**The influence of language dominance and domain-general executive control on semantic context effects**

Jaume Boned1\*, Gemma Cardona1\*, Elizabeth Jefferies2, & Mireia Hernández1

1Cognition and Brain Plasticity Unit, Department of Cognition, Development and Educational Psychology, Institut de Neurociències, Universitat de Barcelona, Barcelona, Spain.

2Department of Psychology, University of York, England, UK.

\* Both authors contributed equally to this study

Address for correspondence

Mireia Hernández

Departament de Cognició, Desenvolupament i Psicologia de l’Educació

Universitat de Barcelona

Passeig de la Vall d’Hebron, 171

08035 Barcelona, Spain

Email: mireiahernandez@ub.edu

**Abstract**

We investigated whether semantic context effects in speech production and comprehension are sensitive to language dominance and whether they involve domain-general executive control. We indexed these effects using semantic blocking within the cyclical semantic paradigm (corresponding to poorer performance in semantically related contexts compared to unrelated contexts) in a study that addressed the limitations of previous research: *(i)* we compared semantic blocking between participants tested in their native language and those tested in a language they were clearly less proficient in (not just the less dominant language), and *(ii)* we examined the involvement of executive control with a non-linguistic (rather than a linguistic) index. Participants in both groups showed equal semantic blocking in production and comprehension. Executive control only predicted the magnitude of semantic blocking in speech production. These results suggest that semantic context effects are insensitive to language dominance, and that effects of executive control arise in production tasks.

**Keywords:** bilingualism, cyclical picture naming, cyclical word-picture matching, executive control, semantic context effects, semantic interference

**Introduction**

In the course of naming an object (e.g., a dress), the semantic representation for that object (i.e., the concept) has to activate the intended lexical representation (i.e., the word /dress/). However, this concept spreads activity to other related concepts (e.g., trousers, shirt, hat, etc.). This means that the lexical representations that are semantically related to the intended item also receive activity from the semantic system (e.g., Caramazza, 1997; Dell, 1986; Levelt et al., 1999; Rapp & Goldrick, 2000). However, semantic errors (e.g., uttering “trousers” for “dress”) are rather scarce in neurologically intact individuals thanks to well-functioning lexical retrieval mechanisms—the nature of which is still hotly debated, with some researchers assuming competition between the intended item and semantically related lexical representations (e.g., La Heij, 1988; Levelt et al., 1999; Roelofs, 2003) and others denying any such lexical competition (e.g., Dell, 1986; Mahon et al., 2007).

Despite its efficiency, speech production can become particularly challenging in a semantically-related context. For instance, the latency for naming “dress” is often slower after naming a semantically related object like “trousers” than after naming an unrelated item (e.g., apple). Semantic context effects are less common in object comprehension. In fact, related concepts most frequently facilitate rather than hamper semantic processing (e.g., McRae & Boisvert, 1998). However, semantic context effects can negatively affect comprehension under particularly demanding conditions (e.g., Campanella & Shallice, 2011; Harvey & Schnur, 2016). The present study aimed to investigate two questions about the capacity to overcome semantic context effects in speech production and comprehension: *(i)* whether these effects are sensitive to language dominance, and *(ii)* whether they involve domain-general executive control. We addressed these questions with a widely used experimental paradigm: the cyclical semantic paradigm.

***The cyclical semantic paradigm***

The speech production variant of this paradigm involves naming pictures of objects in a cyclical manner. In *homogeneous cycles*, pictures are grouped by semantic category (e.g., they are all clothes or they are all tools, etc.). In *heterogeneous cycles*, each picture belongs to a different category. Typically, naming latencies (RTs) are slower in homogeneous cycles than in heterogeneous ones. This context effect, known as *semantic blocking*, arises during cycle 2 and remains constant throughout all subsequent cycles (for the earliest demonstrations of the effect, see Damian et al., 2001; Kroll & Stewart, 1994). There is also a language comprehension variant of the paradigm, in which participants typically perform an auditory word-picture matching task (for other implementations, see e.g., Gardner et al., 2012; Wei & Schnur, 2016). The comprehension variant, however, has been explored only occasionally because semantic blocking is rarely seen in neurologically intact individuals in this paradigm. In contrast to the production variant, semantic blocking in comprehension may arise as early as in the first cycle (e.g., Harvey & Schnur, 2016).

Several different accounts have been proposed to explain semantic blocking effects, most of which focus on speech production. We can divide these accounts into two broad groups, depending on whether they assume lexical competition or not. Typically, lexical competition should increase the difficulty (in terms of RT) of selecting the intended representation as a function of the amount of activation received by other semantically related representations (due to spreading activation between semantically related items, e.g., La Heij, 1988; Levelt et al., 1999; Roelofs, 2003; but see Abdel Rahman & Melinger, 2009). Accordingly, accounts based on lexical competition argue that semantic blocking occurs at the lexical level due to greater competition in homogeneous compared to heterogeneous blocks (e.g., Abdel Rahman & Melinger, 2011; Belke et al., 2005; Damian et al., 2001). This is because continuous exposure to items from the same semantic category maintains the lexical representations of these items which become competitors in homogeneous blocks.

One of the main semantic blocking accounts that has been formulated without assuming lexical competition is that proposed by Navarrete and colleagues (e.g., Navarrete et al., 2010, 2012, 2014, 2016). These authors take the view that a lexical representation is retrieved simply when its level of activation reaches a threshold, regardless of the level of activity for semantically related items (Dell, 1986). This means that, for instance, the retrieval of the word “dress” would not be slower if a semantically related word (e.g., trousers) is almost as activated as “dress” following spreading of activation between semantically related concepts. On this basis, Navarrete et al. account for the particular case of semantic blocking by adopting Oppenheim and colleagues’ idea of “incremental learning”, which proposes continuous adjustments in the connection weights between the conceptual system and the lexicon (Oppenheim et al., 2010). For instance, naming the picture of a dress strengthens the connections between its semantic and lexical representations. Hence, naming “dress” again will be faster in subsequent trials, a phenomenon widely known as repetition priming. At the same time, the connections between that same semantic representation (dress) and the lexical representations of semantically related objects (e.g., trousers) will weaken, which is referred to as incremental weakening. According to Navarrete et al., repetition priming will arise within homogeneous cycles as well as within blocks of heterogeneous trials. This is because the same items repeat across the cycles within both types of blocks. However, in blocks of homogeneous cycles, incremental weakening will also take place and will counteract the repetition priming effect. Therefore, semantic blocking might reflect the way in which repetition priming speeds up naming latencies in heterogeneous but not in homogeneous cycles[[1]](#endnote-1).

Although the repetition priming account was formulated only for production, in principle it could be adapted to comprehension. There is no reason why incremental learning might not operate bidirectionally and, thus, exert its effects also from lexical forms to conceptual representations (but see Navarrete et al.’s (2014; footnote 3) for acknowledgement that these effects might also occur within the conceptual level). Hence, it could also account for semantic blocking in tasks not requiring speech production such as a word-picture matching task, where incremental learning might occur between the auditory word and the concept (or, at least, between the word and the representation of the object’s shape). In spite of this, all accounts of semantic blocking in comprehension have been based on competitive processes at the conceptual level (but see Nathaniel et al., 2018). For instance, it has been proposed that immediately after uttering “dress”, this concept will remain highly activated and will interfere with semantically related concepts (e.g., trousers) in subsequent homogeneous cycles of the same sematic category (e.g. Gardner et al., 2012; Jefferies et al., 2007). The proponents of this account—who are among the few authors using both the production and the comprehension variant of the paradigm—apply the same rationale to both modalities. Other authors who also assume semantic interference at the conceptual level in both modalities argue that this interference transfers to the lexical level during speech production, where it results in lexical competition. Hence, semantic blocking would be the result of semantic interference at the conceptual level in a word-picture matching task, and the result of lexical competition in a picture naming task (e.g., Harvey & Schnur, 2016).

***Language dominance and semantic blocking***

One of the questions we set out to investigate is whether semantic blocking is sensitive to language dominance in speakers with a non-native second language (L2). According to the revised hierarchical model (Kroll & Stewart, 1994), an influential view on bilingual lexical-semantic representations, L2 shares the conceptual system with the first and dominant language (L1; but see Brysbaert & Duyck, 2010). In spite of this, the connectivity between the conceptual system and the lexical representations is likely to be weaker in L2 unless the bilingual speaker has a native-like proficiency in L2 (Kroll et al., 2010). This brings up the question of whether and how semantic context effects differ between L1 and L2. To our knowledge, the only studies that have examined whether L2 modulates semantic context effects have focused on speech production (Calabria et al., 2019; Runnqvist et al., 2012). The only hypothesis put forward to date is that semantic blocking is reduced in L2. Assuming lexical competition, the weaker semantic-to-lexical links in L2 would imply that a given concept (e.g., dress) activates semantically related non-intended lexical representations (e.g., trousers, shirt, hat, etc.) to a lesser degree in L2 compared to L1 (Runnqvist et al.; footnote 2). However, from a lexical competition view, one could also argue that semantic blocking may be greater in L2 due to generally weakened semantic-to-lexical connectivity. Such a weakening may complicate lexical selection because the level of activation of the intended lexical representation is too similar to that of the semantic competitors.

In the study of Calabria et al. (2019), a group of Catalan-Spanish patients with bilingual aphasia and a group of Catalan-Spanish controls completed the cyclical picture naming task twice, in L1 in the first session and in L2 in the second session. The magnitude of the semantic blocking effect did not differ between L1 and L2 in the control participants. However, in the patients with aphasia, the semantic blocking in L2 was greater than that observed in the controls in the last two blocks of the task. Runnqvist et al. (2012) used a different semantic context paradigm (also instantiated in a picture naming task), in which semantically related pictures are presented separated by *n* intervening unrelated pictures (Howard et al., 2006; see also Experiment 4 in Brown, 1981). For instance, participants might be presented with a series of pictures like “…trousers–hammer–dress–orange–table–hat…”. Typically, RTs increase linearly with the ordinal position of the pictures within the same semantic category, an effect known as *cumulative semantic interference (CSI)*. This means that in the sequence above, the time taken to name “dress” would be longer than for “trousers”, and that RTs to name “hat” would be even longer. In the study of Runnqvist et al. (2012; experiments 1, 3 & 5), a group of Catalan-Spanish bilinguals performed the task in L1 and another group in L2. The magnitude of the *CSI* did not differ between the groups.

