Journal of Numerical Cognition

PsychOpen GOLD

Empirical Research

All Roads Lead to Rome: Semantic Priming Between Language and Arithmetic

Sofía Castro[a], Pedro Macizo[bc]

[a] *Institute of Psychology, Jagiellonian University, Kraków, Poland.* **[b]** *Departamento de Psicología Experimental, Facultad de Psicología, Universidad de Granada, Granada, Spain.* **[c]** *Mind, Brain and Behaviour Research Center (CIMCYC), Granada, Spain.*

Journal of Numerical Cognition, 2021, Vol. 7(1), 42–65, https://doi.org/10.5964/jnc.6167

Received: 2020-02-21 • **Accepted:** 2021-02-02 • **Published (VoR):** 2021-03-31

Corresponding Author: Pedro Macizo, Departamento de Psicología Experimental, Facultad de Psicología, Universidad de Granada, Campus de Cartuja, s/n. 18071, Granada, Spain. Phone: +34686103200. E-mail: pmacizo@ugr.es

Supplementary Materials: Data, Materials [see Index of Supplementary Materials]

Abstract

This study evaluated the existence of universal principles of cognition, common to language and arithmetic. Specifically, we analysed cross-domain semantic priming between affirmative sentences and additions, and between negative sentences and subtractions. To this end, we developed and tested a new priming procedure composed of prime sentences and target arithmetic operations. On each trial, participants had to read an affirmative or negative sentence (e.g., "The circle is red", "The square is not yellow") and select, between two images, the one that matched the meaning of the sentence. Afterwards, participants had to solve a one-digit addition or subtraction (e.g., 7 + 4, 6 − 3), either by selecting the correct result between two possible alternatives (Experiment 1), or by verbalizing the result of the operation (Experiment 2). We manipulated the task difficulty of both the sentences and the operations by varying the similarity between the response options for the sentence (Experiment 1 and 2), and the numerical distance between the possible results for the operation (Experiment 1). We found semantic priming for subtractions, so that participants solved subtractions faster after negative versus affirmative sentences, and this effect was modulated by the difficulty of the operation. This is the first study reporting semantic priming effects between language and arithmetic. The outcomes of this work seem to suggest a shared semantic system between both cognitive domains.

Keywords

cross-domain semantic priming, language, arithmetic, task difficulty

In experimental psychology and cognitive science, cognition has been usually understood as a unitary system composed of different cognitive domains, interrelated but functionally independent (Fodor, 1983). In fact, a large number of studies have examined language comprehension (e.g., Brimo, Lund, & Sapp, 2018; Macizo, 2016; Meyer, 2018) and arithmetic processing (e.g., Hinault & Lemaire, 2016; Megías & Macizo, 2016a, 2016b; Schneider et al., 2017) in an independent manner. However, researchers are now wondering if these two cognitive domains might not be as different as it seemed at first, suggesting the idea of a domain-general structuring system that encompasses linguistic and arithmetic processing. This hypothesis makes sense as long as language and arithmetic share characteristics, such as the processing of symbolic representations (e.g., words, numbers) with specific meanings (e.g., concepts, magnitudes), combinatorial rules (e.g., a combination of letters to obtain words, a combination of operands to reach arithmetic results), etc.

One widely used method to evaluate the existence of general principles of human cognition is priming (Branigan & Pickering, 2017). A priming effect occurs when a characteristic stimulus – prime – modulates the processing of a subsequent stimulus – target – (Segal & Cofer, 1960). Recent cross-domain studies have focused on one specific type of

This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International License, CC BY 4.0, which permits unrestricted use, distribution, and reproduction, provided the original work is properly cited.

priming, structural priming, to evaluate integrational processing across language and arithmetic. In structural priming paradigms, a prime stimulus with a particular structure facilitates the processing of a target stimulus that shares the same structure (Bock, 1986, 1989). Using a structural priming paradigm, Scheepers et al. (2011) asked participants to solve an arithmetic operation with a high-attachment or a low-attachment structure (e.g., 80 − (9 + 1) × 5 vs. 80 − 9 + 1 × 5), and to complete, afterwards, an ambiguous sentence (e.g., "The tourist guide mentioned the bells of the church that…"). To finish the sentence, participants could add a relative clause attached to "the bells of the church" (e.g., "the bells of the church that chimed out loudly", high-attachment structure), or to "church" (e.g., "the bells of the church that stood near the town hall"; low-attachment structure). The results revealed that the operations primed the structures participants selected to finish the sentence. In other words, participants completed the sentences using the same structure as the one present in the operation. More recently, Van de Cavey and Hartsuiker (2016) replicated these results in Dutch, and Scheepers, Galkina, Shtyrov, and Myachykov (2019) extended the findings to more complex structures in Russian.

Altogether, these studies show priming from arithmetic to language, but this facilitation effect also occurs from language to arithmetic. For example, in one of their experiments, Scheepers and Sturt (2014) instructed participants to solve an arithmetic operation after indicating if a word compound made sense (e.g., "bankrupt coffee dealer", "organic coffee dealer"). Both the arithmetic and the linguistic stimuli could have a right-branching structure (A [B C], as "bankrupt coffee dealer" or 25 − 4 × 3) or a left-branching structure ([A B] C, as "organic coffee dealer" or 25 × 4 − 3). They found that participants solved more operations correctly when their structure was congruent with the structure of the previous sentence, and this effect was stronger for right-branching structures. To explain this finding, the authors relied on the incremental-procedural account (Scheepers et al., 2011). According to this theory, people prefer to solve operations and read sentences that follow a left-to-right order and, in this manner, their performance is better with stimuli that respect this sequence. Scheepers and Sturt (2014) claimed that, because of the natural tendency of individuals for a left-to-right processing, left-branching arithmetical structures have already an advantage and do not benefit as much from the prime structure as right-branching arithmetical structures (inverse-preference effect)[1].

Alongside with the studies presented above, the idea of domain-general cognitive processing shared by language and arithmetic has been investigated also with children (Zeng, Mao, & Liu, 2018). These authors replicated the structural priming effect between language and arithmetic found by Scheepers et al. (2011) with Mandarin Chinese speakers – both adults and children – but observed that the effect was larger in the latter group. To explain the differences between the groups, the authors relied as well on the incremental-procedural account (Scheepers et al., 2011; Scheepers & Sturt, 2014): Adults would be more willing to change from a left-to-right to a right-to-left processing mode, while children would exhibit a strong tendency to solve operations incrementally due to their lack of expertise with different syntactic structures.

In the field of neuroscience, it is also possible to find evidence supporting the existence of a shared structural system between both cognitive domains. Makuuchi, Bahlmann, and Friederici (2012) reported activation in the pars opercularis when processing linguistic and arithmetic information and, later on, Nakai and Sakai (2014) observed that regions typically related to language processing (e.g., left inferior frontal gyrus, left supramarginal gyrus, bilateral intraparietal sulcus and precuneus) were activated when performing arithmetic tasks that contained a specific tree structure. More recently, Nakai and Okanoya (2018) studied the repetition suppression effect (i.e., reduction of neural activation in a specific brain region due to repeated exposure to specific stimuli) using a cross-linguistic syntactic priming paradigm. The authors presented pairs of sentences and operations that were syntactically congruent or incongruent, and their results revealed smaller activation of specific brain regions (e.g., bilateral inferior frontal gyrus) when the sentences and operations were congruent compared to when they were incongruent.

The evidence summarized in the previous paragraphs supports the idea of shared syntactic processing between language and arithmetic. However, research on the integrational processing of information between these two domains is not limited to syntactic structures, with several studies analysing the commonalities in the processing of semantic information between language and mathematics. For example, Bassok, Pedigo, and Oskarsson (2008) investigated the

1) The inverse-preference effect has been considered also an implicit-learning phenomenon (Pickering & Ferreira, 2008): Individuals would learn more about structures that are less well known to them, which, in turn, would increase the potential effect of the prime structure on the target response.

concept of semantic alignment, defined as the "analogical alignment of semantic and arithmetic relations". The authors suggested that, when computing operations such as "3 girls + 4 boys", an integration between the semantic categories and the arithmetic operation occurs. To test their hypothesis, participants received word pairs followed by two digits with a "+" sign in between. Right after, a target digit was displayed on the screen, and participants had to indicate if the target digit matched any of the digits previously presented. The results revealed a sum effect, meaning that the rejection of a digit that corresponded to the sum of the preceding numbers (e.g., rejection of 7 after the problem 5 tigers + 2 cheetahs) required more time, relative to the condition in which the digit was not the result of the previous addition sentence. However, this effect was only found when the word pairs were categorically related (as in the example with tigers and cheetahs). From these results, the authors concluded that the context provided by the semantic relationship between the prime words determined the activation of addition facts: Only when the context was relational (i.e., categorically related pairs), additions facts were activated.

Guthormsen et al. (2016) further analysed the concept of semantic alignment on an electrophysiological study using Event-Related Potentials (ERPs). They asked participants to respond to semantically aligned and misaligned arithmetic problems (e.g., "6 roses + 2 tulips = ?" vs. "6 roses + 2 vases = ?", respectively), and to verify if a set of semantically aligned and misaligned problem sentences were acceptable ("Twelve roses divided by three vases equals four", "Ten limes plus three bowls equals thirteen"). They focused on two ERP components that are usually elicited by linguistic conceptual disruptions, the N400 and the P600. The authors found a significant N400 effect when participants responded to misaligned additions, and a significant P600 effect when they judged semantically misaligned problem sentences.

The studies conducted by Bassok et al. (2008) and Guthormsen et al. (2016) support integrational semantic processing between language and arithmetic; however, not all evidence points in the same direction. For example, Ronasi, Fischer, and Zimmermann (2018) conducted a pioneer study in which they analysed bi-directional semantic priming effects between subtractions and exception phrases (EPs) with positive quantifiers (e.g., "Everybody except John") and between additions and EPs with negative quantifiers (e.g., "Nobody except John"). Participants had to judge if the sentences were meaningful or meaningless and indicate if the proposed result for the operations was correct. The authors did not find priming effects in any direction, that is, neither from language to arithmetic nor from arithmetic to language. One of the explanations offered by the authors for the lack of priming effects was that EPs with positive and negative quantifiers might not be equivalent to additions and subtractions.

