My pronouns are *they/them:* Talking about pronouns changes how pronouns are understood

Jennifer E. Arnold, Heather C. Mayo, & Lisa Dong

University of North Carolina at Chapel Hill

Address all correspondence to:

Jennifer Arnold

University of North Carolina at Chapel Hill

Dept. of Psychology and Neuroscience, Davie 337B

Chapel Hill, NC 27599-3170

Abstract

The pronoun “they” can be either plural or singular, perhaps referring to an individual who identifies as nonbinary. How do listeners identify whether “they” has a singular or plural sense? We test the role of explicitly discussing pronouns, e.g. “Alex uses they/them pronouns.” In three experiments, participants read short stories like “Alex went running with Liz. They fell down.” Answers to “Who fell down” indicated whether participants interpreted *they* as Alex or Alex-and-Liz. We found more singular responses in discourse contexts that make Alex more available: when Alex was either the only person in the context or mentioned first. Critically, the singular interpretation was stronger when participants heard explicit instructions that Alex uses *they/them* pronouns, even though participants in all conditions had ample opportunity to learn this fact through observation. Results show that the social trend to talk about pronouns has a direct impact on how language is understood.

Keywords: Pronouns, gender identity, nonbinary, discourse

1. Introduction.

In 2019 the American Dialect society voted for *they* as the word of the decade and *(My) Pronouns* as the word of the year. This vote recognizes two related linguistic trends. First, the singular use of *they* is growing in acceptability, in particular as the pronoun of choice for individuals who identify as gender nonbinary. Second, there is a movement to use pronouns to signal gender identity, as in *my pronouns are they/them/theirs.* This means that the speaker wishes for other people to refer to them with the singular pronoun *they*.

For language comprehension, this makes the word *they* ambiguous. Consider a story about Alex (who uses they/them pronouns) and Liz (who uses she/her pronouns): *Alex went running with Liz. They fell down*. Did both fall down, or just Alex? Linguistic ambiguity is not unusual, and the human mind is well adapted to interpreting ambiguous words in context. Yet few ambiguities are accompanied by explicit commentary. Singular *they* offers a unique opportunity to examine how a dispreferred meaning can become more available when the speaker offers explicit guidance, and how this affects an ongoing linguistic change. Our study uses three experiments to test the impact of explicit pronoun-introduction and linguistic context on the interpretation of *they.*

The status of singular they

Contrary to public opinion, the word *they* has had a singular interpretation for centuries (Baron, 2020; McWhorter, 2018). In many English dialects, singular *they* is common for referring to people with unknown gender, such as *someone* or *my future neighbor,* and this usage has occurred in literature as far back as the middle ages (Nunberg, 2016). Singular *they* can be used to maintain anonymity of the referent, as in this excerpt from a newspaper article: "A Wake County resident who was at The Millennium Tour 2020 concert … has tested positive for

coronavirus,…. They were there between 8:20 p.m. and midnight ….." (Aldridge & Jasper, 2020). Camilliere, Izes, Leventhal and Grodner (2019) found that people rated *they* as more natural for referents that were ungendered or ambiguous *(the dentist; Taylor*) vs. gendered (*the actress; Sophia*). People also rated *they* as more natural for referents that were socially distant (*the dentist*) than socially close (*my friend*).

By contrast, a relatively new use of *they* is for reference to known and singular referents (Bjorkman, 2017; Konnelly & Cowper, 2020). Most notably, individuals who identify as nonbinary often prefer to be referred to with *they*, although some prefer neopronouns like *xe* (Bradley, Salkind, Moore, & Teitsort, 2019; Fisher, 2018). The range of singular *they* also appears to be expanding, and can even be used when the referent is known and identifies with a binary gender. For example, the first author observed a Facebook post that said "my spouse… they", even though the spouse does not typically use they/them pronouns.

This expansion of singular *they* is paralleled by the movement to use pronouns as a signal of gender. This raises a question: Does pronoun introduction change the way *they* is understood? Most linguistic meaning is induced from the context, and not discussed explicitly. Indeed, the singular meaning of *they* can be obvious from the context. What role does explicit introduction play, over and above observational evidence?

Research question: what drives the interpretation of 'they'?

Even though singular *they* is common, the plural use of *they* is more common. Thus, singular and plural interpretations likely compete. Current psycholinguistic accounts suggest that comprehenders weigh alternate referents for a pronoun, probabilistically activating likely referents. Activation considers both the pronoun features (e.g., gender and number) and the

discourse context, which makes some referents more available than others (e.g., Arnold et al., 2000; Kehler & Rohde, 2013; MacDonald & MacWhinney, 1990; McDonald & MacWhinney, 1995; Garnham, Oakhill, & Cruttenden, 1992; Sanford & Garrod, 1989).

Two discourse properties that affect pronoun comprehension are the number of competing referents and order of mention (e.g., Arnold et al., 2000; Gernsbacher & Hargreaves, 1988; Jarvikivi et al., 2005). If only one person is mentioned, that person is the likely pronoun referent, e.g. Liz in *Liz went to the store. She bought milk*. If two people are mentioned, the first one (the grammatical subject) is often assumed to be the referent, e.g. Liz in *Liz went to the store with Ana. She bought milk*. Our experiments test how *they* is understood in different discourse contexts within stories about three fictional characters: Alex (*they*), Liz (*she*) and Will (*he).*

In English, people understand pronouns by rapidly applying gender features on the pronouns, e.g. female for *she* and male for *he* (e.g., Arnold, Brown-Schmidt, Eisenband, & Trueswell, 2000; Boland, Acker, & Wagner, 1998; Ehrlich, 1980; MacDonald & MacWhinney, 1990; Osterhout & Mobley, 1995). People also pay attention to gender – or the lack of it -- when interpreting *they*. For an ungendered referent (e.g., *anyone; a runner*), singular *they* is understood as fast as *he* and *she* (Foertch & Gernsbacher, 1997). Comprehension of singular *they/them* is also faster if the referent does not have an assumed gender (*the cyclist)* than if it does (*the mechanic;* Doherty & Conklin, 2017). Similarly, in event-related potential studies, singular *they* or *themselves* can elicit a P600, which is associated with syntactic anomalies (Leventhal et al., 2020; Prasad & Morris, 2020). For example, Leventhal et al. (2020) found a P600 for *Lillian…. they* compared with *Lillian and Paul….they*. These studies suggest that singular *they* is understood easily when the referent is nonspecific or gender neutral, but not when a male or female gender is assumed.