It is worth noting here that a flaw of the *CSI* accounts assuming lexical competition (e.g., Howard et al., 2006) is that the lexical activation of a particular item is unlikely to last long enough to interfere with another one presented several seconds later. Howard et al. (2006) admitted such a flaw themselves, hypothesizing that besides lexical competition, durable changes in the weights of semantic-to-lexical representations must take place throughout the naming task when using the *CSI* paradigm. Oppenheim et al. (2010) elaborated on the hypothesis about the long-lasting changes in the weights of sematic-to-lexical connectivity and proposed the idea of “incremental learning”, which we mention above when discussing the possible origins of the semantic blocking effect. The basic idea of incremental learning is that after naming an item accurately, the semantic-to-lexical link of that item strengthens (leading to semantic priming when the same item is to be named later), whereas the links between the semantic representation of that item and the lexical labels of semantically related items weaken (slowing down naming latencies for subsequent semantically related items). This idea does not necessarily involve the assumption of lexical competition to explain the *CSI*, which could be the result of incremental weakening. Continuing from the above example, the semantic-to-lexical connectivity of “dress” would weaken as a result of previously naming “trousers”, and the weakening of the semantic-to-lexical connectivity of “hat” would be even more evident due to the previous naming of the two items from the same semantic category (“trousers” and “dress”). Hence, as mentioned in the previous sub-section, the adoption of the idea of incremental learning proposed by Oppenheim et al. allowed Navarrete and colleagues (e.g., Navarrete et al., 2010, 2012, 2014, 2016) to account for the semantic blocking effect without the need to assume lexical competition. Semantic priming can facilitate picture naming in homogeneous and heterogeneous blocks, but this facilitation could be neutralized by incremental weakening in homogeneous blocks. This means that the incremental learning approach offers a unified account for both the *CSI* and the semantic blocking effect. This is relevant here because, if one adopts this approach, the study of Runnqvist et al. (2012) can be established as a valid precedent for the current investigation despite the fact that they used the *CSI* paradigm while we have used the cyclical semantic one.

Taken together, the results of Calabria et al. (2019) and Runnqvist et al. (2012) indicate that semantic context effects are insensitive to language dominance in neurologically intact individuals. Critically, however, the L2 proficiency level of the bilinguals in these studies was native-like, even though they considered themselves more dominant in one of the languages. In the current study, we addressed this limitation by comparing the semantic blocking of participants tested in their native language and participants tested in a non-dominant language they were clearly less proficient in.

***Domain-general executive control and semantic blocking***

We also set out to investigate whether domain-general executive control plays a role in the capacity to overcome semantic blocking. Alternatively, this capacity relies on language-specific (or semantic-specific) mechanisms. Existing evidence is mixed in this regard. Evidence that domain-general executive control is involved in semantic context effects comes from an association between semantic blocking in picture naming and the Stroop task. In this task, participants name the color of the ink in which the words denoting the colors are printed. Typically, RTs in incongruent trials (e.g., BLUE printed in red) are longer than those in congruent trials (e.g., BLUE printed in blue), revealing that time is needed to select a target among irrelevant and interfering stimuli. This so-called Stroop interference effect has been widely used to index the general ability to overcome interference from irrelevant stimuli. Crowther and Martin (2014) demonstrated that Stroop interference affected RTs to pictures in homogeneous cycles. This suggested that the executive control mechanisms to overcome semantic context effects are (at least partially) shared with those involved in any task requiring the selection of a target among irrelevant and interfering stimuli. However, the association between semantic blocking and Stroop interference is inconsistent across studies (e.g., Shao et al., 2015). Another issue is that the Stroop task may involve linguistic processes. Hence, any correlation like the one reported by Crowther and Martin (2014) may be primarily driven by linguistic processes rather than by domain-general executive control mechanisms.

By contrast, the results obtained by Alario et al. (2012) suggested that the ability to overcome semantic blocking is language-specific. These authors found no relationship between the semantic blocking in a cyclical picture naming task and the Simon effect, which is an index of domain-general executive control. The widely used “Simon task” (Simon, 1990) has the advantage of not involving linguistic material. A spatial response (i.e., “right” or “left” button) is associated with the color (e.g., “red” or “green”) of a stimulus (e.g., a circle). For instance, participants are instructed to press the right button if the circle is red and left button if the circle is green. Critically, the stimuli appear randomly either on the right or on the left of the screen. Responses are slower in trials in which the position of the stimulus does not match the spatial response (congruent trials; a green circle on the right-hand side of the screen, and a red circle on the left-hand side) compared to trials in which it does (incongruent trials; a green circle on the left-hand side of the screen, and a red circle on the right-hand side). The Simon effect refers to this difference in RTs between incongruent and congruent trials: the smaller it is, the greater the ability to suppress irrelevant and interfering information (the position of the stimulus).

In a similar vein, Calabria et al. (2019) suggested that failure to overcome semantic blocking effects might reflect weaknesses in language-specific control processes. Aphasics in their study showed greater blocking effect compared to controls in the most challenging condition (naming in L2). The authors also indexed participants’ domain-general executive control with the flanker task, which is another widely used task not involving linguistic material that assesses participants’ ability to select a target among irrelevant and interfering stimuli. Participants indicate the direction (right vs. left) of a central arrow while ignoring flanker arrows pointing to either the same (congruent trials) or opposite (incongruent trials) direction, typically showing a conflict effect (hereafter referred to as conflictFT; i.e., RTs in incongruent trials are longer than those in congruent trials). The similar conflictFT in aphasics and controls observed by Calabria et al. suggested that the aphasics’ domain-general executive control was preserved and did not cause group differences in the magnitude of the L2 semantic blocking effect. In addition, conflictFT did not predict the magnitude of L1 or L2 semantic blocking effects, neither in the controls nor in the patients. There was only a correlation between the overall speed in the flanker task (but not conflictFT) and the magnitude of the semantic blocking effect in L2, which the authors argued was simply related to language control challenges in bilinguals rather than to semantically related contexts.

In the current study, we used conflictFT to index participants’ domain-general executive control capacity. If the ability to overcome semantic blocking involves the domain-general executive control system, conflictFT should predict the magnitude of the semantic blocking effect.

***The current study***

The current study included two experiments. In Experiment 1, we tested the role of language dominance and domain-general executive control in speech production (picture naming). We found that the magnitude of the semantic blocking effect was *(i)* similar between participants tested in L1 and participants tested in L2, and *(ii)* predicted by conflictFT. This suggests that the semantic blocking effects in speech production are sensitive to domain-general control processes, but that these do not interact with language dominance. We then wondered whether these results are dependent on the need for lexical selection in picture naming. We addressed this question in Experiment 2 by repeating the study using a speech comprehension variant of the paradigm (auditory word-picture matching). The results of this follow-up experiment showed that language dominance had no effect and that conflictFT did not exert any influence. We will discuss the results of these two experiments in the context of bilingual language processing and the rather complex state-of-the-art regarding the cognitive processes subserving general-domain executive control and semantic context effects.

**Experiment 1 (speech production)**

**Methods**

***Participants***

Forty-eight participants gave written consent to take part in the experiment in exchange for course credits. Ethical approval was granted by the Bioethical Committee of the University of Barcelona (BCUB). All participants had Spanish as their first and dominant language (L1). Half of them performed the task in Spanish (L1 group) and the other half in English (L2 group). The groups did not differ in age (mean age in the L1 group = 22.37 years, SD = 4.37, range = 18-32; mean age in the L2 group = 22.17, SD = 2.66, range = 18-29; *t* = 1.99, *p* < .84) or educational attainment (all participants were students at the University of Barcelona) and had a similar gender distribution (L1 group: 15 females, 9 males; L2 group: 17 females, 7 males). All participants in the L2 group stated that they had acquired English through classroom instruction and had a Common European Framework (CEF) English proficiency level of “lower intermediate” (CEF = B1) or “upper intermediate” (CEF = B2). This proficiency level was estimated through an English vocabulary size test that correlates with the CEF proficiency levels: the Lexical Test for Advanced Learners of English (LexTALE; Lemhöfer & Broersma, 2012). Participants in the L1 group had the expected CEF level of “proficient users of Spanish”, as determined with the Spanish version of the LexTALE (Izura et al., 2014).

***Tasks***

All participants performed a picture naming task variant of the cyclical semantic paradigm, followed by a flanker task (administered with DMDX; Forster & Forster, 2003).

***Cyclical picture naming task.*** The stimuli were 16 pictures belonging to four different semantic categories. The names of the pictures did not overlap phonologically between Spanish and English (animals: rabbit/*conejo*, duck/*pato*, fox/*zorro*, and owl/*búho*; fruits: strawberry/*fresa*, orange/*naranja*, apple/*manzana*, and grapes/*uva*; tools: hammer/*martillo*, screwdriver/*destornillador*, pliers/*alicates*, and drill/*taladro*; clothes: trousers/*pantalón*, hat/*sombrero*, shirt/*camisa*, and dress/*vestido*). Pictures were black and white line drawings from the Snodgrass and Vanderwart (1980) corpus, except for the line drawing of “drill”, which was taken from the internet.

There were four homogeneous and four heterogeneous blocks of cycles. Each block was composed of four cycles, with each cycle containing four pictures. Therefore, each block of cycles included 16 trials, giving a total number of 128 trials per participant. In the homogeneous blocks, each of the four cycles featured all four members of a semantic category (e.g., all clothes). In the heterogeneous blocks, each of the four cycles featured one member of each sematic category (i.e., an animal, a fruit, a tool, and an item of clothing; Figure 1). The task ran continuously within a block, that is, there were no pauses between the cycles. Pauses occurred between the blocks only. The picture names within a cycle were phonologically unrelated (both in the English and Spanish tasks). The pictures were presented in a random fashion within a cycle, but care was taken to ensure that the last picture of a cycle was never the same as the first picture of the next one. The homogeneous and heterogeneous blocks were alternated during the task. We used a Greco-Latin square design to counterbalance the order of the presentation of the blocks among the participants within each language group.