Furthermore, other studies have reported dissociations in the processing of linguistic and arithmetic information. Monti, Parsons, and Osherson (2012) conducted a functional Magnetic Resonance Imaging (fMRI) experiment in which participants had to indicate if pairs of linguistic arguments (e.g., "Z was paid X by Y" and "It was X that Y paid Z") and algebraic logical arguments (e.g., "X minus Y is greater than Z" and "Z plus Y is smaller than X") were semantically equivalent or grammatically correct. During the equivalence task, left perisylvian regions - typically related to language processing - were activated when processing the linguistic arguments, but not when processing the algebraic arguments. From a different field of study, research with neurological patients has also found dissociations between language and arithmetic, with cases of patients with language impairment but relatively preserved calculation abilities (Cappelletti, Butterworth, & Kopelman, 2001; Varley, Klessinger, Romanowski, & Siegal, 2005; but see also Baldo & Dronkers, 2007, for evidence supporting both dissociations and overlap, and Delazer, Girelli, Semenza, & Denes, 1999, for evidence supporting overlap).

Taken together, the studies presented in this section serve as an example of the controversy surrounding the existence of shared processing in language and arithmetic. On one hand, the majority of studies investigating syntactic integration between the two domains support processing overlap (e.g., Scheepers et al., 2011; Scheepers & Sturt, 2014; Van de Cavey & Hartsuiker, 2016), whereas the evidence on common semantic processing is mixed (Bassok, Pedigo, & Oskarsson, 2008; Guthormsen et al., 2016; but see also Monti, Parsons, & Osherson, 2012; Ronasi, Fischer, & Zimmermann, 2018). The heterogeneity in the results of the semantic studies signals the importance of examining the specific conditions under which semantic overlap between language and mathematics might occur.

The Current Study

Recent empirical research suggests the existence of general principles of human cognition that guide the processing of information in different domains, such as language and arithmetic. However, the available evidence is limited, and it is not clear which are the common principles that might underlie human cognition. In a unique study, Ronasi et al. (2018) evaluated cross-domain semantic priming between language and arithmetic using EPs with positive and negative quantifiers (e.g., "Everybody except John", "Nobody except John"), which were theoretically equivalent to additions and subtractions. Nonetheless, their results did not support a common processing of this type of sentences and operations. In the present study, we also analysed cross-domain semantic priming, but focusing on a different type of linguistic structures: affirmative and negative sentences. Specifically, we evaluate the semantic overlap between affirmative sentences and additions, and between negative sentences and subtractions.

When processing a simple affirmative sentence with a Subject-Verb-Complement structure, such as "The square is red", the information provided by the complement is added to the already processed subject. Following this example, the information provided by the adjective "red" (i.e., a colour) would be added to the concept of square, resulting in a final product of square + red (semantic affirmation). On the other hand, when processing an analogous negative sentence (e.g., "The square is not red"), the inclusion of new content limits the information initially processed. This time, "red" would provide also new semantic information to the concept of square, but the particle "no", would block the inclusion of this new information. Thus, the final interpretation of negative sentences implies the removal of red from the final meaning of the sentence (square − red; semantic negation). Similarly, in the arithmetic field, an addition (e.g., 5 + 3) could be conceived as a semantic affirmation while a subtraction (e.g., 5 − 3) could be considered as a semantic negation. Across two experiments, we analysed the relations between the concepts of affirmation and negation and the processing of additions and subtractions using a priming paradigm. In both experiments, participants processed first linguistic material (i.e., they read sentences and selected the image that matched the meaning of the sentence) and then arithmetic information (i.e., they solved additions and subtractions by selecting the correct result or by saying aloud the response). The overall prediction of this study was simple: If there are common semantic principles in language and arithmetic, arithmetic processing will be more efficient when preceded by semantically similar sentences (i.e., additions preceded by affirmative sentences and subtractions preceded by negative sentences).

Experiment 1

The aim of Experiment 1, was to determine whether the processing of affirmative and negative sentences influences the subsequent processing of simple arithmetic operations. We predicted that the resolution of an arithmetic operation would be modulated by the semantic information of the previous sentence. To test our hypothesis, we developed a new paradigm composed of two-stage trials. In Stage 1, participants read an affirmative sentence in Spanish (e.g., "El cuadrado sí es rojo", "The square is red") or a negative sentence (e.g., "El cuadrado no es rojo", "The square is not red"). Afterwards, they had to select, between two images, the one that described the content of the sentence previously read. For example, the sentence "The square is red" could be followed by two coloured shapes: a green square on the left and a red square on the right. In Stage 2, an addition (e.g., 5 + 3) or a subtraction (e.g., 5 − 3) was displayed on the screen, followed by two possible results. This time, participants had to select the correct result for the operation. For example, the addition 5 + 3 could be followed by two possible results: the Number 8 on the left and the Number 6 on the right.

Although it was not the main objective of our study, we explored the influence of task demands on priming effects. The results from the study by Scheepers and Sturt (2014) suggested that priming effects are modulated by task difficulty, reporting larger effects under difficult experimental conditions. A possible explanation of the relationship between task difficulty and the magnitude of priming is as follows: The cognitive load imposed by an easy task is low and thus, it can be successfully performed without the need of structural or semantic facilitation. On the contrary, a difficult task would overload the cognitive system and the benefit of priming would be more evident (see Kootstra & Rossi, 2017, for a similar prediction on structural priming benefits in language attrition). To investigate this idea, we manipulated the task difficulty by varying the similarity (for the sentence) and the numerical distance (for the operations) of the response options. For example, the colour of the two geometric shapes that followed the sentence "The square is red"

could be similar (e.g., a red and a maroon square), or dissimilar (e.g., a red and a blue square), and the two possible results for the operation 5 + 3 could be numerically close (i.e., two numbers in between; e.g., Number 8 and Number 6), or numerically far (i.e., four numbers in between; e.g., Number 8 and Number 4). Sentences with similar response options and operations with results that were numerically close would impose higher cognitive demands, compared to sentences with dissimilar response options and operations with numerically far results.

Overall, we predicted faster responses to additions preceded by affirmative versus negative sentences and faster responses to subtractions preceded by negative versus affirmative sentences. Moreover, if priming effects are influenced by the cognitive load of the task (Scheepers & Sturt, 2014), they would be more evident when the target operation was more demanding (i.e., close numerical distance between the response options). In addition, priming effects would be more likely to be observed when the prime sentence was more easily processed (i.e., dissimilar response options) because the prime would be less demanding and/or more salient during the performance of the task.

Method

Participants

We recruited 61 students from the University of Granada, speakers of Spanish, with an average age of 21 years (M = 20.98, SD = 2.10). A total of 49 participants were females and 12 participants were males. They received either academic credits or money in exchange for their participation. Each participant provided written informed consent before performing the experiment. All procedures used in this study were in accordance with the 1964 Helsinki declaration and its later amendments.

Materials and Design

In Experiment 1 we used a 2 x 2 x 2 x 2 within-participant design. We manipulated the sentence type (2 levels: affirmative, negative), the arithmetic operation (2 levels: addition, subtraction), the sentence similarity (i.e., the similarity between the colour of the correct and the incorrect figure proposed to describe the sentence; 2 levels: similar, dissimilar), and the numerical distance (i.e., the numerical distance between the correct and the incorrect result proposed for the arithmetic operation; 2 levels: close, far). The correspondence between the sentence similarity and the arithmetic operation was counterbalanced across participants. For one group, all additions were preceded by sentences with similar response options (similar condition), and all subtractions were preceded by sentences with very different response options (dissimilar condition), whereas for the other group, all additions were preceded by sentences in the dissimilar condition and all subtractions were preceded by sentences in the similar condition.

The experiment was programmed and controlled with E-Prime 2.0 (Schneider, Eschman, & Zuccolotto, 2002). The experimental task contained a total of 480 trials, each of which had two stages. Each trial was composed of a prime sentence (Stage 1) and a target arithmetic operation (Stage 2). A set of 240 trials contained an affirmative sentence (e.g., “El círculo sí es rojo”, “The circle is red”), and 240 trials contained a negative sentence (e.g., “El cuadrado no es amarillo”, “The square is not yellow”). In Spanish, to create the negative sentences used in our study, the adverb “no” is required. Thus, to match the syntactic structure of negative and affirmative sentences, the adverb “sí” was added to the affirmative sentences (note that all syntactic structures used in the study are plausible in Spanish). After the sentence, two geometric shapes with different colours were presented. We manipulated the similarity between the two response options depending on the proximity (in RGB values) of the figure colours. Thus, 120 trials within each sentence type (affirmative and negative) had two response options that were similar (e.g., one yellow square and one gold square, similar condition), and the other 120 trials had two response options that were very different (e.g., one yellow square and one red square, dissimilar condition). Therefore, we obtained four possible sentence combinations: affirmative/similar, affirmative/dissimilar, negative/similar, and negative/dissimilar. The geometric forms were coloured following the RGB colour codes. We combined 10 different geometric forms with 12 colours, obtaining 120 possible response options (see Figure 1 and Figure 2).

For the target arithmetic operations, we selected 240 additions and 240 subtractions randomly formed by combining one-digit operands from 1 to 9. Operations with a result of zero (e.g., 8 − 8) and subtractions with a negative result (e.g., 2 − 7) were not used. From these combinations, a total of 60 different additions and 35 different subtractions were

selected. The second operand was larger than the first operand (e.g., 2 + 5) in 27 of the additions, smaller than the first operand (e.g., 4 + 3) in 26 additions, and equal (e.g., 5 + 5) in the remaining 7 additions.

Two different numbers were presented as response options after the arithmetic operation, manipulating the numerical distance between the two response options (close and far). Hence, there were 120 additions and 120 subtractions with two close response options (numerical distance = 2; e.g., Number 5 and Number 7), and 120 additions and 120 subtractions whose response options were far from each other (numerical distance = 4; e.g., Number 2 and Number 6). In half of the operations, the correct response was larger than the distractor, and vice versa for all other operations. For Stage 2, there were four operation combinations: addition/close, addition/far, subtraction/close and subtraction/far. For both stages, the correct answer was presented on the right side of the screen in half of the trials and on the left side of the screen on the other half. The sequence of responses (correct-left, correct-right) was randomised between trials. The distribution of the operands and the response options (both as correct responses and as distractors) are shown in the Appendix A (Table A1 and Table A2). The complete set of operations is displayed in the Appendix A (Table A3).

