When two is a crowd: Evidence for asymmetric number representations

Mora Maldonado (mora.maldonado@upf.edu)
Department of Translation and Language Sciences, Universitat Pompeu Fabra

Jennifer Culbertson (jennifer.culbertson@ed.ac.uk)
Center for Language Evolution, The University of Edinburgh

Carmen Saldana (carmen.saldana@ub.edu)
Department of Catalan Philology and General Linguistics, Universitat de Barcelona

Abstract

Research on morphological marking of number across the world's language points to a potential asymmetry: while many languages maintain a distinct form for the singular (as distinct from the plural), fewer languages maintain a distinct form for the dual, and very few distinguish trial or paucal. The presence of these patterns and their asymmetries could indicate that morphological systems are structured around a small set of primitives, including ONE and TWO, which are weighted differently, i.e., one is prioritized over the other. But how these primitives might relate to conceptual representations of numerosity, where small numbers have also been shown to have a privileged status, remains unclear. Further, there has been little work exploring whether representations of small numerosities might differ in their relative weight or importance in numerical categorization tasks. Here, we report the results from two experiments which aim to provide direct evidence for the role of primitives for ONE and TWO in non-linguistic and linguistic categorization. We test participants whose native languages differ in how they express number morphology (English, Japanese, Slovenian). We find strong evidence for a preference to treat 1 distinctly across all populations, and in both task types. The evidence for categorizing 2 as distinct only appears in linguistic tasks, and is generally less strong, even in a population whose language morphologically encodes the dual.

Keywords: morphological number; numerical cognition; conceptual representations; linguistics

Background

Languages differ in how they categorize elements as a function of cardinality. Some languages, such as English, rely on both number words (e.g., *one*, *two*) and number morphology (singular vs. plural) to encode numerical distinctions. Other languages lack a systematic morphological distinction between singular and plural. Japanese, for example, does not morphologically mark number in this way, yet has number terms that allow speakers to refer to cardinalities (Nakanishi & Tomioka, 2004). Yet other languages employ more fine-grained number morphology relative to English. Slovenian, for example, makes a singular–dual–plural distinction, and, although it is much rarer, some languages, like Larike, distinguish singular, dual, trial, and plural (i.e., they have specific markings for sets of two and three) (Corbett, 2000).

Importantly, there appear to be limits on such morphological expression: no language has grammatical markers that systematically encode exact numerosities above four (i.e., no grammatical equivalent of an "exact five") (Corbett, 2000). Relatedly, some languages possess exact number terms only for cardinalities up to four or five, with no exact terms for larger quantities, e.g., Mundurukú (Pica et al., 2004). Crucially, however, no language lacks the ability to express smaller cardinalities altogether (Comrie, 2022).

Taken together, these cross-linguistic patterns suggest that numerosities in the range of 1–4 may be treated linguistically differently from larger numerosities. This distinction is also well established in cognitive science research on numerical representation. Studies on number recognition and naming have shown that small and large quantities are processed differently: small numerosities (typically 1–4) are associated with *subitizing*, whereas larger quantities require counting (Oyama et al., 1981; Trick and Pylyshyn, 1994). These differences have been demonstrated across a number of different tasks and populations (see Feigenson et al., 2004 for a review).[1]

Developmental research on both the existence of numerical concepts and the representations associated with number words further supports this picture. Children acquire *one*, *two*, *three*, and sometimes *four* successively, and are only later able to produce the corresponding cardinality to all number words (i.e., they become what is referred to as Cardinal Principle-knowers, see Carey 2009 and Carey and Barner 2019 for reviews).

Linguistic theories of systems of number terms have exploited the existence of conceptual primitives for small cardinalities (e.g., ONE, TWO, THREE), among other factors, to explain both cross-linguistic and developmental patterns in numerical expression (Denić & Szymanik, 2024; Piantadosi et al., 2012; Xu et al., 2020). However, the link between these conceptual primitives and number morphology is less well established. Morphological theories often rely on a reduced set of morphosemantic primitives that are treated as specific to linguistic representation, and that do not map directly onto numerical primitives (e.g., Harbour, 2014). Yet it remains possible that treating ONE, TWO, and THREE as primitives may account not only for number terms, but also for number morphology. This relies on appropriate assumptions, e.g., about whether these primitives are interpreted as denoting exact cardinality or not, and about their mapping to

[1] There is some debate as to whether very young, pre-verbal infants can reliably distinguish small quantities in the absence of additional cues (see Cantrell and Smith 2013 for a review), but is outside the scope of the present study.