*(Figure 1 about here)*

The procedure started with a familiarization phase. Participants were presented with all the pictures and their corresponding names. Picture names were presented both in a written form (below the picture) and verbally (through headphones) in either English or Spanish (according to the participant’s language group). Participants were able to go back to each previous stimulus. Once the participants had familiarized themselves with the stimuli, the task began. In each trial, a fixation sign (+++) was presented at the center of the screen for 130 milliseconds (ms). After a brief interval of 60 ms, the picture was presented for 2,500 ms or until there was a response. Participants were instructed to name the pictures out loud in the required language (English or Spanish). Naming responses and RTs were recorded online using a voice key and corrected later with the program Check Vocal (Protopapas, 2007).

***Flanker task.*** We used the same variant of the flanker task as the one applied in Calabria et al. (2019). Participants indicated the direction (right vs. left) of a central arrow (→), which was flanked by two arrows on each side pointing to either the same (congruent trials, →→→→→) or the opposite (incongruent trials, ←←→←←) direction. There were 288 trials in total: 216 (75%) congruent trials and 72 (25%) incongruent trials. The task was divided into three blocks, each with the same 75%–25% proportion of congruent and incongruent trials. Before the task, participants performed 32 practice trials with the same congruency proportion as the actual task. In each trial, a fixation cross (+) was presented at the center of the screen for 400 ms. The target arrow and the flankers were presented simultaneously immediately after the fixation cross and remained on the screen until there was a response or up to 1,700 ms. Then, the next trial began. Participants were instructed to respond by pressing the right or left arrow of the keyboard as quickly and accurately as possible. Each participant’s conflictFT wascalculated by subtracting the mean RTs of the correct responses in the congruent trials from the mean RTs of the correct responses in the incongruent trials. A smaller conflictFT indicated more efficient domain-general executive control.

***Data analyses***

In both tasks (cyclical picture naming and flanker tasks), trials classified as an error (i.e., incorrect responses or no responses) were excluded from the analyses. Participants were highly accurate in both the cyclical picture naming (L1 group: 95.5% correct responses, L2 group: 95.7% correct responses) and flanker tasks (L1 group: 97.1% correct responses, L2 group: 97.4% correct responses). We performed the analyses on the RT data, which we cleaned by excluding RTs faster than 250 ms and applying the non-recursive outlier deletion procedure with shifting z-score criterion (Van Selst & Jolicoeur, 1994). This resulted in the loss of only 0.1% of the data from the cyclical picture naming task and 0.03% of the data from the flanker task.

Unless otherwise stated, the analyses of the cyclical picture naming data were carried out using mixed-effect models. Naming RTs were log-transformed to reduce the skewness in their distribution. We referred to this measure as naming logRTs. Naming logRTs was the dependent variable in all the models, which had the maximal random effect structure justified by the data (Barr et al., 2013). We included participant and item as the random effects in all the models. The models that included the factor block type as the fixed effect also had by-participant random slope for block type as a random effect. We used deviation coding for categorical predictor variables (i.e., block type and group) to interpret their main effects. In models including more than one fixed effect, we examined all possible interactions between them in addition to the main effects. Following recent recommendations to assess significance in mixed-effect models (Luke, 2017), we used the Satterthwaite approximation for degrees of freedom (Satterthwaite, 1941) and fitted the models using maximum likelihood (REML).

As for the flanker task, we first checked that participants showed reliable conflictFT (slower RTs in incongruent compared to congruent trials) by conducting an ANOVA with RTs as the dependent variable. The ANOVA included type of trial (congruent vs. incongruent) as the within-subject factor and group (L1 vs. L2) as the between-subject factor. The individual magnitude of the conflictFT was converted into logarithmic values to facilitate the interpretation of the parameter estimates[[2]](#endnote-2). We referred to this measure as log(conflictFT). Then, we used log(conflictFT) to examine whether domain-general executive control contributed to predict the overcoming of the semantic blocking effect at the single trial level. To do this, we included log(conflictFT) as a fixed effect in different mixed effect models of naming logRTs.

The analyses were conducted with the R package (version 3.4.1; R Development Core Team, 2008).

**Results**

Figure 2A-C illustrates the pattern of performance in the cyclical picture naming task in the participants from the L1 group (Figure 2A), participants from the L2 group (Figure 2B) and all collapsed participants (Figure 2C). Table 1 summarizes the RTs.

*(Figure 2 and Table 1 about here)*

As there is generally no semantic blocking in cycle 1 of the picture naming task, this first cycle has been excluded from analyses in many studies (e.g., Shao et al., 2015 – Study 2). A visual inspection of Figure 2A-C suggests that we replicated the absence of semantic blocking in cycle 1. This observation was confirmed with a model including the naming logRTs of the first cycle only as a dependent variable and block type (homogeneous vs. heterogeneous), group (L1 vs. L2) and the interaction between these two variables as the fixed effects. The results showed no main effect of block type (*t* = 1.31, *p* < .19) or group (*t* < 1) and no interaction between block type and group (*t* = 1.06, *p* < .29), indicating that naming latencies in the homogeneous and heterogeneous blocks did not differ in either group of participants. Even though block type and group did not interact, we ran additional t-test comparing the RTs of the homogeneous and heterogeneous blocks in each group of participants separately. Semantic blocking was absent in cycle 1, both in the L1 (mean = -28 ms, SD = 85; *t* = 1.63, *p* < .12) and L2 (mean = 10 ms, SD = 75; *t* < 1; Figure 2A-C, Table 1) groups. Therefore, we only considered cycles 2, 3 and 4 in this experiment. These three cycles were collapsed in all the analyses to avoid having too few observations in the fixed effect combinations in some of the models, which is in fact common practice (for review, see de Zubicaray et al., 2014).

***Semantic blocking and language dominance.*** We tested whether the participants’ capacity of overcoming semantic blocking was affected by language dominance, using a model that had block type (homogeneous vs. heterogeneous) and group (L1 vs. L2) as the fixed effects. Surprisingly, there was no main effect of group (*t <* 1), indicating that participants in the L2 group responded as fast as the participants in the L1 group. There was a main effect of block type (*t* = 10.38, *p* = .0001) because the mean RT was 44 ms longer (SD = 28) for the homogeneous than for the heterogeneous blocks (Figure 2D, Table 1). The lack of a two-way interaction between group and block type (*t* < 1) indicated that participants in the L1 and L2 groups did not show differences in the magnitude of the semantic blocking effect. Hence, language dominance did not affect participants’ capacity to overcome semantic blocking[[3]](#endnote-3).

***Semantic blocking and conflictFT.*** The results of the ANOVA showed a main effect of the type of trial (*F*(1, 46) = 641.47, *MSE* = 668.47, *p* < .0001), which revealed that the responses were slower in the incongruent trials than in the congruent ones. This means the participants showed reliable conflictFT (mean = 134 ms, SD = 36). The lack of any effect involving group (*F*s < 1) indicated that conflictFT was of similar magnitude for participants in the L1 (mean = 135 ms, SD = 42) and L2 (mean = 132 ms, SD = 31) groups. Given the lack of differences between participants in the L1 and L2 groups in their capacity to overcome semantic blocking, we collapsed them all to test the influence of conflictFT on thiscapacity. As mentioned in the data analysis section, we tested this influence by fitting a mixed-effect model to the single trial data from the cyclical picture naming task. Hence, the dependent variable in these models was naming log(RTs) and the log(conflictFT) was used as a fixed effect. Only a conflictFT effect that is particularly evident in contexts where semantic blocking occurs (i.e., in homogeneous rather than in heterogeneous blocks) would allow us to conclude that domain-general executive control aided the overcoming of semantic blocking in this experiment. Thus, we first ran a mixed-effect model with block type (homogeneous vs. heterogeneous) and log(conflictFT) as the fixed effects, seeking a critical two-way interaction between log(conflictFT) and block type. The main effect of log(conflictFT) was marginally significant(*t* = 1.99, *p* < .052), revealing a slight influence of domain-general executive control on participants’ overall naming speed. More importantly, the two-way interaction between log(conflictFT) and block type was significant (*t* = 2.35, *p* < .02). We then ran two separate models, one for each type of block (homogeneous and heterogeneous), with log(conflictFT) as the only fixed effect. The results of these models showed that log(conflictFT) only predicted performance in homogeneous (*t* = 2.27, *p* < .02), but not in heterogeneous blocks (*t* = 1.14, *p* < .26). This means that domain-general executive control was in fact only recruited due to the need to overcome semantic blocking (Figure 3).

*(Figure 3 about here)*

It could be argued that the participants tested in their non-native language were as efficient as those tested in their native language in overcoming semantic blocking because they recruited domain-general executive control to a greater extent. To test this possibility, we complementarily re-ran the last model, including group as an additional fixed effect. The lack of a two-way interaction between log(conflictFT) and group (*t* < 1) suggested that domain-general executive control aided the overcoming of semantic blocking in participants from both language groups similarly.

**Experiment 2 (speech comprehension)**

**Methods**

***Participants***

Forty-nine participants gave written consent to take part in the experiment in exchange for course credits. Of these, 25 performed the task in Spanish (L1 group) and the other 24 in English (L2 group). Ethical approval was granted by the BCUB. The groups did not differ in age (mean age in the L1 group = 20.44 years, SD = 1.8, range = 18-26; mean age in the L2 group = 20.17, SD = 2.09, range = 18-26; *t* < 1) or educational attainment (all participants were students at the University of Barcelona) and had a similar gender distribution (L1 group: 23 females, 2 males; L2 group: 22 females, 2 males). All participants in the L2 group stated that they had acquired English through classroom instruction and had a level of proficiency in English of “lower intermediate” (CEF = B1) or “upper intermediate” (CEF = B2), as measured with LexTALE (Lemhöfer & Broersma, 2012). Participants in the L1 group were determined to be “proficient users of Spanish” with the Spanish version of the LexTALE (Izura et al., 2014).

***Tasks***

All participants performed the auditory word-picture matching task variant of the semantic blocking paradigm (administered with E Prime software 2.0; Psychology Software Tools, Pittsburgh, PA) and then a flanker task (administered with DMDX; Forster & Forster, 2003).