Figure 1

Geometric Shapes Used in Experiment 1 and Experiment 2

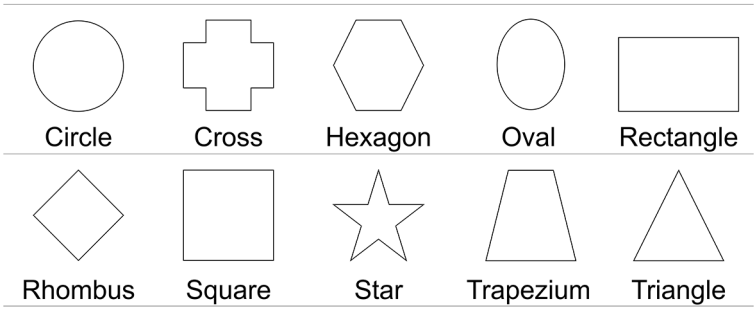


Figure 2

Colour Names, RGB Codes, and Colours Used in Experiment 1 and Experiment 2

Colour Name	RGB Code	Example	Colour Name	RGB Code	Example
Aquamarine	80, 200, 120		Olive Drab	53, 94, 59	
Beige	245, 222, 179		Orange	255, 112, 40	
Dark Blue	18, 10, 143		Light Blue	176, 196, 222	
Pink	255, 192, 203		Purple	197, 75, 140	
Gold	255, 215, 0		Yellow	255, 255, 0	
Red	255, 0, 0		Maroon	128, 0, 0	

The experimental task used in Experiment 1, raw and filtered data, the R scripts used to analyse the data as well as the results from the analyses, are freely accessible at the Open Science Framework (see Supplementary Materials).

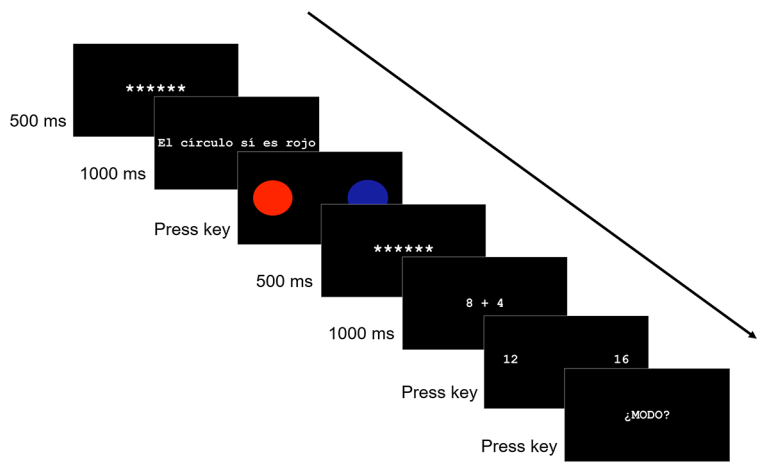
Procedure

Participants performed 12 experimental blocks with 40 trials on each set. Each trial contained a prime sentence that could be affirmative or negative (Stage 1) and a target arithmetic operation that could be an addition or a subtraction (Stage 2). Before each trial, a mask (*******) was displayed on the screen for 500 ms. We used five asterisks instead

of the classical fixation point (+) to avoid priming effect or interference with the sentence structure or the arithmetic operation. Once the mask disappeared, a sentence was presented for 1000 ms and then, two geometric forms were shown, one on the right and the other on the left side of the screen. Both figures matched the geometric form described in the previous sentence, and they differed only in colour. Participants had to choose the figure that matched the meaning of the previous sentence by pressing one of two keys (Z or M). Once they had answered Stage 1, the mask was displayed again for 500 ms, and then an addition or subtraction appeared on the screen for another 1000 ms. After the arithmetic operation, two numbers were shown, one on the right and one on the left side of the screen. Participants had to choose the correct answer for the operation by pressing one of two keys (Z or M). At the end of each trial, the prompt "¿MODO?" ("MODE?") was displayed, and participants had to report how they solved the operation by pressing one key from four possible alternatives (retrieval, counting, transformation, and others). Retrieval strategies comprised the recollection from memory ("When a problem such as 2 + 3 = is presented, you know from memory that 5 is the correct answer"). Non-retrieval strategies included counting ("When a problem such as 2 + 3 = is presented, you count mentally from 2… 3, 4 and 5 to get the answer"), transformation ("When a problem such as 2 + 3 = is presented, you decompose it in other easy problems, e.g., 2 + 2 + 1"), and other strategies different from those explained before (see Figure 3).

Figure 3

Example of an Experimental Trial From Experiment 1



Note. Stage 1: A sentence was displayed on the screen for 1000 ms, followed by two figures. Participants had to select the figure that matched the meaning of the sentence. Stage 2: An addition or a subtraction was displayed on the screen for 1000 ms, followed by two numbers. Participants had to choose the correct answer for the arithmetic operation. Lastly, participants had to press a key to indicate how they solved the operation (retrieval, counting procedure, transformation or other).

To ensure that all participants were familiar with the geometric forms and the colours, all stimuli were displayed on the screen before starting the experiment. White geometric forms were presented one-by-one in the centre of the screen, with their names written below. For the colours, coloured circles were displayed one-by-one in the centre of the screen with the name of each colour written below. Participants pressed the space bar to see the stimuli at their own pace. After participants had seen all stimuli, they completed a practice set to make sure they understood the experimental task (five trials composed of a prime sentence and a target operation). Experiment 1 lasted one hour approximately with small variations depending on the speed of the participants.

Data Cleaning and Analysis

Firstly, we eliminated trials with errors in Stage 1 (i.e., incorrect responses to sentences). After that, we removed univariate outliers in reaction times of Stage 1 following the procedure described by Tabachnick and Fidell (2001): Raw scores were converted to standard scores (*z*-scores) and data points that, after standardization, were 3 *SD* outside the normal distribution, were considered outliers. After removing outliers from the distribution, *z*-scores were calculated

again. The filter was applied in 10 recursive cycles. The percentage of outliers in Stage 1 (sentences) was 6.62%. The same procedure was used for Stage 2 (operations), in which the percentage of outliers was 13.29%.

After data cleaning, reaction times of both Stage 1 and Stage 2 were logarithmically transformed to address right-skewness, and linear mixed-effects models were conducted separately for each stage. Accuracy analyses were conducted using generalized mixed-effects models for binomial data, also separately for each stage. All analyses were implemented using the lme4 package (version 1.1-23, Bates, Mächler, Bolker, & Walker, 2015) in R (version 4.0.0). Estimated marginal means from the models were computed with the package ggeffects (version 0.14.3, Lüdecke, 2018), and plots were created with the package ggplot2 (version 3.3.0, Wickham, 2016). *P*-values for the models and pairwise comparisons were computed with the package lmerTest (version 3.1-2, Kuznetsova, Brockhoff, & Christensen, 2017) and collinearity was checked using the package performance (version 0.6.1, Lüdecke, Makowski, Waggoner, & Patil, 2020). For all models, the optimizer "bobyqa" was selected.

The model of Stage 1 included sentence type (affirmative, negative), sentence similarity (similar, dissimilar) and their interaction as fixed factors. Both categorical predictors were deviated coded (see Appendix B). The maximal model was built with participants and sentences as random intercepts. Random slopes for participants and sentences were included based on the design and hypotheses, and following the recommendations of Barr, Levy, Scheepers, and Tily (2013). In case of non-convergence, we removed recursively the smallest variance component(s) (with value zero or closer to zero) until convergence was achieved, in line with Bates, Kliegl, Vasishth, and Baayen (2018).

The model of Stage 2 included sentence type (affirmative, negative), sentence similarity (similar, dissimilar), operation type (subtraction, addition), numerical distance (close, far), the two-way interactions between sentence type and operation type, between operation type and sentence similarity, and between operation type and numerical distance. It also included the three-way interactions between sentence type, operation type and sentence similarity, and between sentence type, operation type and numerical distance. All categorical predictors were deviated coded (see Appendix B). The arithmetic operations used in this study differed in their problem size, meaning that additions had larger problem sizes than subtractions (see Table A2). Operations with larger problem sizes are linked to slower reaction times than operations with small problem sizes (for a review see Zbrodoff & Logan, 2005). Therefore, to control for the impact of the problem size in our results, we scaled the result of the operation and included it as a covariate in the model. The maximal model included random intercepts for participants, sentences, and operations. Random slopes for participants, sentences, and operations were included based on the experimental design and hypotheses, and following the recommendations of Barr et al. (2013)[2]. In case of non-convergence, the smallest variance component(s) were removed recursively following Bates et al. (2018). The structure of the maximal models of Stage 1 and Stage 2, as well as the models' refit are detailed in the R scripts available at the Open Science Framework (see Supplementary Materials).

Results

Accuracy Analysis

Stage 1. Prime Sentences — The main effect of sentence type was significant, b = 0.85, SE = 0.07, z = 12.31, p < .001, as well as the main effect of sentence similarity, b = 0.58, SE = 0.09, z = 6.65, p < .001. The accuracy was higher for affirmative than negative sentences, 97%, CI [96%, 97%] vs. 93%, CI [91%, 94%], respectively, and when the response options differed from each other compared to when they were similar, 96%, CI [95%, 97%] vs. 93%, CI [92%, 95%], respectively. Moreover, the interaction between sentence type and sentence similarity was significant, b = 0.49, SE = 0.14, z = 3.45, p < .001, with larger differences between the similar and dissimilar condition for affirmative sentences, $M_{similar}$ = 95%, CI [94%, 96%] vs. $M_{dissimilar}$ = 98%, CI [97%, 98%], b = -0.83, SE = 0.12, z = -6.83, p < .001, compared to negative sentences, $M_{similar}$ = 91%, CI [89%, 93%] vs. $M_{dissimilar}$ = 94%, CI [92%, 95%], b = -0.33, SE = 0.10, z = -3.20, p = .001.

2) We conducted additional analyses (Experiment 1 and 2) by running a model for reaction times with the structure of the additions as a new factor (magnitude of the first and second operands of the sum: larger + smaller vs. smaller + larger). In this model, due to the reduced number of observations, only the main effects of sentence and structure, as well as their interaction, were included as fixed effects. Problem size was also included as a covariate. The recommendations of Barr et al. (2013) and Bates et al. (2018) were followed during model specification and model fitting. The results of these analyses revealed that the effect of the structure of the operation was not significant and this variable did not interact with any other factor.

