Le, Khuyen N. & Bale, Alan & Barner, David. 2026. Object-mass nouns specify individuation lexically: Evidence from English and French. *Glossa: a journal of general linguistics* 11(1). pp. 1–25. DOI: https://doi.org/10.16995/glossa.26380

Open Library of Humanities

Object-mass nouns specify individuation lexically: Evidence from English and French

Khuyen N. Le, Department of Psychology, University of California, San Diego, knl005@ucsd.edu

Alan Bale, Department of Classics, Modern Languages and Linguistics, Concordia University, alancbale@gmail.com

David Barner, Department of Psychology, University of California, San Diego, barner@ucsd.edu

In many languages, words in count syntax quantify over countable individuals (e.g., *too many strings*), while mass nouns often don't (e.g., *too much string*). Theories differ in how they characterize nouns that violate this pattern, such as object-mass nouns (e.g. *furniture, clothing*). These nouns exhibit mass syntax, but often quantify by number (Barner & Snedeker 2005). On one hypothesis, the individuation of object-mass nouns is lexically specified, making them semantically like count nouns in that they induce a comparison by number (Bale & Barner 2009). Others argue that, while count nouns always quantify by number, object-mass nouns have different quantification criteria depending on context (Rothstein 2010), including function fulfillment (McCawley 1975). We evaluated these hypotheses by comparing English quantity judgments for object-mass nouns to (1) English judgments for collective count nouns, and (2) French judgments for translations of object-mass nouns. In each case, we found that object-mass nouns behaved like count nouns, and were no more susceptible to contextual effects. These findings support the view that object-mass nouns and count nouns specify individuation to the same extent. We highlight implications for theories about how the mass–count distinction is acquired, and how it arises in languages.

1 Introduction

How does the syntax of a noun phrase impact its meaning? Many languages, including English, make a formal distinction between mass and count syntax. In English, nouns typically used in count syntax (e.g., *a cat; a table*) can be pluralized and modified by cardinal numbers (e.g., *one cat; two cats*) or quasi-cardinal determiners (e.g., *several cats; many cats*). In contrast, nouns typically used in mass syntax (e.g., *some sand; some water*) can't be pluralized without a change in meaning, cannot be modified by cardinal or quasi-cardinal expressions, and instead are modified by terms like little or much (Jespersen 1909; see Grimm & Wahlang 2021 for a more recent analysis). While these grammatical properties of the mass–count distinction are widely agreed upon, its semantic interpretation remains less certain.

According to many prominent proposals, beginning with Quine (1960), but adopted with variations by many linguists (Link 1983; 1998; Bunt 1985; Barker 1998) and psycholinguists (Macnamara 1972; Gordon 1985; Bloom 1994a; 1999; Wisniewski et al. 1996), the mass–count distinction corresponds to a contrast in how nouns quantify over elements in their domain. According to Quine, count nouns, but not mass nouns, "possess built-in modes … of dividing their reference" (1960: 91). Consequently, the denotations of count nouns have atomic and non-divisible minimal parts, or individuals, i.e., members whose sub-portions do not count as instances of the noun (e.g., the leg of a table isn't a *table*). These individuals provide a basis for counting and comparison (e.g., having more tables means having more of these individuals). On this type of account, only mass nouns should allow for "cumulativity of reference", such that combining two referents of a mass noun results in an instance of the same mass noun (e.g., combining water results in *water*), whereas two referents of a count noun cannot be labeled by a singular noun. Relatedly, only mass nouns should permit "divisity of reference" (Cheng 1973), where any part or portion of a referent of a mass noun counts as an instance of the same mass noun (e.g., any sub-portion of sand is *sand*). As a result, there are no atomic minimal parts in the denotation of mass nouns, and therefore no individuals to count or compare.

Against this broad approach, others have argued that there isn't a perfect correlation between cumulativity, divisity, and mass–count syntax (see Gillon 1992; 1999). While many counterexamples are discussed in the literature, most critical to the present study is the claim that some mass nouns, like *furniture*, resist divisity. For example, it is not the case that any part or portion of something labeled as *furniture* also counts as furniture (e.g., the leg of a sofa is not furniture, but is a part of a piece of furniture). Similarly, other so-called "object-mass" nouns (Barner & Snedeker 2005) like *cutlery*, *luggage*, *clothing*, etc. do not appear to allow arbitrary divisity (see also Mihatsch 2016; Erbach et al. 2021).

Observations such as these led Gillon (1992; 1999) to conclude that, while count nouns are grammatically specified to denote a set with atomic minimal parts, the denotations of mass

nouns are linguistically unspecified. According to this hypothesis, mass nouns denote sets with minimal parts if there is evidence in the world that there are linguistically accessible atomic individuals in their denotations (Bunt 1985; Gillon 1992; 1999; see also Katz 1970). In a related proposal, Chierchia (1998; 2010) argued that all mass nouns have minimal parts, but that only object-mass nouns have "stable" atoms like count nouns, such that their denotations can also be compared via counting. On this view, differences in syntax across languages should not affect how a noun quantifies—"Pavarotti's hair is Pavarotti's hair, whether we talk about it in Italian or in English, i.e. whether we get at it through a mass noun or through a count noun. If we don't want semantics to start looking like magic, we have to say that in the real world "hair" and "capello" obviously denote the same stuff." (Chierchia 1998: 88). Meanwhile, whereas nouns like *furniture* denote stable atoms, the minimal parts of other mass nouns, like *mustard*, are only vaguely specified and vary from context to context, making it difficult to count these parts and evaluate them according to number (for related proposals that distinguish between object-mass and substance-mass nouns in terms of the presence or absence of (stable) atomic parts, see Sutton & Filip 2018; Rothstein 2021; Erbach & Schoenfeld 2022).

In an effort to evaluate these claims, Barner and Snedeker (2005) asked adults and 4-year-old children to make quantity judgments for items denoted by object-mass nouns (e.g., *furniture, jewelry*), count nouns (e.g., *shoes, cups*), and substance-mass nouns (e.g., *toothpaste, ketchup*) that differed by number of individuals/portions and summed volume. For example, participants were asked to judge who had "more furniture" between a character who had two large tables and two large chairs, and another who had a greater number of things, but that had a smaller summed volume—e.g., three small tables and three small chairs. For object-mass nouns, both children and adults preferred the character with the greater number of individuals over the character who had the greater volume. Critically, participants made similar judgments for count nouns, but based judgments on volume for substance-mass nouns. Further, these judgments were not based purely on the availability of discrete physical objects in the world (c.f., Gillon 1992; 1999). When presented with items like stones or strings, participants based quantity judgments on number when the questions were phrased using count syntax (e.g., "Who has more strings?") but on mass or volume when phrased with mass syntax (e.g., "Who has more string?"), despite the fact that physical things were salient in both contexts, and the number of atomic particles of string, stone, etc. was constant. Subsequent studies have replicated these findings in English and found similar results in other languages (Barner et al. 2009; Inagaki & Barner 2009; Van Witteloostujin 2013; Inagaki 2014; Van Witteloostuijn & Schaeffer 2014; Hacohen & Schaeffer 2016; Scontras et al. 2017; Lin & Schaeffer 2018; Mohr & Agyepong 2022), with novel mass and count nouns (Barner & Snedeker 2006), and with event-denoting deverbal object-mass nouns (Barner et al. 2008; see Bale & Barner 2018, for review).

Such results suggest an asymmetry between mass and count nouns, such that count nouns receive a uniform semantics, whereas mass nouns do not—some mass nouns individuate while others do not (e.g., see Bale & Barner 2009). Given this, several proposals have sought to preserve symmetrical accounts by arguing that object-mass nouns may actually quantify continuously after all. On one such account, originally described by McCawley (1975), nouns like *furniture* do not denote sets of individuals, but instead represent the function that such things fulfill, such that the denotation of *furniture* is measured and quantified by the degree to which the function of furnishing is fulfilled. As an example, McCawley argues that, "if Fred has 4 chairs, 3 magazine racks, 2 coffee tables, and 1 lamp, and I have 2 chairs, 1 desk, 1 bed, 1 sofa, and 1 table, my 6 pieces of furniture would constitute more furniture than Fred's 10 pieces do" (p. 319). This is because "his" furniture better fulfills the function of "furnishing a space" compared to Fred's (see also Prasada 1999).