***Cyclical auditory word-picture matching task.*** The materials and design of this task were the same as in Experiment 1. In each trial, a spoken word was presented simultaneously with an array of four pictures (a target and three distractors). Participants needed to select the picture that matched the spoken word by pressing one of the four keyboard keys that were horizontally aligned with the position of the pictures on the screen. In the homogeneous blocks, the four pictures presented in each trial belonged to the same semantic category. This means that the four trials of each cycle within a specific semantic category (e.g., clothes) always had the same four pictures (arranged in different orders; e.g., trousers, hat, shirt, and dress), each presented with a different spoken word. Hence, all members of a semantic category acted as a target within a cycle. In the heterogeneous blocks, the four pictures within a cycle belonged to a different semantic category (Figure 4). A native speaker of Spanish and a native speaker of English (both females) had previously recorded the spoken words, which were presented (through headsets) in Spanish to participants in the L1 group and English to participants in the L2 group. The procedure started with the familiarization phase (identical to the one in Experiment 1) followed by a 16-trial practice block (to familiarize the participants with the positions of the response keys). In each trial, the stimuli (spoken word and 4-picture array) were presented for 3,000 ms or until there was a response. The ISI was set to 0 ms and there was no fixation cross from trial to trial, so the next trial began immediately after a response was obtained.

*(Figure 4 about here)*

***Flanker task.*** We used the same task as in Experiment 1.

***Data analyses***

Participants were highly accurate in both the cyclical word-picture matching task (L1 group: 97% correct responses, L2 group: 93% correct responses) and the flanker task (L1 group: 93.4% correct responses, L2 group: 91.8% correct responses). We performed the analyses on the RT data. The data cleaning procedure was the same as in Experiment 1, which resulted in the loss of only 0.11% of the data from the cyclical word-picture tasks and 0.01% of the data from the flanker tasks. All procedures related to data analyses were the same as those used for Experiment 1. That is, unless otherwise stated we used mixed-effect models for the analyses on the cyclical word-picture matching data. The dependent variable in all the mixed models was log-transformed word-picture matching RTs (word-picture matching logRTs). The models’ specifications in terms of fixed and random effects were the same as those used in Experiment 1. As for the flanker task, we used the same ANOVA as in Experiment 1 to examine whether participants showed a reliable conflictFT. We then used the log-transformed conflictFT (log conflictFT) as a fixed factor in a mixed-effect model on picture matching logRTs to examine whether domain-general executive control influenced the overcoming of semantic blocking (see footnote 2).

**Results**

Figure 5 illustrates the pattern of performance in the word-picture matching task in participants from the L1 group (Figure 5A), participants from the L2 group (Figure 5B) and all collapsed participants (Figure 5C). Table 2 summarizes the RTs.

*(Figure 5 and Table 2 about here)*

As mentioned in the Introduction, semantic blocking usually arises in cycle 1 in the speech comprehension variant of the paradigm (in contrast to the speech production variant). Visual inspection of Figure 5A-C seems to indicate that we replicated the presence of semantic blocking in cycle 1. The results of two separate t-tests (one for each language group) confirmed that semantic blocking was present in cycle 1, both in the L1 (mean = 157 ms, SD = 97; *t* = 8.06, *p* < .0001) and L2 (mean = 112 ms, SD = 174; *t* = 3.13, *p* < .005; Figure 5A-C, Table 2) groups. Therefore, we considered all four cycles in this experiment, which were collapsed in all the analyses.

***Semantic blocking and language dominance.*** We ran a model with block type and group as the fixed effects. There was a main effect of block type (*t* = 19.34, *p* < .0001), revealing a mean semantic blocking of 178 ms (SD = 76; Figure 5D, Table 2). There was a main effect of group because, overall, participants tested in their non-native language performed the task 88.48 ms more slowly than those tested in their native language. There was no two-way interaction between type of block and group (*t* < 1), indicating that the magnitude of the semantic blocking effect was similar in both groups. These results suggest that although participants in the L2 group were slower in performing the task, they did not differ from participants in the L1 group in the capacity to overcome semantic blocking[[4]](#endnote-4).

***Semantic blocking and conflictFT.*** The results of the ANOVA showed a main effect of the type of trial (*F*(1, 47) = 409.81, *MSE* = 1,198.42, *p* < .0001), revealing that responses were slower in the incongruent than in the congruent trials. This means that participants showed reliable conflictFT (mean = 141 ms, SD = 49). The lack of any effect involving group (*F*s < 1) indicated that conflictFT was of similar magnitude in the L1 (mean = 148 ms, SD = 46) and L2 (mean = 136 ms, SD = 52) groups. Since the capacity to overcome semantic blocking was not affected by language dominance, we collapsed all participants in the mixed-effect models that we used to explore the influence of log(conflictFT) on this capacity. The results of a model including block type and log(conflictFT) as the fixed factors showed a marginal main effect of conflictFT (*t* = 1.93, *p* < .06). However, conflictFT did not interact with block type (*t* < 1). This suggested that domain-general executive control did not aid the capacity to overcome semantic blocking, merely exerting a slight benefit on the participants’ overall speed during task performance.

**Discussion**

We aimed to investigate two questions related to semantic context effects: whether such effects are sensitive to language dominance and whether they involve domain-general executive control. We first addressed these questions with a speech production variant of the cyclical semantic paradigm (a picture naming task; Experiment 1), which indexes the ability to overcome semantic context effects through the so-called semantic blocking. The results of this experiment indicated that semantic blocking is not affected by language dominance and that it is influenced by domain-general executive control. In a follow-up experiment, we set out to determine whether the lack of effect of language dominance and the influence of domain-general executive control could be replicated in a task that did not require lexical selection. For this purpose, we used a comprehension variant of the cyclical semantic paradigm (a word-picture matching task; Experiment 2). We replicated the lack of effect of language dominance in this follow-up experiment. However, in contrast to Experiment 1, semantic blocking did not engage domain-general executive control. Taken together, our results suggest that *(i)* semantic context effects are insensitive to language dominance, and that *(ii)* the involvement of domain-general executive control in these effects remains elusive. We will now discuss these findings in more detail.

***Semantic context effects are insensitive to language dominance***

One of our particularly interesting findings was that participants tested in L2 were, overall, slower in the word-picture matching task (Experiment 2), but not in the picture naming task (Experiment 1). The similar overall naming speed among the participants in both the L1 and L2 groups may seem surprising, as it is common that speakers name more slowly in L2 (e.g., Strijkers et al., 2013). However, it is in line with the lack of language effect shown by the control group in the study by Calabria et al. (2019). This finding may cast doubt on the notion that English is truly a weak language for participants in the L2 group during Experiment 1. However, participants in the L2 group for Experiment 1 (picture naming) were in fact less proficient in English than those in the L2 group for Experiment 2 (word-picture naming task), according to the LexTALE score (ISDT, range: 0 to 1; higher scores indicate better performance; Huibregtse et al., 2002). This fine-grained score revealed that although participants in the L2 group had an intermediate level of proficiency in English in both experiments, those in Experiment 1 (mean ISDT = 0.24, SD = 0.2) were less proficient than those in Experiment 2 (mean ISDT = 0.36, SD = 0.19; *t* = 2.12, *p* < .04). This observation disqualifies the argument that proficiency could explain the equal naming latencies among participants tested in L1 and those tested in L2.

A more plausible explanation for the equal overall naming latencies between the groups could be that bilingual language control mechanisms were engaged by participants in the L2 group to prevent interference from L1 lexical representations (e.g., the suppression of L1 and/or enhancement of L2 words when naming in L2). We hypothesize that these mechanisms could have enhanced the availability of L2 lexical representations, leading to similar naming latencies in the L1 and L2 groups. The results of several studies on bilingual language control using the so-called language-switching paradigm are in line with this hypothesis. In this paradigm, bilingual participants perform a picture naming task in which they need to switch back and forth between their two languages. An interesting finding in some of these experiments has been the faster overall naming latencies in L2 compared to L1 (e.g., Costa & Santesteban, 2004; Costa et al., 2006). This paradoxical result shows that naming latencies in L2 may speed up in response to bilingual language control needs. In our Experiment 1, these needs would have been elicited by the task of naming in a much weaker language than one’s own L1. The potential involvement of bilingual language control in Experiment 1, elicited by naming in the much weaker language, may have aided participants tested in L2 to overcome semantic blocking. For instance, having L2 lexical representations ready for selection may have reduced the impact of the semantic context in the homogeneous blocks. This means that semantic blocking may in fact have been greater in participants tested in L2 than in those tested in L1 if it were not for the “side effects” of bilingual language control. We were able to address this issue in Experiment 2. As the cyclical auditory word-picture matching task did not require lexical selection, participants in the L2 group did not have to apply bilingual language control to prevent the interference from L1 lexical representations. As expected, the overall RTs of participants tested in L2 were clearly slower than those of participants tested in L1. Despite this, performing the task in L1 or L2 did not affect the magnitude of the semantic blocking effect. This result is in line with data from neurologically intact individuals provided by the only two previous studies investigating this question (Calabria et al., 2019 (control group); Runnqvist et al., 2019). In addition, it adds to previous evidence by showing that semantic context effects are insensitive to language dominance when the L2 has been acquired late and participants have only an

intermediate proficiency level (in contrast to the early L2 acquisition and high L2 proficiency of the participants in the previous two studies).

This insensitivity to language dominance confirms neither of the hypotheses mentioned in the Introduction, which were focused on speech production and assumed lexical competition: one predicted reduced semantic blocking in L2 due to weak activation of lexical competitors (Runnqvist et al. 2012), while the other predicted enhanced semantic blocking in L2 because the intended representation would not be sufficiently activated to win against competitors. The only way that semantic blocking accounts assuming lexical competition (e.g., Abdel Rahman & Melinger, 2011; Belke et al., 2005; Damian et al., 2001; Harvey & Schnur, 2016) might accommodate our finding that semantic blocking is insensitive to language dominance in production and comprehension is by arguing that the spread of lexical-semantic activity occurs similarly in a weak L2 and an L1: this similar spreading would explain the absence of a language-dominance effect in the picture naming task, whereas no language-dominance effect would be expected in word-picture matching because this task does not involve lexical competition. In contrast, the observation that language dominance does not affect semantic blocking in production or comprehension is highly congruent with accounts of this effect that propose a conceptual locus (e.g., Gardner et al., 2012; Jefferies et al., 2007). Indeed, the fact that semantic blocking does not differ between an L1 and an L2 in either modality fits well with the widely accepted idea that conceptual representations are shared across languages (Kroll & Stewart, 1994). This means that the phenomenon at the basis of semantic blocking (interference between semantically related concepts) would occur similarly in L1 and L2. However, it is worth noting that the repetition priming account (e.g., Navarrete et al., 2010, 2012, 2014, 2016) might also accommodate the observation that semantic blocking is insensitive to language dominance in production and comprehension if we assume that incremental learning mechanisms take place bidirectionally at the semantic-lexical interface: it might be argued that these mechanisms are not influenced by the baseline level of semantic-lexical connections (which might be generally weakened in an L2). At the end of the Discussion we will consider how this account might, in fact, accommodate different observations in this study.

