Yining Ye and Jennifer E. Arnold

Learning the statistics of pronoun reference: by word or by category?

Department of Psychology and Neuroscience

The University of North Carolina Chapel Hill

Author note:

We have no known conflict of interest to disclose.
Correspondence concerning this article should be sent to Yining Ye.
Email: yiningy2@email.unc.edu

This research was supported by NSF grant 1917840 to J. Arnold. We are grateful to Emma Caulfield for her work on Experiment 1, and to Eri Kakoki for work on Experiments 2 and 3.

Data availability

Data and materials for all experiments are available at [https://osf.io/racdg/] and [https://arnoldlab.web.unc.edu/publications/supporting-materials/ye-and-arnold23/].

ABSTRACT

Evidence shows that readers tend to follow recently-encountered patterns for interpreting ambiguous pronouns. If recent exposure includes frequent pronouns with prepositional object antecedents (e.g., “Matt went to the library with Ana. She took out a book”), people adapt and are more likely to assign ambiguous pronouns to prepositional antecedents than if they were exposed to pronouns with subject antecedents (Johnson & Arnold, 2022). However, it is unclear how people categorize different referential structures and how broadly they make generalizations. Johnson and Arnold (2022) found that people can learn a referential relationship specific to third-person pronouns and an antecedent’s syntactic or thematic properties. The current study uses this paradigm to test how broadly people generalize categories of different types of pronouns and antecedents. Do people adapt to the behavior of categorizing “he” and “she” as individual words, or as a general class? (Experiment 1). Do people learn about likely antecedents for pronouns separately for different verb constructions (transfer vs. joint-action) and thematic roles, or broadly by grammatical function? (Experiments 2 & 3). Participants were repeatedly exposed to a referential structure of a particular type of pronouns or antecedents, and then were tested on ambiguous pronouns. All experiments showed that pronoun adaptation generalizes to new instances from the broad categorization of pronouns and antecedents.

Keywords: adaptation, pronoun comprehension, referential processing, discourse, exposure

Introduction

In language comprehension, adaptation occurs when people get used to specific linguistic inputs that they encounter in a language environment and thus become faster at processing them or more likely to adopt the frequent structure for ambiguous inputs (e.g., Fine & Jaeger, 2013; Kaan & Chun, 2018; Johnson & Arnold, 2022; Saffran et al., 1996). This raises a question: what units are grouped together for calculating frequency? Adaptation critically requires identifying linguistic categories. For example, does reading a sentence with a relative clause structure ("The defendant examined by the lawyer") result in adaptation to the abstract syntactic structure, or specifically to this structure in sentences using the word "examined"? (Tooley et al., 2019). By examining what contexts lead to adaptation, researchers can probe the types of representations that are stored and categorized together during comprehension.

In this paper we examine a type of adaptation that has received relatively little attention: adaptation to discourse-level referential structures. There is extensive evidence for adaptation at lexical, phonological and syntactic levels. For example, studies found that word frequency plays an important role in one's mental lexical representation. Words of high frequency tend to be recognized and processed faster than low-frequency words, because more exposure leads to a more salient mental representation (e.g., Brysbaert et al., 2018; Monsell et al., 1989). For phonological regularities, infants were found to be sensitive to the artificial syllable sequences as cues for setting word boundaries. They adapted to the distribution of the sequences and used it for future speech segmentation (Saffran et al., 1996). There is also evidence for syntactic adaptation, in which exposure to a syntactic structure facilitates processing of that structure in subsequent utterances. For example, sentences containing a rare or difficult syntactic structure

can become easier to process after repeated exposure to the structure (Fine & Jaeger, 2013; Fine et al., 2013; Jaeger & Snider, 2008; Kaan & Chun, 2018).

In contrast, much less is known about referential adaptation, which involves higher-level processing and more abstract representations that result from the complexity of referential structures. If I encounter a sentence like "Ana went skiing with Matt. He brought red gloves", I am exposed to an example of a pronoun that refers to Matt, who is in the prepositional object position of the previous sentence. Semantically this event includes Ana (the agent) and Matt, who plays the role of comitative (e.g., Stassen, 2000; Arkhipov, 2009). This exposure may be encoded abstractly as "pronoun refers to prior prepositional object", or "pronoun refers to prior comitative". If comprehenders learn about which patterns are most common, they may use these to guide the interpretation of future pronouns.

Evidence suggests that indeed, prior experience with pronouns does modulate later pronoun interpretation. Consider the sentence "William swooked Betty and Kevin brucked him," (Kaiser, 2009). Even though the novel verbs provide little semantic context, structurally the pronoun "him" refers to the subject of the previous clause, William. After reading this structure, participants were more likely to adopt the same structure for a sentence with an ambiguous pronoun ("Stephen tulvered Peter and Diane churbited him") than following other prime types (Kaiser, 2009). This suggests that people represent abstract relations between pronouns and their antecedents. A similar effect was observed by Contemori (2021), who also showed that this effect extends to L2 learners of English.

Kaiser (2009) and Contemori (2021) demonstrated that the interpretation of ambiguous pronouns was modulated by the pronoun-antecedent structure used in the previous item. This could be considered an example of immediate priming, such that interpretation preferences

within each participant changed over the course of the experiment. These effects can also last several days (Contemori et al., 2021). However, such effects are small.

Johnson and Arnold (2022) instead developed a paradigm to elicit stronger effects through repeated exposure. Participants were repeatedly exposed to a series of sentences with pronouns always referring to either the subject (e.g., "Ana went to the library with Will. She…") or the nonsubject (e.g., "Ana went to the library with Will. He…"). Their stimuli used sentences with the structure "X did something with Y", which they term **joint-action** constructions (termed a comitative construction by Arkhipov 2009). Antecedent type (subject vs. nonsubject) was manipulated between subjects. Participants were then tested on sentences that contained ambiguous pronouns to assess their pronoun interpretation preferences. There were 40 exposure sentences, all of the same structure. Since the exposure effect found in previous studies tended to be small, the paradigm starts with a training session consisting of 20 consecutive exposure items, in order to establish a salient representation and thus a more robust exposure effect. Thus, this paradigm examines adaptation as a result of cumulative exposure. They found that participants were more likely to assign an ambiguous pronoun to the subject when the exposure sentences had subject antecedents than when they had nonsubject antecedents.

These findings suggest that exposure to pronoun-antecedent structures may support learning about which types of structures are most frequent. That is, if participants encounter numerous structures where pronouns refer to the prior subject, this may lead to a bias to interpret future pronouns as coreferential with the prior subject. Indeed, there is evidence that subject-antecedent structures are more frequent in natural language (Arnold, 1998; Arnold et al., 2018), and this frequency pattern correlates with a tendency for comprehenders to assign pronouns to subject antecedents.

This line of work provides broad support for theories of pronoun comprehension that rely on probabilistic representations. For example, Arnold's Expectancy hypothesis (1998; 2010; Arnold et al., 2007) suggests that linguistic biases are related to exposure. For example, participants tend to assign pronouns to subject antecedents, and find these structures easier to process (Arnold et al., 2000; Fukumura et al., 2015; Rohde & Kehler, 2014; Stevenson et al., 1994). This correlates with a tendency for speakers and writers to refer more often to subjects overall, which makes reference to subjects highly frequent. Arnold (1998) reported a similar frequency pattern for other contextual biases, including the bias toward goal antecedents, recent referents, focused referents, and parallel syntactic structures. Another example of a probabilistic model comes from Kehler and Rohde (2013, 2019; Kehler et al. 2008; Rohde & Kehler, 2014). They propose that pronoun interpretation is guided by two factors: a) the probability of the referent being mentioned again (similar to Arnold's referential likelihood proposal), and b) the likelihood that speaker would use a pronoun for that referent. Their work focuses primarily on referential likelihood that stems from the semantic context, for example the tendency for implicit causes to be re-mentioned.

However, there are open questions about how people estimate referential likelihood, and what metrics are important for pronoun comprehension. In this paper we use adaptation to examine the nature of referential structures that are stored in memory. If participants use experience to learn about which pronoun-antecedent structures are most frequent, this would require people to categorize inputs. How do they do this?

How specific is adaptation?

Previous research demonstrated that language comprehenders adapt to referential patterns that were defined as a relation between a third person pronoun (he/she/him/her) and its

antecedent, where the antecedents were categorized in terms of syntactic position (Contemori, 2021; Johnson & Arnold, 2022; Kaiser, 2009). In all these studies, the verbs in the exposure and test sentences were different, suggesting that people keep track of the frequencies of abstract pronoun-antecedent structures, but their stimuli leave open questions about precisely how broadly or narrowly these structures are represented.

Anaphor Specificity. First, is adaptation specific to particular anaphor types? For example, exposure to "Ana saw Matt. She…" might provide information about pronouns specifically, supporting the calculation that pronouns tend to refer to the first or subject referent. Alternatively, participants may categorize referential exemplars in terms of as an example of reference in general, and not specific to the anaphor type. Johnson & Arnold (2022, exp. 1b) tested this by presenting participants with exposure sentences using name anaphors (e.g. "Ana went to the library with Will. Ana…." However, this had no effect on the interpretation of ambiguous pronouns. This suggests that people represent pronoun and name anaphors separately, or perhaps find it easier to learn about pronoun-antecedent structures than name anaphor-antecedent structures in a short time.

But even within the class of pronoun anaphors there are numerous possible representations. Exposure to "Ana saw Matt. He…" might provide information about the behavior of "he" specifically, about all third-person singular pronouns, or perhaps about all pronouns (including e.g. "I" and "you"). If adaptation occurs at the level of the entire pronoun class, you would expect that exposure to sentences with I/you anaphors would influence the interpretation of ambiguous he/she pronouns. This was precisely the question asked by Johnson and Arnold (2022, exp. 1c). They used exposure sentences with I and you (e.g., "I went to the

library with you. I…”), but found that this had no impact on the understanding of he/she pronouns.

This suggests that 3rd person pronouns are categorized separately from other anaphor types, but it does not tell us whether this adaptation is represented in terms of the entire class of 3rd person pronouns or is specific to the pronoun words “he” and “she”. Since exposure stories used both “he” and “she” pronouns, which always exhibited the same structure, participants may have separately learned the structure for each word. Thus, our first question is whether exposure generalizes at all from one pronoun to another.

There is reason to expect that learning about pronoun behavior is not specific to individual lexical items. Contemori & Minjarez Oppenheimer (2022) demonstrated that exposure to Spanish pronouns influenced the interpretation of English pronouns, most strongly when the Spanish stimuli used null pronouns; they found a similar but marginal effect for explicit pronouns. Moreover, research has demonstrated numerous contextual biases that guide pronoun comprehension, but no study has demonstrated different effects for he and she pronouns. Thus, we predict that pronoun exposure effects are not lexically specific.

Yet even if pronoun exposure generalizes across the class of third person pronouns, a second question is whether the effect is stronger when the pronoun matches. It has been shown that abstract syntactic priming is stronger when there is also lexical overlap between prime and target (Pickering et al., 2013; Tooley et al., 2014, 2019). This opens the possibility that comprehenders may represent pronoun exemplars in terms of both the abstract class (3rd person pronouns) and the specific lexical item (he vs. she). In addition, he and she pronouns may also be identified based on the gender category. Therefore, the question is whether there is a match advantage of using the same pronoun word that leads to stronger adaptation. If so, we might

predict stronger adaptation when the target matches the specific pronoun used in the exposure sentences than when it does not match. This is the question we ask in experiment 1.

Antecedent Specificity. Second, do people keep track of referential patterns for different categories of antecedents? Previous findings show that people categorize antecedents in terms of syntactic role, so exposure to a structure with a subject antecedent leads to more subject-interpretations of ambiguous pronouns than exposure to a structure with a nonsubject antecedent (Contemori, 2021; Johnson & Arnold, 2022; Kaiser, 2009). But it isn't clear whether such structures are specific to the semantic structure of the antecedent clause.