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

forms in different linguistic domains (see Barner, 2017; Franzon et al., 2019 for discussion).

An important open question concerning the primitives underlying linguistic categorization of cardinality relates to potential asymmetries among these representations. This issue is especially relevant when relying on such primitives to explain number morphology. Typologically, number morphology displays a clear imbalance: singular–plural systems are far more common than systems that also include dual, which are in turn more common than systems with trial or paucal distinctions, the latter being extremely rare (Corbett, 2000; Skirgård et al., 2023).[2] This pattern suggests that if ONE, TWO, and THREE are conceptual primitives, representationally, they may not be equally weighted. Direct evidence that number primitives themselves are asymmetrically represented is currently lacking, beyond asymmetries observed in typology and order of acquisition. Moreover, the nature of these primitives remains poorly understood: While their impact on linguistic systems is clear, it remains an open question whether they are fundamentally linguistic or instead reflect more general cognitive representations.

Following an approach analogous to that of Maldonado et al. (2023) for primitives in the person domain, we ask whether conceptual primitives for ONE and TWO influence categorization in non-linguistic and linguistic tasks. Specifically, we investigate: (a) whether categorization provides evidence for distinct ONE and TWO primitives, insofar as stimuli corresponding to these cardinalities are preferentially mapped onto dedicated categories; (b) whether there is an asymmetry between these primitives that mirrors typological patterns—that is, whether a bias for keeping singular distinct is more pervasive than a bias to keep dual distinct; (c) whether the presence and strength of these biases vary across speakers of languages with different number-morphological systems; and (d) whether the presence and strength of these biases differ between linguistic and non-linguistic tasks.

We present the results of two experiments investigating these questions. The first uses a non-linguistic categorization task to test for the presence of biases to keep either singular or dual distinct across speakers of different languages. The second tests whether these biases are strengthened when the categorization task involves a linguistic mapping.

Non-Linguistic categorization task

In order to tap into potential asymmetries affecting the non-linguistic categorization of quantities, we conducted a cross-language study using an odd-one-out task. Participants are given three cards and must decide which one is the odd one out. Because the task does not require mapping meanings onto any kind of form, it provides a useful method to test non-linguistic categorization (e.g., see Crutch et al., 2009; Landrigan & Mirman, 2018; Lupyan, 2009, for examples of odd-one-out/oddball tasks in other domains).

To explore whether participants' native language influences their categorization preferences, we tested speakers of languages with different number-marking systems: English, Slovenian, and Japanese. These languages differ in how they morphologically encode number, which may affect the salience of particular cardinalities.[3] English has a singular–plural system; Slovenian additionally marks the dual, arguably involving a stronger representation of the number two; Japanese does not consistently mark plural morphologically, relying instead on what has been described as an associative marker (Nakanishi & Tomioka, 2004). Comparing these three languages therefore allows us to examine how grammatical number systems may shape non-linguistic conceptualization of numerical categories.

We hypothesize that, across all three populations, speakers will show a tendency to distinguish one from not-one (a singular bias) and two from not-two (a dual bias). However, we expect the singular bias to be substantially stronger, consistent with the privileged status of singular marking in number morphology across languages. This experiment was preregistered here.

Design and Materials

Participants were presented with three cards and asked to select the card they believed did not belong. Each participant completed only a single such trial. Participants were informed that they would later be asked to explain the strategy underlying their choice. Each card depicted a set of avatars differing in number and color (see Fig. 1). Colors were never repeated across cards within a trial and were sampled from a fixed set of distinct colors.

The experiment included three conditions, which varied in the cardinalities displayed on the cards: 1v2v3, 1v3v5, and 2v4v6. Comparisons across these three conditions allow us to assess asymmetries between singular and dual biases. The contrast between 2v4v6 and 1v3v5 allows us to test the existence and relative strength of each bias. Importantly, the distance between the three cardinalities presented is held constant, and each set includes only even or only odd numbers. The contrast between 1v3v5 and 1v2v3 is less controlled, but provides a more direct test of the potential interaction between ONE and TWO primitives. In particular, it may help us determine whether the singular bias is weaker when a competing dual bias is present.

In all conditions, avatars were spatially arranged in dice-like patterns, which are known to facilitate rapid recognition of cardinalities without counting (i.e., subitizing), including for quantities above four (Benoit et al., 2004). Cards were randomly ordered from left to right for each participant.

