Re-assessing the role of language dominance in n-2 language repetition costs as a marker of inhibition in multilingual language switching

Iring Koch[1], Mathieu Declerck[2], Greta Petersen[1], Daniel Rister[1],

Wolfgang Scharke[1], & Andrea M. Philipp[1]

[1] Institute of Psychology, RWTH Aachen University, Germany

[2] Linguistics and Literary Studies, Vrije Universiteit Brussel, Belgium

10.1037/xlm0001333

Authors' note

Iring Koch, Greta Petersen, Daniel Rister, Wolfgang Scharke, and Andrea M. Philipp, Institute of Psychology, RWTH Aachen University, Aachen, Germany. Mathieu Declerck, Linguistics and Literary Studies, Vrije Universiteit Brussel, Brussel, Belgium.

Correspondence concerning this article should be addressed to Iring Koch, Institute of Psychology, RWTH Aachen University, Jägerstr. 17-19, D-52056 Aachen, Germany. E-mail: koch@psych.rwth-aachen.de.

Abstract

Speaking two or more languages shows bilingual flexibility, but flexible switching requires language control and often incurs performance costs. We examined inhibitory control assessing n-2 repetition costs when switching three languages (L1 [German], L2 [English], L3 [French]). These costs denote worse performance in n-2 repetitions (e.g., L2-L3-L2) than in n-2 non-repetitions (e.g., L1-L3-L2), indicating persisting inhibition. In two experiments (n = 28 in Experiment 1; n = 44 in Experiment 2), n-2 repetition costs were observed, but only for L2. Looking into L2 trials specifically, we found n-2 repetition costs when switching back to L2 from the still weaker L3 but not when returning from the stronger L1, suggesting that L2 is a strong competitor for L3 (requiring L2 inhibition) but less so for L1. Finding n-2 repetition costs supports an inhibitory account of language control in general, but our study shows only partial evidence for the theoretically assumed more specific relation between language dominance and language inhibition (i.e., only for dominance relations with respect to L1 and L3 when switching back to L2). Taken together, the findings thus suggest the need for further refinement of the concept of language dominance and its relation to inhibition.

Keywords: Bilingualism, cognitive flexibility, language switching, language inhibition, n-2 language repetition costs, language dominance

According to survey data of the European Union, about two thirds of the working-age adults (defined as 25-64 years old) knows at least one foreign language (Eurostat 2019; see https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Foreign_language_skills_statistics#Number_of_foreign_languages_known). At a global level, more than half of the world's population knows more than one language (e.g., Grosjean, 2010). However, bilingualism is not an all-or-none phenomenon but represents a complex and graded experience that is governed by a variety of factors, such as age of acquisition of a second language, frequency of use, frequency of code switching, etc. (for recent discussions see, e.g., de Bruin, 2019; Kałamała et al., 2022; Marian & Hayakawa, 2021; Titone & Tiv, 2022). Notably, the ability to communicate in more than one language also requires control processes to allow the speaker to flexibly switch from speaking one language to another depending on current language context and speaker intention (Green, 1998). During the last decades, the cognitive mechanisms underlying this flexible bilingual control have been examined using variants of the language switching paradigm (see Declerck & Koch, 2023, for a recent review).

In the language switching paradigm, participants are generally presented with language-unspecific stimuli, such as pictures or digits, in a naming task. In the most often used cuing-version of language switching, an explicit language cue, such as a national flag, is presented prior to or in parallel with the to-be-named stimulus. Naming performance in terms of reaction time (RT) and error percentages in language switch trials (i.e., trials that require a different language as the previous trial) is compared to that in repetition trials (i.e., trials that require the same language as the previous trial). Performance is typically worse in language switches than in repetitions, thus showing switch costs (e.g., Costa & Santesteban, 2004; Meuter & Allport, 1999; see Declerck & Philipp, 2015, for a review). These switch costs are assumed to be related to the underlying cognitive processes of language selection required in bilingual control (e.g., Green, 1998).

Bilingual control, language dominance, and inhibition

According to a major theoretical account proposed by Green (1998; Green & Abutalebi, 2013), language control is mainly based on inhibition of the currently competing but not intended language. In Green's (1998) inhibitory control model (ICM), there are different levels of control. The highest level refers to the language schema, which represents the instruction or intention to speak in a given language (i.e., it sets the target language). The language schemas compete with each other. At a lower level, the intention to speak a word, such as in a naming task, influences the corresponding lemma, and the ICM assumes that lemma selection is associated with inhibition of the translation-equivalent lemma in the competing language. Beyond this item-specific level, there is also an intermediate level of control based on so-called language tags, which are activated by the corresponding language schema and that serve to inhibit the language tags referring to the lemmas in the competing language. Based on this inhibitory control account, switch costs may represent, to some degree, both the time needed to inhibit the currently irrelevant language and the time to overcome residual inhibition of the current target language that has been inhibited in the previous trial in the context of the other language.

A number of findings are consistent with this inhibitory language control account (e.g., Jackson et al., 2001; Kleinman & Gollan, 2018; Kroll et al., 2008; Meuter & Allport, 1999; Verhoef et al., 2009; see Declerck & Koch, 2023, for a review and discussion). For the present purpose, two findings are most important. The first finding, already reported by Meuter and Allport (1999), refers to the difference in switch costs for the dominant, first language (L1) and the non-dominant, second language (L2), thus showing asymmetric language switch costs (for a review of this effect, see Bobb & Wodniecka, 2013). The second refers to n-2 language repetition costs in contexts that require switching between three languages (Philipp et al., 2007).

Meuter and Allport (1999) found that switch costs were larger for L1 than for L2 (see also Costa & Santesteban, 2004; Philipp et al., 2007, for similar findings). This asymmetry was explained with the ICM (Green, 1998), suggesting that the dominant L1 needs to be suppressed more strongly when activating L2 than vice versa, so that returning (i.e., switching back) to the dominant L1 suffers more strongly from persisting inhibition. However, it has been argued that alternative accounts are possible (e.g., Finkbeiner et al., 2006). Particularly, increased L2 activation and thus increased language competition when returning to L1 might be a viable account, too (Philipp et al., 2007). Moreover, a recent meta-analysis by Gade et al. (2021a, 2021b) based on 73 published studies has found many studies showing asymmetrical switch costs (see Bobb & Wodniecka, 2013; Declerck & Philipp, 2015, for reviews), but there were also many studies that did not find this asymmetry or even found a switch costs pattern in the opposite direction (i.e., larger switch costs for L2 than for L1). The overall lack of consistent evidence for asymmetric switch costs reported by Gade et al. (2021a) was recently confirmed in a re-analysis by Goldrick and Gollan (2023).

The theoretical key for explaining the asymmetric switch costs is the notion of language dominance, so that stronger inhibition should be targeted at the more dominant language. Hence, measures of language dominance should be a predictor for obtaining asymmetric switch costs, but the meta-analysis by Gade et al. (2021a) was not able, based on the published studies, to identify a clear moderator variable (e.g., differences in language dominance). However, finding no clear moderator variable might be due to the large diversity of language pairs used in the studies, which, combined with sometimes lacking or incompletely reported measures of bilingual experience (see de Bruin, 2019, for a discussion), could have resulted in the observed heterogeneity of findings with respect to asymmetric switch costs (see Declerck & Koch, 2023).

A different approach to language inhibition is taken when assessing so-called n-2 language repetition costs in situations that require switching between three languages. For example, Philipp et al. (2007) had eighteen participants switch between their L1 German and two other languages; for fifteen participants the L2 was English and for three it was French, and the L3 varied (English for L2 French speaker, and French, Italian, Spanish, Russian, or Croatian for the L2 English speakers). The critical contrast in this paradigm refers to the n-2 repetition (e.g., languages sequences like L2-L1-*L2*) vs. n-2 non-repetition (e.g., L3-L1-*L2*) across experimental trials. Note that there are no immediate (i.e., n-1) language repetitions in this paradigm. This approach thus focuses on performance in the current trial as a function of the language two trials ago, and this experimental contrast can be calculated separately for L1, L2, and L3. A number of studies found that n-2 language repetitions resulted in worse performance than non-repetitions, indicating n-2 language repetition costs (see Declerck & Koch, 2023, for a recent review). These costs can be explained by assuming that switching away from a language is associated with its inhibition, and this inhibition persists over time, so that there is relatively stronger persisting language inhibition in n-2 repetitions than in n-2 non-repetitions.