On the other hand, *they* is easier to understand with a plural than a singular interpretation (Sanford & Filik, 2007), suggesting that the plural use is primary. In particular, nonbinary *they* is relatively new, and facility with this form depends on personal experience. Konnelly & Cowper (2020) suggest that singular *they* is a change in progress, is accepted by some social groups more than others (e.g., Camilliere et al., 2019).

We aim to better understand this change by measuring variation in the assignment of a singular interpretation to *they*. We use three experiments to capture metalinguistic judgments about the referent of *they* in stories about Alex, who uses *they/them* pronouns. We examine pronoun comprehension in a natural context, where stories refer to known characters. To this end, all stories are about the same three fictional characters, introduced with pictures at the start of the task. The pictures help participants to develop a mental representation of the character. We manipulate 1) whether Alex's pronouns are introduced explicitly or not, and 2) whether the discourse context supports Alex as the referent, manipulating either number of people (Exp. 1) or whether Alex is mentioned first (Exps. 2/3). Note that here we use the terms explicit and implicit to denote properties of the stimuli; it is possible (and indeed likely) that participants were explicitly aware of Alex's gender in both conditions. A second question is whether additional practice with singular *they* can change interpretation patterns within a single experimental session, which we manipulate in experiments 2 and 3. The current paper focuses on the manipulated variables, but analyses of how individual differences pattern with responses are reported in a companion paper (Arnold, Mayo, & Dong 2020).

2. Methods for all three experiments

2.1.Participants

Native English-speaking participants were recruited through Amazon Mechanical Turk and compensated $1.00 (Exp. 1) or $1.25 (Exps. 2 and 3). The HIT was only available in the United States, Australia, and the United Kingdom. There were 63 participants in Exp. 1; 79 in Exp. 2 (plus an additional 15 that were replaced because of a technical problem with the surveys) and 60 in Exp. 3; our primary analysis included 52, 54, and 44 in each experiment. See below for the criteria for exclusion from analysis.

2.2. Methods

Participants completed a Qualtrics questionnaire with three parts: 1) Demographic questions, 2) Instructions, 3) Main task. Materials and other supporting information can be found at https://arnoldlab.web.unc.edu/publications/supporting-materials/arnold-mayo-dong-2020/.

2.2.1 Demographic Questions. The first two questions asked about participants' age, and English proficiency. Participants were not allowed to continue if they were under age 18, reported non-native fluency, or learned English past age 6. Questions about gender, sex, age, race, ethnicity, and education were used to compare samples (see supporting materials). Two questions asked about participants' familiarity with individuals who identify as nonbinary: a) How many individuals do you know who identify as neither male nor female? (None; 1; 2-5; 6 or more); and b) How familiar are you with people who do not identify as either male or female? (0-10 sliding scale).

For Experiment 2 only, participants also completed the Author Recognition Task, given previous findings that print exposure modulates individual biases to assign pronouns to the previous subject (Arnold et al., 2019). We used a version of this task that was designed by Peter Gordon's lab (p.c.), and was developed from previous versions by Acheson et al. (2008) and Stanovich & West (1989). Participants saw a list of 62 true author names intermixed with 64

false names. They were asked to check the box next to all names they recognized, and not to guess. Because we found no effect of print exposure on *they* interpretation, we did not include this measure in Experiment 3.

2.2.2. Instructions.

Prior to the primary task, participants were introduced to three characters: Alex (they/them), Liz (she/her), and Will (he/him), along with pictures of each. Participants were assigned to either the Explicit or Implicit condition. In the Explicit Lists, the pronouns of each character were given, and participants were tested to ensure that they remembered the names and pronouns of each character. In the Implicit Lists, participants just learned and practiced the names. Each name and pronoun were tested in two questions; if participants responded incorrectly they were asked to correct their answer. Participants were told in the consent form "The purpose of this study is to understand how healthy adults understand language in simple stories." Pronouns were not mentioned at all in the instructions except during introduction of the characters in the Explicit lists.

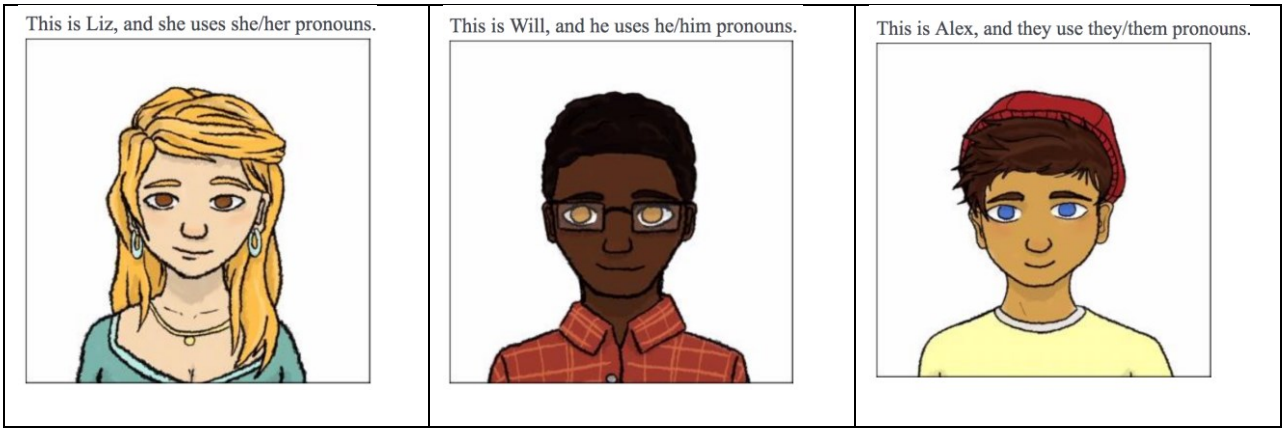

Figure 1. Introductions of the characters for the Explicit lists. Each picture and introductory sentence appeared alone on a separate screen. The Implicit lists were identical except they did not mention pronouns.