There is reason to suspect that antecedents with different semantic roles might be categorized differently, because semantic role has a notable effect on pronoun comprehension. For example, some verbs elicit a strong bias to expect that one person caused the event, a phenomenon termed "implicit causality" (Brown & Fish, 1983). This bias affects pronoun interpretation, for example in "Ana admired Liz because she was a great soccer player", people tend to assign "she" to Liz (e.g., Caramazza et al., 1977; Cozijn et al., 2011; Johnson & Arnold, 2021; Kehler & Rohde, 2013; Kehler et al., 2008). Thus, this semantic context (including the connector word "because") leads to a strong bias toward the object character. By contrast, contexts like "Ana was cleaning with Liz. She…" lead to a strong subject bias (Arnold et al., 2000, 2018; Nappa & Arnold, 2014). If these biases are related to the frequency of pronoun-antecedent structures for each type of context, we might expect that people would only generalize exposure to similar antecedent contexts. That is, perhaps reading "Ana blamed Will because he made a mess" should only affect the interpretation of other implicit-causality scenarios, or perhaps do so most strongly.

Evidence suggests that the strongest version of this hypothesis is not supported. Contemori (2021) used exposure sentences that always followed an implicit causality structure, e.g. “Holly trusted Dean because he never told lies.”[1] Her test stimuli always used the same syntactic structure (a transitive verb matrix clause followed by a subordinate clause), but the events in the prime sentences were always psych verbs with an implicit causality bias, whereas the target stimuli used non-psych transitive verbs followed by “when”, e.g. “Eric stabbed John when he worked in Rome last year.” Interpretation of the ambiguous pronoun was modulated by the prior prime, despite the fact that the semantic context of the prime was different. This shows that exposure effects do not require the antecedent clause to have the exact same semantic category, suggesting that pronoun-antecedent structures are somewhat abstract at the level of the antecedent representation.

On the other hand, we also know that comprehenders can keep track of pronoun-antecedent structures in terms of the semantic role of the antecedent, even when this requires ignoring the syntactic category of the antecedent. Johnson and Arnold (2022, experiment 2) tested whether people could pick out the relationship between a pronoun and an antecedent who is either a source or goal character in a transfer event (see Example 5). Transfer verbs identify a source character who transfers a possession to a goal character. This type of verb usually elicits a semantic bias towards the goal, such that people are more likely to interpret a following pronoun as referring to the goal character who is the endpoint of the described action (e.g., Langlois & Arnold, 2020; Stevenson et al., 1994).

[1] Thank you to Carla Contemori for sharing her stimuli with us.

(5) **Goal-Source condition**: Liz and Ana were attending a concert. Liz received the tickets from Ana and then she waited in line.

Source-Goal condition: Liz and Ana were attending a concert. Liz offered the tickets to Ana and then she waited in line.

In this case, the exposure sentences were manipulated to have pronouns that always referred to either the source or goal. Half of the exposure sentences referred to the subject, and half referred to the non-subject. Results showed that participants gave more goal interpretations following goal-reference exposure items than source-reference exposure items. This shows that they were able to learn the most frequent referential pattern based on the semantic role of the antecedent while ignoring the syntactic properties of the antecedents. When the same exposure sentences were re-organized so that the pronouns always referred to either the subject or nonsubject, participants instead learned this pattern. Thus, these findings also provide strong evidence that participants abstract across multiple exemplars to identify the most frequent pattern.

This raises the possibility that comprehenders categorize pronoun-antecedent structures at multiple levels. For example, when Johnson and Arnold's participants read joint-action constructions like "Ana went to the library with Will. He…", they may have stored that exemplar in terms of its abstract structure (third person pronoun referring to prior subject), in addition to tracking the properties of the antecedent structure – here, an action performed by the subject (Ana) who is the agent in collaboration with a second person who is a comitative introduced in an adjunct phrase (with Will). If people track the frequency of antecedents in terms of their semantic role, we might expect stronger exposure effects when the test stimulus uses the same predicate construction than when it does not. Experiments 2 and 3 test this question by

examining whether joint-action exposure items have a stronger effect on matching joint-action test items or mismatching transfer test items (Exp. 2) and whether transfer exposure items have a stronger effect on matching transfer test items or mismatching joint-action test items (Exp. 3).

The results of these experiments will be important for understanding the role of semantic bias in pronoun comprehension. It is well established that for most predicates, there is a bias to assign the pronoun to the prior subject (Arnold, 1998; Arnold et al. 2018), but some verbs show a relative bias toward the object or prepositional object (Stevenson et al., 1994). These biases can also be modulated by the coherence relations between clauses (Ehrlich, 1980; Kehler et al., 2008). If frequencies are stored on a construction-specific basis, it might suggest that the subject bias has to be learned separately for different verbtypes. But if people generalize frequencies across verbtypes, it would suggest that people may keep track of frequencies at multiple levels of categorization.

The Current study

To examine how fine-grained the characterizations of referential structures are and how generalizable referential adaptation is, the current study employs the paradigm developed by Johnson and Arnold (2022), extended to specifically manipulate the categories of pronouns and antecedents.

We tested whether people make either narrow or broad categorizations of two components of a referential structure: a) pronoun type (Exp. 1) and b) antecedent type (Exps. 2 and 3). For pronouns, on the one hand, a narrow categorization may be lexically based, leading people to identify and memorize a referential structure that is specific to a pronoun word (e.g., *he*), in which case the learning of a referential structure with *he* pronouns would not be extended to interpretations of *she* pronouns because *he* and *she* are two different words. On the other hand,

people making a broad categorization may group different pronouns including *he* and *she* together, as a general category of third-person pronouns. As a result, the learning of a referential structure with either *he* or *she* pronouns should influence the interpretations of the other.

Experiment 1 specifically tested whether people adapt to individual words *he* versus *she* or draw inferences about the behavior of third-person pronouns as a class. In a referential adaptation paradigm, participants were exposed to a series of short stories consisting of sentences containing unambiguous pronouns that identified either a subject-reference or a non-subject-reference referential structure (e.g., "Matt ate French fries with Liz. He spilled ketchup on the table."/ "Liz ate French fries with Matt. He spilled ketchup on the table.") Following the training items, participants were then tested on sentences containing ambiguous pronouns (e.g., "Liz painted the wall with Ana. She used the green paint.").

Each short story was followed by a reference question to assess pronoun interpretation. Following Johnson and Arnold (2022), we asked the reference question indirectly in such a way as to bias participants against the dominant subject interpretation. It is well established that participants tend to follow a subject-assignment strategy, such that in the painting story they prefer to assign "she" to Liz. At the same time, participants have a bias to answer "yes" (Johnson & Arnold, 2022). We therefore always asked about the nonsubject character, e.g. "Did Ana use the green paint?" This leads to greater variation in responses and moves participants away from ceiling. While Johnson and Arnold included both subject- and nonsubject- question types, both had similar effects so here we simplify the experiment by including just nonsubject questions. Participants were instructed to only use the information in the story to answer the question, so if they said "yes" it was interpreted as an assignment of the pronoun to the nonsubject, and if they said "no" it was taken as a subject interpretation of the pronoun.

In a 2x2 design, we manipulated a) the between-subject exposure condition (i.e., whether participants received subject-reference structures or non-subject-reference structures), and b) the within-subject match condition (i.e., whether the pronoun word in testing items matches that in exposure items). We predicted that if participants categorize he/she pronouns together, exposure to just he or just she pronouns would influence both he and she pronoun interpretation. By contrast, if participants categorize pronouns narrowly in terms of specific lexical items, the exposure structures containing *he* pronouns, for example, would not influence the interpretations of *she* pronouns used in testing items. A secondary question is whether the lexical overlap would lead to a "Match Advantage", such that exposure would influence pronoun comprehension in both conditions, but more strongly in the Match condition.

Experiments 2 and 3 instead examined whether people categorize pronoun-antecedent structures broadly or narrowly. People may characterize pronoun-antecedent structures as a function of the verbtype used in the antecedent clause and the thematic roles it imparts to the antecedent. For example, as discussed previously, a transfer verb identifies a goal and a source character and usually elicits a goal bias that people tend to assign an ambiguous pronoun to the goal character. If people make a narrow categorization of antecedents in the learning of a referential structure, their adaptation to referential structures containing a transfer verb should not extend to the comprehension of structures containing other verbtypes. Otherwise, if they make a broad categorization, adaptation to transfer verbs should influence the comprehension of ambiguous pronouns associated with other verb constructions.

Experiments 2 and 3 asked whether people distinguish structures of antecedents specific to one verb construction from structures containing other verb constructions, or generalize the learning of a referential structure across antecedents denoted by different verb constructions. We

manipulated the between-subjects exposure condition (i.e., whether participants receive subject- or non-subject-reference structure), and the within-subjects match condition (i.e., whether the testing items contain verbs matching those in the exposure items). Experiment 2 tested whether adaptation to referential structures containing transfer verbs influences the comprehension of structures containing the joint-action construction. Experiment 3 tested the opposite, that is, whether adaptation to the joint-action construction generalizes to the comprehension of transfer verbs. Again, our secondary question was whether we would observe a Match Advantage such that exposure effects would be stronger when the antecedent type matched between exposure and test stories.

One challenge for testing the match advantage is that we expect that exposure may have a stronger effect on some trial types than others. Considering just the nonsubject questions (the type used in the current experiments), Johnson and Arnold observed an average difference of 11% between the subject and nonsubject exposure conditions for the joint-action trials (Exp. 1a), and the goal-source trials (Exp. 2a), but a 27% difference in the source-goal trials (Exp. 2a). Thus, source-goal verbs exhibit a numerically larger exposure effect. This may be related to the tendency for participants to follow both subject- and goal-assignment strategies for this verbtype (Langlois & Arnold, 2020; Stevenson et al., 1994). Our test sentences in experiments 2 and 3 included both joint-action and source-goal verbs. This counterbalances match condition with verbtype across experiments.

Our studies also included demographic questions for the purpose of characterizing the sample, as well as the Author Recognition Task (Stanovich & West 1989; Moore & Gordon, 2015). This task provides a list of real author names intermixed with fake author names, and the participant indicates which author names they recognize. Participants are asked not to guess, and

those that select too many fake names are considered guessers and not included in the analysis. We included this task for consistency with other studies in this project, and to keep our inclusion criteria constant. For further discussion of this measure and inclusion criteria see Appendix B.

Data availability

Data and materials for all experiments are available at [link here] and [link here].

Experiment 1 (He vs. She)

Methods

Participants. 91 Amazon Mechanical Turk workers participated in exchange for $1.50. The Mturk HIT was only available to participants in the US, Australia, and England, and required participants to have 5000 previous HITS with 95% acceptance rate. Of these, 35 were excluded from the analysis for the following reasons.

Accuracy. Participants were informed that if they did not respond accurately, they would be dismissed from the experiment without pay. For this purpose, we calculated the rate of correct answers to 24 of the unambiguous questions (i.e., content questions and filler reference questions). We only included questions from the first 75% of the stories in this count so that participants would not be dismissed from the study after completing most of it. Out of 24 check questions, participants could miss up to 6 and on the 7th incorrect answer they were dismissed.

For participants who finished, we additionally hand-checked accuracy on all content questions (both filler and target items) and the reference questions for the filler items (n = 76). Participants with less than 75% accuracy were excluded from analysis.

Author Recognition Task Guessing. Participants were instructed not to guess when selecting authors; those who selected foils on more than 33% of their responses were dubbed guessers and excluded from analysis. This metric is used by many studies in our lab, and it typically leads to a

large number of exclusions. The viability of data collected on crowdsourcing platforms critically depends on using strict inclusion criteria (e.g., Chandler & Shapiro, 2016; Miller et al., 2017).

Language disorder. Participants who reported a language disorder were not included in the analysis.

Duplicates. If a participant had the same IP address as a previous participant, we assumed it was a duplicate and did not include the data in our analysis.

On this experiment, we excluded 21 participants for guessing on the Author Recognition Task; 3 for failing our accuracy cutoff; 1 for reporting a language disorder; 4 for both ART guessing and accuracy; 1 for both ART guessing and language disorder; 3 for ART guessing and duplicate entries; 1 for ART, accuracy, and duplicate entries; and 1 for ART guessing, reporting a language disorder, and accuracy. 56 participants were included in the analysis.

Materials and Design.

Participants responded to a survey in Qualtrics, which included background demographic questions and the main story task (see Appendix A). Demographic questions asked about socioeconomic status, age, sex, gender, ethnicity/race, language experience, English language proficiency, and included the Author Recognition Task.

In the main task they read pairs of two-sentence stories, as shown in Table 1, and answered simple questions. The story appeared on one page, and the participants pressed a button to see each question on subsequent pages. All the stories were about four characters: Ana (she); Liz (she); Matt (he), and Will (he). The first sentence mentioned two people, the second mentioned one with a pronoun.