[2] An asymmetry between dual and trial can also be observed in the domain of quantification and number terms. There is evidence for a dual representational primitive, both in languages with dedicated lexical items (e.g., *both* in English) and in languages without such lexical items (e.g., French). There is no parallel evidence for a trial alternative (Chemla, 2007; Jeretič et al., 2024).

[3] We focus on languages that differ in their morphological systems (rather than in their systems of number terms) since we are ultimately interested in how number representation influences number morphology.

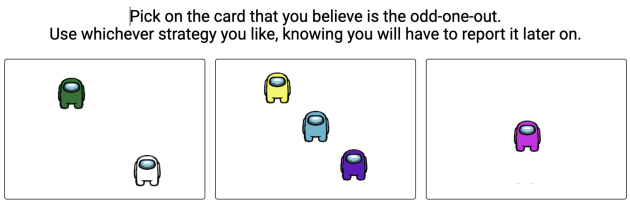

Figure 1: Illustration of an odd-one-out trial in the 1v2v3 condition (English)

After selecting the odd-one-out card, participants were required to provide a written explanation for their choice. At the end of the experiment, participants completed a questionnaire assessing the frequency with which they used their native language. All instructions and other text were in participants' native language.

Participants

A total of 429 participants, native speakers of English, Slovenian, and Japanese, took part in the odd-one-out experiment that lasted approximately 2 minutes. English speakers (n = 119) were recruited entirely through Prolific. Due to the low number of available speakers on Prolific, Japanese participants were recruited through both Prolific (n = 45) and Prime Panels (n = 63). For similar reasons, Slovenian speakers were recruited partly through Prolific (n = 94) and partly via Reddit groups (n = 108), specifically Slovenian-focused communities. Participants recruited trough Prolific and Prime Panels were paid at a rate of £10 the hour. Participants recruited through Reddit were required to complete language-check questions at the beginning of the experiment to ensure that they were fluent speakers of Slovenian. They were not paid for their participation. These different recruitment strategies resulted in the following distribution of participants across conditions: English: [1v2v3: 38], [1v3v5: 40], [2v4v6: 41]; Slovenian: [1v2v3: 31], [1v3v5: 84], [2v4v6: 87]; Japanese: [1v2v3: 37], [1v3v5: 35], [2v4v6: 36].

Results and Discussion

Recall that there are three experimental conditions: 1v2v3, 1v3v5, and 2v4v6. For the analysis, participants' responses were recoded into three quantity-based categories: S (small), M (medium), and L (large). For example, in 1v2v3 the lowest quantity is 1, whereas in 2v4v6 it is 2. Fig. 2 shows the proportion of participants in each language group and condition who selected each category.

We aim to determine whether the smallest quantity category (S) is selected more frequently than the other two options (M and L) when the smallest value is 1 (in conditions 1v3v5 and 1v2v3) than when it is 2 (2v4v6). Such a result would be consistent with a stronger singular bias. We also examine whether this pattern holds independently of language.

To analyze the data, we fit a Bayesian multicategorical

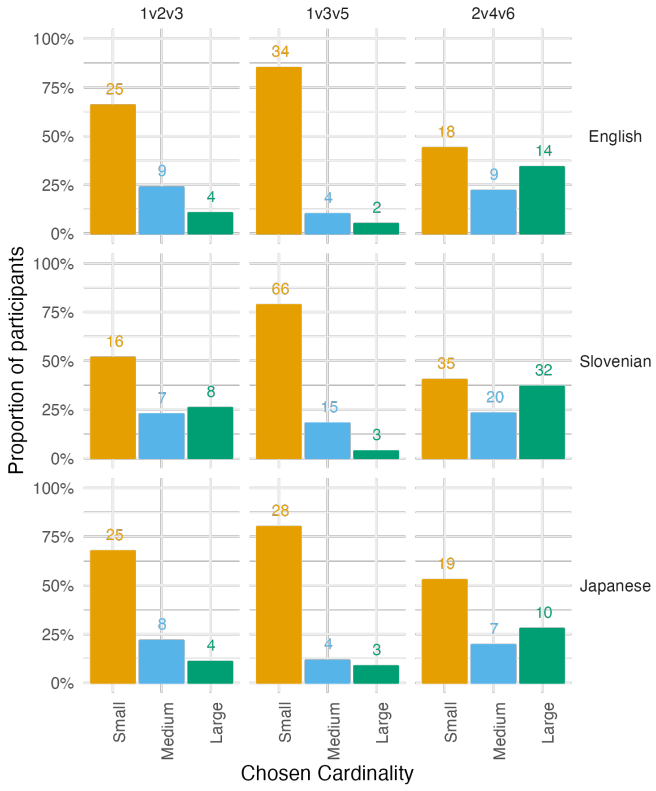

Figure 2: Proportion choice of the small (S), medium (M), and large (L) cardinality as odd-one-out across conditions and language groups. The numbers above the bars indicate the absolute frequencies corresponding to the percentages.