2.2.3. Main Task Materials and Design. The primary task was to read two-sentence stories and answer two questions (see Figure 2 for example). For critical stimuli, the question of interest was designed to reveal whether participants assigned a singular or plural interpretation to the pronoun. For example, following "Liz walked home. She forgot to bring the keys", the critical question asked "Who forgot the keys?" (see Table 1). Participants answered by clicking on either a picture of Liz alone, or a picture of Liz with Alex. During the practice questions, they were instructed to pick the picture that signaled a singular or plural actor: "The next example illustrates the use of pictures as answers. One picture shows Will and Alex together. This means they both did the action together. The other picture shows just Will. This picture means that just Will did the action alone. Pick the best answer". Each question also had a content question, such as "What did Liz do?" (walked home / went to work). These questions were easy to answer if participants were reading the stories, and contributed to our inclusion criteria (see below).

The experiment began with two practice stories, followed by training stories (12 in exps. 1 and 2; 16 in exp. 3), and then 8 critical stimuli intermixed with 15 fillers. The training, critical, and filler stories were not labeled for the participants; instead all stories appeared in a single sequence, one at a time on the screen. All stories were followed by two questions, one "who" question that was answered by selecting either a singular or plural option, and one content question. For three fillers the second sentence did not mention a person, so the "who" question was replaced by a content question.

The training stories served to illustrate the pronouns used by each character, so even in the implicit condition participants had the opportunity to learn them. In experiments 1 and 2, there were four stories about each of the three characters. There was only one character in each story, providing strong evidence that the pronoun should refer to that character (see Table 1).

The appropriate pronouns were consistently used for each character (Liz – she; Alex – they; Will – he).

Experiment 3 tested whether the number of training stories about Alex would improve performance on the critical items, so it included an additional 4 training stories about Alex (total 8 Alex training stories). The first four training stories were identical to those used in experiments 1 and 2, the second four followed a similar format.

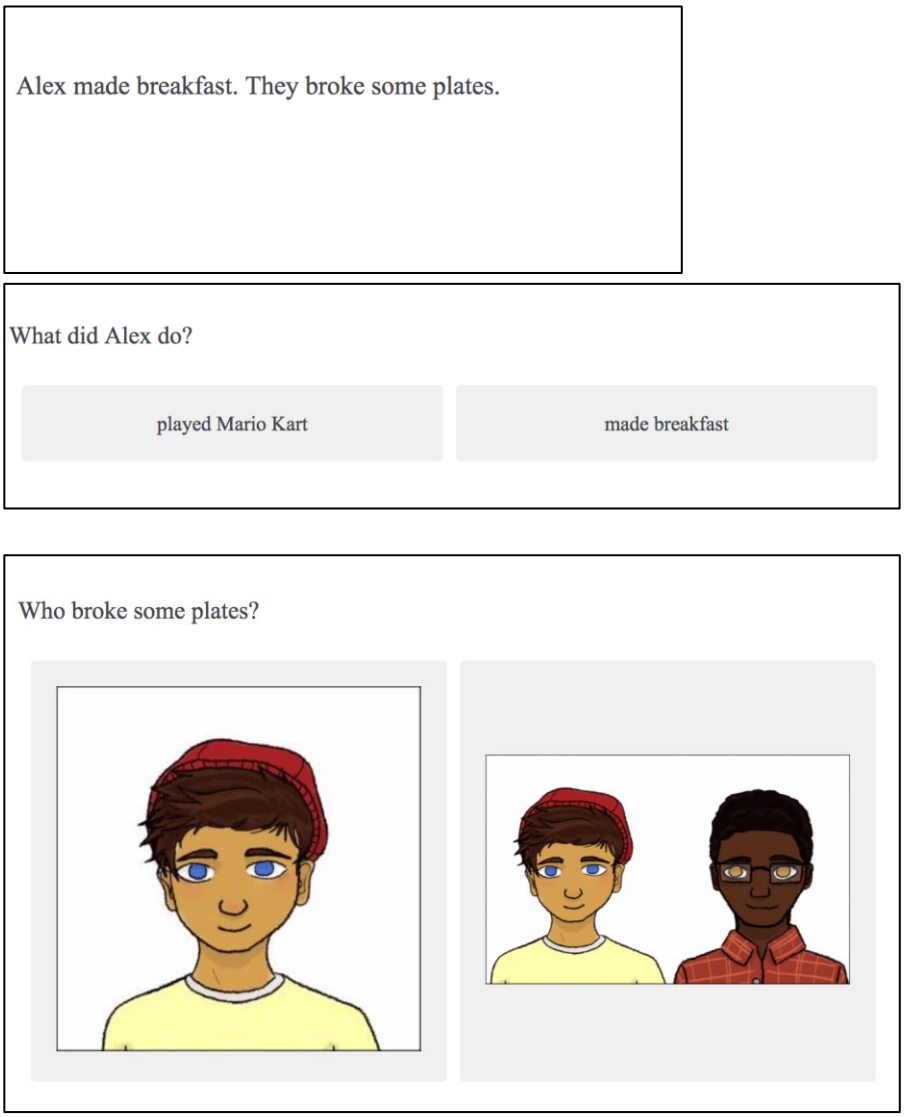

Figure 2. Example critical test story and questions.