Participants were instructed to answer the questions and “only use the information explicitly stated in the sentence.” Feedback on the practice question reinforced this. The practice

story read “Matt ordered pizza with Ana. He ate a pepperoni slice,” and the question asked, “Did Ana eat a pepperoni slice?” If the participant responded yes, they saw: “The correct answer is “no”, because the story talked about what Matt ate but did not say anything about what Ana ate,” and if the participant responded “no” they saw: “Correct, the story talked about what Matt ate but did not say anything about what Ana ate.”

There was a total of 44 items in the survey; 12 critical and 32 exposure stories. The stories were organized so that the first 12 in the list were exposure stories, which established the referential pattern. The remaining 32 were a mix of 12 critical and 20 additional exposure stories.

Our key manipulation was the type of exposure story. All exposure stories had two characters of different gender, so the pronoun was unambiguous. Within each list, all 32 exposure stories followed the same referential pattern, such that the pronoun either always referred to the subject or always referred to the nonsubject character. We also manipulated the order of the names in the first sentence, so that within a given list, all the exposure stories used either “he” or “she” as a pronoun. This resulted in four exposure conditions (subject-reference/he; subject-reference/she; nonsubject-reference/he; nonsubject-reference/she).

The 12 critical stories all included two same-gender characters, so the pronoun was ambiguous. We used these stories to test how participants interpreted the ambiguous pronoun. The key dependent measure was their response to the reference question. E.g., for “Ana had coffee with Liz. She ordered a cappuccino”, we asked “Did Liz order a cappuccino?” The critical reference question always asked about the second-mentioned person, so if participants think that “She” refers to Ana, they should answer “no”.

The 12 critical stories were identical across lists. 6 were about Matt and Will, using the critical ambiguous pronoun “he”, and 6 were about Ana and Liz, using the critical ambiguous

pronoun “she”. This meant that the pronoun in the target story either matched or mismatched the pronouns used in the exposure stories for that list. This match/mismatch category was one of our key predictors in the analysis; the other was whether the exposure stories exhibited subject or nonsubject reference patterns.

All stories were followed by two questions, one reference question and one content question. The reference question for the exposure stories was always unambiguous based on the pronoun’s gender. The content questions were extremely simple to answer if participants were reading, and asked about either what the participants were doing, or what object was involved in the action. See Table 1 for an example. Responses were 2-alternative forced choice, with the correct answer on the top for half the questions, and on the bottom for the other half.

Table 1.
Example exposure story:
Subject-reference / He: Matt ate french fries with Liz. He spilled ketchup on the table.
Nonsubj-reference / He: Liz ate french fries with Matt. He spilled ketchup on the table.
Subject-reference / She: Liz ate french fries with Matt. She spilled ketchup on the table.
Nonsubj-reference / She: Matt ate french fries with Liz. She spilled ketchup on the table.

Content question: What did Liz and Matt eat (or, What did Matt and Liz eat)? (French fries / a Cheeseburger)
Reference question: Did Matt spill ketchup? (Yes/ No)

Example critical story:
Liz painted the wall with Ana. She used the green paint.

Content question: What did Liz and Ana do? (Painted the wall / told jokes)
Reference question: Did Ana use the green paint? (No / Yes)

Procedure

Participants accepted the HIT on Amazon Mechanical Turk and were directed to a link to a Qualtrics survey. The survey began with a consent form, followed by questions about age, whether they spoke English as a native language, and what age they had learned it. If they did not meet our age criterion (18+) or had not learned English as a native language and beginning

before age 7, they were automatically dismissed from the survey. When they completed the survey they received a random code, which they entered in the project page to be paid.

Analytical Approach

Our dependent measure was the selection of "No" responses for the critical items, which signals a subject interpretation of the pronoun. These binary data were analyzed with a mixed-effects logistic regression using SAS proc glimmix with a binary distribution and a logit link. Our primary predictors were exposure condition (subject vs. nonsubject exposure), whether the target story pronoun matched or mismatched the pronoun of the exposure stories, and the interaction between the two. These predictors were effects-coded (subject-reference = 0.5; nonsubject-reference = -0.5; match = 0.5; mismatch = -0.5). Models included random intercepts for participant and item, and maximal random slopes (Match by participant; Match, Exposure, and the interaction by item).

Results and Discussion

Figure 1 demonstrates that the exposure effect obtained for both Match and Mismatch conditions. That is, participants who were only exposed to either "he" or "she" pronouns followed the exposure pattern for interpreting ambiguous target pronouns, regardless of whether the target pronoun was he or she. This answers our primary question, demonstrating that the exposure effect is robust even when the lexical item mismatches. There was no interaction between pronoun and Match status, and in fact numerically the average responses in each condition were exactly the same for match and mismatch conditions. This answers our second question, and suggests there is no match advantage. Together, these findings suggest that people store anaphoric relations in memory as a function of the word class, and not the individual word.

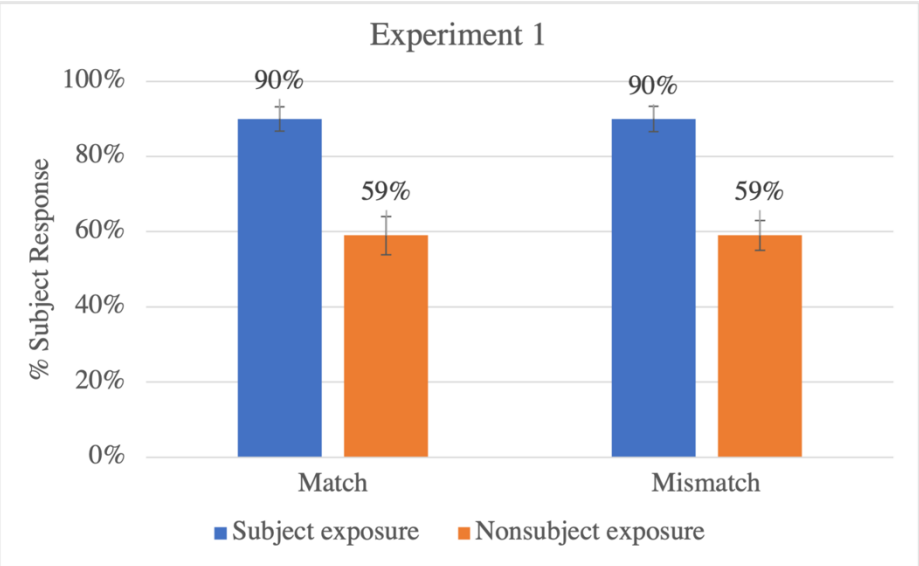

Figure 1. Results from Experiment 1 by condition.

Table 2. Inferential Statistics from Experiment 1.

Effect	Estimate (St. Error)	t Value	p
Intercept	1.82 (0.35)	5.17	<.0001
Subject vs. Nonsubject Exposure	2.46 (0.68)	3.64	0.0006
Pronoun Match vs. Mismatch	-0.04 (0.27)	-0.14	0.8911
Exposure x Match	-0.02 (0.54)	-0.03	0.9745

Table 3. Estimates for the Exposure effect in Match and Mismatch conditions.

Effect	Estimate (St. Error)	t Value	p
Exposure Effect for Match	1.88 (0.67)	2.8	0.0064
Exposure Effect for Mismatch	2.13 (0.68)	3.14	0.0023

Experiment 2 (Transfer vs. Joint-Action)

Experiment 1 suggested that participants considered *he* and *she* pronouns as a general third-person pronoun class, thus making a broad categorization of pronouns. Experiments 2 and 3 then examined the categorization of antecedents as a function of verbtype and the thematic roles that the verb imparts to the antecedent.

Methods

Participants. 204 Amazon Mechanical Turk workers participated; the first 125 received $1.67; we then realized that the experiment was taking too long for that pay rate and increased the pay to $2.5 for the remaining 79 workers. The Mturk HIT was only available to participants in the US, Australia, and England, and required participants to have 5000 previous HITs with 95% acceptance rate. Of these, 86 were excluded from the analysis for the same reasons as Experiment 1 (accuracy, Author Recognition Task guessing, language disorder, and duplicate entries).

For accuracy, this experiment included 35 check questions, out of which participants could miss up to 8 and on the 9th incorrect answer they were dismissed. For participants who completed the study, we additionally hand-checked accuracy on all questions including the content questions for both target and filler items, and the reference questions for the filler items (n = 72).

On this experiment, we excluded 19 participants for guessing on the Author Recognition Task only; 13 for failing our accuracy cutoff only; 4 for both reporting a language disorder and failing accuracy cutoff; 18 for both ART guessing and reporting a language disorder; 30 for both ART guessing and failing accuracy cutoff; and 2 for duplicate entries. 118 participants were included in the analysis.

Materials and Design

Identical to Experiment 1, in this experiment participants responded to a survey in Qualtrics that included background demographic questions and a main story task (see Appendix A). The instructions and the main task took the same form as that in Experiment 1.

There was a total of 44 items in the survey, including 16 critical and 28 exposure stories. The first 12 items were all exposure stories for establishing a referential pattern. The remaining 32 items were a mix of 16 critical and 16 additional exposure stories.

In this experiment, we manipulated the type of referential patterns in the exposure stories. As in Experiment 1, all exposure stories had two characters whose gender was different, making the pronoun unambiguous. Within each list, all exposure stories followed the same referential pattern where the pronoun either always referred to the subject or to the non-subject character. All of the exposure items contained a transfer verb. Half of these used source-goal verbs that denoted the first-mentioned character as the source and the second-mentioned character as the goal; the other half used goal-source verbs that placed the source and goal characters in reverse grammatical positions.

For critical items, the characters were always matched by the same gender so that the pronouns were ambiguous. The key dependent measure was their response to the reference question. As in Experiment 1, the critical question always asked about the second-mentioned character so that a “No” response indicated a subject interpretation.

For the 16 critical items, half contained source-goal verbs (Match condition) and half contained joint-action constructions (Mismatch condition). For the matching transfer verbs we used only source-goal verbs, since these exhibited the largest priming effect (Johnson & Arnold,

2022). Thus, in a 2x2 design we manipulated both exposure (subject vs. nonsubject) and whether the target item matched the exposure predicate type (Match/transfer vs. Mismatch/joint action).

As in Experiment 1, all stories were followed by a content question and a reference question; the order of which came first was counterbalanced. The content questions were assigned to indicate whether participants paid attention to the task, asking about either what the characters were doing, or when/where they were doing something. The reference questions asked whether one of the characters was involved in an action. For exposure items, half of the reference questions asked about the first-mentioned character and the other asked about the second-mentioned character. Across lists, items were matched for the person mentioned in the question, so if the correct answer was yes in the subject-reference list, it was no in the nonsubject-reference list (and vice versa). For critical items, the reference question always asked about the second-mentioned character. For all questions, responses were 2-alternative forced choice. The correct answer was placed on the top for half of the questions and on the bottom for the other half.

Table 4.

Example exposure story:

Subject-reference:

Matt and Liz were studying at the library. Matt retrieved a book from Liz and then he took notes.

Nonsubj-reference:

Matt and Liz were studying at the library. Matt retrieved a book from Liz and then she took a break from studying.

Content question: Where were Matt and Liz studying? (At the library / At the mall)
Reference question: Did Matt take notes? (Yes / No) or Did Matt take a break from studying? (Yes/No)

Example critical stories:

Source-Goal Verb:

Liz and Ana were going to a concert. Liz offered the tickets to Ana and then she waited in line.

Content question: Where were Liz and Ana going? (A concert / The zoo)
Reference question: Did Ana wait in line? (Yes / No)

Joint-Action Constructions:

Will and Matt were watching TV. Will had a snack with Matt and then he cleaned up the table.

Content question: What were Will and Matt doing? (Watching TV / Shopping)
Reference question: Did Matt clean up the table? (Yes / No)

Procedure

The procedure in this experiment was identical to that in Experiment 1.

Analytical Approach

We used the same mixed effects logistic regression analysis with maximal random effects as in Experiment 1, and again the predictors were effects coded. For this experiment, the dependent measure was the frequency of "No" responses to reference question for the critical items, which indicated a subject-interpretation of the pronoun. The predictors were exposure condition (subject vs. non-subject exposure), whether the critical verbs (source-goal verb and joint-action construction) matched or mismatched one of the exposure verbs (source-goal and goal-source verbs), and the interaction between the two.