This result indicates that the similarity between the response options determined to a greater extent the processing of affirmative versus negative sentences. Since negative sentences were harder to perform, participants might have allocated more resources when these sentences were displayed, receiving less impact from similar versus dissimilar response options. However, it is important to take into consideration the small size of the differences when interpreting the interaction. The level of the reported confidence intervals was 95%, and the complete results from this analysis are shown in Table 1.

Table 1

Output of the Generalized Mixed-Effects Model of Accuracy in Stage 1 (Sentence), Experiment 1

Fixed effects	**Estimate**	***SE***	***p***	***z***	**95% CI**
Intercept	2.95	0.10	< .001	29.45	[2.76, 3.15]
Sentence type (SEN)	0.85	0.07	< .001	12.31	[0.72, 0.99]
Sentence similarity (SIM)	0.58	0.09	< .001	6.65	[0.41, 0.75]
SEN x SIM	0.49	0.14	< .001	3.45	[0.21, 0.77]
Random effects	**Variance**	***SD***			
Intercept | Subject	0.53	0.73			
SEN | Subject	0.03	0.19			
SIM | Subject	0.18	0.42			
SEN x SIM | Subject	0.16	0.40			
Intercept | Sentence	0.01	0.11			
SEN | Sentence	0.30	0.55			
SIM | Sentence	0.18	0.42			
SEN x SIM | Sentence	0.91	0.95			

Note. Categorical predictors were deviated coded.

Stage 2. Target Operations — The effect of operation type was significant, b = 1.15, SE = 0.29, z = 3.89, p < .001, as well as the effect of numerical distance, b = 0.49, SE = 0.14, z = 3.48, p < .001. The accuracy was larger for additions than subtractions, 99%, CI [99%, 99%] vs. 97%, CI [96%, 98%], respectively, and under the far condition versus the close condition, 99%, CI [98%, 99%] vs. 98%, CI [97%, 99%], respectively. However, no priming effects in accuracy were observed either for additions or subtractions. Lastly, the accuracy decreased as a function of the problem size, b = -0.54, SE = 0.14, z = -3.79, p < .001. The level of the reported confidence intervals was 95%. The complete results are shown in Table 2.

Table 2

Output of the Generalized Mixed-Effects Model of Accuracy in Stage 2 (Operation), Experiment 1

Fixed effects	**Estimate**	***SE***	***p***	***z***	**95% CI**
Intercept	4.22	0.16	< .001	25.70	[3.90, 4.54]
Sentence type (SEN)	-0.05	0.09	.603	-0.52	[-0.23, 0.14]
Operation type (OP)	1.15	0.29	< .001	3.89	[0.57, 1.72]
Sentence similarity (SIM)	-0.01	0.11	.953	-0.06	[-0.23, 0.22]
Numerical distance (DIS)	0.49	0.14	< .001	3.48	[0.21, 0.76]
SEN x OP	-0.07	0.16	.672	-0.42	[-0.38, 0.24]
OP x SIM	-0.08	0.58	.885	-0.14	[-1.22, 1.05]
OP x DIS	-0.14	0.28	.618	-0.50	[-0.69, 0.41]
SEN x OP x SIM	0.35	0.37	.345	0.95	[-0.38, 1.07]
SEN x OP x DIS	0.12	0.35	.736	0.34	[-0.57, 0.81]
Problem size	-0.54	0.14	< .001	-3.79	[-0.82, -0.26]

Random effects	Variance	*SD*
Intercept | Subject	0.78	0.88
SEN | Subject	0.00	0.00
Problem size | Subject	0.04	0.20
OP x SIM | Subject	1.87	1.37
SEN x OP x SIM | Subject	0.94	0.97
SEN x OP x DIS | Subject	0.98	0.99
Intercept | Sentence	0.00	0.05
SEN | Sentence	0.05	0.23
SIM | Sentence	0.25	0.50
Intercept | Operation	0.00	0.01
OP | Operation	1.34	1.16
DIS | Operation	0.47	0.68
Problem size | Operation	0.01	0.07
OP x SIM | Operation	0.50	0.71
OP x DIS | Operation	1.31	1.14

Note. Categorical predictors were deviated coded, and problem size was scaled. The smallest variance components were removed recursively until convergence was achieved (Bates, Kliegl, Vasishth, & Baayen, 2018). Control variables are shown in grey.

Latency Analysis

Stage 1. Prime Sentences — The model revealed a significant effect of sentence type, b = -0.17, SE = 0.01, t = -12.17, p < .001, with slower reaction times for negative than affirmative sentences, 827 ms, CI [782, 875] vs. 695 ms, CI [658, 735], respectively. The main effect of sentence similarity was also significant, b = -0.18, SE = 0.01, t = -13.87, p < .001, with slower responses for the similar condition compared to the dissimilar condition, 827 ms, CI [783, 874] vs. 695 ms, CI [657, 734], respectively. The interaction between sentence type and sentence similarity did not reach significance (p > .05). The level of the reported confidence intervals was 95%. The complete pattern of results from this analysis is reported in Table 3.

Table 3

Output of the Linear Mixed-Effects Model of Reaction Times in Stage 1 (Sentence), Experiment 1

Fixed effects	Estimate	*SE*	*p*	*t*	95% CI
Intercept	6.63	0.03	< .001	240.44	[6.58, 6.69]
Sentence type (SEN)	-0.17	0.01	< .001	-12.17	[-0.20, -0.15]
Sentence similarity (SIM)	-0.18	0.01	< .001	-13.87	[-0.20, -0.15]
SEN x SIM	0.02	0.02	.406	0.83	[-0.03, 0.06]
Random effects	**Variance**	***SD***			
Intercept | Subject	0.04	0.21			
SEN | Subject	0.00	0.06			
SIM | Subject	0.00	0.04			
SEN x SIM | Subject	0.00	0.03			
Intercept | Sentence	0.00	0.00			
SEN | Sentence	0.03	0.18			
SIM | Sentence	0.01	0.09			
SEN x SIM | Sentence	0.08	0.28			
Residual	0.12	0.34			

Note. Reaction times were logarithmically transformed and categorical predictors were deviated coded.

Stage 2. Target Operations — The main effect of operation type[3] was significant, b = -0.10, SE = 0.02, t = -4.65, p < .001, with faster responses for additions than for subtractions, 511 ms, CI [485, 538] vs. 566 ms, CI [539, 595], respectively. The interaction between sentence type, operation type, and numerical distance was also significant (see

Figure 4), b = 0.06, SE = 0.02, t = 2.77, p = .009. Pairwise comparisons revealed that subtractions were solved faster after negative vs. affirmative sentences, but only when the numerical distance was close (i.e., when the numerical distance between the two response options for the operation was 2, e.g., numbers 2 and 4 as response alternatives), b = -0.02, SE = 0.01, z = -2.17, p = .030. On average, the reaction time for subtractions with close response options preceded by negative sentences was 560 ms, CI [532, 589], and by affirmative sentences was 570 ms, CI [542, 600]. There were no significant differences between additions preceded by affirmative or negative sentences under the close or the far distance condition (i.e., large when the numerical distance between the two response options was 4, e.g., numbers 2 and 6 as response alternatives). Apart from our effects of interest, problem size effect was significant, b = 0.10, SE = 0.01, t = 9.67, p < .001, with an increase in reaction times as a function of the problem size (the classical problem-size effect, Zbrodoff & Logan, 2005). The level of the reported confidence intervals was 95%. The complete pattern of results is present in Table 4.

Table 4

Output of the Linear Mixed-Effects Model of Reaction Times in Stage 2 (Operation), Experiment 1

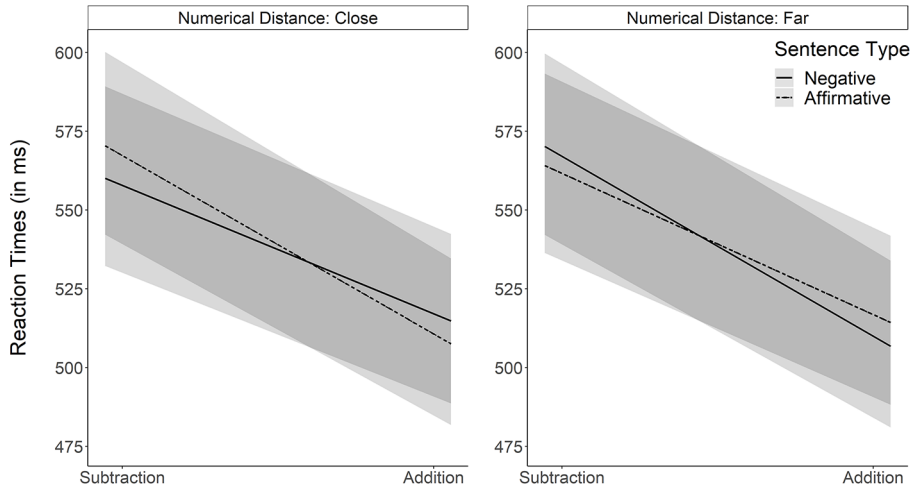
Fixed effects	**Estimate**	***SE***	***p***	***t***	**95% CI**
Intercept	6.30	0.02	< .001	272.78	[6.25, 6.34]
Sentence type (SEN)	0.00	0.01	.711	0.37	[-0.01, 0.01]
Operation type (OP)	-0.10	0.02	< .001	-4.65	[-0.15, -0.06]
Sentence similarity (SIM)	-0.00	0.00	.449	-0.76	[-0.01, 0.01]
Numerical distance (DIS)	0.00	0.00	.802	0.25	[-0.01, 0.01]
SEN x OP	-0.00	0.01	.633	-0.48	[-0.02, 0.01]
OP x SIM	0.02	0.09	.809	0.24	[-0.15, 0.20]
OP x DIS	-0.00	0.01	.604	-0.52	[-0.02, 0.01]
SEN x OP x SIM	0.00	0.02	.939	0.08	[-0.04, 0.05]
SEN x OP x DIS	0.06	0.02	.009	2.77	[0.02, 0.10]
Problem size	0.10	0.01	< .001	9.67	[0.08, 0.13]
Random effects	**Variance**	***SD***			
Intercept | Subject	0.02	0.13			
OP x SIM | Subject	0.12	0.34			
SEN x OP x SIM | Subject	0.01	0.10			
SEN x OP x DIS | Subject	0.01	0.11			
OP | Operation	0.01	0.12			
Residual	0.08	0.28			

Note. Reaction times were logarithmically transformed, categorical predictors were deviated coded, and problem size was scaled. The smallest variance components were removed recursively until convergence was achieved (Bates et al., 2018). Control variables are shown in grey.