The 8 critical test stimuli used stories about Alex (see Figure 2). These stories appeared in one of two critical discourse conditions. In Experiment 1, stories were either about just Alex (1-character condition), or Alex and another character (2-character condition). The two-person stories always introduced Alex in subject position and the other person as an adjunct, in the form "Alex [did something] with Liz/Will." The one-person stories were identical except they did not include the with-phrase that mentioned the second character. In Experiments 2 and 3, all critical stories had two characters, manipulating whether Alex appeared first or second. There were two lists for each experiment, where List 1 and List 2 each had four items in each condition, counterbalanced across lists. Each of the two lists also appeared with both Explicit and Implicit instructions, resulting in a total of four lists (List 1-explicit; List 1-implicit; List 2-explicit, List 2-implicit).

Critical items were intermixed with 15 filler stories, with 3 each in 5 structures: 1) 1 character / no pronoun (e.g., *Liz took a vacation. Sweden was lovely*); 2) 1 character unambiguous pronoun (e.g., *Will cleaned his apartment. He got his clothes dirty*.), 3) Unambiguous pronoun refers to subject (e.g., *Liz danced with Will. She tripped many times.*); 4) Unambiguous pronoun refers to nonsubject (e.g., *Will went shopping with Liz. She bought a suit.);* and 5) Plural they sentences (e.g. *Will ran a marathon with Liz. They took more than five hours to finish.*) The plural-they fillers were all about Will and Liz, so the most plausible interpretation of *they* was as a plural pronoun. These items served to remind participants that *they* also has a plural interpretation. The fillers were pseudorandomly interspersed with the critical items such that there were no clusters of fillers of a particular type and each type of filler was evenly distributed across the experiment.

Table 1. Example Stimuli and Critical Questions

Story type	Example Story	Critical Question	Content Question
Training: Liz	Liz walked home. She forgot to bring her keys.	Who forgot the keys?	What did Liz do? • Walked home • Went to work
Training: Alex	Alex went to the store. They bought some milk.	Who bought milk?	What did Alex do? • Went to the store • Walked home
Training: Will	Will needed to be at work by 9. He ran to the bus stop.	Who ran to the bus stop?	Where did Will need to be? • At work • At home
Critical: One-character	Alex went running. They fell down.	Who fell down?	What did Alex do? • Went running • Had coffee at Starbucks.
Critical: Two-character, Alex first	Alex went running with Liz. They fell down.	Who fell down?	What did Alex and Liz do? • Went running • Had coffee at Starbucks.
Critical: Two-	Liz went running with Alex. They fell down.		

character, Alex second			
Practice	Liz and Alex hiked to the top of the hill for a picnic. They took too long and the ice cream melted.	Who took too long?	What did Liz and Alex do? • Jogged to the park • Hiked up the hill for a picnic

2.2.4. Procedure. Participants accepted the Hit in Amazon Mechanical Turk and began the survey. After providing consent and answering the demographic questions, they read instructions about the survey and answered practice questions. Incorrect answers to the practice questions were corrected and participants were asked to re-read the story and answer again. They were then reminded that too many incorrect answers during the survey would result in the termination of the study and they wouldn't be paid. Then they began the main task.

2.2.5. Inclusion criteria. Participants were automatically excluded from the survey if they made 12 mistakes on the 43 test questions (26%) for experiments 1 and 2, or 15 out of 47 (32%) for experiment 3. These test questions excluded the critical "Who" questions about Alex. We also excluded the last 9 items from our calculation to prevent participants from being kicked out when they were almost done with the study.

Our key question was whether explicitly introducing one's pronouns has an impact on how people interpret singular they. Yet this question would be meaningless if it our comparison group were people who didn't even realize that Alex uses they/them pronouns. Thus, the

strongest test of our hypothesis is to compare the explicit lists with people from the implicit lists who responded successfully to the training questions, so we know that the subjects in both groups were fully aware of Alex's use of they/them pronouns and willing to accept a singular interpretation of they. Participants may have resisted the singular interpretation for multiple reasons: perhaps they speak a dialect of English that does not permit singular they; perhaps they are ideologically opposed to this usage; or perhaps they were inattentive to the story context. Whatever the reason, we used the training set to divide participants into two groups. We labeled "resisters" as participants who made at least one "plural" response. The training stories mentioned only one person, e.g. "Alex saw a worm on the sidewalk. They tried to pick it up," so the discourse context strongly supported a singular interpretation. Some individuals consistently responded "plural" to these items, demonstrating a strong aversion to the singular interpretation. Other participants provided a mix of singular and plural responses. Regardless of the motivation, we limited the "nonresister" group to those who responded singular on 100% of the training stories, demonstrating successful recognition that Alex uses they/them pronouns.

Following other studies in our lab (e.g. Williams & Arnold, 2019), we aimed for an initial sample size of about 30 participants per between-subjects condition, prior to resister exclusions. We excluded data from analysis for participants who reported a language disorder, or for those with an overall accuracy lower than 75% on the full set of filler questions, which included the content questions for all items, plus all questions for fillers, including the last 9 that were excluded from the automatic Qualtrics exclusion criteria. For experiment 2, we also followed our lab's standard policy of excluding participants if more than 33% of their responses on the Author Recognition Task were foils, because it signaled that they did not follow the directions to not guess.

Table 2. Number of participants excluded in each study

Experiment	Total run	Excluded	Usable	Resisters	# in main analysis
1	63	1 (language disorder)	62	2 in explicit 8 in implicit	52
2	79	2 (language disorder) 12 (guessing on ART) 2 (low filler accuracy)	63	2 in explicit 7 in implicit	54
3	61	2 (language disorder)	59	5 in explicit 10 in implicit	44

2.3. Analysis.

Our dependent measure in all experiments was the interpretation of the 8 critical test questions for the stories about Alex, where participants chose either a singular or plural interpretation. Our key predictors were our two manipulated variables: 1) Did participants receive explicit introduction of pronouns or not? and 2) Does the discourse context promote singular interpretation or not? The discourse context variable reflected the singular vs. 2-person context in experiment 1, and the Alex-first vs. Alex-second context in experiments 2 and 3.