model using brms (Bürkner, 2017) predicting the probability of selecting each response category (S, M, and L) as a function of Condition (1v3v5, 1v2v3, 2v4v6) and Language (English, Slovenian, Japanese). The model treated S as the baseline category, and all comparisons are done with respect to L. Both Condition and Language were treatment-coded, with 1v3v5 and English serving as the baseline levels. No random effects were included, as each participant contributed only a single trial. We specified flat priors for the intercept and Student-t prior (DF = 3, $\mu = 0$, $\sigma = 2.5$) for the fixed effects. We report model estimates whenever there is evidence for a difference.

Model results show that, in Condition 1v3v5 in English, the log-odds of choosing either M or L are substantially lower than those of choosing S (M vs. S: $\beta = -2.23$, SE $= 0.54$, 95% CI $[-3.17, -1.38]$, $P(\beta < 0) = 1$; L vs. S: $\beta = -3.07$, SE $= 0.83$, 95% CI $[-4.55, -1.92]$, $P(\beta < 0) = 1$). No reliable differences were observed between English, Slovenian, and Japanese. These results indicate that, across all three languages, participants preferentially selected 1 as the odd-one-out relative to 3 and 5. This is consistent with a singular bias across all groups.

In Condition 2v4v6 in English, the log-odds of choosing S

relative to both M and L decrease substantially compared to 1v3v5 (M vs. S: $\beta = 1.51$, SE $= 0.69$, 95% CI $[0.40, 2.67]$, $P(\beta > 0) = .98$; L vs. S: $\beta = 2.82$, SE $= 0.90$, 95% CI $[1.50, 4.41]$, $P(\beta > 0) = 1$). No differences were observed between English, Slovenian, and Japanese. This indicates that, in Condition 2v4v6, participants in all three language groups behave differently when the smallest available number is 2 rather than 1; they do not exhibit the same bias toward selecting 2 as the odd-one-out in 2v4v6 as they do toward selecting 1 in 1v3v5. To further assess this pattern, we refit the model using an alternative contrast coding that directly compared the three response categories within Condition 2v4v6; this analysis revealed no reliable preference for S over either M or L in this condition (i.e., no evidence for a dual bias). This interpretation is supported by the qualitative responses collected at the end of the experiment, in which participants reported the strategies they used to choose the odd one out. Reported strategies ranged from "fewer avatars", "many", and "middle", to "random," as well as other strategies that do not express precise cardinalities. Taken together, these responses suggest that TWO may not consistently stand out as a conceptual primitive relative to other cardinalities.

Finally, in the English 1v2v3 condition, there is evidence for a higher probability of choosing M over S (i.e., selecting 2 rather than 1 as the odd-one-out) relative to the 1v3v5 condition, but little evidence for a difference between L and S (M vs. S: $\beta = 1.17$, SE $= 0.68$, 95% CI $[0.10, 2.31]$, $P(\beta > 0) = .96$; L vs. S: $\beta = 1.13$, SE $= 0.99$, 95% CI $[-0.40, 2.87]$, $P(\beta > 0) = .88$). Once again, no evidence for differences across languages was observed. These results indicate that participants are more likely to select 2 as the odd-one-out in 1v2v3 than they are to select 3 in 1v3v5. Still, after refitting to compare S, M, and L within 1v2v3, the model indicates that 1 (S) is still much more likely to be chosen as the odd-one-out than either 2 (M) or 3 (L) ($P(\beta < 0) = 1$). A closer examination of participants' qualitative description of their strategies at the end of the experiment suggests that the slight increase in 2 responses in Condition 1v2v3 is related to properties other than primitives, such as the odd–even contrast.