In a related proposal, Rothstein (2010; 2017) argued that although the individuated reading of object-mass noun comparisons is the most salient, this meaning is not lexically encoded and may shift contextually (see also Hampton & Winter 2024; Rothstein & de Oliveira 2020). For example, in a context where two people are moving their furniture into a new place and John needs to move a grand piano, a large sofa, a double bed and a heavy wardrobe (4 pieces), while Bill needs to move four folding chairs, a small table, and a rolled-up mattress (6 pieces), people should judge that John has more furniture, hence basing their judgment on volume of furniture rather than number (Rothstein 2017: 122). Thus, whereas the individuation of count nouns is grammatically specified, the individuation of mass nouns is not, and is more strongly determined by context.

Several studies have attempted to test these ideas by manipulating function fulfillment. To do so, they presented participants with scenes that varied both the number of sub-kinds (e.g., chairs, sofas, etc.) and the number of instances of those sub-kinds (i.e., individual chairs). They then asked whether participants said there was "more" when the number of sub-kinds was greater, and better fulfilled the desires function of, e.g., *furniture*, or if instead quantity judgments were based primarily on the number of discrete individuals. For example, in one study, Gordon and Rodman (2006) placed objects denoted by object-mass nouns in contexts that emphasized their functional role (e.g., items of furniture used to furnish two rooms), and asked English-speaking adults and children to make quantity judgments (e.g., "Which room has more furniture?"). Crucially, each context contrasted the number of sub-kinds with the number of individuals (e.g., one room had a greater number of sub-kinds, while the other had more items of furniture but fewer sub-kinds). They found that adults based judgments on function fulfillment (i.e., choosing the group with more variety over the one with more items) around 30% of the time, and that children did so 55% of the time. In a related study, Grimm and Levin (2012; 2017) presented participants with text descriptions of scenes similar to those used by Gordon and Rodman (2006). When asked to judge between 5 items of one kind (e.g., 5 chairs)

and 5 items that varied in kind (e.g., a sofa, 2 chairs, a coffee table, and a bookcase) participants unsurprisingly based judgments on the number of sub-kinds (since the two sets were equal in the number of individuals; see Hampton & Winter 2024, who also reported that participants base judgments on sub-kind variety when number cannot be used as a basis for judgment – e.g., 4 small bags of the same shape vs. 4 large bags of different shapes). In two follow-up experiments, Grimm and Levin again presented participants with text descriptions of two sets, one that contained more sub-kinds but fewer items (e.g., a sofa, an easy chair, a coffee table and a small bookcase), and the other that contained more items but fewer sub-kinds (one table and 4 chairs). In these studies, participants again based most of their judgments on number. Importantly, there were more judgments based on sub-kind variety when the context placed a focus on function fulfillment compared to a neutral context. For example, in some cases, like *furniture*, the majority responded based on sub-kind variety when function fulfillment was emphasized. Finally, Huang et al. (2022) tested contextual effects on quantification by asking Mandarin speakers to make quantity judgments for Mandarin translations of English object-mass nouns (which appear as bare nouns without grammatical markers). In a substance-focused context (e.g., two monsters grinding up objects and eating them), Mandarin speakers made quantity judgments based on volume, but in an individual-focused context (e.g. two fairies competing to create sets of objects), they made quantity judgments based on the number of things.

One question raised by such studies is whether they demonstrate that mass nouns are especially vulnerable to context effects, as proposed by McCawley (1975) and Rothstein (2010). Alternatively, it is possible that object-mass nouns, like count nouns, specify individuation grammatically, but that all nouns can be coerced contextually, including count nouns. For example, although Gordon and Rodman (2006) found that a function-oriented context made some adults base quantity judgments on function fulfillment for object-mass nouns, they also found the identical result for count nouns (e.g., *toys*, *buildings*). Similarly, although Hampton and Winter (2024) found that nouns classified as "count" led to more quantity judgments based on number than nouns they classified as "object-mass," they did not find that mass nouns were more susceptible to contextual cues. Thus, their study supported the conclusion that some mass nouns may not be object-mass (which is uncontroversial), but did not address the mechanism underlying individuation in those mass nouns that did quantify by number. Finally, although Huang et al. (2022) found that participants made judgments based on volume for translations of English nouns like *furniture*, they did not test whether the same was equally true of words that are count nouns in English. Given this, context effects might be general to both mass and count nouns, and may not weaken the claim that certain nouns (e.g., object-mass nouns or count nouns) specify individuation. This is important, because on most accounts, evidence that pragmatic context can coerce a non-individuated meaning from a count noun would not lead to the abandonment of an individuated semantics for count nouns.

Given these considerations, we conducted two studies that sought to provide a stronger test of the contextual flexibility hypothesis by directly contrasting object-mass nouns with two different kinds of count nouns. In Experiment 1, we asked whether adult participants base quantity judgments on number or variety in contexts that support function fulfillment, and contrasted object-mass nouns to superordinate count nouns (e.g., *tools*, *weapons*). In Experiment 2, we tested English and French speakers with words that are object-mass nouns in English, but whose French translations are count nouns (e.g., *furniture* vs. *meubles*). If individuation is specified for object-mass nouns, as it is for count nouns, then in both experiments there should be no significant differences due to syntax. In contrast, if mass syntax permits greater contextual flexibility than count syntax, then object-mass nouns should be more impacted by context than count nouns in both studies.

2 Experiment 1

Experiment 1 investigated English speakers' quantity judgments for object-mass nouns and collective count nouns in different contexts. All materials, analysis code and data can be found at https://osf.io/5hp8b. A preregistration for both experiments can be found at https://osf.io/3rmnf.

2.1 Methods

2.1.1 Participants

We tested 140 English speakers (90F, 45M, 5 Non-binary; M_{age} = 37.84 [18; 75]; SD_{age} = 12.66) on Prolific.co. We conducted a frequentist *a priori* power analysis using G*Power version 3.1.9.7 (Faul et al. 2007), which found that a sample size of 140 is adequate to achieve 80% power for detecting a medium effect, at a significance criterion of $\alpha = .05$ for a linear regression with 3 predictors (Context, Syntax, and Context * Syntax interaction). This sample size is also adequate to achieve 80% power based on the effects from Huang et al. (2022), using the package *simr* (Green & MacLeod 2016). Finally, because Grimm and Levin (2012; 2017) and Gordon and Rodman (2006) tested separate aspects of the current study, we simulated a combined dataset using their findings. On this simulated dataset, a sample of 140 is also sufficient to achieve 80% power in both a frequentist approach ($p < .05$) and in a Bayesian approach ($BF_{10} > 30$).

All participants resided in the US and had English as a primary language. Nine additional participants were excluded due to self-reported fluency in French ($n = 2$), failed attention checks ($n = 4$) or comprehension checks ($n = 3$). Participants were randomly assigned to either the Neutral Context or the Function-Oriented Context (see **Table 1** for detailed descriptive statistics).

Experiment	n	M_{age}	SD_{age}
Experiment 1 (English speakers)	140 (90F, 45M, 5NB)	37.84 [18; 75]	12.66
Context: Neutral	70 (44F, 24M, 2NB)	40.00 [18; 75]	13.75
Context: Function-Oriented	70 (46F, 21M, 3NB)	35.69 [18; 69]	11.15
Experiment 2 (French speakers)	140 (63F, 75M, 2NB)	35.14 [21; 69]	10.69
Context: Neutral	70 (31F, 38M, 1NB)	34.89 [21; 63]	10.25
Context: Function-Oriented	70 (32F, 37M, 1NB)	35.39 [21; 69]	11.19

Table 1: Descriptive statistics of participants in Experiment 1 & 2.

2.1.2 Materials & Procedure

Participants completed a Truth-Value Judgment task with three characters: Blue Monster, Red Monster, and Penguin.[1] Participants were first randomly assigned to either the Neutral Context ($n = 70$) or the Function-Oriented Context ($n = 70$). Context was manipulated by varying how the two sets of items were introduced (based on Huang et al. 2022). In the Function-Oriented Context, Penguin described a goal specific to the category before the items were introduced. For example, for *furniture*, Penguin said, "I just moved to a new house, and I need some furniture to furnish it. Blue Monster and Red Monster want to give me their furniture." In the Neutral Context, a neutral sentence was used to introduce the category (e.g., "Blue Monster and Red Monster have some furniture."). These prompts remained on the screen until the end of the trial.