Results and Discussion

Figure 2 illustrates the exposure effect obtained for both verb-match and -mismatch conditions. For participants who received subject-interpretation exposure items, the average proportion of subject-interpretations for ambiguous pronouns was 84% for matched verbs vs. 88% for mismatched verbs. For participants who received nonsubject-interpretation exposure items, the average proportion of subject-interpretations for ambiguous pronouns was 52% for matched verbs vs. 72% for mismatched verbs. That is, participants in both conditions exhibited a strong effect of exposure: they were more likely to give a subject interpretation when they had been exposed to pronouns with subject antecedents than when they had been exposed to pronouns with nonsubject antecedents. In addition, the exposure effect for the same verb type (source-goal verb) was stronger than for a different verb type (joint-action construction).

These patterns were confirmed by our analysis (see Table 5). There was a main effect of exposure condition, and a main effect of verb-match condition, and an interaction between the exposure and verb-match conditions. To probe the interaction, we estimated the effect of exposure on both matched- and mismatched-verb conditions. The effect is robust in both conditions (see Table 6).

These findings suggest that referential adaptation is not restricted to contexts using the same predicate type and thematic roles, in that we observed strong exposure effects for both match and mismatch conditions. At the same time, they raise the possibility that there may be a Match Advantage, with stronger exposure for matching conditions. However, an alternative possibility is that the source-goal target items respond more strongly to exposure manipulations than the joint-action target items, which is consistent with what Johnson and Arnold (2022) found.

Experiment 3 tested this possibility by using joint action contexts in the exposure items, thus reversing the verbtypes used in the Match and Mismatch conditions. If there is a Match advantage, this time we should see stronger exposure effects for the joint action contexts. If instead the source-goal verbs are simply more sensitive to exposure effects, we should see a Mismatch advantage in Experiment 3.

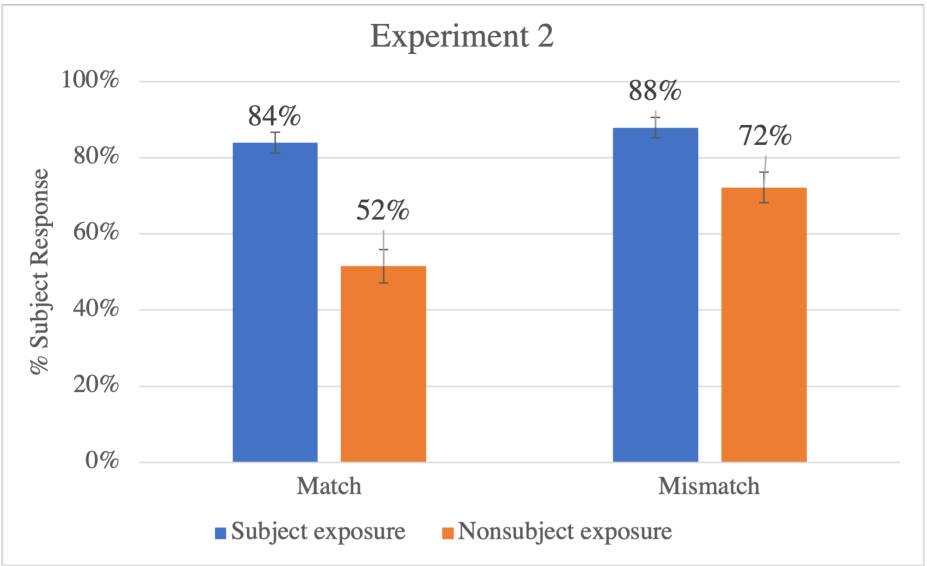

Figure 2. Proportion of subject interpretations by conditions from Experiment 2.

Table 5. Inferential Statistics from Experiment 2

Effect	Estimate (St. Error)	t Value	p
Intercept	1.53 (0.21)	7.28	<.0001
Subject vs. Nonsubject Exposure	1.64 (0.34)	4.85	<.0001
Verb Match vs. Mismatch	-0.83 (0.29)	-2.83	0.0126
Exposure x Match	0.92 (0.31)	2.98	0.0077

Table 6. Estimates for probing interaction in Experiment 2

Effect	Estimate (St. Error)	t Value	p
Priming Effect for Match (Source-Goal Verbs)	2.10 (0.36)	5.77	<.0001
Priming Effect for Mismatch (Joint-Action Constructions)	1.18 (0.38)	3.11	0.0025

Experiment 3 (Joint-Action vs. Transfer)

Methods

Participants. 174 Amazon Mechanical Turk workers participated in exchange for $3.33.[2] The recruitment criteria were the same as in Experiment 2. Of these, 52 were excluded from the analysis for the same reasons as Experiment 2 (accuracy, Author Recognition Task guessing, language disorder, and duplicate entries).

The numbers of items and check questions for accuracy cutoff were the same as those in Experiment 2. For participants who completed the study, we additionally hand-checked accuracy as we did in Experiment 2.

On this experiment, we excluded 14 participants for guessing on the Author Recognition Task only; 10 for failing our accuracy cutoff only; 1 for reporting a language disorder only; 2 for both reporting a language disorder and failing accuracy cutoff; 4 for both ART guessing and reporting a language disorder; 18 for both ART guessing and failing accuracy cutoff; and 3 for duplicate entries. 122 participants were included in the analysis.

Materials and Design

This experiment was the same as Experiment 2, except that all the exposure items used joint-action constructions instead of transfer verbs. As a result, there was a total of 44 items in the survey, including 16 critical items and 28 exposure stories. The critical items consisted of 8 items that contained joint-action constructions and 8 items that contained transfer (source-goal) verbs. The 28 exposure items all contained a joint-action construction. We tested whether exposure to joint-action constructions led to adaptation that generalized to transfer verbs. In this case, the critical items matched the exposure items if they contained joint-action constructions

[2] After finishing data collection for Exp. 2 we concluded that it was taking longer than our initial testing had suggested, so we increased the pay rate for Exp. 3.

and mismatched if they contained transfer verbs. For the content questions, 12 items asked about one of the story characters; 10 asked for the objects involved in characters' activity; 5 asked for when the activity took place and 5 asked for the location. As in Exp. 2, the reference questions always asked about the second character for the critical items. For the exposure items, the reference questions asked about the first and second characters equally, so there were an equal number of yes and no answers. The same question was used for both subject- and nonsubject-reference lists, so within an item the answer (yes or no) was different across lists.

Table 7.

Example exposure story:

Subject-reference:
Matt and Ana were visiting a pet store. Matt pet the cat with Ana and then he pet the dog.

Nonsubj-reference:
Matt and Ana were visiting a pet store. Matt pet the cat with Ana and then she pet the dog.

Content question: Who did Matt pet the cat with? (Ana / Liz)
Reference question: Did Matt pet the dog? (Yes / No)

Example critical story:

Source-Goal Verb:
Liz and Ana were going to a concert. Liz offered the tickets to Ana and then she waited in line.

Content question: Where were Liz and Ana going? (A concert / The zoo)
Reference question: Did Ana wait in line? (Yes / No)

Joint-Action Construction:
Will and Matt were watching TV. Will had a snack with Matt and then he cleaned up the table.

Content question: What were Will and Matt doing? (Watching TV / Shopping)
Reference question: Did Matt clean up the table? (Yes / No)

Procedure

The procedure in this experiment is identical to that in Experiment 2.

Analytical Approach

The analysis and effects-coding of the predictors were the same as in Experiment 2. For this experiment, the predictors were exposure condition (subject vs. non-subject exposure), whether the critical verb constructions (joint-action and source-goal) matched or mismatched the exposure verb constructions (joint-action), and the interaction between the two.

Results and Discussion

Figure 3 illustrates the exposure effect obtained for both verb-match and -mismatch conditions. For participants who received subject reference exposure items, the average proportion of subject-interpretations for ambiguous pronouns was 87% for matched verbs vs. 79% for mismatched verbs. For participants who received non-subject reference exposure items, the average proportion of subject-interpretations for ambiguous pronoun was 66% for matched verbs vs. 42% for mismatched verbs. That is, participants in both conditions showed a strong effect of exposure. In addition, the exposure effect was stronger for one condition than the other. But unlike Experiment 2, the effect was stronger in the Mismatch (transfer verb) condition than the Match (joint action) condition. These patterns were confirmed by our analysis (Table 8), which found a main effect of exposure condition, and a main effect of verb-match condition, and an interaction between the exposure and verb-match conditions.

In sum, Experiment 3 results confirmed the finding that exposure effects generalize from one construction type to another. That is, exposure to a structure such as [pronoun → obj of PP/comitative antecedent] increases the likelihood of assigning a subsequent pronoun to the prepositional object antecedent, even if it has a goal thematic role and participated in a different event type.

These findings also confirmed that there is no pervasive Match advantage, in that exposure effects are not always stronger when the target and exposure event types match. In

experiment 2 we found a Match advantage, but in experiment 3 we found a Mismatch advantage. This pattern instead shows that the exposure effect was always stronger for target items including transfer verbs than for target items including joint-action constructions.

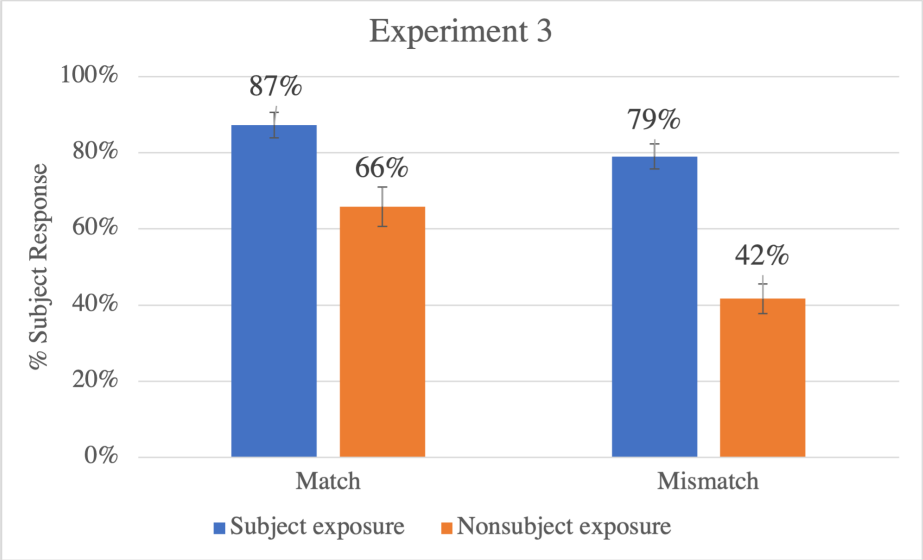

Figure 3. Proportion of subject interpretations by conditions from Experiment 3.

Table 8. Inferential Statistics from Experiment 3

Effect	Estimate (St. Error)	t Value	p
Intercept	1.30 (0.29)	4.49	<.0001
Subject vs. Nonsubject Exposure	2.31 (0.42)	5.48	<.0001
Verb Match vs. Mismatch	1.27 (0.43)	2.95	0.0099
Exposure x Match	-0.97 (0.34)	-2.82	0.0100

Table 9. Estimates for probing interaction in Experiment 3

Effect	Estimate (St. Error)	t Value	p
Priming Effect on Match (Joint-Action Construction)	1.83 (0.47)	3.92	0.0002
Priming Effect on Mismatch (source-goal verb)	2.80 (0.45)	6.29	<.0001

General Discussion

In three experiments, we demonstrated robust effects of referential adaptation, and used this phenomenon to probe the specificity with which people categorize pronoun-antecedent representations. These findings have three critical theoretical implications: 1) they illustrate the robustness of statistical learning at the discourse level, 2) they show that pronoun inputs are categorized at the level of pronoun category and not lexical item, and 3) they show that antecedents are categorized at the level of grammatical category and not separately by thematic role.

First, our findings replicate and extend the findings reported by Johnson and Arnold (2022), and further support the hypothesis that comprehenders store representations of abstract discourse-level structures. While the study of referential adaptation is relatively new, research on pronoun comprehension suggests that comprehenders are guided by the syntactic and thematic roles of potential referents in the context (e.g., Arnold et al., 2000; Kehler et al., 2008; Langlois & Arnold, 2021; Stevenson et al., 1994). The results reported here confirm that people can adapt to the frequency of pronoun-antecedent structures within a relatively short period of time, where that frequency is defined by the syntactic role of the antecedent. When people read stories where pronouns referred to the prior prepositional object, they were relatively more likely to adopt a similar interpretation for new stories with ambiguous pronouns, in contrast with people who read stories where pronouns referred to the prior subject.