To summarize, as predicted, we find evidence for a *singular* bias shared across languages: speakers of different languages appear to use ONE as a salient reference point (or primitive) for categorizing cardinalities. This bias is observed in both the 1v3v5 and 1v2v3 conditions, although it is weaker in the latter. Crucially, we find little evidence for a strategy that reliably distinguishes 2 from quantities greater than 2, suggesting that participants may not represent TWO in a way fully analogous to their representation of ONE. Importantly, this pattern also appears to hold for speakers of Slovenian, a language that grammatically encodes dual number.

We next ask whether this singular bias has the same strength in a linguistic task. Here, one might expect that speakers of languages with a grammatical representation of duality (like Slovenian speakers) might also show a preference for treating 2 differently from quantities greater than 2 (i.e., a *dual* bias).

Linguistic categorization task

In the odd-one-out task, we found evidence only for a singular bias. This raises the question of whether a dual bias might emerge more clearly in a task that explicitly engages linguistic representations. This is an interesting question as it might suggest a domain-specific representation for the TWO primitive. We hypothesized that mapping numerosities to suffix-like labels could encourage participants to rely on linguistic representations. The one-shot design of the task remained the same, but we only target conditions 1v3v5 and 2v4v6, and English and Slovenian speakers.

Because Slovenian has grammatical dual and English does not, one possible prediction is that both groups would show a singular bias in the 1v3v5 condition, but only Slovenians would show a dual bias in the 2v4v6 condition. Specifically, both groups should be most likely to treat 1 as distinct from 3 and 5, while only Slovenians should be more likely to treat 2 as distinct from 4 and 6. This would be consistent with previous developmental results showing that Slovenian speakers process duality different form English speakers (Almoammer et al., 2013). If, instead, a dual bias plays a role even in languages that do not morphologically express duality (as has been explored in Jeretič et al., 2024), we would expect to observe a dual bias in both languages, but only in this task, which explicitly engages linguistic abilities. The experiment was pre-registered here.

Participants

100 native English speakers and 65 Slovenian speakers were recruited via Prolific. The experiment lasted approximately 2 minutes, and all participants were paid at a rate of £10/hour. Condition assignment was random and resulted in the following distribution of participants: English: [1v3v5: 50], [2v4v6: 50]; Slovenian: [1v3v5: 37], [2v4v6: 28].

Design and Materials

Participants were presented with three images, each depicting a different number of avatars. Two images were presented with a label underneath, and the third was unlabeled (see Fig. 3). Participants were asked to choose which of the two available labels they would assign to the unlabeled image. Each participant completed only a single trial. Participants were informed that they would later be asked to explain the strategy underlying their choice. In a debriefing questionnaire at the end of the experiment, participants were asked to explain how they chose the label for the unlabeled meaning.

The experiment had two conditions: 1v3v5 and 2v4v6. As in the odd-one-out experiment, conditions differed in the cardinalities displayed on the cards. In both conditions, the unlabeled image corresponded to the intermediate cardinality (i.e., 3 in the 1v3v5 condition and 4 in the 2v4v6 condition). The design is closely related to the odd-one-out task, but differs in that discontinuous groupings were not allowed. That is, in contrast to the non-linguistic odd-one-out experiment,

Figure 3: Illustration of a labeling trial in the 2v4v6 condition (English).

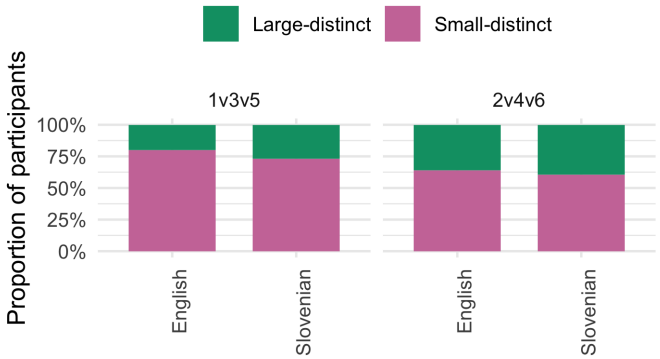

Figure 4: Proportion of participants keeping a distinct label for the small numerosity and for the large numerosity across conditions and languages.

we did not allow participants to linguistically group the small and large numerosities.[4]

Images were identical to those used in the previous odd-one-out experiment, and for each participant were presented in random order on the screen. Labels were constructed from a single root *lur* combined with either a suffix (*-vi*, *-sa*, or *-gu*). Stems and suffixes were selected to ensure they do not form existing words in either English or Slovenian. In Slovenian, dual and plural markers depend on noun class and gender. While none of our suffixes have a one-to-one correspondence with specific Slovenian markers, the vowels *-a* and *-i* are involved in dual or plural marking for certain classes—for example, *stol–stola–stoli* ('chair') and *hiša–hiši–hiše* ('house'). To ensure that no systematic bias was introduced by these forms, suffixes were randomly assigned to participants.