On each trial, Blue Monster and Red Monster each had a set of items of the same category (e.g., furniture). One set had more individual pieces but fewer sub-kinds of items, and the other set had fewer individuals but more sub-kinds. The ratio of items and sub-kinds was fixed across trials, such that one set had 10 individual pieces of the same sub-kind, and the other set had 5 individual pieces of different sub-kinds (**Figure 1**) resulting in a 2:1 ratio in number, and a 5:1 ratio in sub-kinds. Participants saw both a visual display of these sets,[2] and also heard an

[1] These characters (as well as our materials) were designed to be child-friendly because we plan to conduct future work in children using the same stimuli. Furthermore, many past studies of diverse semantic and pragmatic phenomena have used child-friendly methods to assess adults, and find expected results (e.g., Papafragou & Musolino 2003; Sullivan et al. 2019).

[2] One concern with representing these sets visually is that participants might be cued by the visual discrete objects to prioritize a reading by number of individuals. However, previous research finds that making the number of individuals more salient (through a visual display) does not obscure the semantic distinction between mass and count. For example, Barner and Snedeker (2005) used a visual display for mass–count flexible items such as *stone* and *string*, but still found a strong effect of syntax: participants made judgments by number when these nouns were presented in count syntax, but by mass/volume when used in mass syntax (for similar results with novel objects, see Barner & Snedeker 2006). This suggests that seeing discrete physical objects does not force quantity judgments by number.

audio description. We note that while these quantities were larger than the quantities used in other studies (e.g., Barner & Snedeker 2005; Scontras et al. 2017; Hampton & Winter 2024), it is necessary to include sets larger than 2–3 items in order to contrast the number of sub-kinds vs. the number of individuals (i.e., smaller quantities would have prevented us from constructing a set with a smaller number of individual items but more heterogeneous in sub-kinds). Crucially, a large literature in numerical cognition has shown that ratio, rather than absolute magnitude, determines the availability of numerical comparison (e.g., Moyer & Landauer 1967; Dehaene et al. 1998; Brannon 2006; Halberda & Feigenson 2008). The ratio of individuals that we chose (2:1) is actually smaller and less numerically salient than the ratios used in previous studies. Therefore, it is unlikely that the larger absolute magnitude of our sets would lead to an inflated number of quantifying judgments by individuals compared to previous studies. Finally, it is crucial that the ratio used is reliably discriminable, which is why we did not use a ratio like 4:5 (as in, e.g., Grimm & Levin 2012).

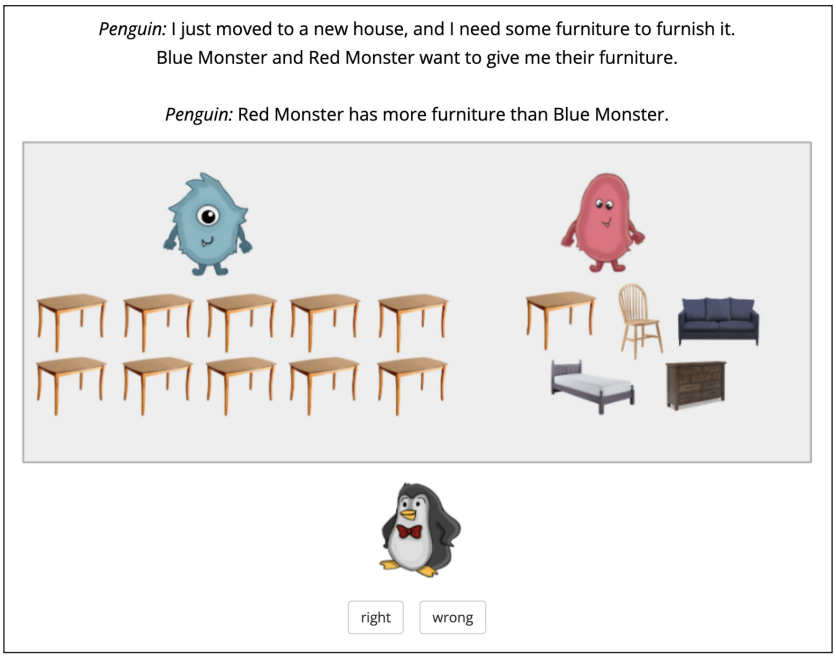

Figure 1: Example trial for *furniture* in the Function-Oriented context in Experiment 1. In the Neutral context, instead of the goal-stating sentence, participants saw the sentence "Blue Monster and Red Monster have some furniture." Before Penguin's quantity judgment appeared, participants also heard an accompanying description of the set "Blue Monster has a table, a chair, a sofa, a bed, and a dresser, and Red Monster has ten tables,"[3] followed by a prompt "Who has more furniture?"

[3] As a reviewer pointed out, our pictorial stimuli did not reflect the true relative sizes of these items. However, previous studies did not find evidence that size impacts judgments for object-mass nouns (e.g., Barner & Snedeker 2005).

After participants were introduced to the items on each trial, they then heard a question (e.g., "Who has more furniture?") and Penguin's reply (e.g., "Blue Monster has more furniture than Red Monster."), then decided whether Penguin was "right" or "wrong." The quantity judgment question, Penguin's response, and the "right or wrong" prompt remained on the screen until participants made their choice. We note that this Truth-Value Judgment (TVJ) format diverged from previous work which had used a two-alternative forced choice (2AFC) task where participants selected which set had "more" (e.g., Barner & Snedeker 2005; Gordon & Rodman 2006; Grimm & Levin 2012; 2017). Huang et al. (2022) argued that a 2AFC task might not capture the possibility that multiple bases of quantity comparison were available, because it forced participants to select the most salient criterion. If one criterion is overwhelmingly salient (e.g., number of individuals), participants might never select less salient criteria even though they can access them. Therefore, a 2AFC task cannot differentiate whether a basis of comparison was entirely unavailable, or merely less salient. Meanwhile, a TVJ task allowed us to explicitly introduce the less salient basis of comparison—in this case, function fulfillment—in some trials. If a functional fulfillment reading is indeed available to participants, then they will respond that a speaker is "right" to say that the set with more variety of sub-kinds of items has "more", even if that set has fewer individuals.

There were 12 critical trials that tested six object-mass nouns (*furniture*, *clothing*, *sports equipment*, *luggage*, *cutlery*, *jewelry*) and six superordinate count nouns (*tools*, *weapons*, *instruments*, *vehicles*, *utensils*, *accessories*). We selected nouns that had been used in previous work (Barner & Snedeker 2005; Gordon & Rodman 2006; Grimm & Levin 2012; 2017), and that reflect "superordinate" categories that consist of multiple "basic-level" items (e.g., *furniture* consists of chairs, tables, beds, etc.; *tools* consists of hammers, wrenches, saws, etc.).

Participants also completed four comprehension trials and two attention check trials. These trials presented two singleton sets (e.g., "Blue Monster has a clock, and Red Monster has a bucket") and Penguin judged which character had a particular item (e.g., "Who has a bucket?"/ Penguin: "Red Monster has a bucket"). In comprehension trials, participants were asked to judge whether Penguin was "right" or "wrong," and they were considered to succeed in these trials if they answered correctly based on the story presented. In attention checks, participants saw and heard additional instructions that specified explicitly to choose either "right" or "wrong." The stipulated response was always opposite of what they would have answered if they followed the story. Participants had to select the response stipulated by the instruction to be considered successful. We excluded participants who failed any of the comprehension ($n = 3$) or attention check trials ($n = 4$). The location of comprehension and attention check trials was distributed evenly across the test trials. To ensure understanding of the task, at the beginning of the study, participants saw two practice trials structured like the comprehension trials. Participants had to pass both practice trials in two tries each in order to proceed. At the end of the study, participants

also completed a demographic questionnaire where they provided information about age, gender, whether English is their first language, and proficiency in other languages.

Trial order was randomized between subjects. The basis of comparison in the critical trials was counterbalanced within each syntactic category (i.e., half of Penguin's statements were made based on number, and half were made based on kinds). We also counterbalanced the correct response for comprehension and attention check trials, and the location of the two sets (left/right).

2.1.3 Analysis

On trials where Penguin judged that the Monster with a greater variety of sub-kinds had "more," participant responses were dummy coded as (1), consistent with function fulfillment, if they judged that Penguin was right, and coded as (0), consistent with number, if they answered that Penguin was wrong. On trials where Penguin judged that the Monster with a greater number of individual items had "more," responses were coded in the opposite manner.

To test whether syntax and context affected quantity judgment, we constructed Bayesian generalized linear mixed-effects models (GLMMs) with "binomial" family and link function "logit" using the package *brms* (Bürkner 2017) to predict responses from Syntax (Mass/Count), Context (Function-Oriented/Neutral), and Syntax * Context interaction. We also included by-subject and by-item random intercepts, and used weakly informative normal priors for slope estimates.[4] We calculated Bayes Factors (BFs) to compare models, and interpreted BFs based on thresholds proposed by Lee and Wagenmakers (2013). We used a Bayesian approach as opposed to a frequentist approach because it allows us to quantify and compare the strength of statistical evidence towards both the null and the alternative hypotheses (Wagenmakers et al. 2008; Wakefield 2013).