Second, the critical findings reported here were that adaptation to pronoun-antecedent structures is generalized across relatively broad categories. Experiment 1 tested whether exposure to “he” pronouns influences interpretation of “she” pronouns and vice versa. We found a strong exposure effect for both matching and mismatching pronouns, and no interaction. This

result contrasts with Johnson and Arnold's (2022) finding that adaptation did not occur when people were exposed to names or I/you pronouns. Together, this suggests that pronoun-antecedent relations are stored with an abstract representation of third-person pronouns as a class, and not in terms of specific lexical items. At the same time, this representation is specific to anaphoric pronouns, given findings that people do not adapt to the discourse relations exhibited by I/you pronouns, which are deictic.

Third, Experiments 2 and 3 tested the antecedent categories involved in adaptation. If referential adaptation involves narrow antecedent categories, people might learn very specific relations like "pronoun refers to the prepositional object that is also a goal in a transfer event". If so, we would not expect exposure to pronouns in the context of one predicate type to influence interpretation of pronouns in another predicate type. But contrary to this hypothesis, we found that exposure generalized across predicate contexts for both Experiments 2 and 3. Exposure to pronouns in the context of transfer events affected pronoun interpretation following both transfer events and joint action predicates (Exp. 2), and exposure to pronouns in the context of joint action predicates affected pronoun interpretation for both predicate types as well (Exp. 3).

This finding suggests that people store pronoun-antecedent relations in terms of an abstract representation of the antecedent type. We already knew that this representation could not be lexically specific, since different verbs were used in each of the stories used by Johnson and Arnold (2022), and the same is true in the current experiments. Experiments 2 and 3 show that this representation also abstracts across thematic role type. In each experiment, participants were presented with numerous exposure items that all followed the same pronoun-antecedent pattern, either "pronoun refers to subject" or "pronoun refers to nonsubject". People appear to have learned this abstract pattern at a level that does not encode the specific thematic role of the

antecedent: when they encountered a new pronoun, they tended to follow this abstract pattern for both antecedents with a matching thematic role and antecedents with a mismatching thematic role.

From these findings it might be tempting to conclude that referential adaptation only encodes antecedents in terms of grammatical role information. However, previous findings suggest that participants can also encode relations that are categorized by thematic role. Johnson and Arnold (2022, Exp. 2b) presented participants with both goal-source and source-goal contexts. The exposure sentences had pronouns that either always referred to the goal, or always referred to the source. In that experiment, grammatical role was not informative for adaptation, because half the antecedents were subjects and half nonsubjects. Their results show that it is possible to learn thematic role-driven patterns. But experiment 2 and 3 results here demonstrate that learning syntactically-conditioned referential patterns is not limited to a single type of semantic context.

Together, these findings suggest that people may be able to encode abstract patterns at multiple levels. As in experiments 2 and 3 here, they may represent pronoun-antecedent patterns in terms of the abstract grammatical role, which is informative in the current context. But if thematic role information is informative, participants may also encode the thematic role of the antecedent too. In the richer input that people gain over a lifetime, we speculate that they may weight each level of abstraction according to its utility in predicting pronoun referents. Thus, for example people may observe that pronouns broadly tend to refer to subjects more than nonsubjects, so this is a highly informative cue that generalizes across many predicate types. But the frequency of this pattern may be stronger for some predicate types than others (Guan & Arnold, 2021). In the case of transfer events, it appears that the goal bias is relatively weak, so it

may not be weighted as strongly as the subject bias in representations of natural language frequencies (Langlois & Arnold, 2022).

Our results also suggest that there is no “Match advantage”, in that people do not follow the exposure context more strongly when the target story matches the context stories in either anaphor or antecedent. In the case of our pronoun manipulation (Exp. 1), this means that people do not encode the specific pronoun word used, and instead encode pronouns as a general class. This is consistent with Contemori and Minjarez Oppenheimer’s (2022) finding that exposure transfers across languages.

At the same time, we observed a stronger effect of exposure on the source-goal target items than on the joint-action target items, replicating the same contrast in Johnson and Arnold’s data. We speculate that this difference results from one of two properties of this story type. First, source-goal contexts are known to elicit a relatively weak subject bias (e.g., Langlois & Arnold, 2021; Stevenson et al., 1994). This has been attributed to the presence of two simultaneous biases: a bias to assign the pronoun to the subject, and a bias to assign the pronoun to the goal. By contrast, the subject bias for joint action contexts is very strong (e.g., Arnold et al., 2018; Nappa & Arnold, 2014). One possibility is that exposure effects are easier to observe when the other contextual constraints are weaker. When the context is very strong (as in the joint action context), people are so strongly biased to select the subject, recent exposure has relatively less of an effect.

A second possibility is that the source-goal context is particularly sensitive to exposure effects because both the subject-exposure and nonsubject-exposure conditions resonate with existing biases. That is, the local exposure may have amplified pre-existing biases. For such verbs, people have both a subject-bias and a goal-bias. When people received subject-exposure

patterns, their subject bias was enhanced and therefore more pronouns were assigned to the subject. In contrast, when people received the nonsubject-exposure pattern, the bias aligned with their goal bias, thus more pronouns were assigned to the goal referents in this case. This possibility suggests there may be an interaction between local exposure and pre-existing biases learned from prior language experience.

Our findings lead us to hypothesize that language exposure results in memories for abstract structures at the discourse level. These representations may take the form of structures like [3rd person pronoun] → [subject antecedent] and [3rd person pronoun] → [nonsubject antecedent], such that each structure is weighted in terms of its probability in the current context. Alternatively, they may take the form of processes, such that the search process for an antecedent for a 3rd person pronoun is weighted toward one type of antecedent over another. Either way, the current evidence for adaptation suggest that people do encode the relationship between pronoun and antecedent in some way.

While our experiments provided consistent evidence for adaptation, it is also notable that responses reflected a strong baseline preference to assign the pronoun to the subject, especially in the joint-action construction. Our critical "no" question was designed to avoid a ceiling effect, and yet the joint-action stories still elicited a strong subject bias, even in the nonsubject-exposure condition. This suggests that people integrate the local frequencies with biases stemming from prior experience. Pronouns tend to refer to the subject (Arnold et al., 2018), leading to a baseline subject bias. When local inputs signal a strong nonsubject bias instead, this shifts pronoun interpretations a little, but does not completely overturn the subject bias. This finding is consistent with evidence that pronoun comprehension in Portuguese is influenced by a mix of the

frequency of null vs. overt pronouns in the local context and the previously-encountered statistics of the language (Fernandes et al., 2018).

On the other hand, the fact that local adaptation occurs at all provides further evidence that recent individual inputs may be prioritized. Participants come to the experiment with a lifetime of experience with English. This provides an overwhelming amount of exposure compared to the exposure provided during the experimental session itself. If people calculate the frequency of pronoun-antecedent relations over their lifetime and weight each token equally, the exposure during the experimental session would be negligible. For example, if we imagine that people hear 20 pronouns a day (likely an underestimate), between ages 5 and 18 they would have heard about 94900 pronouns. If we estimate that 80% of those refer to the prior subject, then adding another 32 pronouns that refer to the nonsubject (as in Experiment 1, nonsubject exposure), this would shift the likelihood of a subject antecedent to 79.97% -- probably not a noticeable difference. This suggests that the inputs in recent experience must be weighted more strongly than previously-encountered inputs. As a result, participants adapt to the local statistics so that they become biased towards the exposure pattern. However, one open question is how long this adaptation lasts. Contemori et al. (2021) used a delayed posttest and found that the effect of exposure can last for several days, so it's possible that with enough input, a new statistical pattern could change people's long-term representations.

Critically, the current findings identify the specificity with which these representations are calculated. Our findings suggest that 3rd person pronouns are represented as a general class, and that antecedents can be represented as a contrast between subjects and nonsubjects, without specifying the thematic role.

But there are also still several unanswered questions. For example, we don't know if pronominal anaphors are categorized together across animacy and number categories. Another open question concerns the grammatical role of the antecedent. Notably, both of the construction types that we investigated contrasted antecedents in subject and prepositional object roles. Other verbtypes instead include two potential antecedents in subject/object roles (e.g., Ana saw Liz). Linguistic theories suggest that subject, object, and prepositional object positions are represented differently, and they tend to be associated with different levels of discourse prominence (Brennan, Friedman, & Pollard, 1987; Kenan & Comrie, 1977). Thus, people may adapt differently to structures with object antecedents vs. prepositional object antecedents. On the other hand, English speakers may instead make a broader classification, dividing antecedents into "subject" (which tends to be the topic), and "other" (including both object and prepositional object). Alternatively, they may not use syntactic role at all in classifying antecedents, but instead may define them in terms of "topic" (which is often assigned to subject) and "non-topic". Future research is needed to test whether adaptation generalizes across these categories or not.

In conclusion, our findings demonstrated that comprehenders can adapt to pronoun-antecedent structures that are frequent in the input, where these structures are easily classified in terms of the syntactic role of the antecedent. This adaptation demonstrates that discourse-level structures are encoded in memory, and further shows that these representations include relatively broad categories of both pronoun and antecedent. We speculate that this experimental effect mimics the way people adapt to the most frequent referential patterns in real-life exposure.

References:

Arkhipov, A. (2009). Comitative as a cross-linguistically valid category. New challenges in typology: Transcending the borders and refining the distinctions, 223-246.

Arnold, J. E. (1998). *Reference form and discourse patterns*. Unpublished doctoral dissertation. Stanford University.

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 eyetracking. *Cognition*, *76*(1), https://doi.org/10.1016/s0010-0277(00)00073-1

Arnold, J. E., Brown-Schmidt, S., & Trueswell, J. (2007). Children's use of gender and order-of-mention during pronoun comprehension. *Language and Cognitive Processes*, *22*(4), 527–565. https://doi.org/10.1080/01690960600845950

Arnold, J. E. (2010). How speakers refer: The role of accessibility. *Language and Linguistics Compass*, *4*(4), 187–203. https://doi.org/10.1111/j.1749-818x.2010.00193.x

Arnold, J. E., Strangmann, I. M., Hwang, H., Zerkle, S., & Nappa, R. (2018). Linguistic experience affects pronoun interpretation. *Journal of Memory and Language*, *102*, 41–54. https://doi.org/10.1016/j.jml.2018.05.002

Brennan, S. E., Friedman, M. W., & Pollard, C. J. (1987). A centering approach to pronouns. *Proceedings of the 25th Annual Meeting on Association for Computational Linguistics*. https://doi.org/10.3115/981175.981197

Brown, R., & Fish, D. (1983). The psychological causality implicit in language. *Cognition*, *14*(3), 237–273. https://doi.org/10.1016/0010-0277(83)90006-9

Brysbaert, M., Mandera, P., & Keuleers, E. (2018). The word frequency effect in word processing: An updated review. *Current Directions in Psychological Science*, *27*(1), 45–50. https://doi-org.libproxy.lib.unc.edu/10.1177/0963721417727521

Caramazza, A., Grober, E., Garvey, C., & Yates, J. (1977). Comprehension of anaphoric pronouns. *Journal of Verbal Learning and Verbal Behavior*, *16*(5), 601–609. https://doi.org/10.1016/s0022-5371(77)80022-4

Chandler, J., & Shapiro, D. (2016). Conducting clinical research using crowdsourced convenience samples. *Annual Review of Clinical Psychology*, *12*(1), 53–81. https://doi.org/10.1146/annurev-clinpsy-021815-093623

Chmielewski, M., & Kucker, S. C. (2020). An MTurk Crisis? Shifts in Data Quality and the Impact on Study Results. *Social Psychological and Personality Science*, *11*(4) 464-473, DOI: 10.1177/1948550619875149

Contemori, C. (2021). Changing comprehenders' pronoun interpretations: Immediate and cumulative priming at the discourse level in L2 and native speakers of English. *Second Language Research*, *37*(4), 573–586. https://doi.org/10.1177/0267658319886644

Contemori, C., Mossman, S., & Ramos, A. K. (2021). Acquiring L2 pronoun interpretation biases: Implicit Learning at the discourse level in L2 English. *Studies in Second Language Acquisition*, 1-8.