The same labels were used for both language populations, all instructions and other text were in participants' native language.

Results and discussion

Recall that participants had to choose whether to keep a distinct label for the smaller or the larger cardinality, with conditions varying in which cardinalities were presented. We coded responses as *small-distinct* and *large-distinct*. In the 1v3v5 condition, *small-distinct* responses are compatible with a singular bias, and in the 2v4v6 condition, with a dual bias. Fig. 2 shows the proportion of participants who kept a distinct label for the smaller versus the larger cardinality, by condition and language.

We fit a Bayesian binomial model predicting the probability of a *small-distinct* response as a function of Condition (1v3v5, 2v4v6) and Language (English, Slovenian). Condition and Language were treatment-coded, with 1v3v5 and English as baseline levels. No random effects were included, as each participant contributed only a single trial. We specified a Student-t prior (df = 6, $\mu = 0$, $\sigma = 1.5$) for the intercept and fixed effects.

[4]This would be a highly unusual grouping choice single non-contiguous categories would be grouped, see Lee (2020) for further evidence that such a grouping is dispreferred in a number morphology learning task.

For the baseline condition (1v3v5, English), the posterior distribution indicates a higher probability of *small-distinct* than *large-distinct* responses (intercept $= 1.33$, SE $= 0.3$, 95% CI $[0.7, 1.9]$, $P(\beta > 0) = 1$). There is little evidence for a difference between English and Slovenian speakers in this condition ($P(\beta > 0) = 0.4$). This reveals that speakers of both languages prefer to maintain a distinct label for 1 and another label for quantities greater than 1, consistent with a singular–plural categorization. This preference is also reflected in what participants report in the debriefing questionnaire at the end of the experiment.

In the 2v4v6 Condition, the estimated log-odds of *small-distinct* decreases relative to the baseline ($\beta = -0.7$, SE $= 0.4$, 95% CI $[-1.5, 0.1]$), with most posterior mass below zero ($P(\beta < 0) = 0.96$). This indicates that participants in Condition 2v4v6 tend to maintain a distinct label for 2 less often than participants in Condition 1v3v5 do for 1. In other words, the dual bias appears weaker than the singular bias. The interaction between Condition and Language is centered near zero ($P(\beta > 0) = 0.59$), suggesting little evidence that the effect of condition differs by language. Thus, Slovenian speakers do not appear to show a stronger dual bias in this task than English speakers, despite having grammatical dual marking.

To evaluate evidence for a dual bias within each language, we re-ran the model using the 2v4v6 English condition as the baseline. The posterior still favors *small-distinct* responses (intercept $= 0.6$, SE $= 0.29$, 95% CI $[0.05, 1.18]$, $P(\beta > 0) = 1$), again with little evidence for a language difference. This suggests that, although weaker than the singular bias observed in Condition 1v3v5, there is evidence for a dual bias in Condition 2v4v6 that appears across languages.

General Discussion

We investigated the role of the number primitives/concepts ONE and TWO across two categorization tasks: one non-linguistic and the other involving mapping of linguistic forms

to meanings. In particular, we asked whether there is an asymmetry in how these concepts are represented, and whether such an asymmetry plays a role in categorization across these two distinct tasks. We also asked whether patterns of behavior were distinct across groups of speakers whose native language differs in how number is morphologically marked: English (singular-plural distinction), Slovenian (singular-dual-plural), Japanese (no morphological marking of number, associative marking only).

This work constitutes a first step toward understanding the role these primitives may play in the linguistic categorization of cardinalities, including morphological number. Typological patterns in morphological number reveal clear asymmetries: the singular–plural distinction—arguably built on a ONE primitive—is far more widespread than the singular–dual–plural distinction, which involves a TWO primitive.