3) According to participants' reports, 83% of additions (SD = 18.4) and 89% of subtractions (SD = 17.3) were solved by retrieval, and this difference was significant, $t(60)$ = -5.78, p < .001. However, when only operations with identical problem sizes were considered (i.e., result 5, 6, 7, or 8), both additions and subtractions were solved in similar proportion by retrieval (M = 88%, SD = 18 vs. M = 87%, SD = 18.5), $t(60)$ = 0.97, p = .033.

Figure 4

Reaction Times for Additions and Subtractions as a Function of the Preceding Sentence and the Numerical Distance in Experiment 1



Discussion

The aim of Experiment 1 was to evaluate the existence of semantic overlap between the processing of linguistic and arithmetic information. To this end, we examined semantic priming effects from sentence comprehension to arithmetic problem solving. We hypothesized that, if language and arithmetic shared semantic features, affirmative sentences would facilitate the processing of additions and negative sentences the processing of subtractions. Furthermore, we explored if task difficulty might modulate the strength of priming effects by manipulating the processing difficulty associated with the prime and the target. Specifically, we manipulated the similarity between the response options for the sentence (the prime) and the numerical distance of the response options for the operation (the target).

The analyses of the sentences (Stage 1) confirmed that negative sentences were more difficult to process than affirmative sentences (e.g., Cheng & Huang, 1980; Kaup, Lüdtke, & Zwaan, 2005; Margolin & Abrams, 2009). Moreover, the data also revealed that participants were sensitive to the sentence difficulty, with slower reaction times and lower accuracy in sentences with similar response options relative to sentences with less similar response alternatives. In addition, the interaction between the two factors in accuracy indicated that the impact of the similarity between the response options was larger in affirmative than in negative sentences, although the differences were rather small.

Nevertheless, our primary hypothesis concerned the possible semantic priming across language and arithmetic. The data obtained in Stage 2 partially confirmed our hypothesis: We found a priming effect for subtractions, with faster reaction times for subtractions preceded by negative versus affirmative sentences. This effect was modulated by the numerical distance, that is, it was only significant under the close condition (i.e., numerical distance between the response options = 2). These results indicate that participants made use of the semantic information present in negative sentences and subtractions (i.e., semantic negation). Thus, the resolution of arithmetic problems benefited from the congruency between the semantic features of sentences and operations (negative sentence, subtraction) compared to when sentences and operations were incongruent (positive sentence, subtraction). Moreover, the fact that the effect was significant only when the task demands for the operation were high (i.e., close condition), shows that priming effects are indeed influenced by the cognitive load associated with the target operation, that is, the benefits are found when the processing of the operation is more difficult. Lastly, the difficulty associated with the sentence did not impact the priming effects in any direction.

However, we did not find priming effects for additions in any of the experimental conditions. One possible explanation for the absence of priming effect for additions might be the methodology used in our experiment: The responses to the operation were recorded once participants had processed it (the operations were displayed for 1000 ms before participants could respond). Therefore, the benefits associated with affirmative sentences might have been present but

dissipated over time (Squire, Shimamura, & Graf, 1987). Taking this issue into consideration, we conducted a second study in which the response to the operations was registered orally.

Experiment 2

The goal of this second experiment was to test whether the way in which the responses for the operations were registered determined the pattern of results found in Experiment 1. In Experiment 2, instead of choosing between two possible response options (two numbers), participants had to say aloud the answer to the arithmetic problem. Therefore, in Stage 2, participants did not perform a decision task between two possible results for the operation and thus, numerical distance (the distance between the two possible results) could not be manipulated in this experiment. We expected to find semantic priming for subtractions and, possibly, for additions after eliminating the delay between the operation presentation and the response of the participants.

Method

Participants

We recruited 25 students from the University of Granada, speakers of Spanish, with an average age of 20 years (M = 20.28, SD = 1.81). A total of 20 participants were females and 5 participants were males. None of the participants took part in Experiment 1. They received either academic credits or money in exchange for their participation. Each participant provided written informed consent before performing the experiment. All procedures used in this study were in accordance with the 1964 Helsinki declaration and its later amendments.

Materials and Design

The materials and design of Experiment 2 were almost identical to Experiment 1. There were two main differences between procedures: the response mode for Stage 2 and the counterbalance across participants. In this new experiment, participants had to say aloud the result of the arithmetical operation instead of choosing manually between two possible response options. In Experiment 2, the same participant received similar and dissimilar sentences followed by additions and subtractions, avoiding the need to counterbalance across participants the sentence similarity and the operation. All other details were the same as those described in Experiment 1. The experimental task used in Experiment 2, raw and filtered data, together with the R scripts used for data analyses and the results from the analyses, are freely accessible at the Open Science Framework (see Supplementary Materials).

Procedure

Participants performed 12 experimental blocks with 40 trials each. As in Experiment 1, each block was formed by a prime sentence (Stage 1) that could be affirmative or negative, and by a target arithmetic operation (Stage 2) that could be an addition or a subtraction. Before each trial, a mask (*******) was displayed on the screen for 500 ms. Once the mask disappeared, a sentence was presented for 1000 ms and then two geometric forms were shown, one on the right and one on the left side of the screen. Again, both figures matched the geometric form described in the previous sentence, and they differed only in colour. Participants had to choose the figure that matched the meaning of the previous sentence by pressing one of two keys (Z or M). Once they had answered, the mask was displayed again for 500 ms, and then an addition or subtraction appeared on the screen. This time, participants had to say aloud the answer to the problem. The arithmetical operation remained on the screen until the participants' response. Response latencies were collected using a microphone ATR 20 with low impedance connected to a PST Serial Response Box (Schneider, 1995) and tape-recorded to eliminate trials with incorrect responses. At the end of each block, the prompt "¿MODO?" ("MODE?") was displayed, and participants had to report how they had solved the operation (as described in Experiment 1). All other details of the procedure were the same as those described in Experiment 1.

Data Cleaning and Analysis

For the latency analyses, erroneous responses were removed and univariate outliers were eliminated following the same procedure as that for Experiment 1 (Tabachnick & Fidell, 2001). In Stage 2, in which auditory responses were collected, data points were excluded if participants produced nonverbal sounds that triggered the voice key, they stuttered or hesitated in producing the response, or they produced something different that the result of the problem. The percentage of outliers was 5.38% in Stage 1, and 7.58% in Stage 2.

After data cleaning, reaction times of Stage 1 and Stage 2 were logarithmically transformed to improve the right skewness. Linear mixed-effects models for reaction times were run separately for sentences and operations (Stage 1 and Stage 2, respectively). Generalized mixed-effects models for binomial data were run for accuracy, also separately for each stage. As in Experiment 1, all categorical predictors were deviated coded (see Appendix B), and the optimizer "bobyqa" was included in all models.

The model for Stage 1 was identical to the model reported in Experiment 1. Sentence type (affirmative, negative), sentence similarity (similar, dissimilar), and their interaction were included as fixed factors. The maximal model was built with participants and sentences as random intercepts. Random slopes for participants and sentences were included based on our design and hypotheses, and following the recommendations of Barr et al. (2013). In case of non-convergence, we removed recursively the smallest variance component(s) (with value zero or closer to zero) until convergence was achieved, in line with Bates et al. (2018).

The model of Stage 2 included, as fixed factors, sentence type (affirmative, negative), sentence similarity (similar, dissimilar), operation type (addition, subtraction), the two-way interactions between sentence type and operation type and between operation type and sentence similarity, and the three-way interaction between sentence type, operation type, and sentence similarity. As in Experiment 1, the problem size was scaled and included as a covariate in the model. Again, the maximal model included random intercepts for participants, sentences, and operations. Random slopes for participants, sentences, and operations were included based on the experimental design and hypotheses, and following the recommendations of Barr et al. (2013). In case of non-convergence, the smallest variance component(s) were removed recursively following Bates et al. (2018). The structure of the maximal models of Stage 1 and Stage 2, as well as the models' refit are detailed in the R scripts available at the Open Science Framework (see Supplementary Materials).

Results

Accuracy Analysis

Stage 1. Prime Sentences — The main effect of sentence type was significant, $b = 0.72$, $SE = 0.14$, $z = 5.00$, $p < .001$, as well as the main effect of sentence similarity, $b = 0.54$, $SE = 0.12$, $z = 4.33$, $p < .001$. Accuracy was higher for affirmative than for negative sentences, 96%, CI [95%, 97%] vs. 93%, CI [91%, 94%], respectively, and under the dissimilar condition compared to the similar condition, 96%, CI [95%, 97%] vs. 93%, CI [92% 95%], respectively. The interaction between sentence type and sentence similarity was not significant ($p > .05$). The level of the reported confidence intervals was 95%. The results from this analysis are shown in Table 5.

Table 5

Output of the Generalized Mixed-Effects Model of Accuracy in Stage 1 (Sentence), Experiment 2

Fixed effects	Estimate	*SE*	*p*	*z*	95% CI
Intercept	2.92	0.12	< .001	23.52	[2.68, 3.16]
Sentence type (SEN)	0.72	0.14	< .001	5.00	[0.44, 1.00]
Sentence similarity (SIM)	0.54	0.12	< .001	4.33	[0.30, 0.78]
SEN x SIM	0.17	0.22	.460	0.74	[-0.27, 0.61]
Random effects	**Variance**	***SD***			
Intercept | Subject	0.32	0.57			
SEN | Subject	0.32	0.57			
SIM | Subject	0.15	0.39			
SEN x SIM | Subject	0.32	0.57			
Intercept | Sentence	0.00	0.06			
SEN | Sentence	0.17	0.41			
SIM | Sentence	0.46	0.68			
SEN x SIM | Sentence	1.05	1.02			

Note. Categorical predictors were deviated coded.