Given our binary dependent measure, we analyzed responses in a mixed effects logistic regression, using SAS proc glimmix, with a binary distribution and a logit link. Models included maximal random effects (Barr, Levy, Scheepers, & Tily, 2013), including random intercepts for

participants and items, and random slopes as appropriate. All predictors were grand-mean centered. Raw data are available here: https://osf.io/fr9az/?view_only.

2.4. Predictions

Our first question is whether explicit introduction of pronouns facilitates the singular interpretation of "they". We predict that it does. Our second question is whether the discourse context changes the availability of the singular interpretation. Following other work on pronoun comprehension, we expect that the singular interpretation should be stronger when there is only one character in the context, and when the nonbinary character was mentioned first in the story. Our final question was whether an increased number of training stories in Exp. 3 would result in a higher rate of singular interpretations for the critical stories, compared to Exp. 2.

3. Results

3.1. Training Items

Table 3. Rate of selecting the singular interpretation for the training Alex sentences.

		Exp. 1	**Exp. 2**	**Exp. 3**	TOTAL
Explicit	100% (Nonresister)	28	30	26	84
	75% singular	0	1	3	4
	25% singular	1	0	1	2
	0% singular	1	1	1	3
Implicit	100% (Nonresister)	24	24	18	66
	75% singular	3	2	2	7
	50% singular	2	1	0	3
	25% singular	1	1	3	5
	0% singular	2	3	4	9

Table 3 shows training performance for the Alex stories in all three experiments. By comparison, in the training stories about Liz and Will, interpretation of he/she pronouns was

singular almost all the time (100% in Exp. 1, 99.8% in Exp. 2, and 99.6% in Exp. 3). Across all three experiments, the total number of nonresisters was 84 for Explicit lists, and 66 for Implicit lists, while the total number of resisters was 9 for Explicit lists and 24 for Implicit lists. The chi-square statistic for this 2x2 comparision is 8.9 ($p = 0.003$), suggesting that explicit introduction of pronouns leads to higher singular interpretation in the training set.

3.2. Critical item results

Given that participants had already answered "singular" on all training questions, we asked how often they had a singular interpretation of the critical question about Alex. In all cases the question asked "Who did [the action]?", where the two possible answers were either Alex, or Alex together with someone else. Results (Table 4 and Figure 3) showed that for all the two-person contexts, singular interpretations were higher following explicit than following implicit pronoun introduction, with the difference between means averaging 25% for all the Alex first and Alex second conditions (range: difference of 15% - 31%). In the one-person context, people responded singular almost all the time.

Table 4. Average selection of the singular interpretation in the critical stories about Alex. Standard error is reported in parentheses.

		Exp. 1	**Exp. 2**	**Exp. 3**
Explicit	1-person context	99% (1%)		
	2-person/Alex first	50% (1%)	37% (7%)	40% (9%)
	2-person/Alex second		30% (7%)	31% (8%)
Implicit	1-person context	98% (8%)		
	2-person/Alex first	19% (6%)	22% (7%)	11% (4%)
	2-person/Alex second		10% (5%)	1% (1%)

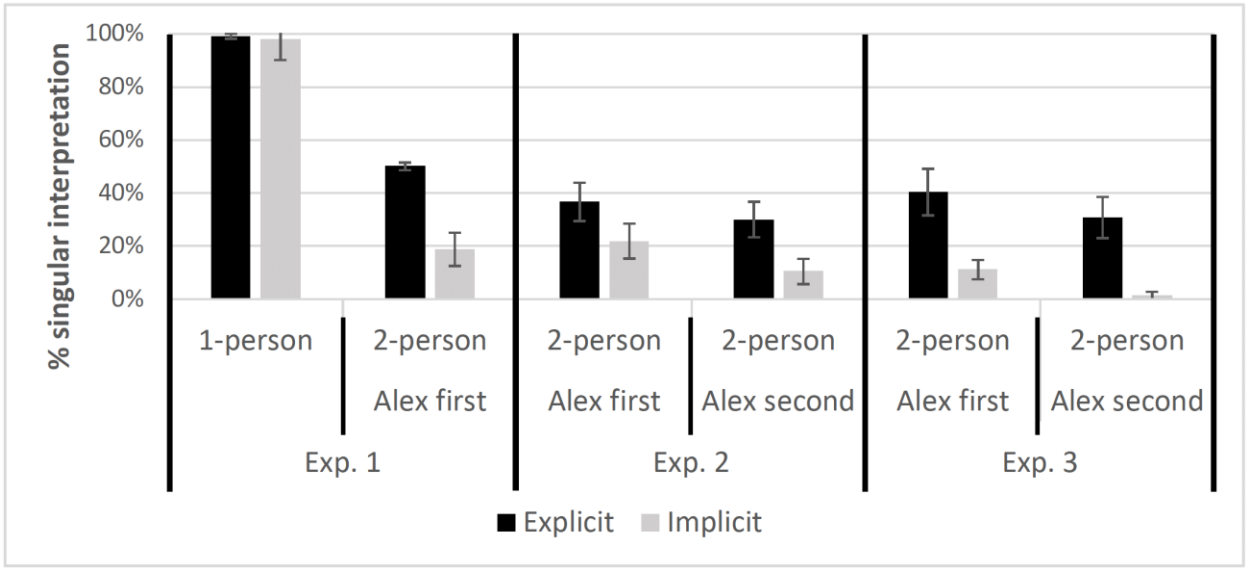

Figure 3. Average percentage of singular interpretations for nonbinary "they" as a function of the pronoun-introduction condition. (explicit vs. implicit) and discourse context (1-person; 2-person Alex first, 2-personAlex second).

We also observed a strong effect of discourse context. In Experiment 1, the singular interpretation was adopted almost all the time when Alex was the only participant in the discourse, but less than half the time when there were two people in the context. In Experiment 1, Alex was always mentioned first in the two-person context. Experiments 2 and 3 further explored the 2-person context by manipulating whether Alex was mentioned first or second. The singular interpretation was more frequent for the Alex-first contexts than the Alex-second contexts in both experiments, with a difference of 9% or 10% between conditions.