Contemori, C., & Minjarez Oppenheimer, N. I. (2022). Changing pronoun interpretations across-languages: Discourse priming in Spanish–English bilingual speakers. *Bilingualism: Language and Cognition*, 1–9. https://doi.org/10.1017/s1366728922000566

Cozijn, R., Commandeur, E., Vonk, W., & Noordman, L. G. M. (2011). The time course of the use of implicit causality information in the processing of pronouns: A visual world paradigm study. *Journal of Memory and Language*, *64*(4), 381–403. https://doi.org/10.1016/j.jml.2011.01.001

Ehrlich, K. (1980). Comprehension of pronouns. *Quarterly Journal of Experimental Psychology*, *32*(2), 247–255. https://doi.org/10.1080/14640748008401161

Fernandes, E. G., Luegi, P., Correa Soares, E., de la Fuente, I., & Hemforth, B. (2018). Adaptation in pronoun resolution: Evidence from Brazilian and European Portuguese. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(12), 1986–2008. https://doi.org/10.1037/xlm0000569

Fine, A. B., & Florian Jaeger, T. (2013). Evidence for implicit learning in syntactic comprehension. *Cognitive Science*, *37*(3), 578–591. https://doi.org/10.1111/cogs.12022

Fine, A. B., Jaeger, T. F., Farmer, T. A., & Qian, T. (2013). Rapid expectation adaptation during syntactic comprehension. *PLoS ONE*, *8*(10). https://doi.org/10.1371/journal.pone.0077661

Fukumura, K., & van Gompel, R. P. (2015). Effects of order of mention and grammatical role on anaphor resolution. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(2), 501–525. https://doi.org/10.1037/xlm0000041

Guan, S., & Arnold, J. E. (2021). The predictability of implicit causes: Testing frequency and topicality explanations. *Discourse Processes*, *58*(10), 943–969. https://doi.org/10.1080/0163853x.2021.1974690

Jaeger, T. F., & Snider, N. (2008). Implicit learning and syntactic persistence: Surprisal and cumulativity. In Love, B. C., McRae, K., & Sloutsky, V. M. (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society*, (pp. 1061–1066). Cognitive Science Society Washington, DC.

Johnson, E., & Arnold, J. E. (2021). Individual differences in print exposure predict use of implicit causality in pronoun comprehension and referential prediction. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.672109

Johnson, E., Arnold, J. E. (2022). Priming discourse structure guides pronoun comprehension, *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Kaan, E., & Chun, E. (2018). Syntactic adaptation. *Psychology of Learning and Motivation*, 85–116. https://doi.org/10.1016/bs.plm.2018.08.003

Kaiser, E. (2009). Effects of anaphoric dependencies and semantic representations on pronoun interpretation. *Anaphora Processing and Applications*, 121–129. https://doi.org/10.1007/978-3-642-04975-0_10

Kehler, A., Kertz, L., Rohde, H., & Elman, J. L. (2008). Coherence and coreference revisited. *Journal of Semantics*, *25*(1), 1–44. https://doi.org/10.1093/jos/ffm018

Kehler, A., & Rohde, H. (2013). A probabilistic reconciliation of coherence-driven and centering-driven theories of pronoun interpretation. *Theoretical Linguistics*, *39*(1-2), 1–37. https://doi.org/10.1515/tl-2013-0001

Kehler, A., & Rohde, H. (2019). Prominence and coherence in a Bayesian theory of pronoun interpretation. *Journal of Pragmatics*, *154*, 63–78. https://doi.org/10.1016/j.pragma.2018.04.006

Keenan, E. L., & Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic inquiry*, *8*(1), 63-99.

Langlois, V. J., & Arnold, J. E. (2020). Print exposure explains individual differences in using syntactic but not semantic cues for pronoun comprehension. *Cognition*, *197*, 104155. https://doi.org/10.1016/j.cognition.2019.104155

Miller, J. D., Crowe, M., Weiss, B., Maples-Keller, J. L., & Lynam, D. R. (2017). Using online, crowdsourcing platforms for data collection in Personality disorder research: The example of Amazon's mechanical turk. *Personality Disorders: Theory, Research, and Treatment*, *8*(1), 26–34. https://doi.org/10.1037/per0000191

Monsell, S., Doyle, M. C., & Haggard, P. N. (1989). Effects of frequency on visual word recognition tasks: Where are they? *Journal of Experimental Psychology: General*, *118*(1), 43–71. https://doi.org/10.1037/0096-3445.118.1.43

Moore, M., & Gordon, P. C. (2014). Reading ability and print exposure: Item response theory analysis of the author Recognition Test. *Behavior Research Methods*, *47*(4), 1095–1109. https://doi.org/10.3758/s13428-014-0534-3

Nappa, R., & Arnold, J. E. (2014). The road to understanding is paved with the speaker's intentions: Cues to the speaker's attention and intentions affect pronoun comprehension. *Cognitive Psychology*, *70*, 58–81. https://doi.org/10.1016/j.cogpsych.2013.12.003

Pickering, M. J., McLean, J. F., & Branigan, H. P. (2013). Persistent structural priming and frequency effects during comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(3), 890–897. https://doi.org/10.1037/a0029181

Rohde, H., & Kehler, A. (2014). Grammatical and information-structural influences on pronoun production. *Language, Cognition and Neuroscience*, *29*(8), 912–927. https://doi.org/10.1080/01690965.2013.854918

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, *35*(4), 606–621. https://doi.org/10.1006/jmla.1996.0032

Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly*, *24*(4), 402. https://doi.org/10.2307/747605

Stassen, L. (2000). And-languages and with-languages. *Linguistic Typology*, *4*(1), 1–54. https://doi.org/10.1515/lity.2000.4.1.1

Stevenson, R. J., Crawley, R. A., & Kleinman, D. (1994). Thematic roles, focus and the representation of events. *Language and Cognitive Processes*, *9*(4), 519–548. https://doi.org/10.1080/01690969408402130

Tooley, K. M., Swaab, T. Y., Boudewyn, M. A., Zirnstein, M., & Traxler, M. J. (2014). Evidence for priming across intervening sentences during on-line sentence comprehension. *Language, Cognition and Neuroscience*, *29*(3), 289–311. https://doi.org/10.1080/01690965.2013.770892

Tooley, K. M., Pickering, M. J., & Traxler, M. J. (2019). Lexically-mediated syntactic priming effects in comprehension: Sources of facilitation. *Quarterly Journal of Experimental Psychology*, *72*(9), 2176–2196. https://doi.org/10.1177/1747021819834247

Appendix A. Stimuli

Table A1. Stimuli from Experiment 1 (she vs. he)

Exposure items for he pronouns (the she pronoun versions were created by replacing "he" with "she" and swapping the Subject- and Nonsubject-Reference conditions)	
Subject-Reference Exposure	**Nonsubject-Reference Exposure**
Will created a playlist with Ana. He danced. Did Ana dance? What did Will and Ana do?	Ana created a playlist with Will. He danced. Did Ana dance? What did Will and Ana do?
Will browsed the shelves at the library with Liz. He checked out a book. Did Will check out a book? What did Will and Liz do?	Liz browsed the shelves at the library with Will. He checked out a book. Did Will check out a book? What did Will and Liz do?
Matt pushed the shopping cart with Liz. He picked up some bread. Did Liz pick up bread? What did Matt and Liz do?	Liz pushed the shopping cart with Matt. He picked up some bread. Did Liz pick up bread? What did Matt and Liz do?
Matt bought a new video game with Ana. He played it immediately. Did Matt play the game immediately? What did Matt and Ana do?	Ana bought a new video game with Matt. He played it immediately. Did Matt play the game immediately? What did Matt and Ana do?
Will cooked some noodles with Ana. He added marinara sauce. Did Will add marinara sauce? What did Ana and Will do?	Ana cooked some noodles with Will. He added marinara sauce. Did Will add marinara sauce? What did Ana and Will do?
Matt baked a cake with Ana. He took a large slice. Did Ana take a large slice of cake? What did Ana and Matt do?	Ana baked a cake with Matt. He took a large slice. Did Ana take a large slice of cake? What did Ana and Matt do?
Will picked out a pair of gloves with Liz. He paid at the register. Did Liz pay at the register? What did Liz and Will do?	Liz picked out a pair of gloves with Will. He paid at the register. Did Liz pay at the register? What did Liz and Will do?
Matt bought tickets for an upcoming concert with Liz. He loved the band that was playing. Did Matt love the band that was playing? What did Liz and Matt do?	Liz bought tickets for an upcoming concert with Matt. He loved the band that was playing. Did Matt love the band that was playing? What did Liz and Matt do?
Matt moved the couch with Ana. He rearranged the pillows. Did Ana rearrange the pillows? What did Matt and Ana move?	Ana moved the couch with Matt. He rearranged the pillows. Did Ana rearrange the pillows? What did Matt and Ana move?
Will found seats at the movie theater with Liz. He brought some popcorn.	Liz found seats at the movie theater with Will. He brought some popcorn.

Did Liz bring popcorn? What did Will and Liz find?	Did Liz bring popcorn? What did Will and Liz find?
Matt baked some cookies with Liz. He ate them all. Did Matt eat all the cookies? What did Matt and Liz bake?	Liz baked some cookies with Matt. He ate them all. Did Matt eat all the cookies? What did Matt and Liz bake?
Will saw giant pandas at the zoo with Ana. He preferred elephants. Did Will prefer elephants? What did Will and Ana see?	Ana saw giant pandas at the zoo with Will. He preferred elephants. Did Will prefer elephants? What did Will and Ana see?
Will painted the wall with Ana. He spilled paint on the floor. Did Will spill the paint? What did Ana and Will paint?	Ana painted the wall with Will. He spilled paint on the floor. Did Will spill the paint? What did Ana and Will paint?
Will hiked up the mountain with Liz. He sat on a bench overlooking the view. Did Liz sit on a bench? What did Will and Liz hike up?	Liz hiked up the mountain with Will. He sat on a bench overlooking the view. Did Liz sit on a bench? What did Will and Liz hike up?
Matt ate french fries with Liz. He spilled ketchup on the table. Did Matt spill ketchup? What did Liz and Matt eat?	Liz ate french fries with Matt. He spilled ketchup on the table. Did Matt spill ketchup? What did Liz and Matt eat?
Matt fed the geese some bread with Ana. He was honked at for being too close. Did Ana get honked at? What did Ana and Matt feed the geese?	Ana fed the geese some bread with Matt. He was honked at for being too close. Did Ana get honked at? What did Ana and Matt feed the geese?
Will drank a glass of wine with Ana. He ordered some cheese. Did Will order cheese? What did Will and Ana do?	Ana drank a glass of wine with Will. He ordered some cheese. Did Will order cheese? What did Will and Ana do?
Matt planned a trip with Ana. He wanted to go by train. Did Ana want to go by train? What did Matt and Ana do?	Ana planned a trip with Matt. He wanted to go by train. Did Ana want to go by train? What did Matt and Ana do?
Will lit up the candles on the cake with Liz. He blew them out. Did Liz blow out the candles? What did Will and Liz do?	Liz lit up the candles on the cake with Will. He blew them out. Did Liz blow out the candles? What did Will and Liz do?
Matt drafted up a thesis with Liz. He thought the writing needed editing. Did Matt think the writing needed editing? What did Matt and Liz do?	Liz drafted up a thesis with Matt. He thought the writing needed editing. Did Matt think the writing needed editing? What did Matt and Liz do?
Matt bought a puzzle with Ana. He organized the pieces by color. Did Matt organize the puzzle pieces? What did Ana and Matt do?	Ana bought a puzzle with Matt. He organized the pieces by color. Did Matt organize the puzzle pieces? What did Ana and Matt do?

