Experiment. Given previous discussion about Lau and Liao (2018) and following work, the semantic contribution of the items is need to be discussed. Even though there are multitude of theoretical discussions on the nature of bare plurals in English, there is no clear agreement on the nature of their mental or behavioral representations.

The present study tested people's a priori tendency to complete sentence fragments that starts with bare plurals in an experimental environment where there is no preceding context. In addition to their base tendency, the present study aimed to test how adjectival modification, coordination, or definite marking affects their base tendency.

To this end, we use a sentence completion task with no forced choice. Participants are asked to provide an appropriate following to sentence fragments that consist of bare plurals, definite plurals, their modified and/or coordinated forms, as well as various different structures as fillers, such as wh-phrases, if clauses, proper names, and adjuncts.

Methods. The experiment consisted of two parts, the first part is the preamble completion part. Participants were provided a sentence fragment and were asked to provide an appropriate completion. In the second part of the experiment, participants were asked to judge sentences using a Likert scale(1-5). The main rationale behind including a judgment task was to make participants be more aware of their completions.

Materials.

Completion Task The first part of the experiment compromised 2 x 2 x 2 design (structure: coordination/list x determiner: bare/definite x modification: modified/not-modified). In total, 4 trials were presented in each of the 8 conditions. Half of the trials contained two noun phrases both of which sharing the same structure. The other half of the experimental trials contained a single noun phrase. All experimental items consisted plural nouns. Example experimental item (1) shows the modified coordinated definite plural condition, and Table 1 shows all manipulations.

(1) The steamy mirrors and blissful journeys ...

Table 1: All experimental conditions.

Condition	Example
Bare Plural	Mirrors
Modified Bare Plural	Steamy mirrors
Coordinated Bare Plurals	Mirrors and journeys
Modified and Coordinated Bare Plurals	Steamy mirrors and blissful journeys
Definite Plural	The mirrors
Definite Modified Plural	The steamy mirrors
Coordinated Definite Plurals	The mirrors and journeys
Modified and Coordinated Definite Plurals	The steamy mirrors and blissful journeys

Noun phrases and adjectives were selected from the original study by Lau and Liao (2018). Noun phrases were selected randomly, and the same adjectives from the original study was used for those noun phrases. Similarly, the coordinands were also kept constant between this study and the original study by Lau and Liao (2018).

To preclude participants from limiting themselves to a small set of context, we included 64 filler sentences. Fillers consisted of structures that are different from the experimental items. They can be categorized into 9 different groups as shown in Table 2.

Table 2: All filler types, examples, and counts.

Filler Type	Example	Count
Relative Clause	The book that kept me up all night	10
Proper Names	The dynamic duo of Lucy and Max	10

Filler Type	Example	Count
Adjuncts	During a starry summer night	10
Wh-phrases	What to do after	10
Quantified NPs	Most challenges	5
Modified Quantified NPs	Few precious chances, which had eluded them for so long,	5
Conditionals	If the world was ruled by cats,	6
Inverted Conditionals	Were I to open a bakery shop,	4
Correlatives	Wherever the party was	4
		Total: 64

Within the whole experiment, the number of preambles that may serve as a plural head was 0.53% of the total items.

Judgment Task   The second part of the experiment consisted two types of sentences: (i) completed versions of the sentences in the first part and (ii) sentences from Frazier et al. (2007) study. Sentences from the first part of the experiment was acquired through a pilot study carried out voluntarily by University of Maryland graduate students. We selected 22 filler sentences and 32 experimental sentences from the pilot study, covering all experimental conditions (except definite ones) and filler types. From Frazier et al. (2007) study, we included 20 items using only two conditions from the experiment, which used bare plurals in the subject positions. We manipulated whether the first sentence is an episodic (2a) or a generic sentence (2b).

(2)   a.   John noticed a cat on my roof last night, cats are skilled climbers.

      b.   Cats are skilled climbers, John noticed a cat on my roof last night.

Procedure.

