]]   (cf. Postal 1966; Helke 1973; Ahn & Kalin 2018)

In this way, the acceptability of *themselves* in contexts like (3b) can share a root cause with the acceptability of *they* in (3a), as they both have ϕ-featureless *they* pronouns that are compatible with singular antecedents. By extension, the plural *-selves* can be understood as again lacking the privative [sg] feature (thereby being referentially compatible with a syntactically [sg] antecedent). Note that this entails that the ϕ-features in a reflexive form may differ from the those of the local antecedent of binding (Ahn 2019 and Collins & Postal 2012 discuss this beyond number features).

The preceding explanation presupposes that each anaphor-internal nominal in a reflexive anaphor in English has its own set of ϕ-features. In fact, these feature bundles need not match one another, as demonstrated by forms like *themself*, exemplified in (3c), well as by *ourself* (and arguably *yourself*).[4] The *themself* form is well-attested, with several examples in the Oxford English Dictionary, and is readily discoverable in corpora.

---

[3]Under this sort of analysis, when using a plural nominal is infelicitous, it is not an issue of entailment or presupposition of a plural feature; rather it should be seen as an issue of scalar implicatures.

[4]This raises the question of why forms like *himselves* are not widely attested; we set this issue aside for now, noting that feature-matching constraints are complex, and reflexive forms containing 3.SG pronominals have been otherwise noted to be constrained differently than other reflexive forms (Ahn 2019).

3 Methods

The present study examines *themself* and *themselves* anteced by different types of singular, un-gendered antecedents, as well as collecting demographic data to explore sociolinguistic variation. In this section, we describe a new large-scale study ($n = 1,127$), disseminated via Twitter ($n = 724$) and Prolific ($n = 403$), and implemented using Qualtrics. The survey measured acceptability of reflexive forms containing a singular *they* pronoun, in conjunction with the social variables that have been explored in related previous work.

In the acceptability survey, linguistic context was manipulated to probe its effect on the acceptability of *them-self/-selves* with respect to seven levels of specificity (given in Table 1). This created 28 stimuli sentences (*7 antecedent types × 2 reflexives types × 2 sentences per condition*); participants rated each one, presented in a randomized order, on a 5-point Likert scale from "*1 – very unnatural*" to "*5 – very natural*." Across all the levels of specificity, each sentence was formed in the same syntactic frame, wherein the antecedent was a singular gender-neutral subject with a post-nominal modifier, and the reflexive anaphor was an argument of the predicate. As a pilot task, there were no distractor items or controls. The full list of stimuli can be accessed in our online repository at https://osf.io/5kr8s/.

Following the acceptability rating task, we collected participants' social demographic information and ideologies about gender and language. The demographic questionnaire asked for participants' age (by birth year) and gender (in both multi-answer and limited-answer form).

The prescriptivist language ideology instrument we used is based on a subset of questions adapted from Hernandez (2020); participants were asked to rate 8 items on a Likert scale from "*1 – Strongly disagree*" to "*5 – Strongly agree*." We developed these questions with the aim to probe whether participants endorsed the general belief that there is a '*correct*' form of grammar, and that deviations from the '*correct*' form are universally '*incorrect*.' We developed a similar survey instrument for binary gender ideology to probe whether participants endorsed the general belief that gender is made up of exactly two, inherent, immutable categories.

For both ideologies of prescriptivist language and binary gender, we calculated a score for each participant based on their answers to the Likert-scale questions. Each participant therefore received a score out of 40 to rate their endorsement of prescriptivist language ideology (8 items × 5 points possible for each item) and a score out of 15 for their endorsement of binary gender ideology (3 items × 5 points possible).

In addition to demographics questions, participants were asked to briefly rate their political views on a 5-point, unnumbered scale from "*Very conservative*" to "*Very liberal*." Finally, participants were asked to report whether they had ever taken a class in linguistics. All questions for survey instruments are in an online repository at https://osf.io/5kr8s/.

4 Results

In this section we discuss the ratings of sentences with *themself* and *themselves*, looking first at the differences by antecedent type, then by social categories (focusing primarily on age and gender; see §2) and social ideological scales.