Critically, persisting overactivation of L2, as a potential alternative account for asymmetric switch costs (see above), would predict better performance in n-2 language repetition sequences (i.e., a n-2 repetition *benefit*) because of increased recency of previous activation. Finding n-2 repetition costs thus supports an account in terms of persisting inhibition. N-2 language repetition costs have been replicated since (e.g., Babcock & Vallesi, 2015; de Bruin et al., 2023; Declerck et al., 2015; Philipp & Koch, 2009), providing converging evidence for the involvement of inhibition in bilingual (or multilingual) language control.

According to the inhibitory control account, the general need to inhibit a competing language should further depend on its dominance, so that more dominant languages should receive stronger inhibition and should thus show larger n-2 repetition costs (cf. Green, 1998). In fact, Philipp et al. (2007) found larger n-2 repetition costs for L1 German compared to L2 and L3. However, the costs for L3 were even higher than those for L2, which is not predicted by the differences in language dominance. Yet, given the moderate sample size (n = 18) combined with heterogeneity of the language sets (with several different L3 languages), it is prudent not to overinterpret such empirical deviations from theoretical predictions.

However, other studies assessing n-2 language repetition costs also failed to find a clear pattern of language-specific difference across L1, L2, and L3 (see Table 1 [adapted from Table 2 in Declerck & Koch, 2023]). From eleven published experiments that reported n-2 language repetition costs as a function of language dominance, not a single study found the theoretically predicted pattern of larger n-2 repetition costs for the more dominant language (i.e., L1 > L2 > L3; note that De Bruin et al., 2023, focused only on L1 vs. L3), two studies clearly contradicted the prediction (finding smaller n-2 repetition costs for L1 than for L2 or L3), and five studies simply did not find differences in n-2 repetition costs across languages.

Hence, even though n-2 language repetition costs provide a clear empirical index of inhibitory control, the observed pattern is not in line with the additional idea that inhibition should be a function of differences in language dominance. Yet, studies often had rather moderate sample size, so that the lack of predicted interaction effects may not be conclusive. The goal of the present study was to provide more evidence on the role of language dominance in language inhibition.

Table 1

Overview of n-2 Language Repetition Costs Studies that Examined the Effect of Language on n-2 Language Repetition Costs. Adapted from Declerck and Koch (2023)

		N-2 Language Repetition Costs		
Study	Type of Stimulus Material	L1 vs. L2	L1 vs. L3	L2 vs. L3
Babcock & Vallesi (2015)	Digits	L1 < L2	L1 < L3	L2 = L3
De Bruin et al. (2023) – Experiment 2	Pictures	n.a.	L1 > L3	n.a.
Declerck, Thoma, Koch, & Philipp (2015)	Digits	L1 = L2	L1 > L3	L2 > L3
Declerck & Philipp (2018) – Picture Naming	Pictures	L1 = L2	L1 = L3	L2 = L3
Declerck & Philipp (2018) – Reading Aloud	Written words	L1 = L2	L1 = L3	L2 = L3
Guo, Liu, Chen, & Li (2013) – Experiment 1	Digits	L1 = L2	L1 > L3	L2 = L3
Guo, Liu, Chen, & Li (2013) – Experiment 2	Digits	L1 = L2	L1 = L3	L2 > L3
Guo, Ma, & Liu (2013)	Digits	L1 = L2	L1 = L3	L2 = L3
Philipp, Gade, & Koch (2007)	Digits	L1 > L2	L1 > L3	L2 < L3
Philipp & Koch (2009) – Experiment 1	Digits	L1 = L2	L1 = L3	L2 = L3
Philipp & Koch (2009) – Experiment 2	Digits & Colors	L1 = L2	L1 = L3	L2 = L3

Note. The larger, smaller, and equal, signs relate to the relative size of n-2 language repetition costs (in RT) between the two indicated languages. The “n.a.” means that the design precluded some specific contrasts. Not all n-2 language repetition costs studies were added to this table, as some of these studies did not report n-2 language repetition costs separately for each language.

The goal of the present study

The present study used the n-2 language repetition paradigm to re-examine the relation between language dominance and language inhibition. Two experiments tested a fairly homogenous sample of participants, with L1 being German, L2 being English, and L3 being

French for all participants. Compared to most previous studies, we had larger sample sizes and thus increased statistical power to detect language-specific differences in n-2 repetition costs. Finally, the specific findings of Experiment 1 were conceptually replicated independently with a larger sample size in Experiment 2, thus further reducing the risk of false positives or of missing a relevant effect. In both experiments, we examined whether n-2 language repetition costs are larger for L1 than for L2, which in turn should be larger than for L3. The possible language sequences are shown in Table 2.

Table 2

Overview of Possible Language Sequences Producing n-2 Language Repetitions vs. Non-Repetitions in the Present Experiments

N-2 Language Sequence		
N-2 Repetition	N-2 Non-Repetition	Expected Empirical Finding
L1 – L2 – L1	L3 – L2 – L1	N-2 Language Repetition Costs for L1
L1 – L3 – L1	L2 – L3 – L1	
L2 – L1 – L2	L3 – L1 – L2	N-2 Language Repetition Costs for L2
L2 – L3 – L2	L1 – L3 – L2	
L3 – L1 – L3	L2 – L1 – L3	N-2 Language Repetition Costs for L3
L3 – L2 – L3	L1 – L2 – L3	

Note. L1 = German, L2 = English, L3 = French

There are two possible n-2 non-repetitions as a control condition for two different n-2 language repetition sequences. For example, for L2 (see the middle two lines in the table), the intervening language in trial n-1 could be either L1 or L3. Hence, the corresponding control conditions would have either L3 or L1 as language in trial n-2, respectively. Note also that in these comparisons, the language switch from trial n-1 to trial n is strictly comparable in n-2 language repetitions and non-repetitions, so that the only difference is in terms of whether the language in trial n-2 repeats in trial n and hence in the recency of returning to a language from which one has switched away earlier. That is, the rationale of assessing n-2 repetition costs is that inhibition triggered when switching from one language to another gets weaker over time, either because of time-based decay of inhibition or of trial-based dissipation (see Declerck & Koch, 2023, for a discussion).

This experimental design allowed us also a more detailed analysis of the relation between language dominance and language inhibition. That is, the two conditions for L1 both entail switching back to L1 from a weaker language (i.e., L2 or L3), and, similarly, for L3 both entail switching back to L3 from a stronger language (i.e., L1 or L2). As can be seen in Table 2, it is specifically for L2 English that we can examine whether it is harder to switch back from a *stronger* language (from L1 back to L2) or from a *weaker* language (from L3 back to L2). Based on the inhibitory account, we would expect that n-2 repetition costs for L2 should be larger when switching back from the weaker L3 because switching from L2 to L3 should require substantial inhibition of the relatively more dominant L2, thus producing strong L2 n-2 repetition costs. In comparison, switching from L2 to L1 should produce less inhibition of L2 because L1 is more dominant in the first place, thus requiring less inhibition of L2 and thus also resulting in smaller L2 n-2 repetition costs. Hence, focusing on L2 English inhibition provides us with a unique opportunity to examine additional predictions derived from the account that inhibition is proportional to differences in language dominance.