2.2 Results & Discussion

We first fitted a uniformly informative Bayesian GLMM that predicted choice of the character with the greater variety of sub-kinds with only by-subject and by-item random intercepts, and then added Syntax and Context as main effects. Neither Syntax ($M_{mass} = 0.18$, $SD_{mass} = 0.39$; $M_{count} = 0.18$, $SD_{count} = 0.38$) nor Context ($M_{function} = 0.22$, $SD_{function} = 0.41$; $M_{neutral} = 0.14$, $SD_{neutral} = 0.35$) were significant predictors (Syntax: $\beta_{mass} = 0.18$, 95% CI [–2.01; 2.38]; Context: $\beta_{function} = 1.29$, 95% CI [–0.05; 2.64]; **Figure 2**). There was strong evidence supporting the null model with only random effects against the model with Syntax + Context main effects ($BF_{10} = 0.05$).

[4] We selected weakly informative priors to remain agnostic in our prediction, as previous research has not tested the specific contrasts used in our studies. Based on suggestions from a reviewer, we reran all studies using two sets of stronger priors. In both cases, the new Bayes Factors mostly suggested the same model preferences. The only exception is that when using the strongest prior for Experiment 1, there was weak evidence supporting the model with Syntax + Context main effects compared to the null model.

Another model found that the Syntax * Context interaction was also not significant (Syntax: β_{mass} = –0.13, 95% CI [–2.38; 2.13]; Context: $\beta_{function}$ = 1.00, 95% CI [–0.35; 2.48]; Syntax * Context: $\beta_{mass*function}$ = 0.52, 95% CI [–0.23; 1.28]). When comparing the interaction model with the model containing just main effects and the null model containing only random effects, there was strong evidence supporting the model with just the main effects (BF_{10} = 0.09), and very strong evidence supporting the null model (BF_{10} < 0.01). These results indicate that participants' quantity judgements did not vary by mass–count syntax, context, or an interaction of these factors. Posterior distributions of all slope estimates were approximately normally distributed (see Supplementary Materials for plots). To investigate whether there were differences between items, we conducted a *post hoc* item analysis for each syntax group and found that both mass and count nouns showed acceptable internal consistency (Mass nouns: Mean inter-item correlation r = 0.40, Cronbach's α = 0.79; Count nouns: Mean inter-item correlation r = 0.33, Cronbach's α = 0.78).[5]

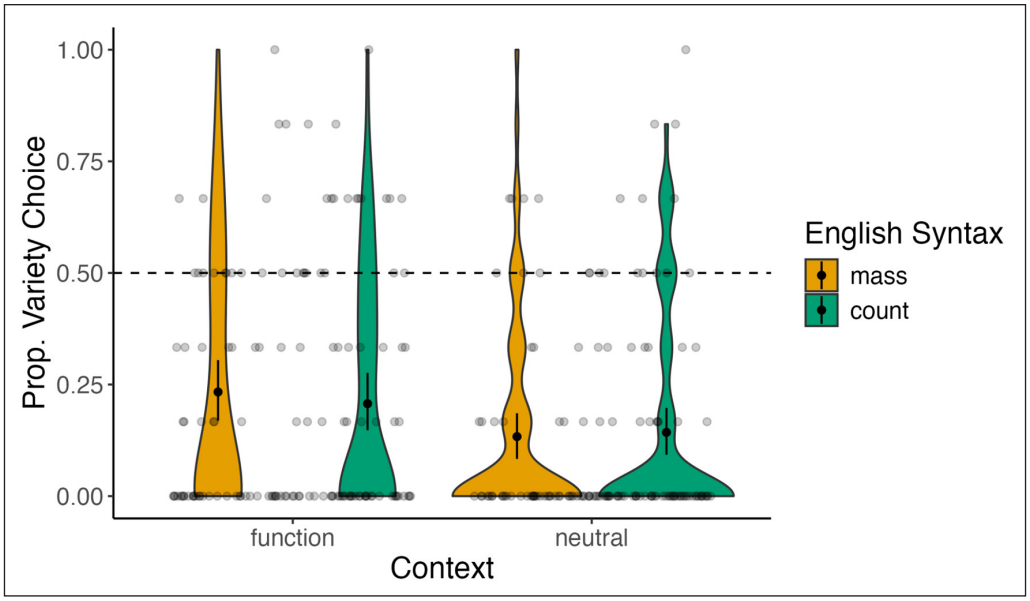

Figure 2: English speakers' responses in Experiment 1. Each dot represents the proportion that a participant chose the character with the greater variety of sub-kinds over the character with the greater number of items. The width of the shaded area of violin plots represents the proportion of the data located there. Error bars represent bootstrapped 95% CIs. Horizontal dashed line indicates chance = ½.

Overall, despite the fact that all trials provided a greater ratio of sub-kinds than number, participants rarely selected the set with greater variety of sub-kinds as having "more" (M = 0.18, SD = 0.38) and instead based most judgments on number (**Figure 2**). Also, given our model

[5] We thank a reviewer for suggesting this analysis.

outputs, participants did not make more judgments based on function fulfillment in contexts that explicitly emphasized function (relative to neutral contexts). Finally, they based quantity judgments on function fulfillment at similar rates for object-mass nouns and collective count nouns, in both contexts. These results cast doubt on the hypothesis that object-mass nouns are more susceptible to contextual effects than count nouns.

Crucially, these results hinge on the assumption that our manipulation of function fulfillment was successful, and that this was achieved by contrasting the number of sub-kinds with the number of individuals. However, it is possible that participants based judgments on number because they did not judge that a greater variety of sub-kinds resulted in greater function fulfillment. To rule out this possibility, we conducted a follow-up experiment with 70 English-speaking participants (43F, 24M, 3 Non-binary; M_{age} = 45.31 [24; 73]; SD_{age} = 13.26) who had not participated in the previous experiment. Participants were tested using stimuli identical to those in the Function-Oriented Context (including both the visual display and the goal statement from Penguin), but instead of asking participants to make quantity judgments, we asked them to judge function fulfillment. For example, in the trial with two sets of furniture, they were asked, "Whose furniture will be better for furnishing a house?" (see Supplementary Materials).[6] If our stimuli had successfully operationalized different levels of function fulfillment, such that the sets with more sub-kinds were better at fulfilling their function compared to the homogeneous sets of the same category, then participants should choose the heterogeneous sets more often than expected by chance in response to the function fulfillment question.

To test this possibility, we fitted a pre-registered Bayesian GLMM predicting participants' responses (Heterogeneous Set Choice: 1, Homogeneous Set Choice: 0) with by-subject and by-item random intercepts, and tested whether the intercept of each Syntax group was greater than 0.[7] We found very strong evidence that participants selected the heterogeneous sets more often than expected by chance for object-mass nouns (β = 5.33; 95% CI [2.30; 8.87], Evidence Ratio = 276.78), and moderate evidence that they did so for superordinate count nouns (β = 2.29; 95% CI [–0.64; 5.35], Evidence Ratio = 9.58). To further investigate the difference between mass and count syntax, we fitted another pre-registered Bayesian GLMM predicting participants' responses with Syntax as a fixed effect, and by-subject and by-item random intercepts. We found no effect of Syntax (β_{mass} = 2.69; 95% CI [–2.51; 7.96]). When comparing this model with a model without the Syntax predictor, there was moderate evidence supporting the model with no

[6] Seven of 12 items in the function fulfillment follow-up experiment were modified to sound more natural. An anonymous reviewer worried that these items may have referred to more specific functions in the follow-up experiment compared to Experiment 1, artificially inflating the rate of function fulfillment judgments (i.e., picking the heterogeneous set) in the follow-up experiment. However, a *post hoc* analysis comparing responses in the two experiments found that whether an item was modified or not had little impact. There was no significant main effect of prompt modification on function fulfillment responses, nor interaction effects involving this factor.

[7] This analysis effectively tests whether participants choose the heterogeneous set more than chance = ½, for each Syntax group.

Syntax predictor ($BF_{10} = 0.49$). In summary, we find strong evidence that, when directly asked about function fulfillment, participants interpreted the sub-kind variety illustrated in our stimuli as evidence for better function fulfillment for both mass and count nouns.