Stage 2. Target Operations — The analysis of Stage 2 revealed a significant interaction between sentence type and operation type, b = 0.30, SE = 0.15, z = 1.99, p = .046. Accuracy was higher for subtractions after negative versus affirmative sentences, 94%, CI [92%, 96%] vs. 93%, CI [91%, 95%], respectively; and for additions after affirmative versus negative sentences, 94%, CI [91%, 95%] vs. 93%, CI [90%, 95%], respectively. However, it is important to note that the difference in accuracy was only 1% for both subtractions and additions (that is, a difference of 2 trials out of 120), and caution should be applied when interpreting it as semantic priming. The level of the reported confidence intervals was 95%. The results from this analysis are presented in Table 6.

Table 6

Output of the Generalized Mixed-Effects Model of Accuracy in Stage 2 (Operation), Experiment 2

Fixed effects	Estimate	*SE*	*p*	*z*	95% CI
Intercept	2.64	0.16	< .001	16.32	[2.32, 2.96]
Sentence type (SEN)	0.01	0.09	.948	0.06	[-0.17, 0.18]
Operation type (OP)	-0.10	0.13	.468	-0.73	[-0.36, 0.17]
Sentence similarity (SIM)	0.13	0.09	.164	1.39	[-0.05, 0.30]
SEN x OP	0.30	0.15	.046	1.99	[0.00, 0.60]
OP x SIM	-0.16	0.15	.285	-1.07	[-0.45, 0.13]
SEN x OP x SIM	0.57	0.33	.083	1.73	[-0.07, 1.22]
Problem size	0.07	0.07	.376	0.89	[-0.08, 0.21]
Random effects	**Variance**	***SD***			
Intercept | Subject	0.60	0.77			
SEN | Subject	0.05	0.22			
Similarity | Subject	0.04	0.21			
Problem size | Subject	0.02	0.13			
SEN x OP x SIM | Subject	0.32	0.57			
SIM | Sentence	0.11	0.33			
Intercept | Operation	0.00	0.00			
OP | Operation	0.04	0.20			
SEN x OP | Operation	0.15	0.38			

Note. Categorical predictors were deviated coded and problem size was scaled. The smallest variance components were removed recursively until convergence was achieved (Bates et al., 2018). Control variables are shown in grey.

Latency Analysis

Stage 1. Prime Sentences — We found significant main effects of sentence type, b = -0.13, SE = 0.02, t = -8.06, p < .001, and sentence similarity, b = -0.18, SE = 0.01, t = -14.04, p < .001, but no interaction between the two factors (p > .05). Reaction times were larger for negative than for affirmative sentences, 934 ms, CI [876, 996] vs. 822 ms, CI [771, 876], respectively, and when the response options were similar compared to when they were dissimilar, 958 ms, CI [899, 1020] vs. 803 ms, CI [754, 855], respectively. The level of the reported confidence intervals was 95%. The results from this analysis are reported in Table 7.

Table 7

Output of the Linear Mixed-Effects Model of Reaction Times in Stage 1 (Sentence), Experiment 2

Fixed effects	**Estimate**	***SE***	***p***	***t***	**95% CI**
Intercept	6.78	0.03	< .001	214.21	[6.72, 6.84]
Sentence type (SEN)	-0.13	0.02	< .001	-8.06	[-0.16, -0.10]
Sentence similarity (SIM)	-0.18	0.01	< .001	-14.04	[-0.20, -0.15]
SEN x SIM	0.04	0.03	.144	1.47	[-0.01, 0.09]
Random effects	**Variance**	***SD***			
Intercept | Subject	0.02	0.15			
SEN | Subject	0.00	0.04			
SEN | Sentence	0.04	0.20			
SEN x SIM | Sentence	0.11	0.34			
Residual	0.11	0.33			

Note. Reaction times were logarithmically transformed. Categorical predictors were deviated coded.

Stage 2. Target Operations — The results revealed a significant effect of operation type[4], b = -0.27, SE = 0.05, t = -5.82, p < .001, with faster reaction times for additions than for subtractions, 977 ms, CI [897, 1063] vs. 1279 ms, CI [1170, 1397], respectively. However, the interaction between sentence type and operation type was not significant (p > .05), neither the interaction between sentence type, operation type, and sentence similarity (p > .05). In other words, we did not find priming effects from sentence processing to arithmetic problem solving. Lastly, apart from our main effects of interest, the problem size effect was significant, b = 0.20, SE = 0.03, t = 7.65, p < .001, with an increase in response latencies when the problem size increased. The level of the reported confidence intervals was 95%. The results of this analysis are displayed in Table 8.

4) As in Experiment 1, participants reported retrieval as the main strategy used to solve both additions (M = 70%, SD = 22.5) and subtractions (M = 81%, SD = 18), and the difference between the two types of operations was significant, $t(24)$ = -6.07, p < .001. When analysing only operations with identical problem size, additions were solved more frequently by retrieval (M = 82%, SD = 20.9) than subtractions (M = 74%, SD = 24.4), and this difference was significant, $t(24)$ = 3.32, p = .003. This goes in line with previous literature indicating that additions are solved by retrieval more frequently than additions (e.g., Ashcraft, 1992, 1995; Campbell, 1995; Geary, Frensch, & Wiley, 1993; see Hinault & Lemaire, 2016, for a recent review of strategies used in cognitive arithmetic).

Table 8

Output of the Linear Mixed-Effects Model of Reaction Times in Stage 2 (Operation), Experiment 2

Fixed effects	**Estimate**	***SE***	***p***	***t***	**95% CI**
Intercept	7.03	0.04	< .001	186.49	[6.96, 7.11]
Sentence type (SEN)	0.01	0.01	.348	0.94	[-0.01, 0.02]
Operation type (OP)	-0.27	0.05	< .001	-5.82	[-0.36, -0.18]
Sentence similarity (SIM)	0.01	0.01	.516	0.65	[-0.01, 0.02]
SEN x OP	-0.01	0.01	.360	-0.92	[-0.04, 0.01]
OP x SIM	-0.01	0.02	.723	-0.36	[-0.04, 0.03]
SEN x OP x SIM	-0.03	0.03	.305	-1.03	[-0.08, 0.02]
Problem size	0.20	0.03	< .001	7.65	[0.15, 0.25]
Random effects	**Variance**	***SD***			
Intercept | Subject	0.03	0.17			
OP | Subject	0.01	0.08			
SIM | Subject	0.00	0.01			
Problem size | Subject	0.00	0.04			
OP x SIM | Subject	0.00	0.02			
SEN | Sentence	0.00	0.05			
SIM | Sentence	0.00	0.05			
OP | Operation	0.04	0.22			
Problem size | Operation	0.01	0.08			
SEN x OP x SIM | Operation	0.00	0.07			
Residual	0.06	0.25			

Note. Reaction times were logarithmically transformed. Categorical predictors were deviated coded and problem size was scaled. The smallest variance components were removed recursively until convergence was achieved (Bates et al., 2018). Control variables are shown in grey.

Discussion

The objective of Experiment 2 was to address one possible limitation of Experiment 1, that is, the delay between the presentation of the operation and the codification of the response, and to evaluate again the possible existence of semantic priming between language and arithmetic. To this end, we registered the oral response of participants to additions and subtractions, since we expected this measure to be more sensitive to possible semantic priming effects than the manual response employed in Experiment 1. Participants were more accurate when solving subtractions after negative sentences, and additions after affirmative sentences. However, the difference in accuracy based on the preceding sentence was only 1% for both types of operations and, therefore, we suggest caution when interpreting this effect as semantic priming. We did not find priming effects in reaction times for subtractions or additions. The results from this experiment, together with Experiment 1, are further discussed in the next section.

General Discussion

In recent years, behavioural and neuroimaging studies have investigated the existence of general principles of cognition shared across different cognitive domains, such as language and arithmetic. Some studies have found commonalities between language and mathematics regarding the processing of syntactic information (e.g., Makuuchi, Bahlmann, & Friederici, 2012; Nakai & Okanoya, 2018; Scheepers et al., 2011; Scheepers & Sturt, 2014; Van de Cavey & Hartsuiker, 2016), and also semantic information (Bassok et al., 2008; Guthormsen et al., 2016). The evidence supporting integrative syntactic processing seems to be robust. However, the studies evaluating shared semantic processing in language and arithmetic reveal mixed results. Some studies support the hypothesis of semantic overlap in both cognitive domains (Bassok et al., 2008; Guthormsen et al., 2016), while others report dissociations (e.g., Monti et al., 2012), or null results (Ronasi et al., 2018).

In the current study, we sought to extend the evidence about the possible domain-general semantic system shared between language and arithmetic by evaluating the impact of semantic affirmation and semantic negation in the processing of sentences and in the resolution of arithmetic problems (additions and subtractions). To accomplish this goal, we developed a new paradigm employing affirmative and negative sentences as primes, and additions and subtractions as targets. We expected to find priming effects from affirmative sentences to additions and from negative sentences to subtractions. Moreover, we explored whether task difficulty modulates the occurrence of these priming effects. To this end, we manipulated the cognitive load of both the prime and the target by modifying the similarity between the two response options for the sentences (Experiment 1 and 2), and by presenting numerically close and far response options for the operations (Experiment 1). We expected to find stronger priming effects when the operation was harder to perform (i.e., close condition), and when the cognitive load associated with the sentence was low (i.e., dissimilar condition).

The results from Experiment 1 revealed semantic priming for subtractions: participants solved subtractions faster when they were preceded by negative versus affirmative sentences. However, the effect was modulated by the processing difficulty of the arithmetic operations, that is, it was significant only when the subtraction was hard to perform (i.e., close numerical distance between the possible results). This result goes in line with the incremental-procedural account (Scheepers et al., 2011; Scheepers & Sturt, 2014): Operations that demand more cognitive resources and are more difficult to perform benefit from the prime sentence to a greater extent than easy arithmetic problems. In Experiment 1, operations with close response options were more difficult to process (e.g., they were associated with a lower response accuracy) than those with far distance. Thus, the cognitive load associated with the operation seemed to modulate the priming effect, in the sense that it was observed only for subtractions with close versus far response options. However, priming effects from affirmative sentences to addition problems were not observed in Experiment 1.