A recent study by de Bruin et al. (2023) used a complementary approach. In two experiments, they examined language intrusions with a setup relying on three languages and found that the number of intrusions when performing in L2 was higher for L3 than for L1 (see Tomoschuk et al., 2021, for related evidence). This finding suggests that accessibility of L1 translation equivalents was lower than that of L3 equivalents (de Bruin et al., 2023, p. 8) and that

L1 was inhibited more strongly than L3 during L2 performance. Moreover, in their Experiment 2, de Bruin et al. (2023) used a variant of the n-2 repetition costs paradigm testing English-French-Spanish trilinguals. Specifically, the authors designed their trial sequences in a way that focused on conditions in which L2 (French) is performed in trial n-1 to see whether inhibition is stronger for L1 (English, i.e., L1 – L2 – L1 vs. L3 – L2 – L1) than for L3 (Spanish, i.e., L3 – L2 – L3 vs. L1 – L2 – L3), as suggested already by their findings of reduced L1 intrusions. They found somewhat larger n-2 repetition costs for L1 than for L3 (24 ms vs. 7 ms; $p = .049$ for the interaction contrast), and the 7 ms costs for L3 were not significant. However, due to their design, specifically focusing on two conditions out of the six conditions described in Table 2 above (i.e., line 1 and line 6), these authors did not report n-2 repetition costs for L2. In the present study, we used all six comparisons listed in Table 2 and first report overall n-2 repetition costs for each of the three languages. In the next step, we focused on the two different conditions for L2 n-2 repetition costs (i.e., lines 3 and 4 in Table 2) to see if it is harder to switch back to L2 from a more dominant language (L1) or from a less dominant language (L3). This is a complementary approach to that taken by de Bruin et al. (2023). Moreover, we examined these specific effects in Experiment 1 and then replicated the main findings in Experiment 2, thus adding further credibility to the robustness of the findings.

Experiment 1

Based on the inhibitory control model, we expected larger n-2 language repetition costs for L1 than for L2, which in turn might be larger than that for L3. We also examined the more specific question, which is unique to a design with three languages, whether performance in L2 would be differentially affected by whether the previous language required the more dominant L1 or the less dominant L3.

In addition, we manipulated the response-cue interval (RCI), which varied randomly between 600 ms and 1600 ms. In non-linguistic task switching (see Kiesel et al., 2010; Koch et al., 2018; Koch & Kiesel, 2022, for reviews), it has been found that task-switch costs decrease with longer RCIs, suggesting a process of dissipation of task-set activation, which should facilitate switching tasks but also reduce the benefit of task repetitions (see Horoufchin et al., 2011). In language switching, Ma et al. (2016; see also de Bruin & Xu, 2023) showed that language switch costs decreased with increasing RCI particularly for L1. The present experiment assessed n-2 language repetition costs using language switch trials only (i.e., there were no immediate [n-1] language repetitions, so that we did not assess traditional "switch costs"), but we could derive from Ma et al.'s (2016) findings that switching should be generally easier with longer RCI. Such an RCI effect could have two different causes. First, with a short RCI, the previous language task schema is still highly activated, producing more competition in language switches. Second, the upcoming task might still be inhibited, which likewise increases competition. With longer RCI, both influences should get weaker, thus leading to overall better performance. Based on this reasoning, we also examined whether n-2 repetition costs would get smaller with longer RCI.

Methods

Participants. Thirty participants were tested in Experiment 1, but data from two participants were lost due to voice-key malfunction, which resulted in more than 40% trials without response registration. The remaining twenty-eight participants (24 female, 4 male; mean age = 21.75 years, SD = 4.11 years, age range 18-35) were mostly psychology students of RWTH Aachen University (with a curriculum generally in German), who received partial course credit.

The participants were German native speakers. English is taught in school for all pupils in Germany, and English reading and comprehension is often required in psychology classes (despite the curriculum being overall in German). French is a common second foreign language in Germany and (if chosen from other alternatives like Latin or Spanish) is typically learned for at least four years in school. Their proficiency level was assessed with subjective ratings on a 7-point scale for L2 English and L3 French with respect to hearing and reading comprehension and speaking and writing ability. Moreover, as an objective proficiency measure for lexical knowledge, we included the LexTale for L1 German and L2 English (Lemhoefer & Broersma, 2012) as well as L3 French (Brysbaert, 2013). Note that it may be problematic to compare LexTale scores across languages, because these tasks were not developed to compare language proficiency across languages, but the LexTale differences aligned well with the differences in the subjective self-ratings in our sample. Together, the descriptive data clearly suggest that German is the L1, English the L2, and French the L3 in our sample (see Table 3).

Table 3

Description of Language Proficiency across German, English, and French (Means [SD]) in Experiment 1 (N = 28).

Proficiency Measure	L1 (German)	L2 (English)	L3 (French)
Lextale	87.3 (8.6) range = 63-98	73.3 (11.5) range = 49-96	47.4 (6.6) range = 35-67
Subjective self-rating (7-point scale)			
Writing	--	5.14 (0.88)	3.25 (1.48)
Speaking	--	4.78 (0.92)	2.96 (1.23)
Hearing	--	5.19 (0.98)	2.96 (1.14)
Reading	--	5.54 (0.82)	3.46 (1.38)

To determine the sample size of the current study, we performed a power analysis with Gpower (Faul et al., 2007). Considering the ability to detect pairwise differences between n-2 language repetition costs (employing the logic of a paired *t* test), the results showed that 28 participants would be needed to detect medium effect sizes (d_z between .5. and .6) with a power of .80. A sample size of 28 exceeds the sample size of most previous studies that focused on n-2 language repetition costs (e.g., Declerck et al., 2015; Philipp et al., 2007).

Stimuli and Task. The experiment was programmed in PsychoPy3 and conducted in November and December 2021 in the lab. Participants were seated inside a sound-insulated cabin. The experimenter sat outside the cabin and could hear the participants via headphones. The stimuli were the digits 1-9, which were presented individually in white on a black background (approx. 6 mm high) at the center of a 17-inch screen. The German, British, and French national flags (approx. 25 mm x 38 mm) served as language cues. These cues were surrounded by a white margin to be clearly demarcated from the black background. In each trial, four identical flags were presented 8 cm to the left and right and 5.3 cm above and below the target digits, measured from center to center. The target digit was thus enclosed by the cues. The onset of the vocal naming responses was recorded by a voice key but was also recorded and offline checked for accuracy.

Procedure. All procedures were in line with the Helsinki 2013 declaration. Participants first gave their informed consent and signed a data protection document, then they first filled in the subjective rating scales regarding their language proficiency in English and French followed by the LexTale tasks (with counterbalanced order for the three languages). Then, for the main experiment, the participants were instructed that their task would be to name individually presented single Arabic digits from 1-9 in German, English, or French as indicated by the national flags. They were instructed to respond as fast and correctly as possible.

In an individual trial, the language cue was first presented for 100 ms (i.e., cue-stimulus interval) followed by the target digit, which both remained on the screen until the participants responded by vocally naming it. A language could never repeat from one trial to the next. The subsequent response-cue interval (RCI) varied randomly (600 ms vs. 1,600 ms), and each RCI level occurred equally often in each experimental block.

The experiment started with a practice block of 27 trials, in which each of the nine digits was named once in each of the three languages, and performance was not recorded in these practice trials. Then, five blocks with 108 trials each followed, in which each digit was named four times in each language, n-2 languages repetitions and non-repetitions were about equally frequent, and the RCI levels were about equally frequent. Hence, in theory (i.e., disregarding the fact that the first two trials of a block cannot be coded as a function of n-2 language condition) there are about 90 trials overall per participant for assessing the n-2 repetition costs for each of the three languages, and still about 45 when considering potential interactions with RCI. Participants had the opportunity for a short break between experimental blocks. An immediate repetition of a language (i.e., from trial n-1 to trial n) and of a target digit was not possible. It was also controlled that n-2 language repetitions could not co-occur with n-2 digit repetitions.

Design. The dependent variables were RT and error percentages. For the first analysis, the independent variables were language (English, French, German), n-2 language repetition (n-2 repetition vs. n-2 non-repetition), and RCI (600 ms vs. 1,600 ms). For the second analysis, only data from L2 English trials were taken and analyzed with the independent variables n-1 language (L1 German vs. L3 French), n-2 language repetition, and RCI.