Although Experiment 1 finds no difference between superordinate nouns used in mass vs. count syntax, one limitation of this experiment is that it did not control completely for differences in the lexical meanings of words in each category. For example, it's possible that differences in how participants conceptualize furniture and tools might mask the influence of syntax on quantification. This is also a concern with prior studies that used the same contrast between object-mass nouns and superordinate count nouns in one language (e.g., Gordon & Levin 2012; Hampton & Winter 2024). The strongest possible test of whether mass syntax makes nouns more vulnerable to context effects would be to present participants with object-mass nouns, but used in count syntax, to test whether they are less flexible as count nouns than as mass nouns. Although no language that we know of allows this form of alternation for object-mass nouns, some languages, like French, allow translations of English object-mass nouns to be used in count syntax (e.g., *furniture* vs. *meubles*). With this in mind, we conducted Experiment 2 in French. By comparing results of French speakers in Experiment 2 and English speakers from Experiment 1, we controlled for conceptual differences between words while still testing for effects of syntax. We reasoned that if mass–count syntax has an effect on quantification, then French speakers should be more likely to quantify by number than English participants.

3 Experiment 2

Experiment 1 provided preliminary support for the hypothesis that object-mass nouns specify individuation just as strongly as count nouns. In Experiment 2, we conducted a second test of this hypothesis, by investigating how French speakers interpret translations of English object-mass nouns, which in French occur in count syntax. We reasoned that if nouns used in mass syntax are more context sensitive than similar nouns used in count syntax, then French speakers should be more likely to quantify by number than English participants when context favors function fulfillment. All materials, analysis code and data can be found at https://osf.io/5hp8b. A preregistration for both experiments can be found at https://osf.io/3rmnf.

3.1 Methods

3.1.1 Participants

We tested 140 French speakers (63F, 75M, 2 Non-binary; $M_{age} = 35.16$ [21; 69]; $SD_{age} = 10.69$) on Prolific. This sample size was selected to match the sample size of English-speaking participants. All participants had French as a primary language. An additional 10 participants were excluded due to failing attention checks ($n = 2$) or comprehension checks ($n = 8$). Because most participants (104 out of 140) rated themselves as proficient in English (at least 6/10 in speaking and/or comprehending English), we did not carry out an exclusion with this criterion.

We reasoned that if there is an effect of syntax (i.e., mass vs. count nouns quantify differently, and/or differed in their sensitivity to contexts), then even if French speakers draw on their knowledge of English, this should not affect their responses for French count nouns. Participants were randomly assigned to either the Neutral Context (n = 70) or the Function-Oriented Context (n = 70; see **Table 1** for detailed descriptive statistics). Data from French participants were compared to English data from Experiment 1.

3.1.2 Materials & Procedure

Materials were identical to Experiment 1, but translated to French by two native French speakers. Audio was recorded by a third native French speaker. To equate the task with Experiment 1, we presented French participants with translations of both English object-mass and count nouns. Because comparing French and English speakers' interpretation of English object-mass nouns is critical to answering our research question, we only report results for these nouns, though data for superordinate count nouns do not differ from mass nouns and are compatible with other findings.[8] Participants followed the same procedure as Experiment 1. The only difference was in the demographic questionnaire, where they were asked to confirm that French (not English) was their first language.

3.1.3 Analysis

We used the same coding scheme and analysis approach in Experiment 1, but with Bayesian generalized linear mixed-effects models that predicted English and French participants' responses to English object-mass nouns. We included Language (English/French), Context (Function-Oriented/Neutral), and Language * Context interaction as fixed effects.[9] Note that Language is a proxy for Syntax in these models, since we only analyzed nouns that were mass in English and count in French.

3.2 Results & Discussion

We first fitted a uniformly informative Bayesian GLMM that predicted choice of the character with the greater variety of sub-kinds, and then added Language and Context as main effects. Only Context was a significant predictor (Language: β_{French} = 0.40, 95% CI [–0.35; 1.18]; Context: β_{function} = 1.75, 95% CI [1.01; 2.55]). There was very strong evidence supporting the model with Language + Context main effects against both the null model and a *post hoc* model with only Language as a main effect (BF_{10} > 30). In another model including an interaction effect of Language and Context, only the main effect of Context was significant (Language: β_{French} =

[8] See Supplementary Materials for a combined analysis of English and French speakers' judgements of both English object-mass nouns and superordinate count nouns.

[9] This analysis diverged from the preregistered models, which we erroneously specified to include a predictor for English Syntax (mass/count) despite only analyzing English object-mass nouns.

–0.07, 95% CI [–1.21; 1.10]; Context: $\beta_{function}$ = 1.31, 95% CI [0.23; 2.41]; Language * Context: $\beta_{French*function}$ = 0.88, 95% CI [–0.66; 2.42]). When comparing the interaction model with the model containing only the Language + Context main effects, there was strong evidence supporting the model with just the main effects (BF_{10} = 0.06). These results suggested that participants' quantity judgments varied based on context, with participants in the Function-Oriented context making more quantity judgments by function fulfillment compared to participants in the Neutral context. However, quantity judgments did not vary based on language or an interaction between language and context. Posterior distributions of all slope estimates were approximately normally distributed (see Supplementary Materials for plots).

Overall, French participants, like English participants, preferred to base quantity judgments on number, and were less likely to select the more heterogeneous set as having "more" in both contexts (M = 0.21, SD = 0.41; $M_{function}$ = 0.31, $SD_{function}$ = 0.46; $M_{neutral}$ = 0.11, $SD_{neutral}$ = 0.32; **Figure 3**). While there was an overall effect of Context, with the function-oriented context resulting in more quantification by function fulfillment, French speakers and English speakers were just as likely to be influenced by context, again suggesting no impact of mass–count syntax. Further, as in Experiment 1, we found no evidence of important item effects, with acceptable internal consistency for French speakers' judgments of translated object-mass nouns (Mean inter-item correlation r = 0.33, Cronbach's α = 0.75) and superordinate count nouns (Mean inter-item correlation r = 0.33, Cronbach's α = 0.79).

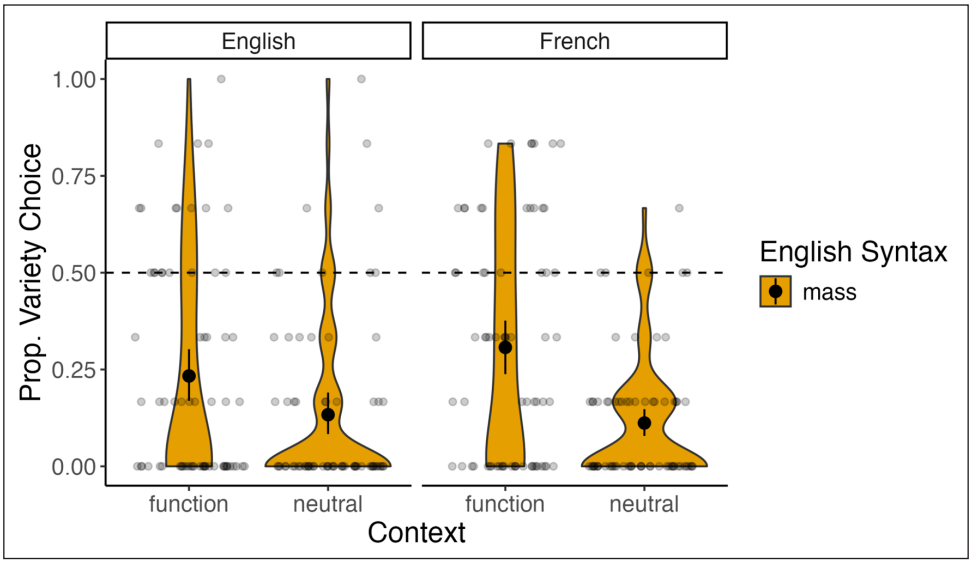

Figure 3: English and French speakers' responses in Experiment 2. Each dot represents the proportion that a participant chose the character with the greater variety of sub-kinds over the character with the greater number of items. The width of the shaded area of violin plots represents the proportion of the data located there. Error bars represent bootstrapped 95% CIs. Horizontal dashed line indicates chance = ½.