Will washed the dishes with Liz. He dried the utensils Did Will dry the utensils? What did Liz and Will do?	Liz washed the dishes with Will. He dried the utensils. Did Will dry the utensils? What did Liz and Will do?
Matt painted a beautiful picture with Liz. He framed it on the wall. Did Liz frame the picture on the wall? What did Liz and Matt do?	Liz painted a beautiful picture with Matt. He framed it on the wall. Did Liz frame the picture on the wall? What did Liz and Matt do?
Will went to a wedding with Ana. He brought some flowers. Did Ana bring flowers? What did Ana and Will do?	Ana went to a wedding with Will. He brought some flowers. Did Ana bring flowers? What did Ana and Will do?
Will made some breakfast with Ana. He wanted bacon and eggs. Did Will want bacon and eggs? What did Ana and Will make?	Ana made some breakfast with Will. He wanted bacon and eggs. Did Will want bacon and eggs? What did Ana and Will make?
Will raked the leaves into a pile with Liz. He jumped into them. Did Liz jump into the pile of leaves? What did Liz and Will rake?	Liz raked the leaves into a pile with Will. He jumped into them. Did Liz jump into the pile of leaves? What did Liz and Will rake?
Matt made sandwiches with Liz. He preferred whole-grain bread. Did Matt prefer whole-grain bread? What did Liz and Matt make?	Liz made sandwiches with Matt. He preferred whole-grain bread. Did Matt prefer whole-grain bread? What did Liz and Matt make?
Matt pet the cat with Ana. He wanted to pet the dog instead. Did Ana want to pet the dog? What did Matt and Ana pet?	Ana pet the cat with Matt. He wanted to pet the dog instead. Did Ana want to pet the dog? What did Matt and Ana pet?
Will grilled meat with Liz. He added some corn on the cob. Did Will add corn on the cob? What did Will and Liz grill?	Liz grilled meat with Will. He added some corn on the cob. Did Will add corn on the cob? What did Will and Liz grill?
Matt gave a presentation with Ana. He controlled the computer. Did Matt control the computer? What did Ana and Matt give?	Ana gave a presentation with Matt. He controlled the computer. Did Matt control the computer? What did Ana and Matt give?
Will set up a picnic in the park with Ana. He packed sandwiches. Did Ana pack sandwiches? What did Ana and Will set up in the park?	Ana set up a picnic in the park with Will. He packed sandwiches. Did Ana pack sandwiches? What did Ana and Will set up in the park?
Matt got the mail with Liz. He used scissors to open a package. Did Liz use scissors? What did Liz and Matt get?	Liz got the mail with Matt. He used scissors to open a package. Did Liz use scissors? What did Liz and Matt get?

Critical
Ana cleaned up with Liz. She swept the floor with the broom. Did Liz sweep the floor? What did Ana and Liz do?
Will had a snack with Matt. He ate some potato chips. Did Matt eat the potato chips? What did Will and Matt do?
Liz painted the wall with Ana. She used the green paint. Did Ana use the green paint? What did Liz and Ana do?
Matt went to the library with Will. He read the book on world history. Did Will read the book on world history? What did Matt and Will do?
Ana had coffee with Liz. She ordered a cappuccino. Did Liz order the cappuccino? What did Ana and Liz do?
Ana played a board game with Liz. She rolled the dice. Did Liz roll the dice? What did Ana and Liz do?
Liz assembled a bed with Ana. She used the hammer. Did Ana use the hammer? What did Liz and Ana do?
Liz planted flowers with Ana. She watered the seeds. Did Ana water the seeds? What did Liz and Ana do?
Matt had dinner with Will. He wanted some chicken. Did Will want some chicken? What did Matt and Will do?
Matt read with Will. He read the book on aliens. Did Will read the book? What did Matt and Will do?
Will went to the bookstore with Matt. He bought the new book on sailing. Did Matt buy the book on sailing? What did Will and Matt do?
Will made dinner with Matt. He asked for the salt. Did Matt ask for the salt? What did Will and Matt do?

Table A2. Stimuli from Experiment 2 (transfer vs. joint-action)

Filler	
Subject-Reference Exposure	**Nonsubject-Reference Exposure**
Ana loaned a backpack to Matt and then she grabbed some snacks. Did Ana grab any snacks? Who did Ana loan a backpack to?	Ana loaned a backpack to Matt and then he filled it with snacks. Did Ana fill the backpack with snacks? Who did Ana loan a backpack to?
Ana threw the ball to Will and then she shielded her eyes. Did Will shield his eyes? What were Ana and Will doing?	Ana threw the ball to Will and then he tagged the runner out. Did Will tag the runner out? What were Ana and Will doing?
Matt handed the salt to Ana and then he made a salad. Did Matt make a salad? When did Matt hand the salt to Ana?	Matt handed the salt to Ana and then he seasoned the chicken. Did Matt season the chicken? When did Matt hand the salt to Ana?
Will brought the skates to Ana and then he put his gloves on. Did Ana put gloves on? What did Ana borrow?	Will brought the skates to Ana and then she put them on. Did Ana put the skates on? What did Ana borrow?
Liz passed the rake to Matt and then she jumped in the pile of leaves. Did Liz jump in the pile of leaves? Who did Liz pass the rake to?	Liz passed the rake to Matt and then he raked the entire front yard. Did Liz rake the entire front yard? Who did Liz pass the rake to?
Liz sent an article to Will and then she started working on an outline. Did Will start work on an outline? What were Liz and Will doing?	Liz sent an article to Will and then he took notes on it. Did Will take notes on the article? What were Liz and Will doing?
Liz snatched the frisbee from Matt and then she slipped on some mud. Did Liz slip on some mud? What were Liz and Matt playing?	Liz snatched the frisbee from Matt And then he decided not to play anymore. Did Liz decide not to play anymore? What were Liz and Matt playing?
Liz received a note from Will and then she giggled. Did Will giggle? When were Liz and Will passing notes?	Liz received a note from Will and then he looked the other way. Did Will look the other way? When were Liz and Will passing notes?
Matt retrieved a book from Liz and then he took notes. Did Matt take notes? Where were Matt and Liz studying?	Matt retrieved a book from Liz and then she took a break from studying. Did Matt take a break from studying? Where were Matt and Liz studying?
Will took the fries from Liz and then he ate them all up. Did Liz eat all the fries? Where were Will and Liz eating?	Will took the fries from Liz and then she wiped her hands with a napkin. Did Liz wipe her hands with a napkin? Where were Will and Liz eating?

Ana got the ball from Matt and then she made a basket. Did Ana make a basket? What were Ana and Matt doing?	Ana got the ball from Matt and then he ran down the court. Did Ana run down the court? What were Ana and Matt doing?
Ana accepted a ride from Will and then she offered to pay for parking. Did Will offer to pay for parking? Who did Ana accept a ride from?	Ana accepted a ride from Will and then he paid for parking, too. Did Will pay for parking? Who did Ana accept a ride from?
Matt mailed the Christmas card to Liz and then he bought more stamps. Did Matt buy more stamps? Who did Matt mail the Christmas card to?	Matt mailed the Christmas card to Liz and then she read it. Did Matt read the Christmas card? Who did Matt mail the Christmas card to?
Will gave the credit card to Liz and then he browsed the magazines. Did Liz browse the magazines? What did Will give to Liz?	Will gave the credit card to Liz and then she got in line to check out. Did Liz get in the line to check out? What did Will give to Liz?
Ana offered the broom to Matt and then she tidied up the kitchen. Did Matt tidy up the kitchen? What did Ana offer?	Ana offered the broom to Matt and then he tidied up the kitchen. Did Matt tidy up the kitchen? What did Ana offer?
Ana loaned a novel to Will and then she organized her comic books. Did Will organize his comic books? What were Ana and Will doing?	Ana loaned a novel to Will and then he read it. Did Will read the novel? What were Ana and Will doing?
Matt brought the ball to Ana and then he pet the dog. Did Matt pet the dog? When did Matt bring the ball to Ana?	Matt brought the ball to Ana and then she threw the ball to the dog. Did Matt throw the ball to the dog? When did Matt bring the ball to Ana?
Will offered some wine to Ana and then he tried the appetizer. Did Ana try the appetizer? Where were Will and Ana having dinner?	Will offered some wine to Ana and then she took a sip. Did Ana take a sip? Where were Will and Ana having dinner?
Liz gave a stack of dishes to Matt and then she swept the floor. Did Liz sweep the floor? When were Liz and Matt cleaning up?	Liz gave a stack of dishes to Matt and then he put them in the dishwasher. Did Liz put the dishes in the dishwasher? When were Liz and Matt cleaning up?
Liz passed the soap to Will and then she turned on the hose. Did Liz turn on the hose? What were Liz and Will cleaning?	Liz passed the soap to Will and then he soaped up the car. Did Liz soap up the car? What were Liz and Will cleaning?
Matt accepted the gift from Ana and then he unwrapped it. Did Matt unwrap the gift? Who did Matt accept the gift from?	Matt accepted the gift from Ana and then she smiled. Did Matt smile? Who did Matt accept the gift from?
Will snatched the guest list from Ana and then he added ten friends to the list.	Will snatched the guest list from Ana and then she refused to help any further.

Did Ana add ten friends to the guest list? What were Will and Ana doing?	Did Ana refuse to help any more? What were Will and Ana doing?
Liz borrowed a hat from Will and then she put it on. Did Will put the hat on? Who did Liz borrow a hat from?	Liz borrowed a hat from Will and then he put on his jacket. Did Will put on his jacket? Who did Liz borrow a hat from?
Liz got the instructions from Will and then she read them out loud. Did Liz read the instructions out loud? What did Liz get?	Liz got the instructions from Will and then he laid all the pieces on the floor. Did Liz lay all the pieces on the floor? What did Liz get?
Matt received a check from Liz and then he paid the landlord. Did Matt pay the landlord? What were Matt and Liz doing?	Matt received a check from Liz and then she checked her bank account. Did Matt check his bank account? What were Matt and Liz doing?
Will took the remote from Liz and then he changed the channel. Did Will change the channel? Who did Will take the remote from?	Will took the remote from Liz and then she went to get a beer. Did Will go to get a beer? Who did Will take the remote from?
Ana borrowed the book from Matt and then she looked up a reference. Did Matt look up a reference? Who did Ana borrow the book from?	Ana borrowed the book from Matt and then he turned his essay in. Did Matt turn his essay in? Who did Ana borrow the book from?
Ana retrieved the fishing pole from Will and then she prepared the bait. Did Will prepare the bait? Where were Ana and Will fishing?	Ana retrieved the fishing pole from Will and then he got a drink from the cooler. Did Will get the drink from the cooler? Where were Ana and Will fishing?
Critical	
Ana brought the coupons to Liz and then she wrote a shopping list. Did Liz write a shopping list? What did Ana bring to Liz?	
Ana sent a text to Liz and then she took a screenshot. Did Liz take a screenshot? Who did Ana send a text to?	
Liz offered the tickets to Ana and then she waited in line. Did Ana wait in line? Where were Liz and Ana going?	
Liz gave the notebook to Ana and then she wrote the results section. Did Ana write the results section? Who did Liz give the notebook to?	
Matt handed the flowers to Will and then he walked up to the counter. Did Will walk up to the counter? What did Matt hand to Will?	
Matt threw the phone to Will and then he wrote down the directions. Did Will write down the directions? Who did Matt throw the phone to?	
Will loaned a spare key to Matt and then he took the bus home.	

Did Matt take the bus home? Where were Will and Matt arriving?
Will passed the popcorn to Matt and then he drank some soda. Did Matt drink some soda? When did Will pass the popcorn to Matt?
Ana had coffee with Liz and then she ordered a sandwich. Did Liz order a sandwich? When did Ana have coffee with Liz?
Ana played a board game with Liz and then she rolled the dice. Did Liz roll the dice? Where were Ana and Liz sitting?
Liz assembled a bed with Ana and then she bought some new sheets. Did Ana buy some new sheets? What were Liz and Ana doing?
Liz planted flowers with Ana and then she took a picture. Did Ana take a picture? Who did Liz plant the flowers with?
Matt had dinner with Will and then he wanted some snacks. Did Will want some snacks? Where were Matt and Will meeting?
Matt went to the library with Will and then he read a book on world history. Did Will read the book on world history? Where was Will reading?
Will read with Matt and then he wrote a summary. Did Matt write a summary? Where was Will reading?
Will had a snack with Matt and then he cleaned up the table. Did Matt clean up the table? What were Will and Matt doing?