Transparency and openness. We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study. The data will be made available [after manuscript acceptance] at https://www.psycharchives.org/. Data were analyzed using SPSS. This study's design and its analysis were not pre-registered.

Results

For data analyses, we discarded the first two trials in each block because for those trials it is not possible to define n-2 language repetition status. Then we deleted trials with vocal RT < 100 ms and "non-response" trials because those trials most likely reflect voice-key malfunction. This resulted in a loss of 4.6% of the trials. Next, we discarded trials that were preceded by an error in trial n-2 or trial n-1 (6.1% of the trials) because the classification of n-2 language repetitions is dubious when the sequence included incorrect responses. We then excluded RT outliers (1.9% of the remaining trials), which we defined as RT larger than 3 SD or smaller than -3 SD of the individual (i.e., participant-based) mean RT. For the remaining trials, we analyzed error percentages as a function of experimental condition. Finally, for RT analysis, we also excluded trials with an error in the current trial (i.e., trial n). Table 4 shows mean RT and error percentage as a function of n-2 language repetition, RCI, and language.

We submitted the RT data to a repeated-measures analysis of variance (ANOVA). The ANOVA yielded a significant effect of RCI, $F(1,27) = 19.877$, $p < .001$, $\eta_p^2 = .424$, showing longer RT for the short RCI compared to long RCI (1138 ms vs. 1110 ms), and of language, $F(2,54) = 31.323$, $p < .001$, $\eta_p^2 = .537$, showing that RT was shortest for German (1026 ms), followed by French (1170 ms) and English (1178 ms). The interaction of RCI and language was non-significant, $F(2,54) = 0.528$, $p > .59$, $\eta_p^2 = .019$, showing that the influence of RCI was similar for all three languages.

Table 4

Mean RT (ms [SE]) and Mean Error Percentage (SE) as a Function of Language, N-2 Language Repetition, and RCI in Experiment 1

	N-2 Language Repetition		
Language and RCI	N-2 Repetition	N-2 Non-Repetition	N-2 Repetition Costs
RT (ms [SE])			
L1 (German)			
RCI = 600 ms	1043 (34)	1040 (31)	3
RCI = 1600 ms	1008 (29)	1012 (31)	-4
L2 (English)			
RCI = 600 ms	1210 (34)	1178 (35)	32
RCI = 1600 ms	1175 (33)	1148 (34)	27
L3 (French)			
RCI = 600 ms	1182 (29)	1179 (30)	3
RCI = 1600 ms	1157 (30)	1161 (32)	-4
Error (% [SE])			
L1 (German)			
RCI = 600 ms	2.9 (0.6)	3.7 (0.8)	-0.8
RCI = 1600 ms	3.1 (0.7)	2.5 (0.7)	-0.6
L2 (English)			
RCI = 600 ms	3.7 (0.7)	4.2 (0.5)	-0.5
RCI = 1600 ms	3.3 (0.7)	2.4 (0.6)	0.9
L3 (French)			
RCI = 600 ms	3.6 (0.7)	2.6 (0.6)	1.0
RCI = 1600 ms	2.5 (0.6)	3.0 (0.6)	-0.5

RT for n-2 language repetitions were 9 ms longer than for n-2 language switches (1129 ms vs. 1120 ms), but the main effect of n-2 language repetition was not significant, $F(1,27) = 2.304$, $p < .141$, $\eta_p^2 = .079$. N-2 language repetition also did not interact with RCI, $F(1,27) = 0.520$, $p = .477$, $\eta_p^2 = .019$, suggesting that the influence of RCI was similar for n-2 language repetitions and non-repetitions. Importantly, n-2 language repetition interacted with language, *F*(2,54) = 4.870, *p* = .011, η_p^2 = .153, showing 30 ms language repetition cost for English, *t*(27) = 3.258, *p* = .003, d_z = 0.616, but virtually no n-2 language repetition costs for French and German (both -1 ms). The three-way interaction was non-significant, *F*(2,54) = 0.001, *p* = .999, η_p^2 < .001, suggesting, again, that the differential n-2 repetition costs for the three languages, with n-2 repetition costs only for English, was the same regardless of whether the RCI was short or long.

The ANOVA on error percentages revealed a significant main effect of RCI, *F*(1,27) = 5.690, *p* < .024, η_p^2 = .174, showing higher error percentages with short RCI than with long RCI (3.4% vs. 2.8%), consistent with the RT data. All other effects were non-significant, all *F*s < 1, all *p*s > .49, all η_p^2 < .018.

Next, we report the analyses using only L2 English trials to examine whether n-2 language repetition costs are affected by whether participants switched back from the more dominant L1 German (i.e., L2-*L1*-L2) or from the less dominant L3 French (i.e., L2-*L3*-L2). In this ANOVA, we included RCI and n-2 language repetition using only L2 trials, and we replaced language in the current trial (German, English, French) now with language in the preceding trial n-1 (L1 German vs. L3 French) to focus on the influence of language dominance on n-2 repetition costs in L2 specifically (see Table 5).

The ANOVA on RTs yielded a significant effect of n-2 language repetition, *F*(1,27) = 11.692, *p* < .01, η_p^2 = .302. The main effect of the n-1 language (L1 German vs. L3 French) was also significant, *F*(1,27) = 7.893, *p* < .01, η_p^2 = .226, showing 33 ms longer RTs when switching back and forth from L1 German (i.e., *L1*-L2 sequences) compared to trials preceded by L3 French (1193 ms vs. 1160 ms), suggesting that L1 indeed produced more competition than L3 when returning to L2 English.

Table 5

Mean RT (ms [SE]) and Mean Error Percentage (SE) for L2 (English) Naming Trials (ms [SE]) as a Function of Previous-Trial (N-1) Language (L1 vs. L3), N-2 Language Repetition, and RCI in Experiment 1

	N-2 Language Repetition		
N-1 Language and RCI	N-2 Repetition	N-2 Non-Repetition	N-2 Repetition Costs
RT (ms [SE])			
N-1 L1 (German)			
RCI = 600 ms	1226 (34)	1207 (39)	19
RCI = 1600 ms	1159 (34)	1181 (36)	-22
N-1 L3 (French)			
RCI = 600 ms	1192 (36)	1139 (33)	53
RCI = 1600 ms	1189 (35)	1119 (33)	70
Error (% [SE])			
N-1 L1 (German)			
RCI = 600 ms	2.8 (0.8)	4.8 (0.8)	-2.0
RCI = 1600 ms	3.2 (0.8)	2.7 (0.7)	0.5
N-1 L3 (French)			
RCI = 600 ms	4.7 (1.1)	3.6 (0.9)	1.1
RCI = 1600 ms	3.4 (0.8)	2.0 (0.7)	1.4

Importantly, the interaction of n-2 language repetition and n-1 language was significant, too, $F(1,27) = 9.202$, $p < .01$, $\eta_p^2 = .254$, showing larger n-2 language repetition costs for L3 French as preceding language (62 ms, $t(27) = 5.653$, $p < .001$, $d_z = 1.068$) than for L1 German as preceding language, for which there were virtually no overall n-2 repetition costs (-1 ms). This suggests that the interaction of n-2 language repetition and language in the main analysis was driven only by English naming trials that comprised an alternation with the even weaker L3 French, whereas switching back and forth with L1 German did not entail any extra costs based on lingering inhibition of L2 when having to name a picture in L1 in-between.

There was also a main effect of RCI, $F(1,27) = 8.197$, $p < .01$, $\eta_p^2 = .233$, showing 29 ms longer RT for the short RCI compared to long RCI (1191 ms vs. 1162 ms). The interaction of RCI and n-2 language repetition costs was not significant, $F(1,27) = 0.825$, $p > .37$, $\eta_p^2 = .030$. Likewise, the interaction of RCI with n-1 language was not significant, $F(1,27) = 3.696$, $p = .065$, $\eta_p^2 = .120$, just like the three-way interaction, $F(1,27) = 2.919$, $p = .099$, $\eta_p^2 = .098$. Note that there was an unexpected pattern for short vs. long RCI when the previous language was German, but this pattern was numerically reversed in the error percentages (see below).