We tested the reliability of these patterns with a mixed-effects logistic regression, using SAS proc glimmix. Inferential statistics are presented in Table 5. In Experiment 1, the effect of discourse context was significant ($p < .001$), although the effect of Explicit vs. Implicit pronoun introduction did not quite reach significance ($p = .091$). This experiment also revealed that all variation occurred in the 2-person context, while singular interpretations were at ceiling in the 1-

person context. In Experiments 2 and 3, we observed significant effects of both discourse context and explicitness. There were no interactions.

Table 5. Inferential statistics

	Estimate (SE)	95% CI	t	p
Experiment 1				
Explicit vs. Implicit introduction of pronouns	1.34 (0.78)	-0.22 - 2.91	1.71	0.091
Discourse context (1 vs. 2 person)	5.47 (0.72)	4.03 - 6.90	7.61	<.0001
Explicitness x Discourse context	-0.91 (1.46)	-3.82 - 1.99	-0.63	0.533
Experiment 2				
Explicit vs. Implicit introduction of pronouns	1.41 (0.66)	0.09 - 2.73	2.15	0.037
Discourse context (Alex first vs. Alex second)	0.85 (0.33)	0.08 - 1.62	2.58	0.034
Explicitness x Discourse context	-0.79 (0.63)	-2.04 - 0.45	-1.25	0.211
Experiment 3				
Explicit vs. Implicit introduction of pronouns	3.12 (1.08)	1.00 - 5.25	2.9	0.004
Discourse context (Alex first vs. Alex second)	1.45 (0.63)	0.16 - 2.74	2.3	0.029
Explicitness x Discourse context	-1.41 (1.15)	-3.67 - 0.85	-1.23	0.22

Our second question was whether singular interpretation is higher when people have had more practice. In experiment 3, we provided 8 training stories about Alex, instead of the 4 training stories used in experiments 1 and 2. However, experiment 3 had no more singular responses than experiment 2, and in fact the average rate of singular responses was numerically lower for experiment 3 (21%) than experiment 2 (25%). When we entered both experiments into a single analysis, we again observed main effects of the discourse context and explicitness, but no effect of experiment 2 vs. 3, nor any interactions (see supporting materials for statistical detail).

We also asked whether individuals vary in how often they adopt the singular interpretation, based on 1) how many nonbinary individuals they know, and 2) how familiar they are with nonbinary individuals. However, there were no effects of question 1 (how many nonbinary people they know). While greater familiarity predicted more singular responses for Experiment 1, this effect was not replicated in the other experiments. For experiment 2, we also tested individual differences in print exposure, using the author recognition task, and found no effect. For a full report of the individual difference analyses, see Arnold, Mayo, & Dong (2020).

General Discussion

Our critical finding is that explicitly introducing Alex's pronouns promotes the singular interpretation of *they*. This finding emerged in our main analysis, which was restricted to the "nonresister" participants who clearly understood that Alex uses they/them pronouns. While it is not surprising that people can learn from our training examples that Alex uses they/them pronouns, it is notable that the explicit introduction of pronouns had a significant effect over and above this exposure effect. We also saw an effect of explicit introduction on training performance: we had more "resisters" in the implicit than explicit conditions. Unsurprisingly, discourse context mattered too: singular interpretation was more frequent when Alex was the only character or the first character in the context. However, reading an additional four training stimuli in Experiment 3 had no effect.

Our experimental task mimics the newly-popular social convention of stating one's pronouns explicitly, and in doing so demonstrates that this practice has a direct effect on how pronouns are understood. To understand why this practice matters, we consider the cognitive

processes known to be involved in the representation and use of gender during pronoun interpretation.

Gender critically has representations at multiple levels, both conceptual and linguistic. Pronouns are words that are marked linguistically for gender, including feminine (*she*), masculine (*he*), and ungendered (*it, they*). They are also marked for number, where *she/he/it* are singular, but *they* can be used with either singular or plural meanings. These features help determine how pronouns map onto referents in the world, which can be categorized in terms of gender at a conceptual level. While some languages have broader systems of lexical gender (e.g, Spanish or German), here we focus on the use of pronouns in English, where gender is primarily used to match the conceptual gender of animate referents.

We couch our interpretation within Ackerman (2019)'s framework, which proposes three levels of representation of gender (see Figure 4). Linguistic gender features are represented on the Feature tier. Conceptual (nonlinguistic) gender falls on two different tiers. The Exemplar tier is a multidimensional representation of features of gender expression, based on culturally-relevant gender features (e.g., body shape, clothing, etc.). These exemplars form the input to the gender categories on the Category tier. For most people, features are clumped into bimodal representations of male and female, while the space outside these categories accounts for cases of unknown gender, ambiguous gender, or nonbinary gender.

Research suggests that *they* is perceived as ungendered (Camilliere et al., 2019; Foertch & Gernsbacher, 2007; Prasad & Morris 2020; Leventhal et al., 2020). Konnelly & Cowper (2020) suggest that the most innovative users of singular *they* find it acceptable even with a gendered referent (e.g., *mother),* but we have observed that these cases tend to occur when the identity of the referent is backgrounded or suppressed. This suggests that for many language

users, *they* is only used when the conceptual representation of gender is either absent or very weak. Thus, on the Akerman model, the Feature tier represents *they* without masculine or feminine features, and at the Category tier it picks out referents that fall between the feminine and masculine categories.