Table A3. Stimuli from Experiment 3 (joint-action vs. transfer)

Filler	
Subject-Reference Exposure	**Nonsubject-Reference Exposure**
Ana baked a cake with Matt and then she took a large slice. Did Ana take a large slice of cake? Who did Ana bake a cake with?	Ana baked a cake with Matt and then he took a large slice. Did Ana take a large slice of cake? Who did Ana bake a cake with?
Ana twirled around with Will and then she collapsed in exhaustion. Did Will collapse in exhaustion? What were Ana and Will doing?	Ana twirled around with Will and then he collapsed in exhaustion. Did Will collapse in exhaustion? What were Ana and Will doing?
Matt brought out a puzzle with Ana and then he organized the pieces by color. Did Matt organize the puzzle pieces by color? What did Matt bring out?	Matt brought out a puzzle with Ana and then she organized the pieces by color. Did Matt organize the puzzle pieces by color? What did Matt bring out?
Will set up a picnic in the park with Ana and then he ate some sandwiches. Did Ana eat sandwiches? What did Ana do in the park?	Will set up a picnic in the park with Ana and then she ate some sandwiches. Did Ana eat sandwiches? What did Ana do in the park?
Liz painted a beautiful picture with Matt and then she framed it. Did Liz frame the picture on the wall? Who did Liz paint a picture with?	Liz painted a beautiful picture with Matt and then he framed it. Did Liz frame the picture on the wall? Who did Liz paint a picture with?
Liz bought a cake with Will and then she put it in the fridge. Did Will put the cake in the fridge? What were Liz and Will doing?	Liz bought a cake with Will and then he put it in the fridge. Did Will put the cake in the fridge? What were Liz and Will doing?
Liz ate french fries with Matt and then she spilled ketchup on the table. Did Liz spill ketchup on the table? What were Liz and Matt eating?	Liz ate french fries with Matt and then he spilled ketchup on the table. Did Liz spill ketchup on the table? What were Liz and Matt eating?
Liz grilled meat with Will and then she added some corn on the cob. Did Will add some corn on the cob? When did Liz and Will grill meat?	Liz grilled meat with Will and then he added some corn on the cob. Did Will add some corn on the cob? When did Liz and Will grill meat?
Matt drafted up a thesis with Liz and then he thought the writing needed editing. Did Matt think the writing needed editing? What were Matt and Liz doing?	Matt drafted up a thesis with Liz and then she thought the writing needed editing. Did Matt think the writing needed editing? What were Matt and Liz doing?
Will scored a goal with Liz and then he yelled in excitement. Did Liz yell in excitement? Where were Will and Liz practicing shooting?	Will scored a goal with Liz and then she yelled in excitement. Did Liz yell in excitement? Where were Will and Liz practicing shooting?
Ana planned a trip with Matt and then she bought the train tickets.	Ana planned a trip with Matt and then he bought the train tickets.

Did Ana buy the train tickets? What were Ana and Matt doing?	Did Ana buy the train tickets? What were Ana and Matt doing?
Ana saw giant pandas with Will and then she took a picture. Did Will take a picture of the pandas? Who did Ana see giant pandas with?	Ana saw giant pandas with Will and then he took a picture. Did Will take a picture of the pandas? Who did Ana see giant pandas with?
Matt picked up the mail with Liz and then he used scissors to open a package. Did Matt use scissors to open a package? Who did Matt pick up the mail with?	Matt picked up the mail with Liz and then she used scissors to open a package. Did Matt use scissors to open a package? Who did Matt pick up the mail with?
Will washed the dishes with Liz and then he dried the utensils. Did Liz dry the utensils? What were Will and Liz doing?	Will washed the dishes with Liz and then she dried the utensils. Did Liz dry the utensils? What were Will and Liz doing?
Ana described the study with Matt and then she summed up the main points. Did Matt sum up the main points? What were Ana and Matt doing?	Ana described the study with Matt and then he summed up the main points. Did Matt sum up the main points? What were Ana and Matt doing?
Ana looked for a novel with Will and then she started reading the comics. Did Will start reading the comics? What were Ana and Will doing?	Ana looked for a novel with Will and then he started reading the comics. Did Will start reading the comics? What were Ana and Will doing?
Matt walked the dog with Ana and then he unleashed the dog. Did Matt unleash the dog? Where were Matt and Ana going?	Matt walked the dog with Ana and then she unleashed the dog. Did Matt unleash the dog? Where were Matt and Ana going?
Will drank some wine with Ana and then he tried the appetizer. Did Ana try the appetizer? Where were Will and Ana having dinner?	Will drank some wine with Ana and then she tried the appetizer. Did Ana try the appetizer? Where were Will and Ana having dinner?
Liz picked out a pair of gloves with Will and then she paid at the register. Did Liz pay at the register? What were Liz and Matt shopping for?	Liz picked out a pair of gloves with Will and then he paid at the register. Did Liz pay at the register? What were Liz and Matt shopping for?
Liz soaped up the car with Will and then she turned on the hose. Did Liz turn on the hose? What were Liz and Will cleaning?	Liz soaped up the car with Will and then he turned on the hose. Did Liz turn on the hose? What were Liz and Will cleaning?
Matt pet the cat with Ana and then he pet the dog. Did Matt pet the dog? Who did Matt pet the cat with?	Matt pet the cat with Ana and then she pet the dog. Did Matt pet the dog? Who did Matt pet the cat with?
Will made the guest list with Ana and then he added ten friends to the list. Did Ana add ten friends to the guest list? What were Will and Ana doing?	Will made the guest list with Ana and then she added ten friends to the list. Did Ana add ten friends to the guest list? What were Will and Ana doing?

Liz baked a pie with Will and then she put it in a box. Did Will put the pie in a box? Who did Liz bake a pie with?	Liz baked a pie with Will and then he put it in a box. Did Will put the pie in a box? Who did Liz bake a pie with?
Liz read the instructions with Will and then she wrote their names on the paper. Did Liz write their names on the paper? What were Liz and Will doing?	Matt wrote a check with Liz and then she paid the landlord. Did Matt pay the landlord? What were Matt and Liz doing?
Matt wrote a check with Liz and then he paid the landlord. Did Matt pay the landlord? What were Matt and Liz doing?	Matt wrote a check with Liz and then she paid the landlord. Did Matt pay the landlord? What were Matt and Liz doing?
Will found a funny channel with Liz and then he dropped the remote. Did Will drop the remote? Who did Will find a channel with?	Will found a funny channel with Liz and then she dropped the remote. Did Will drop the remote? Who did Will find a channel with?
Ana stirred up the paint with Matt and then she spilled paint on the floor. Did Matt spill paint on the floor? Who did Ana stir up the paint with?	Ana stirred up the paint with Matt and then he spilled paint on the floor. Did Matt spill paint on the floor? Who did Ana stir up the paint with?
Ana bought some flowers with Will and then she wrote a card. Did Will write a card? What were Ana and Will attending?	Ana bought some flowers with Will and then he wrote a card. Did Will write a card? What were Ana and Will attending?
Critical	
Ana brought the coupons to Liz and then she wrote a shopping list. Did Liz write a shopping list? What did Ana bring to Liz?	
Ana sent a text to Liz and then she took a screenshot. Did Liz take a screenshot? Who did Ana send a text to?	
Liz offered the tickets to Ana and then she waited in line. Did Ana wait in line? Where were Liz and Ana going?	
Liz gave the notebook to Ana and then she wrote the results section. Did Ana write the results section? Who did Liz give the notebook to?	
Matt handed the flowers to Will and then he walked up to the counter. Did Will walk up to the counter? What did Matt hand to Will?	
Matt threw the phone to Will and then he wrote down the directions. Did Will write down the directions? Who did Matt throw the phone to?	
Will loaned a spare key to Matt and then he took the bus home. Did Matt take the bus home? Where were Will and Matt arriving?	

Will passed the popcorn to Matt and then he drank some soda. Did Matt drink some soda? When did Will pass the popcorn to Matt?
Ana had sandwiches with Liz and then she ordered dessert. Did Liz order dessert? When did Ana have sandwiches with Liz?
Ana played a board game with Liz and then she turned on the TV. Did Liz turn on the TV? Where were Ana and Liz sitting?
Liz assembled a bed with Ana and then she bought some new sheets. Did Ana buy some new sheets? What were Liz and Ana doing?
Liz planted flowers with Ana and then she took a picture. Did Ana take a picture? Who did Liz plant the flowers with?
Matt had dinner with Will and then he wanted to go out for drinks. Did Will want to go out for drinks? Where were Matt and Will meeting?
Matt went to the library with Will and then he read a book on world history. Did Will read the book on world history? What were Matt and Will doing?
Will read with Matt and then he wrote a summary. Did Matt write a summary? Where was Will reading?
Will had a snack with Matt and then he cleaned up the table. Did Matt clean up the table? What were Will and Matt doing?

Appendix B

This study collected data on Amazon Mechanical Turk. The viability of crowd-sourced data is contingent on using strict inclusion criteria, given that some participants may not genuinely attend to the task, or may even be computer programs (Chimielwski & Kucker, 2020). This study follows the same protocol used for numerous studies in our lab, combining several exclusion criteria.

One of our criteria comes from the Author Recognition Task (ART), which asks participants to identify the real authors they know out of a list of real and fake authors. This is a proxy measure of print exposure (Stanovich & West 1989), and tends to correlate with both reading skill (Moore and Gordon, 2015) and spoken language processing (Arnold et al., 2018). The instructions tell people not to guess, so when participants select a high number of fake authors, it suggests that they are not reading the directions. We calculate the percentage of total responses that are fake authors; if this number is greater than 33% of the total number of responses selected, we judge the responder to be a guesser. The instructions ask the participant to read all the author names; for this reason the button to proceed to the next item does not appear for 60 seconds, but this information is not shared with the participants. If participants are clicking randomly and not reading, this constraint may lead some participants to continue clicking on additional names. Other reasons for excluding data from analysis are described in the manuscript, and include a) being a non-native speaker of English; b) below 75% accuracy on the unambiguous filler questions; c) reporting a language disorder, or d) duplicate submission.

Typically when this task is used with in-lab participants, the rate of guessing is very low. By contrast, with Amazon Mechancial Turk participants, our guessing metric correlates with other indicators that the responder is not attending. For example, in Experiment 1 there were a total of 165 participants who began the study, but only 91 finished. Of the other 74, 38 exited the task before the ART task. The rate of ART guessing is much higher for the 36 people who did the ART but then didn't finish for some other reason (72%) than for the 91 finishers (33%).

The same story emerges when we look at the rate of exclusion for non-ART-guessing reasons: of the 127 people who did the ART (both finishers and nonfinishers), those who did not guess on the ART were excluded for some other reason only 30% of the time, whereas those who did guess were excluded for some other reason 70% of the time. Thus, guessing on the ART correlates with other indicators of unreliable performance.

Nevertheless, one concern is that a very high number of participants are excluded for guessing on the ART, and that incorrect answers may in some cases reflect the participant's true print exposure in cases where knowledge of authors is very low. For this reason we re-ran each of our analyses with the entire set of participants, including people who guessed on the ART but were not excluded for other reasons (language disorder, inaccuracy, nonnative status, duplicate submission). These results are reported in Tables B1-B3.

Table B1. Inferential Statistics from Experiment 1 including ART guessers who were not excluded for some other reason (N = 77).

Effect	Estimate (St. Error)	t Value	p
Intercept	0.9 (0.33)	2.75	0.0078
Subject vs. Nonsubject Exposure	2.00 (0.63)	3.18	0.0023
Verb Match vs. Mismatch	-0.22 (0.24)	-0.89	0.3759
Exposure x Match	-0.25 (0.48)	-0.52	0.6016

Table B2. Inferential Statistics from Experiment 2 including ART guessers who were not excluded for some other reason (N = 137).

Effect	Estimate (St. Error)	t Value	p
Intercept	1.16 (0.22)	5.44	<.0001
Subject vs. Nonsubject Exposure	2.01 (0.36)	5.57	<.0001
Verb Match vs. Mismatch	-0.90 (0.29)	-3.14	0.0063
Exposure x Match	1.02 (0.30)	3.42	0.0024

Table B3. Inferential Statistics from Experiment 3 including ART guessers who were not excluded for some other reason (N = 136).

Effect	Estimate (St. Error)	t Value	p
Intercept	1.06 (0.27)	3.89	0.0004
Subject vs. Nonsubject Exposure	2.22 (0.41)	5.45	<.0001
Verb Match vs. Mismatch	1.23 (0.40)	3.10	0.0074
Exposure x Match	-0.85 (0.30)	-2.84	0.0054