In Experiment 2, we evaluate whether the delay between the presentation of the operation and the codification of the response in Experiment 1 (manual response) determined the absence of priming effects with addition problems. In Experiment 1, responses to the operations were recorded only when the operation disappeared from the screen, which might have caused a reduction in priming effects. To address this possibility, in Experiment 2, instead of selecting the correct arithmetic result between two possible results (i.e., a decision task), participants were instructed to say aloud the correct response to the math problem. The results of Experiment 2 revealed that subtractions were solved more accurately after negative versus affirmative sentences, and additions after affirmative versus negative sentences. This pattern of results would suggest the presence of semantic priming between language and arithmetic during the resolution of both additions and subtractions. However, we honestly recognize, as mentioned in the Results section of Experiment 2, that the magnitude of this effect was very small to reach a definitive conclusion. In fact, regarding reaction times, we did not find priming effects from sentence processing to arithmetic problem solving for any of the two types of operations (additions and subtractions). The absence of priming in reaction times for subtractions in Experiment 2 can be explained by taking into consideration the impact of task difficulty in the priming effects evaluated in our study: Priming for subtractions seems to be present only when the resolution of the operation imposes higher demands on participants (i.e., close condition) compared to the resolution of easy problems (i.e., far condition) as observed in Experiment 1. However, task difficulty (i.e., numerical distance) was not manipulated in Experiment 2, so this could have been the reason why the priming effect was not captured in this experiment[5].

Returning to the pattern of results found in Experiment 1, the data revealed that semantic priming was modulated by the difficulty of the subtraction problems. However, it remains to be explained why priming effects with additions were not observed in Experiment 1, regardless of whether they were easy or difficult to resolve. Although more studies on semantic priming for simple additions need to be conducted to address this question properly, we propose two possible explanations: First, the results of Experiment 1 revealed that participants were more efficient when selecting the correct response for additions compared to subtractions, both in terms of reaction times and accuracy. The one-digit additions used in our experiments might have been too simple to benefit from the prime sentence, even when the

5) The differences in variance of response times between the experiments (184 ms vs. 421 ms in Experiment 1 and 2, respectively) could underlie the absence of effects in the latency analyses of Experiment 2. This difference in response variability could be due to changes in the response type (button press vs. oral responses) or sample size (n = 61 vs. n = 25) between Experiment 1 and 2, respectively.

numerical distance between the response options was close. A second possibility comes from the differential processing of affirmative and negative sentences (Kaup, Lüdtke, & Zwaan, 2006; Kaup, Yaxley, Madden, Zwaan, & Lüdtke, 2007; Kumar, Padakannaya, Mishra, & Khetrapal, 2013). For priming to occur, the prime concept must be activated (the sentences in the current study). In Experiment 1, this happened for negative sentences but perhaps not for affirmative sentences because they may be the "default processing mode" (Christensen, 2009) and as such they would not require any particular mental operation. In fact, the results found in Experiment 1 (Stage 1, sentence processing) confirmed that affirmative sentences were processed faster and more accurately than negative sentences (see also Cheng & Huang, 1980; Kaup et al., 2005; Margolin & Abrams, 2009). Thus, priming would occur only when the prime involved the extra operation of semantic negation (negative sentences)[6]. We acknowledge that more empirical research is needed to explore these two hypotheses. For example, regarding the first explanation, the processing of simple additions with one-digit operands could be compared with the resolution of more complex additions (e.g., additions with two-digit operands). In relation to the second account, the processing of simple affirmative sentences (the default mode of processing) could be contrasted with other affirmative sentences implying greater complexity.

Another aspect of the study that needs to be discussed is why the difficulty of the sentence did not impact priming effects in any of our experiments. This result might suggest that higher demands in the processing of the prime do not translate into increased difficulties in the processing of the target. However, another reason could be that the demands imposed by the difficult condition when processing the sentences (i.e., similar response options) were not high enough to influence the subsequent operation. In fact, the overall accuracy when participants processed the sentences was very high regardless of whether they belonged to the easy or difficult condition (i.e., dissimilar or similar response options, respectively).

To sum up, the results obtained in our study suggest the existence of shared semantics between language and arithmetic. This evidence is restricted to the case of semantic information that involves negation (negative sentences and subtractions). Moreover, the accuracy results from Experiment 2 offer (limited) evidence suggesting that the overlap in semantic processing between language and arithmetic could be extended to additions, but further research is needed to extract a more solid conclusion on the relationship between affirmative sentences and additions.

Funding: This research was supported by the Spanish Ministry of Science and Innovation. Reference of the project: PID2019-111359GB-I00 / SRA (State Research Agency /10.13039/501100011033) awarded to Pedro Macizo.

Acknowledgments: We thank Dr. Alba Casado (Institute of Psychology, Jagiellonian University, Kraków) for her assistance in data analysis.

Competing Interests: The authors have declared that no competing interests exist.

Data Availability: For this article, two datasets are freely available (Castro & Macizo, 2021).

Supplementary Materials

The experimental tasks used in Experiments 1 and 2, raw and filtered data, the R scripts used to analyse the data and the results from the analyses, are freely accessible at https://osf.io/h5k4y

Index of Supplementary Materials

Castro, S., & Macizo, P. (2021). *Supplementary materials to "All roads lead to Rome: Semantic priming between language and arithmetic"* [Data, code, and materials]. OSF. https://osf.io/h5k4y

6) We are very grateful to an anonymous reviewer of a previous version of this study for suggesting this second possible account for the pattern of results obtained in Experiment 1.

References

Ashcraft, M. H. (1992). Cognitive arithmetic: A review of data and theory. *Cognition, 44*, 75-106. https://doi.org/10.1016/0010-0277(92)90051-I

Ashcraft, M. H. (1995). Cognitive psychology and simple arithmetic: A review and summary of new directions. *Mathematical Cognition, 1*(1), 3-34.

Baldo, J. V., & Dronkers, N. F. (2007). Neural correlates of arithmetic and language comprehension: A common substrate? *Neuropsychologia, 45*(2), 229-235. https://doi.org/10.1016/j.neuropsychologia.2006.07.014

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255-278. https://doi.org/10.1016/j.jml.2012.11.001

Bassok, M., Pedigo, S. F., & Oskarsson, A. T. (2008). Priming addition facts with semantic relations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(2), 343-352. https://doi.org/10.1037/0278-7393.34.2.343

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2018). *Parsimonious mixed models.* arXiv. https://arxiv.org/abs/1506.04967v2

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48. https://doi.org/10.18637/jss.v067.i01

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology, 18*(3), 355-387. https://doi.org/10.1016/0010-0285(86)90004-6

Bock, K. (1989). Closed-class immanence in sentence production. *Cognition, 31*(2), 163-186. https://doi.org/10.1016/0010-0277(89)90022-X

Branigan, H. P., & Pickering, M. J. (2017). An experimental approach to linguistic representation. *Behavioral and Brain Sciences, 40*, Article e282. https://doi.org/10.1017/S0140525X16002028

Brimo, D., Lund, E., & Sapp, A. (2018). Syntax and reading comprehension: A meta-analysis of different spoken-syntax assessments. *International Journal of Language & Communication Disorders, 53*(3), 431-445. https://doi.org/10.1111/1460-6984.12362

Campbell, J. I. (1995). Mechanisms of simple addition and multiplication: A modified network-interference theory and simulation. *Mathematical Cognition, 1*(2), 121-164.

Cappelletti, M., Butterworth, B., & Kopelman, M. (2001). Spared numerical abilities in a case of semantic dementia. *Neuropsychologia, 39*(11), 1224-1239. https://doi.org/10.1016/S0028-3932(01)00035-5

Cheng, C.-M., & Huang, H.-J. (1980). The process of verifying affirmative and negative sentences against pictures. *Memory & Cognition, 8*(6), 573-583. https://doi.org/10.3758/BF03213777

Christensen, K. R. (2009). Negative and affirmative sentences increase activation in different areas in the brain. *Journal of Neurolinguistics, 22*(1), 1-17. https://doi.org/10.1016/j.jneuroling.2008.05.001

Delazer, M., Girelli, L., Semenza, C., & Denes, G. (1999). Numerical skills and aphasia. *Journal of the International Neuropsychological Society, 5*(3), 213-221. https://doi.org/10.1017/S1355617799533043

Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology.* Cambridge, MA, USA: MIT Press.

Geary, D. C., Frensch, P. A., & Wiley, J. G. (1993). Simple and complex mental subtraction: Strategy choice and speed-of-processing differences in younger and older adults. *Psychology and Aging, 8*(2), 242-256. https://doi.org/10.1037/0882-7974.8.2.242

Guthormsen, A. M., Fisher, K. J., Bassok, M., Osterhout, L., DeWolf, M., & Holyoak, K. J. (2016). Conceptual integration of arithmetic operations with real-world knowledge: Evidence from event-related potentials. *Cognitive Science, 40*(3), 723-757. https://doi.org/10.1111/cogs.12238

Hinault, T., & Lemaire, P. (2016). What does EEG tell us about arithmetic strategies? A review. *International Journal of Psychophysiology, 106*, 115-126. https://doi.org/10.1016/j.ijpsycho.2016.05.006

Kaup, B., Lüdtke, J., & Zwaan, R. A. (2005). Effects of negation, truth value, and delay on picture recognition after reading affirmative and negative sentences. In B. G. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the 27th Annual Conference of the Cognitive Science Society* (pp. 1114–1119). Mahwah, NJ, USA: Lawrence Erlbaum.