Completion Task   The experiment was run online in PCIbex platform (Drummond, 2013; Zehr & Schwarz, 2018). Each participant session took approximately 25 minutes. Participants provided demographic information and gave informed consent to participate in the experiment. A practice and instructions sessions preceded the experiment, in which participants were provided with a set of sentence fragments. Instruction sections included a comprehension question to confirm that they understood the task. This section also included two trials with bare plural noun phrases as sentence fragments, after they were probed to complete them, they were also provided other possible completions which featured both existential and generic examples. In this section, they were told to not repeat themselves while completing sentences and provide vivid completions. They were also instructed about the character (15) and time limit (25 seconds). In the practice section, they were asked to complete 7 sentence fragments.

Each trial began with 200 ms blank screen, and followed by a full sentence presentation in the center of the screen that stayed in the screen for 25 seconds. In this time period, they were expected to write the sentence completiong and press ENTER key. During the practice item, a warning message in red font appeared if they did not respond within 25 seconds or write more than 15 characters, including whitespaces. Sentence fragments and instructions were presented in black 16-point case Helvetica font on a white background.

Items were randomly paired into 8 different lists according to a balanced Latin-square design (Bradley, 1958). Balanced Latin-square made sure that each item would appear in either one of the condition and not any other for a participant. It also removes the immediate carry-over effects: A condition will precede another condition exactly once. The items in these lists were shuffled as well.

Judgment Task The experiment was run online in PCIbex platform (Drummond, 2013; Zehr & Schwarz, 2018). Each participant session took approximately 25 minutes. Participants were directed from the previous task, and greeted by a welcome screen. Following the initial screen, the instruction and the practice screen were presented, in which participants judge **10?** sentences. In the instructions screen, participants were told to judge sentences as fast and accurately as possible. They were also instructed about time limit (25 seconds).

Each trial began with 200 ms blank screen, and followed by a asterisk in the middle of the screen as a fixation point, that stayed in the screen for 500 ms and presented in black 16-point case Helvetica font on a white background.

Following blank screen and fixation point, the full sentence was presented in the center of the screen and stayed on the screen for 25 seconds. In this time period, they were expected to read the sentence and judge it using the Likert (1-7) scale presented in the screen. During the practice item, a warning message in red font appeared if they did not respond within 25 seconds. Sentences and instructions were presented in black 16-point case Helvetica font on a white background.

Participants. A total of 40 participants will be recruited through Prolific. They will be paid **5?** US dollars. Written informed consent was obtained from all participants. All participants are intented to be English native speakers from United States, and participants that are not native English speakers will be excluded. Participants who answered the comprehension question in the instructions wrongly will be excluded as well. In order to detect bots, we included a dialect recognition task following Maher (2023), in which participants will listen an audio recording and select the dialect used in the audio. Participants that failed to correctly recognize the dialect task will be exclude, since they will be treated as bots. Participants that complete sentence in an inappropriate or repeating way, such as completions in the same manner all the time or completions as random strings, will also be excluded.

References.

Bradley, J. V. (1958). Complete counterbalancing of immediate sequential effects in a latin square design. *Journal of the American Statistical Association*, *53*(282), 525–528.

Drummond, A. (2013). *Ibex farm* [https://spellout.net/ibexfarm].

Frazier, L., Majewski, H., Hirotani, M., & Rayner, K. (2007). Processing bare plurals and indefinites: Evidence from eye movements. *University of Massachusetts Occasional Papers in Linguistics*, *33*(1), 4.

Lau, E., & Liao, C.-H. (2018). Linguistic structure across time: Erp responses to coordinated and uncoordinated noun phrases. *Language, Cognition and Neuroscience*, *33*(5), 633–647.

Maher, Z. (2023). *Knowledge and processing of morphosyntactic variation in African American Language and Mainstrean American English* [Doctoral dissertation, University of Maryland, College Park].

Zehr, J., & Schwarz, F. (2018). PennController for Internet Based Experiments (IBEX).