Antecedent types   We originally hypothesized that *themself* would be rated lower for less specific antecedents, and *themselves* would be rated lower for more specific antecedents. Our pilot data, however, did not confirm this. Figure 1 shows that, on average, the ratings of *themself* and *themselves* do not show difference in statistical significance for antecedents that were quantified univer-
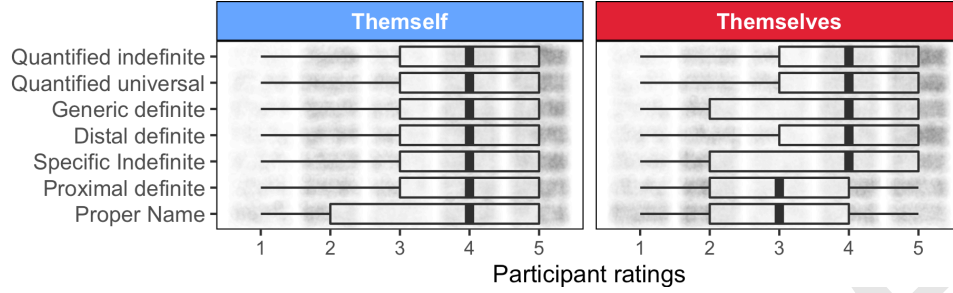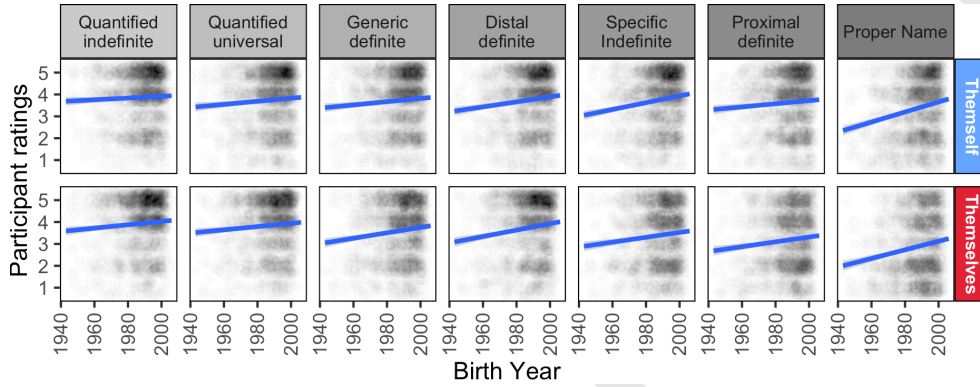
Figure 1: Ratings by antecedent type

Figure 2: Birth year and ratings by antecedent type

sals or definite specifics. For all other antecedent types, *themself* was rated higher than *themselves* (counter to our hypothesis, and contra Ackerman et al. 2018). We also found that proper names and specific antecedents were rated lower for *themselves* than *themself* (in a logistic regression, $p < 0.001$).

Relating also back to Hypothesis 4, it does not appear to be the case that there is a continuous decline in acceptability across all seven antecedent types. The fact that we did not find significant differences in the ratings of each type of antecedent instead suggests that this seven-point scale of specificity and definiteness may in fact be *too* fine-grained; we discuss this possibility in greater detail in §5 below.

Social categories: age and gender For social variables, we predicted that age would predict ratings for more definite antecedents (particularly proper names), and that age would have an affect on ratings of *themself* versus *themselves*. In a logistic regression we found no interaction between age and *-self/-selves* that predicted acceptability ratings. There was, however, an interaction between age and antecedent type: older participants were significantly more likely to rate *them*-anaphors lower for proper names ($p < 0.0001$), as well as for indefinite ($p < 0.01$) and definite specific ($p < 0.01$) antecedents. There were no three-way interactions between age, *-self/-selves*, and antecedent type.