We first employed a non-linguistic odd-one-out task, based on the assumption that these number concepts, and the biases they may give rise to, could manifest even in the absence of language. Evidence of such effects would suggest that these concepts are not exclusively linguistic in nature. We tested speakers of different languages under the assumption that these primitives should influence performance even for speakers whose languages do not grammatically encode them. We found evidence for a singular bias in this task: participants tended to categorize cards depicting a cardinality of one as belonging to a different category than cards depicting other cardinalities, even when all cardinalities were small (Condition 1v2v3), and across different ratio structures. This pattern was observed across languages, including Japanese, which does not have a singular-plural productive distinction. In this experiment, however, we found no clear evidence for a dual bias: a TWO primitive didn't seem to be influencing categorization behavior.

We then adapted our odd-one-out design so that it included a labeling task in order to test whether a dual bias might emerge when language is involved. Such a bias might be expected to appear more clearly in speakers of languages that grammatically encode dual number, such as Slovenian (see Almoammer et al., 2013). In this labeling experiment, we once again found evidence for a singular bias: participants tended to maintain a separate label for ONE and another label for NOT-ONE, and often described their strategy as marking a singular–plural distinction. We also found evidence for a weaker dual bias. Participants were more likely to maintain one distinct label for TWO and another for quantities greater than TWO than to group 2 and 4 together and contrast them with 6.

Crucially, we found no quantitative differences between English and Slovenian, suggesting that both populations exhibit similar biases in the linguistic task, despite only one of the languages having systematic grammatical dual marking.

Our results are consistent with the idea that conceptual primitives involved in representing numerosities may not be weighted equally (i.e., ONE is weighted more than TWO). That conceptual primitives might carry different weights has also been shown in other domains relevant to morphology. For example, in the domain of person, Maldonado et al. (2023) use a non-linguistic card sorting task to show that participants tend to sort cards involving the speaker (in a previously shown interaction) together. In other words, cards with the speaker were treated as more similar to each other than cards with other speech act participants (e.g., the addressee and others). This is consistent with the idea that a SPEAKER primitive is more heavily weighted, or carries more importance, compared to other primitives, like ADDRESSEE. Nevertheless, some important caveats remaining around the interpretation of our results, which point to clear directions for future research.

First, although the quantitative results do not differ across languages, there are notable differences in participants' qualitative reports. Most Slovenian speakers explicitly described making a TWO vs. MORE-THAN-TWO distinction in the 2v4v6 condition. By contrast, most English speakers did not, instead referring to color-based or arrangement-based strategies. Incidentally, this suggests that English participants are not taking the labels to pick up cardinalities, unlike Slovenian participants.

Second, the Slovenian sample is relatively small, which may limit statistical power. A larger sample might yield stronger effects for this group. Moreover, Slovenian speakers recruited via Prolific may not all speak varieties with the most systematic use of the dual (e.g., Central Slovenian). This factor was not controlled for in the present study and should be addressed in future work. As a result, some Slovenian participants may not differ substantially from English speakers in their linguistic behavior.

Nevertheless, it is striking that a dual-like pattern emerges even for English speakers. This raises the question of whether English may nevertheless exhibit a form of dual bias, or reflect a TWO primitive within its linguistic system, as observed above. Indeed, English quantifiers such as *both* or *either* specifically select for sets of cardinality 2, suggesting that the concept TWO operates somehow productively in the language (in a way that other numerosities do not). Alternatively, English speakers might be sensitive to other linguistically relevant distinctions with vague thresholds, such as the quantificational contrast between *few* and *many*; in the context of this experiment, this distinction may simply happen to align with the 2 vs. 4/6 contrast. A potential follow-up study manipulating the distance between numerosities might be able to test this possibility.

Note, however, that our results indicate that when the task is non-linguistic, there is little evidence for a dual bias across all three language populations tested. Such a bias may thus be primarily linguistic in nature, or considerably weaker outside of language. Our results do not provide direct evidence for a *conceptual* primitive for TWO (differing from Jeretič et al., 2024). Future research should examine additional languages with different number systems to determine whether

this asymmetry is robust and whether it can help explain the typological asymmetry observed in morphological number systems.

Acknowledgements

We thank Milica Denic, Wataru Uegaki, and Marie-Léa Le Clainche for their help with translations and data gathering, as well as the CogSci reviewers for their interesting observations. This project has received funding from the Agence Nationale de la Recherche (grant agreement ANR-24-CE28-746 to Mora Maldonado) and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 757643 to Jennifer Culbertson).

References

Almoammer, A., Sullivan, J., Donlan, C., Marušič, F., Žaucer, R., O'Donnell, T., & Barner, D. (2013). Grammatical morphology as a source of early number word meanings. *Proceedings of the National Academy of Sciences*, *110*(46), 18448–18453.