It is important to acknowledge here that we cannot rule out the possibility that participants in the L2 group performed the cyclical word-picture matching task (Experiment 2) by translating the L2 auditory words into L1. This might be possible because direct connectivity between L2 lexical representations and the conceptual system may be limited even for speakers with a high L2 proficiency level (e.g., Kotz & Elston-Güttler, 2004; Silverberg & Samuel, 2004)—it seems to depend on their particular history of L2 acquisition (e.g., Ferré et al., 2006). Note that the similar overall RTs between the L1 and L2 groups in the picture naming task (Experiment 1) makes unlikely that participants retrieved the response in L1 and then translated it into L2. This extra step would have slowed down the performance of the participants tested in L2.

***The involvement of domain-general executive control in semantic context remains elusive***

We found an association between participants’ performance in the flanker task and the cyclical picture naming task. First, participants with slower naming times tended to be, overall, slower in the flanker task as well. This is consistent with previous studies (Shao et al., 2012, 2013) showing a correlation between naming RTs and performance in a non-linguistic domain-general executive control task (i.e., the stop signal task; Logan & Cowan, 1984). We interpret this finding simply as consistency in the participants’ processing speed across the tasks. More interestingly, conflictFT predicted participants’ performance in the homogeneous blocks of the cyclical picture naming task. This does not mean that overcoming semantic context effects in speech production is fully regulated by the domain-general executive control system, but seems to indicate that it becomes involved, at least to a certain extent. This finding is consistent with the observation of Crowther and Martin (2014) that the Stroop interference correlated with participants’ performances in the homogeneous blocks of the cyclical picture naming task. As we used a non-linguistic executive control task, our results extend those of Crowther and Martin (2014) by showing that overcoming semantic context effects does in fact involve mechanisms of control that are domain-general. However, this finding is at odds with the results of Calabria et al. (2019), which is the only study, to our knowledge, that has examined the association between conflictFT and the magnitude of the semantic blocking effect in a cyclical picture naming task. They only found a correlation between overall RTs in the flanker task and the magnitude of the semantic blocking effect in L2 when collapsing the performance of controls and patients. However, it is unclear whether controls showed reliable semantic blocking in either L1 or L2 or whether patients showed it in L1, since Calabria et al. did not report statistical significance for these effects. Nevertheless, the authors suggested that such a correlation was likely to be driven by patients’ deficient bilingual language control.

Our findings in Experiment 1 (picture naming task) are also at odds with those in our Experiment 2 (word-picture matching task) and those reported in Alario et al. (2012), who found no relationship between the Simon effect and the semantic blocking effect in a picture naming task. These inconsistent findings underscore the elusiveness of the relationship between domain-general executive control and semantic context effects. The cognitive mechanisms underlying different executive control tasks are not yet fully understood, and this is likely to contribute to the diversity of results in the literature. For instance, differences in the type of cognitive mechanisms engaged by the Simon task and the flanker task may explain the contrasting results between our Experiment 1 and Alario et al.’s study. The Simon task is thought to mainly engage selective inhibition—i.e., executive control mechanisms that lower the levels of activation of competing responses (e.g., Ridderinkhof, 2002). In contrast, the flanker task is proposed to engage response selection, a decision-making mechanism that accumulates information from the environment until a threshold needed to select a response is reached (e.g., Hübner et al., 2010).

In this scenario, it is useful to try to elucidate the cognitive mechanisms at play in the cyclical semantic paradigm. Examination of the distributional patterns of responses may be of use in this matter. Typically, the magnitude of the Simon effect is reduced for longer response times (calculated across both congruent and incongruent trials) compared to faster overall response times (across both congruent and incongruent trials). This reduction in the Simon effect for response time has been associated with the engagement of selective inhibition. In contrast, the conflictFT (flanker task) typically increases with response time, which has been associated with the engagement of response selection (see Hübner and Töbel (2019) for an analysis of the similarities and differences between the flanker task and the Simon task based on distributional patterns). The results of an examination of this kind are reported in detail in Appendix A. In short, similarly to what we observed for the conflictFT,the magnitude of the semantic blocking effect increased with RTs (in both experiments). Therefore, it seems that the cyclical semantic paradigm relies more on response selection than on selective inhibition, which makes the use of the flanker task preferable to the Simon task in order to establish whether semantic blocking relies on domain-general executive control (at least according to the rationale based on distributional patterns). In this vein, one would expect that the increase in the magnitude of the effects with RTs (i.e., delta slopes) to correlate between the flanker task and semantic blocking. Appendix A shows that we did not find this correlation. However, this observation needs to be interpreted with caution because the tasks were not designed with the aim of examining associations between their distributional patterns. For instance, the steepness of the slope may vary between a blocked (cyclical paradigm) and a randomized (flanker task) design.

Another difficulty is that the locus of the semantic blocking effect remains a matter of debate. On the one hand, perhaps accounts assuming lexical competition (e.g., Belke et al., 2005; Howard et al., 2006) or semantic interference at the conceptual level (e.g., Gardner et al., 2012; Jefferies et al., 2007) would provide the most straightforward interpretation for the fact that conflictFT predicted RTs in the homogeneous blocks of the picture naming task: resolving competition or interference between semantically related lexical or conceptual representations is particularly difficult and needs to be aided (at least partly) by domain-general executive control mechanisms. The absence of a relationship between conflictFT and semantic blocking in the word-picture matching task would fit well with lexical accounts (since they assume competition only at lexical level) but it would be problematic for conceptual accounts. On this issue, one could argue that the perceptual load was probably higher in the word-picture matching task. Every trial of that task required the processing of four pictures and an auditory word, which involved a considerable amount of perceptual processing compared to the trials in the picture naming task. The effects of perceptual load on conflict resolution have been widely studied in the context of visual search paradigms (e.g., Wei et al., 2013). In our view, it is unclear how perceptual demands could have exactly overridden the influence of domain-general mechanisms in the comprehension variant of the cyclical semantic paradigm. However, we cannot categorically rule out that task differences in the perceptual load could have explained why conflictFT played a role only in the picture naming task. A related issue is that semantic blocking in the word-picture matching task may have been mainly driven by visual similarity: all pictures being visible in word-picture matching trials may increase the chances of visual similarity effects occurring in this task. This possibility, however, was ruled out by the results of the complementary analyses that we report in Appendix B, which indicated that the visual similarity effects were minimal in both experiments. First, although visual similarity was higher in the homogeneous blocks than in the heterogeneous ones, the median visual similarity value was rather low for both types of blocks. Second, both the degree of visual similarity and the magnitude of the semantic blocking effect varied across the semantic categories, but not in parallel. That is, higher visual similarity among items from a particular category did not necessarily involve greater semantic blocking for the former category.

On the other hand, Navarrete and colleagues’ account—which is based on the view that lexical selection does not involve competition (Dell, 1986)—could accommodate different observations in our study. First, as noted above, distributional pattern analyses suggest that overcoming semantic blocking relies on reaching a threshold for selection (as opposed to mechanisms lowering the level of activation of competitors). Second, there was a reduction in RTs from cycle 1 to cycle 2 which was more pronounced in heterogeneous than homogeneous blocks (mean differential RT reduction between type of blocks = -48 ms in picture naming (Experiment 1), and -45ms in word-picture matching (Experiment 2); Figure 2C and 5C). Such a speeding up could be due to repetition priming, which would be reduced in homogeneous relative to heterogeneous cycles because of incremental weakening occurring in the homogeneous ones. Here one may wonder whether the facilitation driven by repetition priming in the linguistic tasks is influenced by domain-general facilitation effects. We took a very preliminary step in testing this prediction with the only facilitation effect that we could compute in the current design of the flanker task: trial-to-trial adaptation (faster RTs to trials that are of the same type (congruent or incongruent) as the immediately previous one; mean adaptation effect in Experiment 1 = 17.68 ms, SD = 26.69, *t* = 4.58, *p* < .0001; mean adaptation effect in Experiment 2 = 19.31 ms, SD = 35.57, *t* = 3.79, *p* < .0001). A correlation between the trial-to-trial adaptation effect (flanker task) and the differential RT reduction from cycle 1 to cycle 2 between type of blocks would have confirmed this hypothesis. However, this correlation was not present in either Experiment 1 (picture naming; *r* = -0.03, *p* < .79) or Experiment 2 (word-picture matching; *r* = -0.06, *p* < .69). This is congruent with Freund and Nozari (2018)’s finding of trial-to-trial adaptation effects within a linguistic task and within a spatial task, but no transfer of this adaptation across the two tasks. Interestingly, some authors consider trial-to-trial adaptation effects to be task-specific (e.g., Funes et al., 2010). This is because they do not seem to transfer across different domain-general executive control tasks, unlike other types of adaptation effects, such as contextual adaptation (modulation of RTs as a function of the proportion of congruent trials). However, an undeniable flaw of our preliminary approach was that switch costs are very likely to contribute to the trial-to-trial adaptation effects of the flanker task. Therefore, specific task designs should be developed to address whether facilitation effects driven by semantic priming in the cyclical semantic paradigm are influenced by domain-general adaptation effects. Lastly, the fact that the conflictFT predicted RTs in the homogeneous but not in the heterogeneous blocks of the picture naming (but not the word-picture matching) task could be interpreted as domain-general executive control aiding conflict resolution at the late post-lexical stages. Navarrete and colleagues take the view that a lexical representation is retrieved when its level of activation reaches a threshold (Dell, 1986; as opposed to lexical competition), but they argue that competition may happen at the post-lexical level if more than one lexical representation reaches the threshold for lexical retrieval. In this case, post-lexical filtering mechanisms exclude the non-target representations so that only the relevant one is articulated. According to this proposal, referred to as the response exclusion hypothesis (Janssen et al., 2008; Mahon et al., 2007), the semantic category could be one of the filtering mechanisms used to prioritize one word for articulation among the various alternatives that have reached the threshold for selection. Therefore, words that do not match the semantic category of the target are excluded. Although the response exclusion hypothesis was formulated for the Stroop and picture word interference contexts, it could be also applied to the semantic blocking effects. If the semantic category is one of the criteria to exclude candidates for articulation, this process may be particularly difficult for homogeneous (compared to heterogeneous) blocks in a cyclical semantic paradigm. This means that homogeneous blocks would entail conflict at the post-lexical level. A tentative possibility is that the resolution of this conflict relies, at least to some extent, on domain-general executive control. For example, Dhooge and Hartsuiker (2010) reported an experimentally based argument that the post-lexical filtering introduced in the response exclusion hypothesis could in fact correspond to an error monitoring system, while Nozari et al. (2019) discussed whether and how error monitoring may recruit domain-general executive control.