We also predicted effects of participant gender on ratings. Non-binary participants ($n = 139$) rated all *them-* anaphors higher than women ($n = 651$) or men ($n = 321$) did (logistic regression,
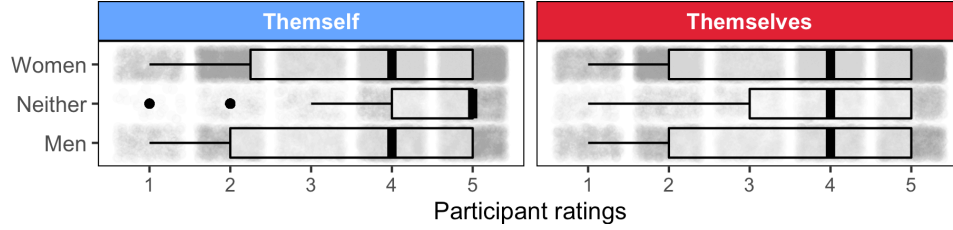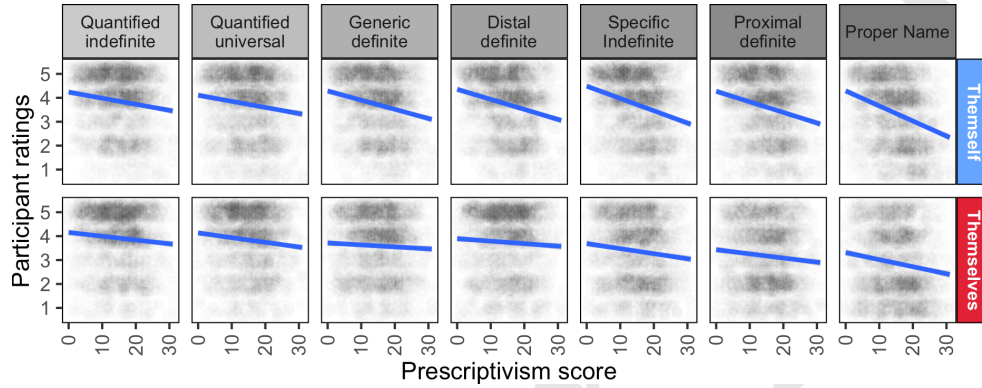
Figure 3: Ratings by participant gender

Figure 4: Prescriptivism and ratings by antecedent type

$p < 0.001$).[5] Non-binary participants also rated *-self* higher than *-selves* ($p < 0.001$). There was no interaction between participant gender and antecedent type.

Ideological beliefs   We predicted that participants who endorsed prescriptivist language ideology would rate sentences with *themself* lower, as well as rating sentences with more definite antecedents lower; this was borne out. Using the score of endorsement of prescriptivist ideology, we found that participants with a higher prescriptivism score rated sentences lower overall ($p < 0.0001$); higher prescriptivism correlated with lower ratings of sentences with *themself*

---

[5]Some participants ($n = 16$) opted not to be included in any gender group in our analyses.
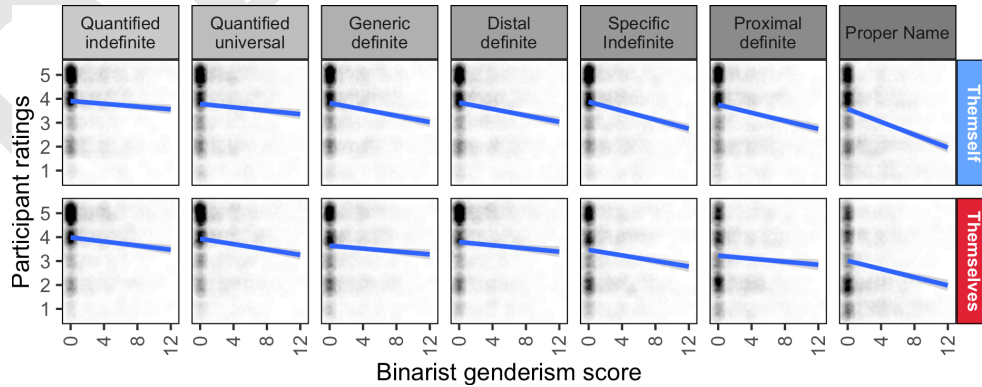

Figure 5: Binarist genderism and ratings by antecedent type

Hypothesis	Finding
H1: Nonbinary raters: *-self* > *-selves*	Supported
H2: More prescriptive raters: *-selves* > *-self*	Supported; prescriptivism scores correlated with greater effects on more specific antecedent types
H3: Younger raters: *-self* > *-selves*	Partially supported; effect found only for indefinites, distal definites, and proper names
H4: Specific antecedents: *-self* > *-selves*; vice-versa for lower specificity	Partially supported; effect found only for proper names and proximal definites

Table 2: Summary table of hypotheses and findings for forms of singular *they* anaphors

($p < 0.0001$), and also with lower ratings of sentences with all antecedents aside from quantified universals ($p < 0.05$ for all antecedent groups).

Summary   Table 2 summarizes these results alongside our original hypotheses.