Kaup, B., Lüdtke, J., & Zwaan, R. A. (2006). Processing negated sentences with contradictory predicates: Is a door that is not open mentally closed? *Journal of Pragmatics, 38*(7), 1033-1050. https://doi.org/10.1016/j.pragma.2005.09.012

Kaup, B., Yaxley, R. H., Madden, C. J., Zwaan, R. A., & Lüdtke, J. (2007). Experiential simulations of negated text information. *Quarterly Journal of Experimental Psychology, 60*(7), 976-990. https://doi.org/10.1080/17470210600823512

Kootstra, G. J., & Rossi, E. (2017). Moving beyond the priming of single-language sentences: A proposal for a comprehensive model to account for linguistic representation in bilinguals [Peer commentary on "An experimental approach to linguistic representation," by H. P. Branigan and M. J. Pickering]. *Behavioral and Brain Sciences, 40*, Article e297. https://doi.org/10.1017/S0140525X17000450

Kumar, U., Padakannaya, P., Mishra, R. K., & Khetrapal, C. L. (2013). Distinctive neural signatures for negative sentences in Hindi: An fMRI study. *Brain Imaging and Behavior, 7*(2), 91-101. https://doi.org/10.1007/s11682-012-9198-8

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software, 82*(1), 1-26. https://doi.org/10.18637/jss.v082.i13

Lüdecke, D. (2018). Ggeffects: Tidy data frames of marginal effects from regression models. *Journal of Open Source Software, 3*(26), Article 772. https://doi.org/10.21105/joss.00772

Lüdecke, D., Makowski, D., Waggoner, P., & Patil, I. (2020). *performance: Assessment of Regression Models performance* (R package version 0.6.1). Retrieved from https://easystats.github.io/performance/

Macizo, P. (2016). Phonological coactivation in the bilinguals' two languages: Evidence from the color naming task. *Bilingualism: Language and Cognition, 19*(2), 361-375. https://doi.org/10.1017/S136672891500005X

Makuuchi, M., Bahlmann, J., & Friederici, A. D. (2012). An approach to separating the levels of hierarchical structure building in language and mathematics. *Philosophical Transactions of the Royal Society of London: Series B: Biological Sciences, 367*, 2033-2045. https://doi.org/10.1098/rstb.2012.0095

Margolin, S. J., & Abrams, L. (2009). Not may not be too difficult: The effects of negation on older adults' sentence comprehension. *Educational Gerontology, 35*(4), 308-322. https://doi.org/10.1080/03601270802505624

Megías, P., & Macizo, P. (2016a). Simple arithmetic: Electrophysiological evidence of coactivation and selection of arithmetic facts. *Experimental Brain Research, 234*, 3305-3319. https://doi.org/10.1007/s00221-016-4728-z

Megías, P., & Macizo, P. (2016b). Activation and selection of arithmetic facts: The role of numerical format. *Memory & Cognition, 44*, 350-364. https://doi.org/10.3758/s13421-015-0559-6

Meyer, L. (2018). The neural oscillations of speech processing and language comprehension: State of the art and emerging mechanisms. *The European Journal of Neuroscience, 48*(7), 2609-2621. https://doi.org/10.1111/ejn.13748

Monti, M. M., Parsons, L. M., & Osherson, D. N. (2012). Thought beyond language: Neural dissociation of algebra and natural language. *Psychological Science, 23*(8), 914-922. https://doi.org/10.1177/0956797612437427

Nakai, T., & Okanoya, K. (2018). Neural evidence of cross-domain structural interaction between language and arithmetic. *Scientific Reports, 8*(1), Article 12873. https://doi.org/10.1038/s41598-018-31279-8

Nakai, T., & Sakai, K. L. (2014). Neural mechanisms underlying the computation of hierarchical tree structures in mathematics. *PLoS One, 9*(11), Article e111439. https://doi.org/10.1371/journal.pone.0111439

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin, 134*(3), 427-459. https://doi.org/10.1037/0033-2909.134.3.427

Ronasi, G., Fischer, M. H., & Zimmermann, M. (2018). Language and arithmetic: A failure to find cross cognitive domain semantic priming between exception phrases and subtraction or addition. *Frontiers in Psychology, 9*, Article 1524. https://doi.org/10.3389/fpsyg.2018.01524

Scheepers, C., Galkina, A., Shtyrov, Y., & Myachykov, A. (2019). Hierarchical structure priming from mathematics to two-and three-site relative clause attachment. *Cognition, 189*, 155-166. https://doi.org/10.1016/j.cognition.2019.03.021

Scheepers, C., & Sturt, P. (2014). Bidirectional syntactic priming across cognitive domains: From arithmetic to language and back. *Quarterly Journal of Experimental Psychology, 67*(8), 1643-1654. https://doi.org/10.1080/17470218.2013.873815

Scheepers, C., Sturt, P., Martin, C. J., Myachykov, A., Teevan, K., & Viskupova, I. (2011). Structural priming across cognitive domains: From simple arithmetic to relative clause attachment. *Psychological Science, 22*(10), 1319-1326. https://doi.org/10.1177/0956797611416997

Schneider, M., Beeres, K., Coban, L., Merz, S., Schmidt, S., Stricker, J., & De Smedt, B. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: A meta-analysis. *Developmental Science, 20*(3), Article e12372. https://doi.org/10.1111/desc.12372

Schneider, W. (1995). *Micro experimental laboratory*. Pittsburgh, PA, USA: Psychology Software Tools.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime User's Guide*. Pittsburgh, PA, USA: Psychology Software Tools, Inc.

Segal, S. J., & Cofer, C. N. (1960). The effect of recency and recall on word association [Abstract]. *The American Psychologist, 15*, 451.

Squire, L. R., Shimamura, A. P., & Graf, P. (1987). Strength and duration of priming effects in normal subjects and amnesic patients. *Neuropsychologia, 25*(1), 195-210. https://doi.org/10.1016/0028-3932(87)90131-X

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics*. Needham, MA, USA: Allyn & Bacon.

Van de Cavey, J., & Hartsuiker, R. J. (2016). Is there a domain-general cognitive structuring system? Evidence from structural priming across music, math, action descriptions, and language. *Cognition, 146*, 172-184. https://doi.org/10.1016/j.cognition.2015.09.013

Varley, R. A., Klessinger, N., Romanowski, C. A. J., & Siegal, M. (2005). Agrammatic but numerate. *Proceedings of the National Academy of Sciences of the United States of America, 102*(9), 3519-3524. https://doi.org/10.1073/pnas.0407470102

Wickham, H. (2016). *Ggplot2: Elegant Graphics for Data Analysis* (2nd ed.). Cham, Switzerland: Springer.

Zbrodoff, N. J., & Logan, G. D. (2005). What everyone finds: The problem-size effect. In J. I. D. Campbell (Ed.), *The handbook of mathematical cognition* (pp. 331–345). New York, NY, USA: Psychology Press.

Zeng, T., Mao, W., & Liu, R. (2018). Structural priming from arithmetic to language in Chinese: Evidence from adults and children. *Quarterly Journal of Experimental Psychology, 71*(7), 1552-1560. https://doi.org/10.1080/17470218.2017.1340968

Appendices

Appendix A: Description of the Arithmetic Stimuli Used in Experiment 1 and Experiment 2

Table A1

Distribution (in %) of the Two Operands for Both Experiment 1 and Experiment 2

	Addition		Subtraction	
Number	**As first operand**	**As second operand**	**As first operand**	**As second operand**
1	6.67	6.67	–	23.33
2	6.67	8.33	1.67	19.17
3	11.67	8.33	5.00	15.00
4	10.00	13.33	9.17	12.50
5	13.33	11.67	11.67	10.83
6	13.33	15.00	14.17	10.83
7	11.67	11.67	18.33	5.00
8	13.33	13.33	20.83	3.33
9	13.33	11.67	19.17	–

Note. A total of 240 additions and 240 subtractions were used in Experiment 1 and Experiment 2, formed by combining digits from 1 to 9. A horizontal line within a cell indicates that the digit was not included in any operation.

Table A2

Distribution (in %) of Numbers as Correct Results and as Distractors in Experiment 1

	Addition		Subtraction	
Number	**As correct response**	**As distractor**	**As correct response**	**As distractor**
1	–	0.833	20.00	25.00
2	–	–	18.33	17.50
3	–	3.33	25.00	15.00
4	–	3.33	11.67	11.67
5	1.67	1.67	10.00	15.00
6	5.00	5.83	7.50	10.00
7	10.00	6.67	5.00	5.00
8	8.33	6.67	2.50	0.83
9	10.00	7.50	–	–
10	13.33	10.83	–	–

	Addition		Subtraction	
Number	**As correct response**	**As distractor**	**As correct response**	**As distractor**
11	10.00	6.67	–	–
12	10.00	11.67	–	–
13	8.33	11.67	–	–
14	8.33	5.83	–	–
15	6.67	5.00	–	–
16	3.33	1.67	–	–
17	3.33	3.33	–	–
18	1.67	1.67	–	–
19	–	3.33	–	–
20	–	1.67	–	–
21	–	–	–	–
22	–	0.83	–	–

Note. A total of 240 additions and 240 subtractions were used in this study. The presentation of each operation was followed by two numbers, one indicated the correct response to the operation and the other acted as a distractor. A horizontal line within a cell indicates that the digit was not included as correct response and/or as distractor.

Table A3

List of the Operations Used in Experiment 1 and Experiment 2

Additions						Subtractions			
1 + 5	3 + 5	4 + 8	6 + 2	7 + 7	8 + 9	2 – 1	6 – 1	7 – 6	9 – 3
1 + 6	3 + 6	5 + 2	6 + 4	7 + 8	9 + 1	3 – 1	6 – 2	8 – 1	9 – 4
1 + 8	3 + 7	5 + 3	6 + 5	7 + 9	9 + 2	3 – 2	6 – 3	8 – 2	9 – 6
1 + 9	3 + 8	5 + 4	6 + 6	8 + 1	9 + 3	4 – 1	6 – 4	8 – 3	9 – 7
2 + 4	3 + 9	5 + 5	6 + 7	8 + 2	9 + 4	4 – 2	6 – 5	8 – 4	9 – 8
2 + 5	4 + 1	5 + 6	6 + 8	8 + 4	9 + 5	4 – 3	7 – 1	8 – 5	
2 + 6	4 + 3	5 + 7	6 + 9	8 + 5	9 + 6	5 – 1	7 – 2	8 – 6	
2 + 7	4 + 4	5 + 8	7 + 2	8 + 6	9 + 7	5 – 2	7 – 3	8 – 7	
3 + 3	4 + 6	5 + 9	7 + 4	8 + 7	9 + 8	5 – 3	7 – 4	9 – 1	
3 + 4	4 + 7	6 + 1	7 + 6	8 + 8	9 + 9	5 – 4	7 – 5	9 – 2	

Appendix B: Coding of Categorical Predictors in Experiment 1 and Experiment 2

Table B.1

Coding of Categorical Predictors in Experiment 1 and Experiment 2

Experimental condition	**-0.5**	**0.5**
Sentence type	Negative sentence	Affirmative sentence
Operation type	Subtraction	Addition
Sentence similarity	Similar	Dissimilar
Numerical distance	Close	Far

Note. Numerical distance was manipulated only in Experiment 1.

Journal of Numerical Cognition (JNC) is an official journal of the Mathematical Cognition and Learning Society (MCLS).

leibniz-psychology.org

PsychOpen GOLD is a publishing service by Leibniz Institute for Psychology (ZPID), Germany.

PsychOpen GOLD