4 General Discussion

In two experiments, we asked whether individuation of object-mass nouns is best explained as the product of grammatical specification, or instead arises from contextual coercion that is specific to the mass category. In particular, we tested whether, for object-mass nouns like *furniture*, participants prefer to base quantity judgments on the degree to which items fulfill a desired function, like "furnishing" (e.g., McCawley 1975; Gordon & Rodman 2006; Grimm & Levin 2012) or instead make judgments on the basis of number, as they do for related count nouns (e.g., Barner & Snedeker 2005; Bale & Barner 2009). In both experiments, we manipulated function fulfillment by presenting participants with a choice between two sets: one that had a greater number of individuals, and another that had fewer individuals, but a greater number of sub-kinds that would afford greater function fulfillment. We also manipulated whether function fulfillment of the sets was emphasized or not. Using this approach, we found three main results. First, both English and French speakers based quantity judgments overwhelmingly on number. This was despite the fact that the ratio of sub-kinds across each pair of sets (5:1) was much greater than the ratio of individuals (2:1), and despite the fact that participants judged the set with the great number of sub-kinds to better fulfill the intended function than the set with more individuals. Second, syntax did not affect judgments when comparing English object-mass nouns with either English count superordinate nouns (Experiment 1) or French count nouns that were translations of English object-mass nouns (Experiment 2). Finally, while there was some evidence that participants were more likely to base judgments on sub-kind variety when nouns were presented in a function-oriented context, there was no significant difference in how context affected judgments for English object-mass nouns relative to their French count noun translations, or English superordinate count nouns. These results suggest that, although context may sometimes inform how language users make quantity judgments (as per McCawley 1975), participants nevertheless base judgments mostly on number for object-mass nouns and related count nouns. Also, to the extent that context does impact judgments, it does so equally for nouns used in mass and count syntax. Thus, we did not find evidence for the idea that object-mass nouns are more contextually flexible than equivalent count nouns (Rothstein 2017).

As noted in the Introduction, several previous accounts argue that mass nouns like *furniture* have denotations that include atomic individuals, much like count nouns (e.g., Katz 1970; Bunt 1985; Gillon 1992; 1999; Chierchia 1998; 2010; Barner & Snedeker 2005; 2006; Bale & Barner 2009; Rothstein 2010; 2017; Landman 2020). Most accounts agree that the comparative morpheme *more* must be underspecified with respect to its dimension of comparison to account for the wide range of possible measurements such as number, mass, length, and volume (e.g., *more books*, *more dirt*, *more string*, *more water*) as well as abstract dimensions relevant to actions

and events (e.g., *more walking*, *more hockey*, *more hope*; see Barner et al. 2008). However, past accounts differ in their treatment of object-mass nouns. According to some, only count nouns strictly force a numerical comparison, and while object-mass nouns may be biased towards numerical comparison, they do not impose it grammatically and are highly sensitive to context (McCawley 1975; Rothstein 2017; Landman 2020). In contrast, others propose that both count nouns and object-mass nouns specify comparison by number grammatically (Barner & Snedeker 2005; Bale & Barner 2009; Wellwood 2015). Consistent with the latter class of accounts, our data reveal that object-mass nouns are no more sensitive to context than count nouns. The contextual sensitivity of measurement may therefore have more to do with the flexible nature of the comparative morpheme than it has to do with the grammar of individuation (see Schwarzschild 2006; Wellwood 2015; Solt 2018; Bale & Schwarz 2020; Bale et al. 2021).

One important question that remains unanswered by this study and others in the literature is why some superordinate nouns that denote collections of individuals are expressed as mass nouns (e.g., *furniture*, *clothing*) while others are not (e.g., *vehicles*, *weapons*). Relevant to this question, previous studies (Gordon 1985; Barner & McKeown 2005) report that when a novel word is presented in ambiguous syntax, the number of objects referred to impacts whether an object-denoting noun is acquired as mass or count. For example, in a study by Barner and McKeown (2005), when children were told, "Look at the garn" while only one object was present, they often inferred that *garn* was a singular count noun and later pluralized it when referring to multiple objects of the same type. However, when the same utterance was used to refer to multiple objects, children interpreted it as a mass noun—since it lacked plural morphology—and did not pluralize it when referring to multiple objects. Given the fact that the referents of superordinates like *furniture* tend to appear together in groups (Wisniewski et al. 1996), when children hear an expression like "look at the furniture," they might be more likely to interpret it as having plural reference, and therefore being a mass noun. Perhaps children make this inference even if the noun used is a count noun in the speaker's dialect, and they intended to refer to only one piece of furniture (e.g., as might occur in French). In contrast, if the referents of superordinate terms like *instrument* are more likely to occur in isolation (e.g., a guitar), then children might be more likely to assume that these expressions are singular count nouns. In sum, our proposal is that syntactic ambiguity combined with referential ambiguity where children could assume that a singular noun refers to a plurality might result in misclassification of count nouns as mass nouns.

Related to this idea, future work should also explore why some languages feature more object-mass nouns than others. Because syntactic ambiguity should generally lead to classification errors, either towards mass or count nouns, cross-linguistic variability in syntactic ambiguity has

also been investigated as a predictor of object-mass noun presence. One possibility, proposed by Erbach (2020), is that ambiguity might lead to more count nouns, and to fewer object-mass nouns. This idea is supported by findings from early corpus work (Bloom 1994b) that showed children misclassifying mass nouns as count when they denoted solid physical entities, resulting in errors like "broccolis," "bacons," and "furnitures." However, it is also possible that syntactic ambiguity creates opportunities for children to classify individuating expressions as mass nouns. If this is the case, it may be that languages with richer number morphology are less likely to feature object-mass nouns. For example, in English, which features multiple object-mass nouns, differentiating between mass and count usages often hinges upon a single morpheme (e.g. "Look at the hair/hairs"), and some determiners (e.g., *the*) are mass–count neutral. By contrast, Italian marks plural information on both determiners and root nouns (e.g., "Guarda la sedia"/"Guarda le sedie"). Because plural information is more salient in Italian, count nouns may be less likely to be mis-classified by children as object-mass nouns, leading to fewer of them in Italian relative to English.[10]

Finally, in addition to understanding the question of how words might become object-mass nouns, future studies investigating the interface between mass–count syntax and semantics should broaden the scope of words included in research on this question. For example, whereas superordinate object-mass nouns like *furniture* are somewhat rare and idiosyncratic, in some languages, like English, mass nouns that refer to individuated phenomena can be freely generated via processes of nominalization from verbs that denote punctual events (e.g., *too much kicking/jumping*). Such cases are important because they allow additional ways to directly control lexical semantics while manipulating syntax (e.g., to compare *jumps* vs. *jumping*), and because they allow an abstraction away from the particularities of words like *furniture*. While work in English has found that deverbal nouns (e.g., *a lot of jumping*) quantify by number to the same extent as their count counterparts (e.g., *a lot of jumps*; Barner et al. 2008), future studies should explore this phenomena from a cross-linguistic perspective, and in language acquisition.

In summary, we presented two studies measuring English and French speakers' quantity judgments of English object-mass nouns and collective count nouns. We found that in both

[10] Note that these claims are difficult to test regarding classifier languages for multiple reasons. Some argue that all nouns in classifier languages are mass (meaning that there are many object-mass nouns; e.g., Chierchia 1998). Others argue that there are no mass or count nouns in classifier languages (because mass and count categories are characterized morpho-syntactically, in contrast to one another; Bale & Barner 2018). Others argue that classifier languages have a mass–count distinction in the sortal/mensural distinction (e.g., Doetjes 1996). Finally, some argue that the mass–count distinction is semantic/conceptual rather than morpho-syntactic, and therefore nouns that individuate are *de facto* count (e.g., Lin & Schaeffer 2018). Our own view is that these languages do not have a morpho-syntactic mass–count distinction comparable to the one made in languages like English or Italian, and that distinct terminology should be used to analyze conceptual distinctions made elsewhere in the language.

languages, participants made quantity judgments mostly by number of individuals as opposed to number of kinds, and that there was no effect of syntax on quantity judgments (within English and cross-linguistically). These findings are compatible with the hypothesis that some mass nouns lexically specify individuation, and are no more susceptible to effects of context than similar count nouns. Therefore, while all count nouns individuate, mass nouns can refer to either individuated (object-mass nouns) or non-individuated (substance-mass nouns) phenomena.

Data Availability/Supplementary Files

All materials, analysis code, data, and supplementary files can be found at https://osf.io/5hp8b

Supplementary file 1: Materials. Stimuli and web pages presenting the experiments. https://osf.io/ajfvg

Supplementary file 2: Analysis code. Analysis pipeline. https://osf.io/p8k4r

Supplementary file 3: English Data. Anonymized data from English speakers. https://osf.io/v8z6r

Supplementary file 4: French Data. Anonymized data from French speakers. https://osf.io/7ezgp

Supplementary file 5: Appendix. Combined analysis of English and French speakers' responses for both English object-mass and superordinate count nouns & Posterior distributions and trace plots of posterior samples from Bayesian models https://osf.io/uxg3b

Ethics and consent

This study was approved by the University of California, San Diego Human Research Protections Program (Protocol #171652).