The ANOVA on error percentages revealed no significant n-2 language repetitions costs ($F < 1$) and no significant effect of n-1 language ($F < 1$). Their interaction was not significant either, $F(1,27) = 3.798$, $p = .062$, $\eta_p^2 = .123$, but it reflected a non-significant trend consistent with RT, with small n-2 language repetition costs (1.3%) for L3 French as preceding language but an even smaller n-2 language repetition benefit (0.7%) for L1 German in trial n-1.

Again consistent with the RT data, there was a significant main effect of RCI, $F(1,27) = 7.345$, $p = .012$, $\eta_p^2 = .214$, showing more errors for short RCI than for long RCI (4.0% vs. 2.8%). The interaction of RCI and n-2 language repetition was not significant, $F(1,27) = 1.729$, $p = .200$, $\eta_p^2 = .060$, just like the interaction of RCI and n-1 language ($F < 1$) and the three-way interaction, $F(1,27) = 1.042$, $p > .31$, $\eta_p^2 = .037$.

Discussion

Experiment 1 set out to examine n-2 language repetition costs as a function of language dominance and of RCI. In the analysis with all three languages, we found only small and non-significant n-2 repetition costs of 9 ms, but there was an interaction with language, indicating n-2 repetition costs only for L2 English trials. That is, while n-2 repetition costs were larger for L2 than for L3, they were also larger for L2 than for L1, for which we observed no significant n-2 repetition costs at all. So, the absence of n-2 repetition costs for L1 German was an unexpected finding.

In our second analysis, we focused on the novel contrast of n-2 language repetition costs for L2 English as a function of whether the preceding language, and thus the most recent competitor, was the stronger L1 German or the weaker L3 French. Here we found that n-2 repetition costs for L2 English were quite large when the intervening language was L3 (i.e., L2-*L3*-L2), whereas there were virtually no n-2 repetition costs when participants switched from L1 back to L2 (i.e., L2-*L1*-L2). This pattern is actually in line with the prediction from an inhibitory framework: When naming in L3, competition from the stronger L2 needs to be resolved, resulting in n-2 repetition costs based on the persistence of L2 inhibition triggered in the L3 naming trials. In contrast, when naming in L1, which is most dominant, it may be less critical to inhibit L2 in order to perform in L1, which in turn did not result in n-2 language repetition costs.

Notably, we found a consistent effect of RCI in terms of generally improved performance with longer RCI. This general influence can be explained by time-based dissipation of activation of the previous language schema. However, this general RCI effect was the same for all three languages, and, more importantly, it did not influence the n-2 language repetition costs. Note that a somewhat similar RCI manipulation (though with three RCI levels) in a study by Ma et al. (2016) had an effect on switch costs and mixing costs, but this study used only two languages and thus did not assess n-2 language repetition costs when using three languages. Hence, our results complement their finding, showing that short RCI generally seems to increase competition in naming in different languages, but there was no indication that n-2 repetition costs decreased with long RCI, suggesting that the type of inhibition that is measured with n-2 repetition costs does not dissipate quickly, at least not within the range of RCIs employed in the present study.

In Experiment 2, we aimed at replicating our finding that n-2 language repetition costs occurred only for L2 English. Given that RCI did not interact with language or n-2 repetition costs in Experiment 1, we decided to omit this variation and used a constant medium RCI (1100 ms) instead. Moreover, given that we aimed at replicating an unexpected finding, we increased sample size (from n = 28 to n = 44) to achieve more power for this conceptual replication, which was an exact replication but without the RCI variation and a constant RCI instead.

Experiment 2

Method

***Participants*.** In Experiment 1, in the RT data we obtained interaction effects that were by convention large effects. For the interaction of n-2 language repetition and language we found a $\eta_p^2 > .15$. For the second analysis (English trials only), we found $\eta_p^2 > .25$ for the interaction of n-2 language repetition and previous (n-1) language (i.e., L1 vs. L3). If we conceptualize this two-way interaction as the difference of the two n-2 language repetition costs and use the paired t-test procedure, Gpower (Faul et al., 2007) suggested that 44 participants should give a power exceeding .90 for replicating this effect even for medium effect sizes. Hence, 44 native German speakers (30 female, 14 male; mean age = 23.07) participated in the study. They were mostly psychology students of RWTH Aachen University, who received partial course credit. Their language profile with respect to L1 German, L2 English, and L3 French was similar to that in Experiment 1. According to LexTale scores and self-ratings, L2 English was clearly stronger than L3 French (see Table 6).

Table 6

Description of Language Proficiency across German, English, and French (Means [SD]) in Experiment 2 (N = 44).

Proficiency Measure	L1 (German)	L2 (English)	L3 (French)
LexTale	88.4 (5.7) range = 71-100	78.4 (9.6) range = 63-98	48.2 (8.4) range = 34-79
Subjective self-rating (7-point scale)			
Writing	6.64 (0.71)	5.16 (1.06)	2.41 (1.48)
Speaking	6.91 (0.29)	5.20 (0.94)	2.43 (1.57)
Hearing	6.89 (0.32)	5.68 (0.97	2.61 (1.63)
Reading	6.89 (0.32)	5.89 (1.00)	3.14 (1.56)

***Stimuli, Task, and Procedure*.** The stimuli and task were identical to Experiment 1. The only difference referred to the RCI, which varied randomly in Experiment 1 but was kept constant at 1100 ms in Experiment 2, giving again about 90 trials for assessing overall the n-2 repetition costs for each of the three languages per participant.

***Design*.** For the first analysis, the independent variables were language (English, French, German) and n-2 language repetition (n-2 repetition vs. n-2 non-repetition). For the second analysis, only data from L2 English trials were used and analyzed with the independent variables n-1 language (L1 German vs. L3 French) and n-2 language repetition. The dependent variables were RT and error percentages.

Results

The data were analyzed as in Experiment 1. We discarded trials with apparent voice-key malfunctions (1.0% of the trials) and trials preceded by an error in trial n-2 or n-1 (9.1% in total). For the analyses of RTs (see Table 7), we focused on the remaining correct trials and excluded 1.6% of trials as RT outliers.

Table 7

Mean RT (ms [SE]) and Mean Error Percentage (SE) as a Function of Language and N-2 Language Repetition in Experiment 2

	N-2 Language Repetition		
Language	N-2 Repetition	N-2 Non-Repetition	N-2 Repetition Costs
RT (ms [SE])			
L1 (German)	943 (22)	941 (23)	2
L2 (English)	1067 (28)	1043 (30)	24
L3 (French)	1111 (33)	1113 (35)	-2
Error (% [SE])			
L1 (German)	3.2 (0.5)	4.2 (0.5)	-1.0
L2 (English)	4.3 (0.5)	4.2 (0.6)	0.1
L3 (French)	3.5 (0.5)	3.6 (0.5)	-0.1

The ANOVA on RT revealed an effect of language, $F(2, 86) = 44.842$, $p < .001$, $\eta_p^2 = .510$, indicating that L1 German responses were much faster than responses in L2 English or L3 French (942 ms vs. 1055 ms vs. 1112 ms). The effect of n-2 repetition was just not significant, $F(1, 43) = 3.578$, $p = .065$, $\eta_p^2 = .077$. Yet, the interaction of n-2 repetition with language was significant, $F(2, 86) = 5.335$, $p = .007$, $\eta_p^2 = .11$, indicating n-2 language repetition costs only for L2 English trials (24 ms), thus replicating and confirming the results of Experiment 1.

An ANOVA on the error percentages revealed no effect of language, $F(2, 86) = 1.569$, $p = .214$, $\eta_p^2 = .035$, of n-2 repetition, $F(1, 43) = 1.600$, $p = .213$, $\eta_p^2 = .036$, and also no significant interaction, $F(2, 86) = 1.788$, $p = .174$, $\eta_p^2 = .040$.

In the next step, we again focused on L2 English trials and examined whether returning to L2 English was harder when switching back from the even weaker L3 French than from the stronger L1 German (see Table 8).