**Conclusion**

Our results showed an association between domain-general cognitive control and semantic blocking in production but not in comprehension, and that semantic blocking in neither modality was sensitive to language dominance. While multiple interpretations of these results may be possible, it is interesting to note that an account rejecting lexical competition or semantic interference at the conceptual level (which have long been proposed by many authors) could accommodate the results: semantic blocking derives from incremental strengthening and weakening of semantic-lexical connections (in both directions) regardless of the general baseline strength of these connections, which explains why blocking did not differ between L1 and L2; domain-general executive control could have been recruited by post-lexical mechanisms of speech error monitoring, which is why it only exerted an influence on production. A relevant avenue for future research might be to test this specific interpretation.

**Acknowledgments**

Albert’s amazing way to embrace science prevails as Mireia’s essential inspiration.

**Funding**

MH was supported by the Ramón y Cajal Research Program of the Spanish Ministry of Science, Innovation, and Universities (RYC-2016-19477). EJ was supported by the European Research Council (FLEXSEM-771863).

**Disclosure of interest.** The authors report no conflict of interest.

**References**

Abdel Rahman, R., & Melinger, A. (2009). Semantic context effects in language production: A swinging lexical network proposal and a review. *Language and Cognitive Processes, 24*(5), 713–734. https://doi.org/10.1080/01690960802597250.

Abdel Rahman R, & Melinger A. (2011). The dynamic microstructure of speech production: Semantic interference built on the fly. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(1), 149–161. https://doi.org/10.1037/a0021208

Alario, F.-X., Ziegler, J. C., Massol, S., & de Cara, B. (2012). Probing the link between cognitive control and lexical selection in monolingual speakers. *L’Année psychologique*, *112*(4), 545–559. doi: https://doi.org/10.4074/S0003503312004010.

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.

Belke, E., Meyer, A. S., & Damian, M. F. (2005). Refractory effects in picture naming assessed in a semantic blocking paradigm. *The Quarterly Journal of Experimental Psychology Section A, 58*(4), 667–692. https://doi.org/10.1080/02724980443000142

Brown, A. S. (1981). Inhibition in cued retrieval. *Journal of Experimental Psychology: Human Learning and Memory, 7*(3), 204–215. http://dx.doi.org.sire.ub.edu/10.1037/0278-7393.7.3.204

Brysbaert, M., & Duyck, W. (2010). Is it time to leave behind the Revised Hierarchical Model of bilingual language processing after fifteen years of service? *Bilingualism: Language and Cognition 13*(3), 359–371. https://doi.org/10.1017/S1366728909990344

Calabria, M., Grunden, N., Serra, M., García-Sánchez, C., & Costa, A. (2019). Semantic processing in bilingual aphasia: Evidence of language dependency. *Frontiers in Human Neuroscience*, *13:205*. https://doi.org/10.3389/fnhum.2019.00205

Campanella, F., & Shallice, T. (2011). Refractoriness and the healthy brain: A behavioural study on semantic access. *Cognition, 118*(3), 417–431. https://doi.org/10.1016/j.cognition.2010.08.005

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology, 14*(1), 177−208. https://doi.org/10.1080/026432997381664

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language*, *50*(4), 491–511. https://doi.org/10.1016/j.jml.2004.02.002

Costa, A., Santesteban, M., & Ivanova, I. (2006). How do highly proficient bilinguals control their lexicalization process? Inhibitory and language-specific selection mechanisms are both functional. *Journal of Experimental Psychology: Learning Memory and Cognition*, *32*(5), 1057–1074. https://doi.org/10.1037/0278-7393.32.5.1057

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson’s method. *Tutorials in Quantitative Methods for Psychology, 1*(1), 4–45.

Crowther, J. E., & Martin, R. C. (2014). Lexical selection in the semantically blocked cyclic naming task: the role of cognitive control and learning. *Frontiers in Human Neuroscience*, *8:9*. https://doi.org/10.3389/fnhum.2014.00009

Damian, M. F., Vigliocco, G., & Levelt, W. J. M. (2001). Effects of semantic context in the naming of pictures and words. *Cognition*, *81*(3), 77–86. https://doi.org/10.1016/S0010-0277(01)00135-4

de Zubicaray, G., Johnson, K., Howard, D., McMahon, K. (2014). A perfusion fMRI investigation of thematic and categorical context effects in the spoken production of object names. *Cortex, 54,* 13–149. https://doi.org/10.1016/j.cortex.2014.01.018

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93*(3), 283−321. http://dx.doi.org/10.1037/0033-295X.93.3.283

Dhooge, E., & Hartsuiker, R. J. (2010). The distractor frequency effect in picture-word interference: Evidence for response exclusion. *Journal of experimental psychology. Learning, memory, and cognition*, *36*(4), 878–891. https://doi.org/10.1037/a0019128

Ferré, P., Sánchez-Casas, R., & Guasch, M. (2006). Can a horse be a donkey? Semantic and form interference effects in translation recognition in early and late proficient and nonproficient Spanish–Catalan bilinguals. *Language Learning, 56*(4), 571–608. https://doi.org/10.1111/j.1467-9922.2006.00389.x

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments & Computers*, *35*(1), 116–124. https://doi.org/10.3758/BF03195503

Freund, M., & Nozari, N. (2018). Is adaptive control in language production mediated by learning? Cognition, 176, 107–130. https://doi.org/10.1016/j.cognition.2018.03.009.

Funes, M. J., Lupiáñez, J., & Humphreys G. (2010). Sustained vs. transient cognitive control: Evidence of a behavioral dissociation. *Cognition 114,* 338–347. https://doi.org/10.1016/j.cognition.2009.10.007

Gardner, H. E., Lambon Ralph, M. A., Dodds, N., Jones, T., Ehsan, S., & Jefferies, E. (2012). The Differential Contributions of pFC and Temporo-parietal Cortex to Multimodal Semantic Control: Exploring Refractory Effects in Semantic Aphasia. *Journal of Cognitive Neuroscience, 24*(4), 778–793. <https://doi.org/10.1162/jocn_a_00184>

Harvey, D. Y., & Schnur, T. T. (2016). Different loci of semantic interference in picture naming vs. word-picture matching tasks. *Frontiers in Psychology*, *7:710*. https://doi.org/10.3389/fpsyg.2016.00710

Howard, D., Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: experimental and computational studies. *Cognition, 100*(3), 464–482. https://doi.org/10.1016/j.cognition.2005.02.006

Hübner, R., Steinhauser, M., & Lehle, C. (2010). A dual-stage two-phase model of selective attention. *Psychological Review, 117*(3), 759–784. https://doi.org/10.1037/a0019471

Hübner, R., & Töbel, L. (2019). Conflict resolution in the Eriksen flanker task: Similarities and differences to the Simon task. *PLoS One, 14*(3):e0214203. https://doi.org/10.1371/journal.pone.0214203

Huibregtse, I., Admiraal, W., & Meara, P. (2002). Scores on a yes–no vocabulary test: Correction for guessing and response style. *Language Testing, 19*, 227–245. https://doi:10.1191/0265532202lt229oa

Izura, C., Cuetos, F., & Brysbaert, M. (2014). Lextale-Esp: A test to rapidly and efficiently assess the Spanish vocabulary size. *Psicológica*, *35*, 49–66. https://doi.org/10.7498/aps.63.104215

Janssen, N., Schirm, W., Mahon, B. Z., & Caramazza, A. (2008). Semantic interference in a delayed naming task: Evidence for the Response Exclusion Hypothesis. Journal of Experimental Psychology: Learning, Memory, and Cognition, *34*(1), 249–256. https://doi.org/ 10.1037/0278-7393.34.1.249

Jefferies, E., Baker, S. S., Doran, M., & Lambon Ralph, M. A. (2007). Refractory effects in stroke aphasia: A consequence of poor semantic control. *Neuropsychologia*, *45*(5), 1065–1079. https://doi.org/10.1016/j.neuropsychologia.2006.09.009

Kotz, S. A., & Elston-Güttler, K. (2004). The role of proficiency on processing categorical and associative information in the L2 as revealed by reaction times and event-related. *Journal of Neurolinguistics, 17*(2–3), 215–235. https://doi.org/10.1016/S0911-6044(03)00058-7

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33*(2), 149–174. https://doi.org/10.1006/jmla.1994.1008

Kroll, J. F., Van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The Revised Hierarchical Model: A critical review and assessment. *Bilingualism: Language and Cognition, 13*(3), 373–381. https://doi.org/10.1017/S136672891000009X

La Heij, W. (1988). Components of Stroop-like interference in picture naming. *Memory and Cognition,* *16,* 400−410. https://doi.org/10.3758/BF03214220

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, *44*(2), 325–343. https://doi.org/10.3758/s13428-011-0146-0

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*(1), 1−75. https://doi.org/10.1017/S0140525X99001776

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review, 91,* 295–237. http://dx.doi.org/10.1037/0033-295X.91.3.295

Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior research methods, 49*(4), 1494–1502. http://dx.doi.org/10.3758/s13428-016-0809-y

Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., & Caramazza, A. (2007). Lexical selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 503–535. https://doi.org/10.1037/0278-7393.33.3.503