Funding Information

This research was supported in part by a SSHRC Insight Grant (#435-2023-0699) to AB.

Acknowledgements

We thank the participants, and also thank David Baillot, Julie Burelle, Sophia Daluraya, Christine Kwon, Jingyi Liu, Chelsea Moreno, and Jolene Nesnas for assistance with creating the stimuli.

Competing interests

The authors have no competing interests to declare.

Authors' contribution

The authors made the following contributions. Khuyen N. Le: Conceptualization, Methodology, Formal analysis, Investigation, Data Curation, Writing – Original Draft, Writing – Review & Editing, Visualization, Project administration; Alan Bale: Conceptualization, Methodology, Writing – Review & Editing; David Barner: Conceptualization, Methodology, Resources, Writing – Review & Editing, Supervision, Funding acquisition.

Author Note

Part of these findings were presented at the 2024 California Meeting on Psycholinguistics and the 47th Annual Meeting of the Cognitive Science Society.

References

Bale, Alan & Barner, David. 2018. Quantity judgment and the mass–count distinction across languages: Advances, problems, and future directions for research. *Glossa: a journal of general linguistics* 3(1). 63. https://doi.org/10.5334/gjgl.536

Bale, Alan & Schwarz, Bernhard. 2020. Proportional readings of many and few: the case for an underspecified measure function. *Linguistics and Philosophy* 43(6). 673–699. https://doi.org/10.1007/s10988-019-09284-5

Bale, Alan & Schwarz, Bernhard & Shanks, David. 2021. Monotonicity revisited: mass nouns and comparisons of purity. *Journal of Semantics* 38(4). 681–708. https://doi.org/10.1093/jos/ffab012

Bale, Alan C., & Barner, David. 2009. The interpretation of functional heads: Using comparatives to explore the mass/count distinction. *Journal of Semantics* 26(3). 217–252. https://doi.org/10.1093/jos/ffp003

Barker, Chris. 1998. Partitives, double genitives and anti-uniquenes. *Natural Language & Linguistic Theory* 16(4). 679–717. https://doi.org/10.1023/A:1005917421076

Barner, David & Libenson, Amanda & Cheung, Pierina & Takasaki, Mayu. 2009. Cross-linguistic relations between quantifiers and numerals in language acquisition: Evidence from Japanese. *Journal of Experimental Child Psychology* 103(4). 421–440. https://doi.org/10.1016/j.jecp.2008.12.001

Barner, David & McKeown, Rebecca. 2005. The syntactic encoding of individuation in language and language acquisition. *BUCLD 28: Proceedings of the 28th Annual Boston University Conference on Language Development*, 61–72. Cascadilla Press Somerville, MA.

Barner, David & Snedeker, Jesse. 2005. Quantity judgments and individuation: Evidence that mass nouns count. *Cognition* 97(1). 41–66. https://doi.org/10.1016/j.cognition.2004.06.009

Barner, David, & Snedeker, Jesse. 2006. Children's early understanding of mass–count syntax: individuation, lexical content, and the number asymmetry hypothesis. *Language Learning and Development* 2(3). 163–194. https://doi.org/10.1207/s15473341lld0203_2

Barner, David & Wagner, Laura, & Snedeker, Jesse. 2008. Events and the ontology of individuals: Verbs as a source of individuating mass and count nouns. *Cognition* 106(2). 805–832. https://doi.org/10.1016/j.cognition.2007.05.001

Bloom, Paul. 1994a. Syntax-semantics mappings as an explanation for some transitions in language development. In Levy, Yonata (ed.), *Other children, other languages: Issues in the theory of language acquisition,* 41–75. Lawrence Erlbaum Associates, Inc.

Bloom, Paul. 1994b. Possible names: The role of syntax-semantics mappings in the acquisition of nominals. *Lingua* 92. 297–329. https://doi.org/10.1016/0024-3841(94)90345-X

Bloom, Paul. 1999. The role of semantics in solving the bootstrapping problem. In Jackendoff, Ray S. & Bloom, Paul & Wynn, Karen (eds.), *Language, logic, and concepts: Essays in memory of John Macnamara,* 285–309. The MIT Press. https://doi.org/10.7551/mitpress/4118.003.0015

Brannon, Elizabeth M. 2006. The representation of numerical magnitude. *Current opinion in neurobiology* 16(2). 222–229. https://doi.org/10.1016/j.conb.2006.03.002

Bunt, Harry C. 1985. *Mass terms and model-theoretic semantics*. New York: Cambridge University Press.

Bürkner, Paul. C. 2017. brms: An R package for Bayesian multilevel models using Stan. *Journal of statistical software* 80. 1–28. https://doi.org/10.18637/jss.v080.i01

Cheng, Chung-ying. 1973. Comments on Moravcsik's Paper. In Hintikka, Kaarlo J. J. & Moravcsik, Julius M. E. & Suppes, Patrick (eds.), *Approaches to Natural Language: Proceedings of the 1970 Stanford Workshop on Grammar and Semantics*, 286–288. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-010-2506-5_14

Chierchia, Gennaro. 1998. Plurality of Mass Nouns and the Notion of "Semantic Parameter." In Rothstein, Susan (ed.), *Events and Grammar*, 53–103. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-011-3969-4_4

Chierchia, Gennaro. 2010. Mass nouns, vagueness and semantic variation. *Synthese* 174. 99–149. https://doi.org/10.1007/s11229-009-9686-6

Dehaene, Stanislas & Dehaene-Lambertz, Ghislaine & Cohen, Laurent. 1998. Abstract representations of numbers in the animal and human brain. *Trends in neurosciences* 21(8). 355–361. https://doi.org/10.1016/S0166-2236(98)01263-6

Doetjes, Jenny. 1996. Mass and count: Syntax or semantics? *Proceedings of Meaning on the HIL*. 34–52.

Erbach, Kurt. 2020. Predicting object mass nouns across languages. *Proceedings of the Linguistic Society of America* 5(1). 228–241. https://doi.org/10.3765/plsa.v5i1.4698

Erbach, Kurt & Schoenfeld, Aviv. 2022. Object mass nouns and subkind countability. *Glossa: a journal of general linguistics* 7(1). https://doi.org/10.16995/glossa.5788

Erbach, Kurt & Sutton, Peter R. & Filip, Hana & Byrdeck, Kathryn. 2021. Object mass nouns as arbiter for the mass/count category. In Kiss, Tibor & Pelletier, Francis J. & Husić Halima (eds.), *Things and stuff: The semantics of the count-mass distinction*, 167–192. Cambridge University Press. http://doi.org/10.1017/9781108937979.008

Faul, Franz & Erdfelder, Edgar & Lang, Albert-Georg & Buchner, Axel. 2007. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods* 39. 175–191. https://doi.org/10.3758/BF03193146

Gillon, Brendan S. 1992. Towards a common semantics for English count and mass nouns. *Linguistics and Philosophy* 15(6). 597–639. https://doi.org/10.1007/BF00628112

Gillon, Brendan S. 1999. The lexical semantics of English count and mass nouns. In Viegas, Evelyne (ed.), *Breadth and Depth of Semantic Lexicons*, 19–37. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-017-0952-1_2

Gordon, Peter. 1985. Evaluating the semantic categories hypothesis: The case of the count/mass distinction. *Cognition* 20(3). 209–242. https://doi.org/10.1016/0010-0277(85)90009-5

Gordon, Peter & Rodman, Alison. 2006. Functional quantification of object mass nouns in children and adults. In *30th Annual Boston University Conference on Language Development*. Boston, MA.

Green, Peter & MacLeod, Catriona J. 2016. SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution* 7(4). 493–498. https://doi.org/10.1111/2041-210X.12504

Grimm, Scott & Levin, Beth. 2012. Who has more furniture? An exploration of the bases for comparison. In *Mass/Count in Linguistics, Philosophy and Cognitive Science Conference*. Paris, France.

Grimm, Scott & Levin, Beth. 2017. Artifacts: Reference, countability, and categorization. In *North East Linguistic Society (NELS)* 47. 55–64.