Table 8

Mean RT (ms [SE]) and Mean Error Percentages (SE) for L2 (English) Naming Trials (ms [SE]) as a Function of Previous-Trial (N-1) Language (L1 vs. L3) and N-2 Language Repetition in Experiment 2

	N-2 Language Repetition		
N-1 Language	N-2 Repetition	N-2 Non-Repetition	N-2 Repetition Costs
RT (ms [SE])			
N-1 L1 (German)	1059 (27)	1054 (31)	5
N-1 L3 (French)	1075 (30)	1031 (30)	44
Error (% [SE])			
N-1 L1 (German)	4.8 (0.6)	4.8 (0.7)	0.0
N-1 L3 (French)	3.8 (0.5)	3.5 (0.7)	0.3

The ANOVA on RT revealed a significant effect of n-2 language repetition, $F(1, 43) = 16.939$, $p < .001$, $\eta_p^2 = .283$, but not of n-1 language, $F(1, 43) = 0.154$, $p = .697$, $\eta_p^2 = .004$. Critically, the interaction was significant, $F(1, 43) = 12.742$, $p = .001$, $\eta_p^2 = .229$. The 44 ms n-2 language repetition costs when returning from L3 French were clearly significant, $t(43) = 6.124$, $p < .001$, $d_z = 0.923$, whereas the 5 ms effect when returning from L1 German was not significant, $t(43) = 0.586$, $p = .561$, $d_z = 0.088$.

The ANOVA on the error percentages did not show a significant effect of n-2 language repetition, $F(1, 43) = 0.146$, $p = .704$, $\eta_p^2 = .003$. The effect of n-1 language was significant, $F(1, 43) = 6.523$, $p = .014$, $\eta_p^2 = .132$, showing overall 1.1% more errors when returning from L1 German to L2 than when returning from L3 French to L2, but the interaction was non-significant, $F(1, 43) = 0.118$, $p = .732$, $\eta_p^2 = .003$.

Discussion

Experiment 2 replicated and confirmed the results of Experiment 1. We omitted the RCI manipulation to simplify the design and double the number of trials in each experimental condition, and we increased sample size (from n = 28 to n = 44). Again, we found n-2 repetition costs only for L2 English. Similar to Experiment 1, we also found that for English trials, there was lingering inhibition produced by switching from L2 to L3 French and back to L2, whereas no such effect was found when switching from L2 to L1 German and back to L2. This finding suggests that inhibition was mainly implemented when switching back from L3 to L2, which we found across both experiments.[1]

General Discussion

The present study examined the role of language inhibition in speech production in three different languages. To this end, we employed a cued language-switching paradigm, in which participants named digits in their L1 German, in their L2 English, or in their L3 French, and each trial required a language switch. With this paradigm, we assessed n-2 language repetition costs, which represent a marker of lingering inhibition caused by inhibition of the just performed language when switching to a different language (e.g., L2-L3-L2), so that switching back to the just inhibited language should impair performance relative to n-2 non-repetitions. In Experiment 1, the RCI was manipulated to examine whether activation or inhibition of languages dissipates

[1] For the sake of completeness, we also analyzed the L1 German and L3 French trials in a similar way in both experiments. Because these analyses are theoretically less conclusive than that for L2 English, we report them in the appendix. For Experiment 1 we aggregated the data across RCI to yield the same 2x2 data format as in Experiment 2. In the appendix we also report the ANOVA results, essentially confirming that there was overall little inhibition of L1 German and L3 French.

over time. Experiment 2 was aimed at replicating the main findings of Experiment 1 in conditions with constant RCI.

Experiment 1 showed that performance was generally better with long RCI than with short RCI, independent of the effects of language and of language sequence. Furthermore, in both Experiment 1 and 2, overall performance was generally better in L1 German than in the other two languages. However, clear n-2 language repetition costs were only found for L2 English. More specifically, the L2 n-2 repetition costs were only observed when participants switched back from L3 French, whereas switching back to L2 from L1 German did not produce significant n-2 language repetition costs in L2. This pattern was replicated across two experiments.

In Experiment 1, the general RCI effect suggests that activation of the previously activated language schema dissipates over time, so that with long dissipation time (i.e., long RCI) there is less competition when activating the new language schema in a switch trial. Such an account is consistent with models of non-linguistic task switching, in which dissipation of task-set activation is widely assumed (see Allport et al., 1994; Altmann & Gray, 2008; Horoufchin et al., 2011; see also Kiesel et al., 2010; Koch & Kiesel, 2022; Koch et al., 2018, for general reviews). By and large, the RCI-related findings in Experiment 1 are also in line with findings reported by Ma et al. (2016) on the influence of RCI manipulations on switch costs and mixing costs when switching between two languages. They found larger costs with shorter RCI. However, the effect of RCI, which we attribute to dissipation of language schema activation, was similar for all three languages and for both n-2 language repetitions and non-repetitions and thus represents a general effect, suggesting that the functional locus of the RCI effect differs from that of the other variables, to which we turn now.

The Relation of Language Inhibition and Language Dominance

The present two experiments used a methodology, measuring n-2 language repetition costs, that is designed to assess the after-effect of inhibition triggered when switching from one language to a different language (see Philipp et al., 2007). It has been debated to which degree some major empirical markers of language competition, such as language switch costs or mixing costs, can also be accounted for by non-inhibitory accounts, such as by persisting activation of the previous target language (see, e.g., Koch et al., 2010, for a review of inhibition in task switching also referring to language switching). In comparison to those other empirical markers, n-2 repetition costs are less open to non-inhibitory alternative accounts because persisting activation would rather predict a n-2 repetition benefit based on lingering activation of the switched-away language. Yet, the typical finding of n-2 repetition costs, which was observed in both of the experiments of the current study, clearly speaks in favor of an inhibitory after-effect (see Declerck & Koch, 2023). The existence of n-2 language repetition costs thus clearly supports theoretical accounts that propose the contribution of inhibitory control in bilingual language production (e.g., Declerck et al., 2015; Green, 1998; Kleinman & Gollan, 2018; Meuter & Allport, 1999; Philipp et al., 2007). So, the present study further confirms accounts of inhibitory language control in general.

However, the ICM of Green (1998) also makes a more specific prediction, which is that the more dominant language is the stronger competitor and thus requires a stronger inhibitory control input to enable the weaker language as a target language. Based on this account, we expected stronger inhibition for L1 German than for L2 English, which in turn would show stronger inhibition than L3 French. In both Experiment 1 and 2 we only found n-2 language repetition costs for L2, thus seemingly violating the more specific prediction. Note that previous n-2 language repetition costs studies, as summarized in Table 2 in the introduction section, are overwhelmingly inconsistent with the predicted pattern of a clear dominance relation with respect to the observed n-2 repetition costs (i.e., L1 > L2 > L3).

Yet, we would like to point out that our data are consistent with the prediction based on language dominance if we look at our specific analysis referring to L2. This more specific analysis focused on whether L2 receives more inhibitory control input (measured as after-effects in terms of n-2 language repetition costs) when switching back from the stronger L1 or the weaker L3. Here, we found large n-2 language repetition costs when switching back from the weaker L3, for which L2 should be a strong competitor, but we did not find evidence for inhibitory aftereffects when switching back from the stronger L1, for which L2 presumably does not require much inhibition. This observation is in line with the more specific prediction of the ICM regarding language dominance and inhibition. It should also be noted that this empirical effect is specific to the L2, for which we can tease apart the influence of switching back from the stronger L1 and from the weaker L3, and it is notable that a similar modulation of language inhibition as a function of the immediately preceding language was not found for L1 German and L3 French (as reported in the appendix).