McRae, K., & Boisvert, S. (1998). Automatic semantic similarity priming. Journal of Experimental Psychology: Learning, Memory, and Cognition, 24(3), 558–572. http://dx.doi.org/10.1037/0278-7393.24.3.558

Nathaniel, U., Thompson, H. E., Davies, E., Arnold, D., Hallam, G., Stampacchia, S., Smallwood, J., & Jefferies, E. (2018). When comprehension elicits incomprehension: Deterioration of semantic categorisation in the absence of stimulus repetition. *Quarterly journal of experimental psychology, 71*(9), 1817–1843. https://doi.org/10.1080/17470218.2017.1363793

Navarrete, E., Del Prato, D., & Mahon, B. Z. (2012). Factors determining semantic facilitation and interference in the cyclic naming paradigm. *Frontiers in Psychology*, *3:38*. https://doi.org/10.3389/fpsyg.2012.00038

Navarrete, E., Del Prato, P., Peressotti, F., & Mahon, B. Z. (2014). Lexical retrieval is not by competition: Evidence from the blocked naming paradigm. *Journal of memory and language*, *76*, 253–272. https://doi.org/10.1016/j.jml.2014.05.003

Navarrete, E., Mahon, B. Z., & Caramazza, A. (2010). The cumulative semantic cost does not reflect lexical selection by competition. *Acta Psychologica, 134*(3), 279–289. https://doi.org/10.1016/j.actpsy.2010.02.009

Navarrete, E., Mahon, B. Z., Lorenzoni, A., & Peressotti, F. (2016). What can written-words tell us about lexical retrieval in speech production? *Frontiers in Psychology*, *6:1982*. https://doi.org/10.3389/fpsyg.2015.01982

Nozari, N., & Hepner, C. R. (2018). To select or to wait? The importance of criterion setting in debates of competitive lexical selection. Cognitive Neuropsychology, 1–15. http://doi.org/10.1080/02643294.2018.1476335.

Nozari, N., Martin, C. D., & McCloskey, N. (2019). Is repairing speech errors an automatic or a controlled process? Insights from the relationship between error and repair probabilities in English and Spanish. *Language, Cognition and Neuroscience, 43*(9), 1230–1245. https://doi.org/10.1080/23273798.2019.1637007.

Oppenheim, G., M., Dell, G. S., & Schwartz, M. F. (2010). The dark side of incremental learning: A model of cumulative semantic interference as learning. *Cognition, 114*(2), 227–252. https://doi.org/10.1016/j.cognition.2009.09.007

Protopapas, A. (2007). Check Vocal: A program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods*, *39*(4), 859–862. https://doi.org/10.3758/BF03192979

R Core Team. (2018). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review, 107*(3), 460–499. http://dx.doi.org/10.1037/0033-295X.107.3.460

Ridderinkhof, K. R. (2002). Activation and suppression in conflict tasks: Empirical clarification through distributional analyses. In W. Prinz & B. Hommel (Eds.), Attention and performance XIX: Common mechanisms in perception and action (pp. 494–519). New York: Oxford University Press.

Roelofs, A. (2003). Goal-referenced selection of verbal action: Modelling attentional control in the Stroop task. *Psychological Review, 110*(1), 88−125. https://doi.org/10.1037/0033-295X.110.1.88

Runnqvist, E., Strijkers, K., Alario, F.-X., & Costa, A. (2012). Cumulative semantic interference is blind to language: implications for models of bilingual speech production. *Journal of Memory and Language, 66*(4), 850−869. https://doi.org/10.1016/j.jml.2012.02.007

Satterthwaite, F. E. (1941). Synthesis of variance. *Psychometrika, 6*(5), 309–316. https://doi.org/10.1007/BF02288586

Shao, Z., Meyer, A. S., & Roelofs, A. (2013). Selective and nonselective inhibition of competitors in picture naming. *Memory & Cognition, 41*(8)*,* 1200–1211. http://dx.doi.org/10.3758/s13421-013-0332-7

Shao, Z., Roelofs, A., Martin, R. C., & Meyer, A. S. (2015). Selective inhibition and naming performance in semantic blocking, picture-word interference, and colot-word Stroop tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(6), 1806–1820. http://dx.doi.org/10.1037/a0039363

Shao, Z., Roelofs, A., & Meyer, A. S. (2012). Sources of individual differences in the speed of naming objects and actions: The contribution of executive control. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 65*(10)*,* 1927–1944. http://dx.doi.org/10.1080/17470218.2012.670252

Silverberg, S., & Samuel, A. G. (2004). The effects of age of acquisition and fluency on processing second language words: Translation or direct conceptual access? *Journal of Memory and Language, 51*(3), 381–398. https://doi.org/10.1016/j.jml.2004.05.003

Simon, J. R. (1990). The effects of an irrelevant directional cue on human information processing. In R. W. Proctor & T. G., Reeve (Eds.), *Stimulus-response compatibility: An integrated perspective* (p. 31-86). Amsterdam, North-Holland: Elsevier.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human learning and Memory, 6*(2), 174–215. http://dx.doi.org/10.1037/0278-7393.6.2.174

Strijkers, K., Baus, C., Runnqvist, E., FitzPatrick, I., & Costa, A. (2013). The temporal dynamics of first versus second language production. *Brain & Language 127*(1), 6–11. http://dx.doi.org/10.1016/j.bandl.2013.07.008

Van Selst, M., & Jolicoeur, P. (1994). A solution to the effect of sample size on outlier elimination. *Quarterly Journal of Experimental Psychology*, *47A*, 631–650. https://doi.org/10.1080/14640749408401131

Wei, P., Kang, G., & Zhou, X. (2013). Attentional selection within and across hemispheres: implications for the perceptual load theory. *Experimental brain research*, *225*(1), 37–45. https://doi.org/10.1007/s00221-012-3346-7

Wei T., & Schnur, T. T. (2016). Long-Term Interference at the Semantic Level: Evidence from blocked-cyclic picture matching. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(1), 149–157. http://dx.doi.org/10.1037/xlm0000164

**Appendix A. Distributional pattern analyses**

We examined whether the magnitude of the conflictFT and the semantic blocking effects increased with RTs. Cognitive models about the flanker task (e.g., the dual-stage two-phase (DSTP) model; Hübner et al., 2010), explain this increase as the reflection of a late (and slow) response selection strategy driven by the identity (target vs. flankers) of the arrows taking over from an early perceptual strategy (right vs. left) in some trials.

To examine whether the conflictFT increased as a function of RT in our experiments, we divided the RTs of each flanker task (Experiment 1 and Expeiment 2) into three tertiles, which we ordered from the fastest (tertile 1) to the slowest (tertile 3). This division was performed for incongruent and congruent trials separately. This means that each tertile contains one third of RTs to incongruent trials and one third of the RTs to congruent trials. *T*-tests comparing RTs to incongruent and RTs to congruent trials showed a reliable conflictFT in every tertile of the flanker task in Experiment 1 (mean conflictFT in tertile 1 = 111.65, SD = 30.52; mean conflictFT in tertile 2 = 131.9, SD = 31.18, mean conflictFT in tertile 3 = 155.65, SD = 58.25; *p*s < .0001) and Experiment 2 (mean conflictFT in tertile 1 = 108.37, SD = 33.52; mean conflictFT in tertile 2 = 136.67, SD = 47.35, mean conflictFT in tertile 3 = 175.71, SD = 80.85; *p*s < .0001). Further *t*-tests comparing every other tertile in the magnitude of the conflictFT indicated that, as expected, this magnitude increased across tertiles (*p*s < .0004).

We next found that, in both experiments, the magnitude of the semantic blocking effect also increased across tertiles, which were calculated for homogeneous and heterogeneous blocks separately. *T*-tests comparing RTs between homogeneous and heterogeneous blocks showed a reliable semantic blocking effect in every tertile of Experiment 1 (mean semantic blocking effectin tertile 1 = 16.27, SD = 27.82; mean semantic blocking effectin tertile 2 = 33.44, SD = 24.18, mean semantic blocking effectin tertile 3 = 79.96, SD = 67.8; *p*s < .0002) and Experiment 2 (mean semantic blocking effectin tertile 1 = 65.73, SD = 42.26; mean semantic blocking effectin tertile 2 = 138.9, SD = 70.04, mean semantic blocking effectin tertile 3 = 331.37, SD = 151.9; *p*s < .0001). Further *t*-tests comparing every other tertile in the magnitude of the semantic blocking effect confirmed a significant increase across the three tertiles of each experiment (*p*s < .0004).

Additionally, we examined whether the delta plot slopes of the flanker tasks and semantic blocking tasks were associated (“delta” being the difference in RTs between trials in homogeneous and heterogeneous blocks in the cyclical tasks, and between incongruent and congruent trials in the flanker tasks). The delta plot slopes indicate the increase in the magnitude of a given effect across two successive quantiles taking into consideration overall RTs (Ridderinkhof, 2002). This means that we calculated delta plot slopes as: delta (tertile *y*) – delta (tertile *x*) / mean (tertile *y*) – mean (tertile *x*). We made this calculation for every possible slope. That is, we had the delta plot slope from tertile 1 to tertile 2 in the cyclical paradigm (Experiment 1: mean = 0.181, SD = 0.283; Experiment 2: mean = 0.416, SD = 0.249) and the flanker task (Experiment 1: mean = 0.189, SD = 0.174; Experiment 2: mean = 0.239, SD = 0.252), from tertile 2 to tertile 3 in the cyclical paradigm (Experiment 1: mean = 0.246, SD = 0.327; Experiment 2: mean = 0.485, SD = 0.244) and the flanker task (Experiment 1: mean = 0.129, SD = 0.218; Experiment 2: mean = 0.192, SD = 0.251), and from tertile 1 to tertile 3 in the cyclical paradigm (Experiment 1: mean = 0.228, SD = 0.243; Experiment 2: mean = 0.462, SD = 0.2) and the flanker task (Experiment 1: mean = 0.152, SD = 0.133; Experiment 2: mean = 0.209, SD = 0.22). There was no correlation between the corresponding slopes of the cyclical paradigm and the flanker task in either experiment (*r*s < .068).