5   Discussion

The results of this pilot study suggest that referential properties of the antecedent do play a role in the acceptability of *themself* and *themselves*. Our initial assumption that *themselves* would be a generally preferred form may not be correct for the majority of our participants, for whom *themself* appears to be more widely acceptable; we suggest that this may be due to a lack of a plural feature on the singular antecedents, which is in turn compatible with the *-self* form.

A tempting analysis may be that 'notionally singular' nominals are what exhibit lower ratings with *themselves* compared to *themself* (See Fig. 1). This would serve to explain why proper names and proximal definites (which always pick out a single, familiar individual) behave differently than, for example, quantified universal antecedents like *everyone* (which syntactically may be singular, but may be notionally *not* singular). This idea of notional number is insufficient, however, because *someone* antecedents (very much notionally singular) pattern with *everyone*: it is higher rated with *themselves*, and higher rated by "conservative" speakers. Thus, if notional number matters, it must also interact with definiteness/specificity: "definite, specific, and notionally singular" may be a description that picks out only proximal definites and proper names.

This suggests that syntactic values of ϕ-features may be less relevant than context, when deciding whether to use *-self* or *-selves*. Instead, we analyze this as being derived by the *self* nominal being able to merge with a valued number feature, which is checked against the antecedent for interpretive compatibility (confirming similar results in Conrod 2019 and Ahn 2019), rather than being valued over the course of the syntactic derivation via e.g. Agree with the antecedent. Further evidence for the lack of complete syntactic ϕ-matching in anaphor morphosyntax is that there are feature-mismatches possible between the *they* pronoun (lacking a [sg] feature) and *-self* (being specified as [sg]). We leave it to planned future work to explore the nature of which feature bundles can co-occur (or not) between the antecedent, anaphoric pronoun, and anaphoric *-self* nominal.

In our initial hypotheses, we expected that the pragmatic effects of antecedent on the acceptability of both *themself* and *themselves* would be detectable through a continuous grade of acceptability across all types of antecedents; instead, however, we found that the seven antecedent types do not exhibit a clear continuous effect on the ratings. Instead, what is minimally clear is that proximal definites and proper names stand out as different (Fig.1). Further discussion on modelling the effects of antecedent type (as categorical) are in Ahn & Conrod (*in prep.*).

In addition, these results exhibit quite a bit of inter-participant variability; participants' age and gender both had an effect on how acceptable they rated the sentences, suggesting that the use of *themselves/themself* is a sociosyntactically active variable. This is consistent with what we expect from a variable undergoing change. Notably, singular *they* has been consistently used with indefinites and non-specific definites antecedents throughout the Modern English period and earlier (see Bjorkman 2017's overview). Moreover, an Apparent Time analysis of the negative correlation between age and ratings supports proposals that singular *they* is currently undergoing language change. In particular, Conrod's empirical study finds that speakers with innovative grammars more readily accept singular *they* with definite specific antecedents. The present findings confirm this: effects across both singular *they* anaphor forms reaffirm that DST is undergoing change, not singular *they* in general. In addition, what might otherwise be called "free variation" between *themself/themselves* is in fact similarly a sociosyntactically determined, warranting both variationist and generative analyses.

That gender correlates with acceptability of *themself* (and indeed higher ratings of all forms of the anaphor) is also expected under a LVC approach. What is interesting about these findings, however, is that the inclusion of a nonbinary gender category in the demographics questionnaire allowed us to identify the actual leaders of the change in progress: women and men rated sentences similarly, while nonbinary respondents rated "innovative" variants much higher.[6] Because many decades of variationist sociolinguistics research have operated in a mostly-binary view of gender identity, there is not much precedent leading us to expect that nonbinary language users might lead language change more generally (though see Eckert 2014 for discussion of how binary approaches to gender identity have been complicated throughout this literature). Because of the close ties between the use of *they* as a singular gender-neutral pronoun and the conception of nonbinary identities, we do not necessarily expect that this pattern will generalize to other sociolinguistic variables; however, we hope that future sociosyntactic work takes an expansive approach to gender in order to empirically test this. Konnelly & Cowper (2020) and Conrod (2019) both discuss the connection between *they* and non-binary approaches to gender further than we have space to do here.