Our finding of differential L2 inhibition as a function of whether the preceding language was the weaker L3 or the stronger L1 complements recent findings reported by de Bruin et al. (2023) on trilingual language production. They found that there were more L3 intrusions than L1 intrusions when speaking in L2. They also examined n-2 language repetition costs in their Experiment 2, finding small costs for L1 (24 ms) and no significant costs for L3 (7 ms), but they did not focus on L2 because they designed their trial sequences in a way that focuses on analyzing performance when L2 was the target language in the preceding (n-1) trial (see our

Table 2). Hence, they found n-2 repetition costs for L1 (which was English in their study), even though their design was not suited to examine the rank order of inhibition from L1, L2, and L3. This leaves us with the question of why we did not find significant n-2 repetition costs for L1 German in our study.

In the present two experiments, L1 German was clearly the strongest language. The participants indicated German as their native language and they performed well in the German variant of the LexTale task (Lemhöfer & Broersma, 2012), confirming the self-report data. Given that we used a fairly typical design to examine n-2 language repetition costs and also that digits were used in many previous studies as stimulus material (see Table 1), combined with the fact that our findings replicated very well across two experiments, we can offer only a speculative account why we did not find n-2 repetition costs for L1.

The absence of n-2 repetition costs for L1 suggests that inhibition was not stronger for L1 n-2 repetition trials than for L1 n-2 non-repetition trials, even though it still does not rule out some general inhibition in all L1 trials (see Grange et al., 2013, for modeling of n-2 repetition costs in the domain of task switching). Finding that performance in L1 was overall faster than in L2 and L3 differs from findings of Philipp et al. (2007), who also used cued digit naming (but there were diverse L2 and L3 and other methodological differences) and who actually found longer RTs in n-2 language repetition conditions for L1 German compared to their L2 and L3 languages. Yet, given that we have two independent (and larger) samples showing the absence of n-2 repetition costs for L1 German, with a well-controlled and similar language experience, it is possible that the relative shortest RT in L1 was due to the specific combination of languages (L1 German, L2 English, and L3 French) that we used in our two experiments. For instance, Philipp et al. (2007) categorized their three languages in L1, L2, and L3, and L1 was mostly German, but for L2 and L3 there were diverse language combinations, and given their comparatively small sample size, it was not possible to examine the influence of language combination. Likewise, Declerck et al. (2015) used German-Turkish bilinguals in a switching task involving naming in English as well, which is again not comparable to the present study. Moreover, de Bruin et al. (2023) used English-French-Spanish trilinguals in their Experiment 2, and their trial sequence was specifically designed to focus on L1 and L3 rather than on L2, so that procedural differences beyond the different type of samples make a comparison difficult. In future studies, it will be important to explore whether the degree of inhibition depends on the specific language combinations or other aspects of the design, such as the size of the stimulus set or the type of task (e.g., digit naming vs. picture naming).

However, it is important to remember that the novel analysis focused on L2 trials revealed data that are entirely consistent with the idea that language dominance is related to the degree to which a potentially competing language is inhibited. In fact, here, L2 was the strongest competitor for L3, more strongly so than L1, so that performing in L3 French required specific inhibition of L2 English. From third-language learning research, it is known that two foreign languages can produce particularly strong mutual competition, presumably because they have both a late onset of acquisition (at school age) and a common context of use (e.g., the classroom) and possibly share other characteristics of second languages (see, e.g., Kałamała et al., 2023; Marian & Hayakawa, 2021; Tomoschuk et al., 2021, for discussion of relevant factors defining kinds of bilingual experience). So, there are multiple retrieval cues that can cause retrieval interference and thus specific competition among L2 and L3 that is not shared by L1 (see, e.g., the L2 status hypothesis in third language acquisition, e.g., Bardel & Falk, 2021; see Puig-Mayenco et al., 2020, for a review). Hence, strong interference between L2 and L3 could have been expected, and the specific pattern of our n-2 repetition costs for L2 in both experiments is consistent with the idea that strong competitors receive stronger reactive inhibition (e.g., de Bruin et al., 2023; Goldrick & Gollan, 2023; see Declerck & Koch, 2023, for a review).

Conclusions

The present findings suggest that persisting inhibition in a trilingual naming task can be measured in terms of n-2 language repetition costs. The novel finding of stronger n-2 language repetition costs in L2 when switching back from the weaker L3 compared to when switching back from the stronger L1 supports a relation between language-specific strength of competition, with more competition from the more dominant language, and the degree of inhibition required to overcome this competition. This finding was replicated across two cued language switching experiments. However, apart from this finding specifically focusing on L2, overall we did not find specific evidence for trial-based inhibition of L1, whereas evidence for persisting inhibition of L2 was found. Hence, while generally supporting an inhibitory account of language control, the present study also suggests a need for further refinement of the concept of language dominance and its relationship to inhibition.

References

Allport, A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umilta & M. Moscovitch (Eds.), *Conscious and nonconscious information processing: Attention and performance XV* (pp. 421–452). Cambridge, MA: MIT Press.

Altmann, E. M. & Gray, W. D. (2008). An integrated model of cognitive control in task switching. *Psychological Review, 115*, 602-639.

Babcock, L., & Vallesi, A. (2015). Language control is not a one-size-fits-all languages process: evidence from simultaneous interpretation students and the n-2 repetition cost. *Frontiers in Psychology*, *6*, 1622.

Bardel, C., & Falk, Y. (2021). L1, L2 and L3: Same or different? *Second Language Research*, *37*(3), 459–464.

Bobb, S. C., & Wodniecka, Z. (2013). Language switching in picture naming: What asymmetric switch costs (do not) tell us about inhibition in bilingual speech planning. *Journal of Cognitive Psycholology, 25,* 568–585.

Brysbaert, M. (2013). LEXTALE_FR: A fast, free, and efficient test to measure language proficiency in French. *Psychologica Belgica, 53*, 23-37.

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*, 491–511.

de Bruin, A. (2019). Not all bilinguals are the same: A call for more detailed assessments and descriptions of bilingual experiences. *Behavioral Sciences, 9,* 33.

de Bruin, A., Hoversten, L. J., & Martin, C. D. (2023). Interference between non-native languages during trilingual language production. *Journal of Memory and Language*, *128*, 104386.

de Bruin, A., & Xu, T. (2023). Language switching in different contexts and modalities: Response-stimulus interval influences cued-naming but not voluntary-naming or comprehension language-switching costs. *Bilingualism: Language and Cognition, 26(2),* 402-415.

Declerck, M., & Koch, I. (2023). The concept of inhibition in bilingual control. *Psychological Review, 130(4*), 953–976.

Declerck, M., & Philipp, A. M. (2015). A review of control processes and their locus in language switching. *Psychonomic Bulletin & Review*, *22*, 1630-1645.

Declerck, M., & Philipp, A. M. (2018). Is inhibition implemented during bilingual production and comprehension? n-2 language repetition costs unchained. *Language, Cognition and Neuroscience*, *33*, 608-617.

Declerck, M., Thoma, A., Koch, I., & Philipp, A. M. (2015). Highly proficient bilinguals implement inhibition - Evidence from N-2 repetition costs when switching between three languages. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *41*, 1911-1916.

Faul, F., Erdfelder, E., Lang, A.G., & Buchner, A. (2007) G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

Finkbeiner, M., Almeida, J., Janssen, N., & Caramazza, A. (2006). Lexical selection in bilingual speech production does not involve language suppression. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 1075–1089.

Gade, M., Declerck, M., Philipp, A. M., Rey-Mermet, A., & Koch, I. (2021a). On the existence of asymmetrical switch costs and reversed language dominance. *Journal of Cognition, 4*, 55.

Gade, M., Declerck, M., Philipp, A. M., Rey-Mermet, A., & Koch, I. (2021b). Correction: Assessing the evidence for asymmetrical switch costs and reversed language dominance effects – A meta-analysis. *Journal of Cognition, 4,* 60, pp. 1–2.

Goldrick, M., & Gollan, T. (2023). Inhibitory control of the dominant language: Reversed language dominance is the tip of the iceberg, *Journal of Memory and Language, 130,* 104410.

Grange, J.A., Juvina, I., & Houghton, G. (2013). On costs and benefits of n–2 repetitions in task switching: Towards a behavioral marker of cognitive inhibition. *Psychological Research, 77*, 211–222.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1*, 213–229.