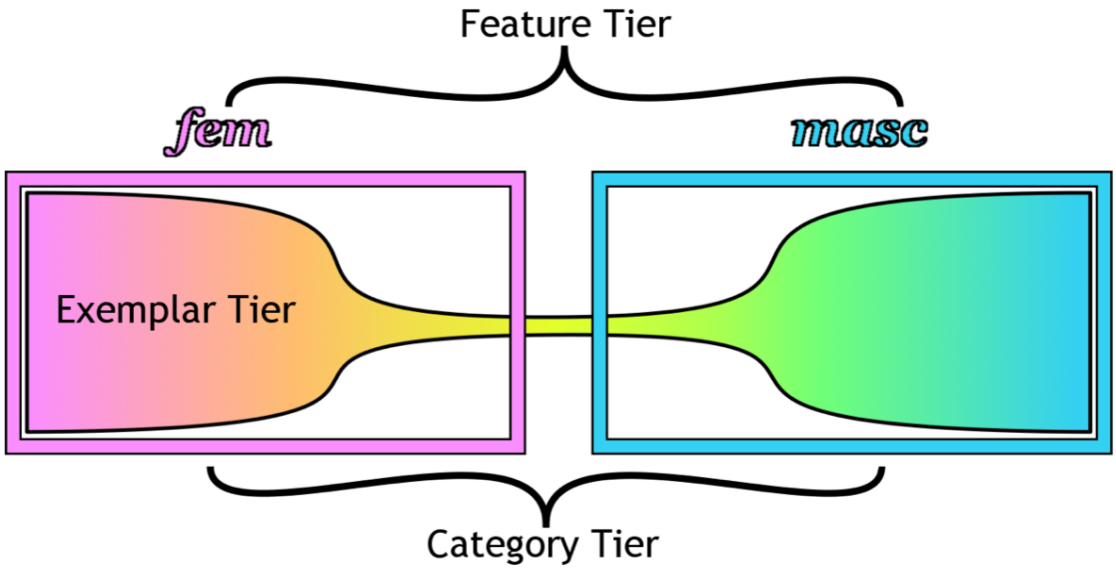

Figure 4. Ackerman (2019)'s schema of three gender tiers: The Exemplar Tier, the Category Tier, and the Feature tier. (copied from Ackerman, 2019, p. 22).

Under the above model, the comprehension of *they* requires the identification of a specific referent. The set of available referents is constrained by the gender and number features of the pronoun, and could include singular/ungendered or plural referents. This presents a classic case of lexical ambiguity. We theorize that encountering an instance of *they* leads to a frequency-weighted competition between singular and plural meanings, similar to other words with multiple senses or meanings (e.g., Duffy, Morris, & Rayner, 1988). It simultaneously leads to activation of potential referents, weighted by the contextual support for each. In our example

Alex went to the store. They…., the plural meaning is relatively more activated due to frequency, but Alex is the most available discourse referent, supporting the singular interpretation.

On this model, there are three possible mechanisms by which pronoun-introduction affects interpretation. First, it may affect linguistic representations: hearing "Alex uses they/them pronouns" might remind the listener of the singular use of *they* and thereby strengthen the singular lexical representation, increasing its level of baseline activation. Second, it may affect conceptual representations by directing the participant's attention to the fact that some people identify outside the gender binary, effectively broadening the ungendered gap in the category tier. Third, it may situate Alex clearly on the exemplar tier. Either of the second two possibilities would strengthen the representation of Alex outside of binary gender categories, supporting the match with *they*. A question for future work is whether introducing Alex's pronouns facilitates the interpretation of *they* for a new nonbinary individual; if so it would support a more general shift in the use of *they* (the first two possibilities here).

Explicit instructions about pronouns may also modulate the semantic features associated with *they.* Usage statistics would suggest that singular *they* bears a socially-distant semantic feature (Camilliere et al., 2019), while the nonbinary use requires a socially-close interpretation. These variants also may compete during interpretation. Explicit pronoun instructions establish expectations that *they* can refer to a specific individual, and not just when the speaker wishes to de-emphasize the identity and gender of the referent.

Importantly, our experiment shows that when people read that Alex uses they/them pronouns, it does more than simply help them categorize Alex as nonbinary. We limited our analysis to people who responded perfectly to the Alex training questions, which means that all participants were equally aware of Alex's nonbinary status. Yet even within this sample, people

who had received the explicit instructions were more likely to assign a singular interpretation. This showed up particularly in the 2-person context, where there was a plausible competing plural interpretation.

Our experiment highlights variation as a function of discourse features, but singular *they* is also more common for some groups than others (Camilliere et al., 2019; Konnelly & Cowper, 2020; Nichols, Chun, & Almor, 2019). Individuals likely vary in both the baseline weight of singular *they*, and their experience of gendered exemplars and the corresponding conceptual categories relating to gender. Individuals may also differ in the degree to which singular *they* is semantically marked for social distance. In principle, individual facility with singular *they* should change as a function of repeated exposure to discourse contexts where the singular interpretation is supported, as in our explicit condition and our supporting discourse contexts. However, we did not observe any difference between experiments 2 and 3, so merely adding 4 training sentences is not enough. We also found little effect of our measures of familiarity with nonbinary individuals (see Arnold, Mayo, & Dong, 2020), but this is likely due to limitations in our coarse-grained questions, and lack of variability in our sample. At this point in time, nonbinary singular *they* is relatively new to many people, so the biggest influence on interpretation is a direct discussion of its meaning.

In sum, singular *they* has been around for centuries, but its frequency and range of uses is currently expanding. This expansion parallels a movement to adopt a more fluid concept of gender, which has entered public awareness through the trend of using pronouns as a symbol of gender identity. It is likely that these changes are influenced by multiple forces, one of which is talking about pronouns. In short, people who hear that Alex uses *they*/*them* pronouns are more likely to assume that *they* refers to Alex.

Acknowledgements

We are grateful to Darith Klibanow for drawing the pictures of the characters in our survey. This project was partially funded by a Stephenson-Lindquist grant from the Dept. of Psychology at UNC Chapel Hill, and partially funded by NSF grants 1651000 and 1917840 to J. Arnold.

Open Practices Statement

None of the experiments reported in this article were formally preregistered. The data are available on a permanent third-party archive, https://osf.io/fr9az/?view_only.

References

Acheson, D. J., Wells, J. B., & MacDonald, M. C. (2008). New and updated tests of print exposure and reading abilities in college students. Behavior Research Methods, 40, 278–289. http://dx.doi.org/10.3758/brm.40.1.278.