One final question that remains is whether the interspeaker variation we observed is fully explained by speaker demographics. Because *themself* contains an apparent mismatch in features, it may be surprising that *themself* was rated better than *themselves* overall. If the ratings could be fully explained by speaker variables, we might expect this result to be an indication that our respondents were overall somewhat young and "innovative". It is also possible, however, that there exist different grammars in the speech community that are related to, but not completely determined by, speaker demographic categories; if this is the case, then we should expect in future work to be able to identify discrete groups of speakers who pattern similarly in terms of how they rate *themself/themselves* with various antecedents. See Ahn & Conrod *in prep.* for further explorations of this, using clustering analysis on ratings data to uncover and hypothesize about grammatical differences.

6 Conclusion

Our large-scale study on the acceptability of singular-*they* anaphors (*themself/themselves*) identifies some social and linguistic sources of variation. Age, gender, and ideologies correlated with how sentences with singular-*they* anaphors were rated generally, and that correlation was stronger

---

[6]For reasons of space, this analysis doesn't explore effect of gender orientation (e.g., trans men vs cis men). See Conrod (2019) exploration in this vein.

with proximal definite and proper name antecedents. Participants showed a slight preference for *themself* across contexts with singular antecedents, and deeper investigation revealed that *themselves* was rated higher for indefinite and quantified singular antecedents than for other singular antecedents.

As these aggregated findings wash out important details of interspeaker variation, further work is necessary to identify the properties of grammars that yield these results. Even at this level of analysis, we see that the grammatical number of the anaphor-internal pronoun and of the antecedent are not deterministic for the number of the *-self* nominal. The system that determines when such number features must (or need not) match is a sociolinguistically active variable among English speakers.

References

Ackerman, Lauren, Nick Riches & Joel Wallenberg. 2018. Coreference dependency formation is modulated by experience with variation in human gender. Presented at the 2018 Annual LSA Meeting: `https://osf.io/jdcf8/`.

Ahn, Byron. 2019. Features, identity, and 'yourself'. In *NELS 49*, vol. 1, 15–24.

Ahn, Byron & Kirby Conrod. In preparation. Grammatical variation: themself and themselves (*working title*). ms., Princeton University & Swarthmore College.

Ahn, Byron & Laura Kalin. 2018. What's in a (English) reflexive? In *NELS 48*, vol. 1, 1–13.

Bjorkman, Bronwyn. 2017. Singular *they* and the syntactic representation of gender in English. *Glossa* 2(1). 80. doi:10.5334/gjgl.374.

Block, Trevor. 2019. *Interspeaker variation in the syntactic processing of referential singular they*: Simon Fraser University MA thesis.

Bradley, Evan. 2020. The influence of linguistic and social attitudes on grammaticality judgments of singular 'they'. *Language Sciences* 78. doi:10.1016/j.langsci.2020.101272.

Collins, Chris & Paul Postal. 2012. *Imposters*. Cambridge, MA: MIT Press.

Conrod, Kirby. 2019. *Pronouns raising and emerging*: University of Washington dissertation.

Eckert, Penelope. 2014. The problem with binaries: Coding for gender and sexuality. *Language and Linguistics Compass* 8(11). 529–535.

Helke, Michael. 1973. On reflexives in English. *Linguistics* 11(106). 5–23. doi:10.1515/ling.1973.11.106.5.

Hernandez, Ellis. 2020. *Pronouns, prescriptivism, and prejudice: Attitudes toward the singular 'they', prescriptive grammar, and nonbinary transgender people*: Purdue University dissertation.

Konnelly, Lex & Elizabeth Cowper. 2020. Gender diversity and morphosyntax: An account of singular *they*. *Glossa: a journal of general linguistics* 5(1). 40. doi:10.5334/gjgl.1000.

Oxford English Dictionary. 2022. Themself, pron. In *OED Online*, `http://www.oed.com/viewdictionaryentry/Entry/334443` (accessed March 15, 2022). Oxford University Press.

Postal, Paul M. 1966. On so-called 'pronouns' in English. In Francis P. Dinneen (ed.), *Report of the 17th annual round table meeting on linguistics and language studies*, 177–206. Georgetown University Press.

Sheydaei, Iman. 2021. Gender identity and nonbinary pronoun use: exploring reference strategies for referents of unknown gender. *Gender and Language* 15(3). 369–393.

Weinreich, Uriel, William Labov & Marvin Herzog. 1968. *Empirical foundations for a theory of language change*. University of Texas Press.

Wiltschko, Martina. 2008. The syntax of non-inflectional plural marking. *Natural Language & Linguistic Theory* 26(3). 639–694.