**Appendix B. Visual similarity**

We collected visual similarity ratings through an online survey from 53 volunteers (13 males, 40 females; mean age = 25 years, SD = 6.76). They were each presented with two pictures side-by-side and rated how visually similar they were on a 5-point scale (1 = *not at all similar*, 5 = *very similar*). As ratings were not normally distributed, we analyzed them using the Wilcoxon signed-rank test. As expected, pictures in the homogeneous blocks were rated as more similar than those in the heterogeneous blocks (*V* = 21826, *p* < .0001). However, the level of visual similarity was rather low for both the homogeneous and heterogeneous blocks, with a median (Mdn) of 2 (*not similar)* and 1 (*not at all similar*), respectively (Figure B.1A). Hence, it is unlikely that visual similarity accounted for the semantic blocking effect, at least entirely. Nevertheless, we sought further evidence by examining if the pattern of variability in terms of semantic blocking across the categories paralleled that in terms of visual similarity.

To this end, we first assessed whether the semantic categories varied in visual similarity through paired comparisons with the Wilcoxon signed-rank test and using Bonferroni correction for multiple comparisons (α threshold = .008). The results showed that all the semantic categories differed between one another in visual similarity (*p*s < .0001), except for clothes and animals (*p* > .38). Pictures in the category of fruits were the most visually similar, followed by pictures of tools, clothes and animals (Figure A.1B). Then, in each experiment, we ran eight separate mixed-effect models to compare the magnitude of the semantic blocking effect between each semantic category (Bonferroni’s α threshold = .008)—the potential visual similarity issue is more relevant in the word-picture matching task (Experiment 2) where all pictures appear at once but we performed the same analyses on picture naming data (Experiment 1) for the sake of completeness. These models used block type and semantic category as the fixed effects and participants, item, and by-participant random slope for block type as the random effects. The results are reported below.

***Experiment 1*.** The magnitude of the semantic blocking effect was the largest for tools (mean = 108 ms, SD = 93), followed by fruits (mean = 53 ms, SD = 54), clothes (mean = 27 ms, SD = 62) and animals (mean = -5 ms, SD = 48). The results of the mixed models revealed that neither the comparison between clothes and animals (*p* > .45) nor the comparison between fruits and tools reached the threshold for significance (*p* > .02; all remaining *p* values were < .0001). The results of the additional models (one per semantic category with block type as the fixed factor) showed that the semantic blocking effect was significant in all the categories (Bonferroni’s α threshold = .01; *p* values < .003), except for animals (*t* < 1). Two observations indicated that visual similarity was not (at least entirely) behind the semantic blocking. First, semantic blocking was not significantly different between tools and fruits, even though pictures of fruits were more visually similar than pictures of tools. Second, although the pictures of animals and clothes did not differ in terms of their visual similarity ratings, the semantic blocking effect was significant only for clothes.

***Experiment 2*.** The magnitude of the semantic blocking effect was the largest in the category of tools (mean = 448 ms, SD = 150), followed by animals (mean = 168 ms, SD = 99), fruits (mean = 117 ms, SD = 87), and clothes (mean = 59 ms, SD = 91). The results of the mixed models revealed that all the categories differed significantly between one another (*p* values < .0008). The results of the additional models (one per semantic category with block type as the fixed factor) showed that the semantic blocking effect was significant in each category (Bonferroni’s α threshold = .01; *p* values < .0001). Three observations suggested that visual similarity did not account for semantic blocking (at least entirely). First, semantic blocking was smaller for fruits than animals, even though visual similarity was higher among pictures of fruits compared to pictures of animals. Second, although pictures of tools were less visually similar than pictures of fruits, semantic blocking was much larger for the category of tools than for the category of fruits. Lastly, the categories of animals and clothes did not differ in terms of their visual similarity ratings, but they did in the magnitude of the semantic blocking effect (animals > clothes).

**Table 1.** RTs (SD) in the cyclical picture naming task (Experiment 1).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Cycle 1 | Cycle 2 | Cycle 3 | Cycle 4 | **Total (Cycles 2 to 4)** |
| **L1 group** |  |  |  |  |  |
|  |  |  |  |  |  |
| Homogeneous blocks | 790 (101) | 733 (85) | 726 (124) | 729 (126) | **729 (105)** |
| Heterogeneous blocks | 818 (85) | 694 (93) | 690 (122) | 665 (95) | **683 (98)** |
| *Semantic blocking* | *-28 (85)* | *39 (59)* | *36 (53)* | *64 (68)* | ***46 (32)*** |
|  |  |  |  |  |  |
| **L2 group** |  |  |  |  |  |
|  |  |  |  |  |  |
| Homogeneous blocks | 810 (124) | 750 (106) | 740 (113) | 727 (112) | **738 (104)** |
| Heterogeneous blocks | 800 (113) | 711 (110) | 687 (98) | 694 (115) | **697 (103)** |
| *Semantic blocking* | *10 (75)* | *39 (41)* | *53 (39)* | *33 (76)* | ***41 (24)*** |
|  |  |  |  |  |  |
| **All participants** |  |  |  |  |  |
|  |  |  |  |  |  |
| Homogeneous blocks | 800 (112) | 742 (96) | 733 (117) | 728 (118) | **734 (104)** |
| Heterogeneous blocks | 809 (99) | 703 (101) | 689 (110) | 680 (105) | **690 (100)** |
| *Semantic blocking* | *-9 (82)* | *39 (50)* | *44 (47)* | *48 (73)* | ***44 (28)*** |

**Table 2.** RTs (SD) in the cyclical auditory word-picture matching task (Experiment 2).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Cycle 1 | Cycle 2 | Cycle 3 | Cycle 4 | **Total (Cycles 1 to 4)** |
| **L1 group** |  |  |  |  |  |
|  |  |  |  |  |  |
| Homogeneous blocks | 1,147 (122) | 1,142 (146) | 1,148 (157) | 1,190 (160) | **1,157 (129)** |
| Heterogeneous blocks | 990 (113) | 960 (95) | 985 (106) | 980 (105) | **979 (97)** |
| *Semantic blocking* | *157 (97)* | *182 (109)* | *163 (105)* | *210 (116)* | ***178 (64)*** |
|  |  |  |  |  |  |
| **L2 group** |  |  |  |  |  |
|  |  |  |  |  |  |
| Homogeneous blocks | 1,258 (198) | 1,235 (147) | 1,239 (119) | 1,258 (150) | **1,248 (132)** |
| Heterogeneous blocks | 1,146 (177) | 1,054 (124) | 1,032 (129) | 1,048 (151) | **1,070 (134)** |
| *Semantic blocking* | *112 (174)* | *181 (110)* | *207 (84)* | *210 (118)* | ***178 (87)*** |
|  |  |  |  |  |  |
| **All participants** |  |  |  |  |  |
|  |  |  |  |  |  |
| Homogeneous blocks | 1,201 (171) | 1,187 (152) | 1,192 (146) | 1,224 (157) | **1,201 (137)** |
| Heterogeneous blocks | 1,066 (166) | 1,006 (119) | 1,008 (119) | 1,014 (132) | **1,024 (124)** |
| *Semantic blocking* | *135 (141)* | *181 (108)* | *184 (96)* | *210 (116)* | ***178 (76)*** |

**Figure legends**

**Figure 1.** Example of a homogeneous and heterogeneous cycle in thecyclical picture naming task (Experiment 1).

**Figure 2.** Participants’ performances in the cyclical picture naming task (Experiment 1). **(A) – (C)** Naming latencies (in ms) in participants from the L1 group, participants from the L2 group, and all participants as a function of block type (homogeneous vs. heterogeneous) and cycle. Error bars represent standard errors (SE). Cousineau-Morey corrections were applied for within-subject designs (Cousineau, 2005). **(D)** Magnitude of the total semantic blocking effect (collapsing across cycles 2 to 4) shown by participants from the L1 group, participants from the L2 group and all collapsed participants. Error bars represent SE of the mean in the bars showing the effect in the L1 and L2 groups, but represent the SE with Cousineau-Morey corrections for within-subject design in the bar showing the effect in all the participants.

**Figure 3.** Scatter plot showing the relationship between conflictFT and semantic blocking in the picture naming task.For sake of visualization, we illustrate the greater influence of the conflictFT on RTs in homogeneous blocks (compared to RTs in heterogeneous blocks) with raw effects.

**Figure 4.** Example of a homogeneous and heterogeneous cycle in theauditory word-picture matching task.

**Figure 5.** Participants’ performances in the cyclical auditory word-picture matching task (Experiment 2). **(A) – (C)** Response latencies (in ms) in participants from the L1 group, participants from the L2 group, and all participants as a function of block type (homogeneous vs. heterogeneous) and cycle. Error bars represent SE with Cousineau-Morey corrections for within-subject designs (Cousineau, 2005). **(D)** Magnitude of the total semantic blocking effect (collapsing across cycles 1 to 4) shown by participants from the L1 group, participants from the L2 group, and all collapsed participants. Error bars represent SE of the mean in the bars showing the effect in the L1 and L2 groups, but represent the SE with Cousineau-Morey corrections for within-subject design in the bar showing the effect in all the participants.

**Figure A.1.** **(A)** Boxplots comparing the visual similarity ratings of pictures in the homogeneous and heterogeneous blocks. **(B)** Boxplots comparing the visual similarity ratings between the semantic categories.

**Footnotes**

1. For an alternative interpretation of Oppenheim et al.’s incremental learning that would lead to semantic interference in homogeneous blocks, see Nozari and Hepner (2018). [↑](#endnote-ref-1)
2. It is not strictly necessary to rescale continuous predictors in mixed linear models, but it is often useful to better understand the magnitude of the effect of the predictor on the dependent variable. All the results in this study (including those of experiments 1 and 2) remained the same when using the raw values of the conflictFT. [↑](#endnote-ref-2)
3. Semantic blocking is not characterized by a linear increase across blocks. Therefore, there is no reason to expect that language dominance may exert its effect only in some particular cycles. Nevertheless, we conducted complementary analyses with three separate t-tests, each comparing the magnitude of the semantic blocking effect in the L1 and L2 groups in each cycle (cycles 2, 3 and 4). The results showed no between-group differences in any of the cycles (*p* values > .14). [↑](#endnote-ref-3)
4. Due to the same rationale mentioned in footnote 1, we did not expect between-group differences in the magnitude of the semantic blocking effect across the cycles. We confirmed this lack of between-group differences with four separate t-tests, one for each cycle (*p* values > .16). [↑](#endnote-ref-4)