Ackerman, L. (2019). Syntactic and cognitive issues in investigating gendered coreference. Glossa, 4, 117. doi:10.5334/gjgl.721/

Aldridge, B., & Jasper, S. (2020). Coronavirus live updates: Here's what to know in North Carolina on March 19. The News and Observer; March 19, 2020. https://www.newsobserver.com/news/coronavirus/article241320026.html#storylink=cpy

Arnold, J. E., Eisenband, J. G., Brown-Schmidt, S., & Trueswell, J. C. (2000). The rapid use of gender information: evidence of the time course of pronoun resolution from eye tracking. Cognition, 76, B13-B26.

Arnold, J. E., Mayo, H., & Dong, L. (2020). Individual differences (or the lack of them) in comprehension of singular *they.* UNC Language Processing Lab, Technical Report #3.

Baron, D. (2020). *What's your pronoun? Beyond He and She*. New York: Liveright publishing.

Barr, D. J., Levy, R., Scheepers, C. & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Languag*e, *68*, 255- 278. doi:10.1016/j.jml/2012.11.001

Bjorkman, B. (2017). Singular they and the syntactic representation of gender in English. G*lossa: a Journal of General Linguistics, 2*, 80, 1–13.

Boland, J. E., Acker, M. T., & Wagner, L. (1998). The use of gender features in the resolution of pronominal anaphora (Cognitive Science Technical Reports No. #17). The Ohio State University Center for Cognitive Science.

Bradley, E. D., Salkind, J., Moore, A., & Teitsort, S. (2019). Singular 'they' and novel pronouns: gender-neutral, nonbinary, or both? *Proc Ling Soc Amer* 4. 36:1-7.https://doi.org/10.3765/plsa.v4i1.4542

Camilliere, Izes, Leventhal & Grodner (2019) Pragmatic and Grammatical Factors that License Singular they. Paper presented at Xprag, November 2019, Berlin.

Doherty, A. & Conklin, K. (2017). How gender-expectancy affects the processing of "them." The Quarterly Journal of Experimental Psychology, 70(4), 718-735.

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language 27*, 429-446.

Ehrlich, K. (1980). Comprehension of pronouns. *Journal of Experimental Psychology*, *32*, 247-255. doi:10.1080/14640748008401161

Fisher, J. R., (2018). On Neopronouns. Retreived on June 24, 2020 from https://medium.com/@JessieRaeFisher/on-neopronouns-eec46705c5dc.

Foertsch, J. & Gernsbacher, M. (1997). In search of gender neutrality: Is singular they a cognitively efficient substitute for generic He? Psychol Sci., 8(2), 106-111. Fowler, HR., Aaron, JE. The Little, Brown handbook. 3rd ed.. Scott, Foresman; Glenview, IL: 1983.

Garnham, A., Oakhill, J., & Cruttenden, H. (1992). The role of implicit causality and gender cue in the interpretation of pronouns. Language and Cognitive Processes*, 73*, 231-255.

Gernsbacher, M. A. & Hargreaves, D. J. (1988). Accessing sentence participants: The advantage of first mention. Journal of Memory and Language, 27, 699-717.

Järvikivi, J., van Gompel, R. P. G., Hyönä, J., & Bertram, R. (2005). Ambiguous pronoun resolution: contrasting the first-mention and subject-preference accounts. *Psychological Science*, *16*, 260-264.

Kehler, A., & Rohde, H. (2013). A probabilistic reconciliation of coherence-driven and centering-driven theories of pronoun interpretation. *Theoretical Linguistics, 39*, 1-37. doi: 10.1515/tl-2013-0001.

Konnelly, L., & Cowper, E. (2020). Gender diversity and morphosyntax: An account of singular *they*. *Glossa: A Journal of General Linguistics*, *5*(1), 40. DOI: http://doi.org/10.5334/gjgl.1000

Leventhal, O., Camilliere, S., Chen, Peiyao, & Grodner, D. (2020). Using ERPs to investigate the processing of singular *they*. Poster presented at the CUNY conference on human sentence processing, Amherst, MA (virtual).

MacDonald, M. C., & MacWhinney, B. (1990). Measuring inhibition and facilitation from pronouns. *Journal of Memory & Language*, 29, 479-492.

McDonald, J. L., & MacWhinney, B. (1995). The time course of anaphor resolution: Effects of implicit verb causality and gender. *Journal of Memory and Language, 34*(4), 543-566.

McWhorter, J. (2018). Call Them What They Wants. The Atlantic, Sept. 4, 2018. https://www.theatlantic.com/ideas/archive/2018/09/the-new-they/568993/

Nichols, A., Chun, E. & Almor, A. (2019). Non-binary gendered reference in LGBTQ+ English: Implications of singular 'they' for sentence processing. Poster presented at the CUNY conference on human sentence processing, Boulder, Co.

Nunberg, G. (2016). Everyone Uses Singular 'They,' Whether They Realize It Or Not. Heard on Fresh Air, NPR, January 13, 2016. Retrieved on June 24, 2020 from https://www.npr.org/2016/01/13/462906419/everyone-uses-singular-they-whether-they-realize-it-or-not.

Osterhout, L. & Mobley, L. (1995). Event-Related potentials elicited by failure to agree. Journal of Memory and Language, 34, 739-773.

Prasad, G., & Morris, J. (2020, January 24). The P600 for singular "they": How the brain reacts when John decides to treat themselves to sushi. https://doi.org/10.31234/osf.io/hwzke

Sanford, A. J., & Garrod, S. C. (1989). What, when and how?: Questions of immediacy in anaphoric reference resolution. Language and Cognitive Processes, 4, 235-262.

Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. Reading Research Quarterly, 24, 402–433.

Williams, E. & Arnold, J. E. (2019). Priming discourse structure guides pronoun comprehension. Poster, CUNY conference on human sentence processing, University of Colorado (March 2019).