Barner, D. (2017). Language, procedures, and the non-perceptual origin of number word meanings. *Journal of child language*, *44*(3), 553–590.

Benoit, L., Lehalle, H., & Jouen, F. (2004). Do young children acquire number words through subitizing or counting? *Cognitive development*, *19*(3), 291–307.

Bürkner, P.-C. (2017). Brms: An r package for bayesian multilevel models using stan. *Journal of statistical software*, *80*, 1–28.

Cantrell, L., & Smith, L. B. (2013). Open questions and a proposal: A critical review of the evidence on infant numerical abilities. *Cognition*, *128*(3), 331–352.

Carey, S. (2009). Where our number concepts come from. *The Journal of philosophy*, *106*(4), 220.

Carey, S., & Barner, D. (2019). Ontogenetic origins of human integer representations. *Trends in cognitive sciences*, *23*(10), 823–835.

Chemla, E. (2007). French both: A gap in the theory of antipresupposition. *Snippets*, *15*, 4–5.

Comrie, B. (2022). The arithmetic of natural language: Toward a typology of numeral systems. *Macrolinguistics*, *10*(1), 1–35.

Corbett, G. G. (2000). *Number*. Cambridge University Press.

Crutch, S. J., Connell, S., & Warrington, E. K. (2009). The different representational frameworks underpinning abstract and concrete knowledge: Evidence from odd-one-out judgements. *Quarterly journal of experimental psychology*, *62*(7), 1377–1390.

Denić, M., & Szymanik, J. (2024). Recursive numeral systems optimize the trade-off between lexicon size and average morphosyntactic complexity. *Cognitive Science*, *48*(3), e13424.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in cognitive sciences*, *8*(7), 307–314.

Franzon, F., Zanini, C., & Rugani, R. (2019). Do non-verbal number systems shape grammar? numerical cognition and number morphology compared. *Mind & Language*, *34*(1), 37–58.

Harbour, D. (2014). Paucity, abundance, and the theory of number. *Language*, *90*(1), 185–229.

Jeretič, P., Gonzalez, A., Bassi, I., Yatsushiro, K., & Sauerland, U. (2024). Dual as a core concept and the pronounceability of alternatives. *Proceedings of the Amsterdam Colloquium*, 181–187.

Landrigan, J.-F., & Mirman, D. (2018). The cost of switching between taxonomic and thematic semantics. *Memory & cognition*, *46*(2), 191–203.

Lee, N. (2020). Learning (im)possible number syncretisms: Investigating innate featural representations. *Proceedings of the 50th Annual Meeting of the North East Linguistic Society*.

Lupyan, G. (2009). Extracommunicative functions of language: Verbal interference causes selective categorization impairments. *Psychonomic bulletin & review*, *16*(4), 711–718.

Maldonado, M., Zaslavsky, N., & Culbertson, J. (2023). Evidence for a language-independent conceptual representation of pronominal referents. *Proceedings of the annual meeting of the cognitive science society*, *45*(45).

Nakanishi, K., & Tomioka, S. (2004). Japanese plurals are exceptional. *Journal of East Asian Linguistics*, *13*(2), 113–140.

Oyama, T., Kikuchi, T., & Ichihara, S. (1981). Span of attention, backward masking, and reaction time. *Perception & Psychophysics*, *29*(2), 106–112.

Piantadosi, S. T., Tenenbaum, J. B., & Goodman, N. D. (2012). Bootstrapping in a language of thought: A formal model of numerical concept learning. *Cognition*, *123*(2), 199–217.

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an amazonian indigene group. *Science*, *306*(5695), 499–503.

Skirgård, H., Haynie, H. J., Blasi, D. E., Hammarström, H., Collins, J., Latarche, J. J., Lesage, J., Weber, T., Witzlack-Makarevich, A., Passmore, S., Chira, A., Maurits, L., Dinnage, R., Dunn, M., Reesink, G., Singer, R., Bowern, C., Epps, P., Hill, J., ... Gray, R. D. (2023). Grambank reveals the importance of genealogical constraints on linguistic diversity and highlights the impact of language loss. *Science Advances*, *9*(16).

Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? a limited-capacity preattentive stage in vision. *Psychological review*, *101*(1), 80.

Xu, Y., Liu, E., & Regier, T. (2020). Numeral systems across languages support efficient communication: From approximate numerosity to recursion. *Open Mind*, *4*, 57–70.