Grimm, Scott & Wahlang, Aeshaan. 2021. Determining Countability Classes. In T. Kiss, F. J. Pelletier, & H. Husić (Eds.), *Things and Stuff: The Semantics of the Count-Mass Distinction*, 357–376. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108937979.017

Hacohen, Aviya & Schaeffer, Jeannette. 2016. On the mass/count distinction in Hebrew: Language acquisition and language change. *Journal of Child Language Acquisition and Development* 4(2). 34–61. https://hdl.handle.net/11245.1/58e31300-d16a-4e14-94f7-7a27a9ce6002

Halberda, Justin & Feigenson, Lisa. 2008. Developmental change in the acuity of the "number sense": The approximate number system in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology* 44(5). 1457–1465. https://doi.org/10.1037/a0012682

Hampton, James A. & Winter, Yoad. 2024. Object mass nouns and comparative judgements. *Proceedings of Sinn Und Bedeutung* 28. 355–368. https://doi.org/10.18148/sub/2024.v28.1128

Huang, Aijun & Li, Jingjing & Meroni, Luisa. 2022. Grammatical and contextual factors affecting the interpretation of superordinate collectives in child and adult Mandarin. *Linguistics* 60(4). 933–972. https://doi.org/10.1515/ling-2022-0048

Inagaki, Shunji. 2014. Syntax–semantics mappings as a source of difficulty in Japanese speakers' acquisition of the mass–count distinction in English. *Bilingualism: Language and cognition* 17(3). 464–477. https://doi.org/10.1017/S1366728913000540

Inagaki, Shunji & Barner, David. 2009. Countability in absence of count syntax: Evidence from Japanese quantity judgments. In Inagaki, Shunji & Hirakawa, Makiko & Setsuko, Arita & Hirakawa, Yahiro & Morikawa, Hiromi & Nakayama, Mineharu & Sirai, Hidetosi & Tsubakita, Jessika (eds.), *Studies in Language Sciences 8: Papers from the Eighth Annual Conference of the Japanese Society for Language Sciences*. Tokyo: Kurosio.

Jespersen, Otto. 1909. *A modern english grammar: On historical principles. I–VII*. Einar Munksgaard.

Katz, Jerrold J. 1970. Interpretative semantics vs. generative semantics. *Foundations of language,* 220–259. https://www.jstor.org/stable/25000452

Landman, Fred. 2020. *Iceberg semantics for mass nouns and count nouns*. Springer International Publishing. https://doi.org/10.1007/978-3-030-42711-5

Lee, Michael D. & Wagenmakers, Eric-Jan. 2013. *Bayesian cognitive modeling: A practical course.* Cambridge University Press. https://doi.org/10.1017/CBO9781139087759

Lin, Jing & Schaeffer, Jeannette. 2018. Nouns are both mass and count: Evidence from unclassified nouns in adult and child Mandarin Chinese. *Glossa: a journal of general linguistics* 3(1). 54. https://doi.org/10.5334/gjgl.406

Link, Godehard. 1983. The Logical Analysis of Plurals and Mass Terms: A Lattice Theoretical Approach. In Bäuerle, Rainer & Schwarze, Christoph & von Stechow, Arnim (eds.), *Meaning, Use, and Interpretation of Language*, 302–323. Walter de Gruyter. https://doi.org/10.1515/9783110852820.302

Link, Godehard. 1998. *Algebraic semantics in language and philosophy*. Stanford, CA: Center for the Study of Language and Information.

Macnamara, John. 1972. Cognitive basis of language learning in infants. *Psychological Review* 79. 1–13. https://doi.org/10.1037/h0031901

McCawley, James D. 1975. Lexicography and the Count-Mass Distinction. *Annual Meeting of the Berkeley Linguistics Society* 1. 314–321. https://doi.org/10.3765/bls.v1i0.2335

Mihatsch, Wiltrud. 2016. Collectives, object mass nouns and individual count nouns: Nouns between lexical and inflectional plural marking. *Lingvisticæ Investigationes* 39(2). 289–308. https://doi.org/10.1075/li.39.2.05mih

Mohr, Susanne & Agyepong, Dorothy. P. 2022. The cultural adaptation of quantity judgment tasks in Ghanaian English and Akan. *Contemporary Journal of African Studies* 9(2). 120–146. https://doi.org/10.4314/contjas.v9i2.5

Moyer, Robert S. & Landauer, Thomas K. 1967. Time required for Judgements of Numerical Inequality. *Nature* 215. 1519–1520. https://doi.org/10.1038/2151519a0

Papafragou, Anna & Musolino, Julien. 2003. Scalar implicatures: Experiments at the semantics–pragmatics interface. *Cognition* 86(3). 253–282. https://doi.org/10.1016/S0010-0277(02)00179-8

Prasada, Sandeep. 1999. Names of things and stuff: An Aristotelian perspective. In Jackendoff, Ray S. & Bloom, Paul & Wynn, Karen (eds.), *Language, logic, and concepts: Essays in memory of John Macnamara*, 119–146. The MIT Press. https://doi.org/10.7551/mitpress/4118.003.0010

Quine, Willard V. O. 1960. *Word and object*. MIT Press.

Rothstein, Susan. 2010. Counting and the mass/count distinction. *Journal of semantics* 27(3). 343–397. https://doi.org/10.1093/jos/ffq007

Rothstein, Susan. 2017. *Semantics for Counting and Measuring*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9780511734830

Rothstein, Susan. 2021. Counting and Measuring and Approximation. In Filip, Hana (ed.), *Countability in Natural Language*, 217–251. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316823774.010

Rothstein, Susan & de Oliveira, Roberta P. 2020. Comparatives in Brazilian Portuguese: counting and measuring. In Moltmann, Friederike (ed.), *Mass and count in linguistics, philosophy, and cognitive science*, 141–158. John Benjamins Publishing Company. https://doi.org/10.1075/lfab.16.07rot

Schwarzschild, Roger. 2006. The Role of Dimensions in the Syntax of Noun Phrases. *Syntax* 9(1). 67–110. https://doi.org/10.1111/j.1467-9612.2006.00083.x

Scontras, Gregory & Davidson, Kathryn & Deal, Amy R. & Murray, Sarah E. 2017. Who has more? The influence of linguistic form on quantity judgments. *Proceedings of the Linguistic Society of America* 2(41). 1–15. https://doi.org/10.3765/plsa.v2i0.4097

Solt, Stephanie. 2018. Proportional comparatives and relative scales. *Proceedings of Sinn Und Bedeutung* 21(2). 1123–1140. Retrieved from https://ojs.ub.uni-konstanz.de/sub/index.php/sub/article/view/188

Sullivan, Jessica & Boucher, Juliana & Kiefer, Reina J. & Williams, Katherine & Barner, David. 2019. Discourse Coherence as a Cue to Reference in Word Learning: Evidence for Discourse Bootstrapping. *Cognitive Science* 43(1). e12702. https://doi.org/10.1111/cogs.12702

Sutton, Peter & Filip, Hana. 2018. Restrictions on subkind coercion in object mass nouns. *Proceedings of Sinn Und Bedeutung* 21(2). 1195–1214. https://ojs.ub.uni-konstanz.de/sub/index.php/sub/article/view/193

Van Witteloostuijn, Merel. 2013. *A cross-linguistic study of the mass–count distinction in number marking languages: A comparison of English, Hebrew, Dutch and French.* Utrecht: Utrecht University Master thesis.

Van Witteloostuijn, Merel & Schaeffer, Jeannette. 2014. On the acquisition of the mass–count distinction in Dutch. *Poster presentation at the International Association for the Study of Child Language (IASCL)*.

Wagenmakers, Eric-Jan & Lee, Michael & Lodewyckx, Tom & Iverson, Geoffrey J. 2008. Bayesian Versus Frequentist Inference. In Hoijtink, Herbert & Klugkist, Irene & Boelen, Paul A. (eds.), *Bayesian Evaluation of Informative Hypotheses. Statistics for Social and Behavioral Sciences,* 181–207. Springer. https://doi.org/10.1007/978-0-387-09612-4_9

Wakefield, Jon. 2013. *Bayesian and frequentist regression methods*. New York: Springer. https://doi.org/10.1007/978-1-4419-0925-1

Wellwood, Alexis. 2015. On the semantics of comparison across categories. *Linguistics and Philosophy* 38. 67–101. https://doi.org/10.1007/s10988-015-9165-0

Wisniewski, Edward J. & Imai, Mutsumi & Casey, Lyman. 1996. On the equivalence of superordinate concepts. *Cognition* 60(3). 269–298. https://doi.org/10.1016/0010-0277(96)00707-X