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, *25*, 515-530.

Grosjean, F. (2010). *Bilingual: Life and reality.* Harvard University Press.

Guo, T., Liu, F., Chen, B., & Li, S. (2013). Inhibition of non-target languages in multilingual word production: Evidence from Uighur-Chinese-English trilinguals. *Acta Psychologica, 143*, 277–283.

Guo, T., Ma, F., & Liu, F. (2013). An ERP study of inhibition of non-target languages in trilingual word production. *Brain and Language, 127*, 12–20.

Horoufchin, H., Philipp, A. M., & Koch, I. (2011). The dissipating task-repetition benefit in cued task switching: Task-set decay or temporal distinctiveness? *Journal of Experimental Psychology: Human Perception and Performance, 37*, 455-472.

Jackson, G. M., Swainson, R., Cunnington, R., & Jackson, S. R. (2001). ERP correlates of executive control during repeated language switching. *Bilingualism: Language and Cognition, 4(2),* 169-178.

Kałamała, P., Chuderski, A., Szewczyk, J., Senderecka, M., & Wodniecka, Z. (2023). Bilingualism caught in a net: A new approach to understanding the complexity of bilingual experience. *Journal of Experimental Psychology: General, 152*(1), 157–174.

Kiesel, A., Wendt, M., Jost, K., Steinhauser, M., Falkenstein, M., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching: A review. *Psychological Bulletin, 136*, 849–874.

Kleinman, D., & Gollan, T. H. (2018). Inhibition accumulates over time at multiple processing levels in bilingual language control. *Cognition, 173*, 115-132.

Koch, I., Gade M., Schuch S., & Philipp A. M. (2010). The role of inhibition in task switching: A review. *Psychonomic Bulletin & Review, 17*, 1–14.

Koch, I., & Kiesel, A. (2022). Task switching: Cognitive control in sequential multitasking. In A. Kiesel, L. Johannsen, I. Koch & H. Müller (Eds.), *Handbook of human multitasking* (pp. 85-144). Springer. https://doi.org/10.1007/978-3-031-04760-2_3

Koch, I., Poljac, E., Müller, H., & Kiesel, A. (2018). Cognitive structure, flexibility, and plasticity in human multitasking—An integrative review of dual-task and task-switching research. *Psychological Bulletin*, *144*, 557-583.

Kroll, J. F., Bobb, S. C., Misra, M. M., & Guo, T. (2008). Language selection in bilingual speech: Evidence for inhibitory processes. *Acta Psychologica*, *128*, 416-430.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, *44*, 325-343.

Ma, F., Li, S., & Guo, T. (2016). Reactive and proactive control in bilingual word production: An investigation of influential factors. *Journal of Memory and Language*, *86*, 35-59.

Marian, V., & Hayakawa, S. (2021). Measuring bilingualism: The quest for a "bilingualism quotient". *Applied Psycholinguistics*, *42*(2), 527–548.

Meuter, R. F. I., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, *40*, 25–40.

Philipp A. M., Gade M., & Koch I. (2007). Inhibitory processes in language switching: Evidence from switching language-defined response sets. *European Journal of Cognitive Psychology*, *19*, 395–416.

Philipp, A. M., & Koch, I. (2009). Inhibition in language switching: What is inhibited when switching between languages in naming tasks? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 1187–1195.

Puig-Mayenco, E., González Alonso, J., & Rothman, J. (2020). A systematic review of transfer studies in third language acquisition. *Second Language Research*, *36*(1), 31–64.

Titone, D. A., & Tiv, M. (2022). Rethinking multilingual experience through a systems framework of bilingualism. *Bilingualism: Language and Cognition.* Advance online publication.

Tomoschuk, B., Duyck, W., Hartsuiker, R., Ferreira, V., & Gollan, T. (2021). Language of instruction affects language interference in the third language. *Bilingualism: Language and Cognition, 24(4)*, 707-718.

Verhoef, K. M. W., Roelofs, A., & Chwilla, D. J. (2009). Role of inhibition in language switching: Evidence from event-related brain potentials in overt picture naming. *Cognition, 110*, 84–99.

Appendix

Table A1

Mean RT (ms [SE]) and Mean Error Percentages (SE) for L1 (German) Naming Trials (ms [SE]) as a Function of Previous-Trial (N-1) Language (L1 vs. L3) and N-2 Language Repetition in Experiment 2

	N-2 Language Repetition		
N-1 Language	N-2 Repetition	N-2 Non-Repetition	N-2 Repetition Costs
Experiment 1			
RT (ms [SE])			
N-1 L2 (English)	1030 (32)	1017 (29)	13
N-1 L3 (French)	1023 (31)	1035 (32)	-12
Error (% [SE])			
N-1 L2 (English)	3.6 (0.7)	3.7 (0.9)	-0.1
N-1 L3 (French)	2.5 (0.7)	2.6 (0.6)	-0.1
Experiment 2			
RT (ms [SE])			
N-1 L2 (English)	937 (20)	921 (21)	16
N-1 L3 (French)	948 (24)	960 (26)	-12
Error (% [SE])			
N-1 L2 (English)	4.3 (0.7)	5.2 (0.6)	-0.9
N-1 L3 (French)	2.2 (0.4)	3.3 (0.6)	-1.1

Note. The data for Experiment 1 are aggregated over RCI.

For the L1 German trials in Experiment 1, the ANOVA on RT revealed that the main effects of language sequence of n-1 language were non-significant, $F < 1$, as was the interaction, $F(1, 27) = 2.936$, $p = .098$, $\eta_p^2 = .098$. The same analysis for the error rates also showed clearly no interaction, $F < 1$, and no main effect of language sequence, $F < 1$, but a just significant main effect of n-1 language, $F(1, 27) = 4.306$, $p = .048$, $\eta_p^2 = .138$, suggesting higher error rates when switching from L2 to L1 than from L3 to L1.

For the L1 German trials in Experiment 2, the ANOVA on RT did not show an effect of n-2 language repetition, $F < 1$, but of n-1 language, $F(1, 43) = 9.380$, $p = .004$, $\eta_p^2 = .179$, showing higher RT in L1 German when coming from L3 French than from L2 English (954 ms vs. 929 ms). The interaction was significant, too, $F(1, 43) = 5.582$, $p = .023$, $\eta_p^2 = .115$, showing 16 ms n-2 repetition costs when coming from L2 English to L1 German but a n-2 repetition benefit of 12 when coming from L3 French. However, the main effect of n-1 language and the interaction is difficult to interpret because the ANOVA on error rates revealed a clear speed-accuracy tradeoff. Specifically, we found a main effect of n-2 repetition, $F(1, 43) = 4.415$, $p = .042$, $\eta_p^2 = .093$, showing an overall 1.0% benefit (not a cost) of n-2 repetition, regardless of whether coming from L2 or L3, $F < 1$ for the interaction. The error rates also show a main effect of n-1 language, $F(1, 43) = 20.130$, $p = .001$, $\eta_p^2 = .319$, showing lower error rates when coming from L2 English compared to L3 French (2.7% vs. 4.7%), which is the opposite of what was found in RT. Hence, the focused analysis on L1 German as a function of language in trial n-1 showed a speed-accuracy tradeoff pattern that does not invite any clear interpretation.

Table A2

Mean RT (ms [SE]) and Mean Error Percentages (SE) for L3 (French) Naming Trials (ms [SE]) as a Function of Previous-Trial (N-1) Language (L1 vs. L2) and N-2 Language Repetition in Experiment 2

	N-2 Language Repetition		
N-1 Language	N-2 Repetition	N-2 Non-Repetition	N-2 Repetition Costs
Experiment 1			
RT (ms [SE])			
N-1 L1 (German)	1159 (30)	1171 (29)	-12
N-1 L2 (English)	1182 (31)	1167 (32)	15
Error (% [SE])			
N-1 L1 (German)	3.0 (0.7)	2.1 (0.6)	0.9
N-1 L2 (English)	3.0 (0.7)	3.4 (0.7)	-0.4
Experiment 2			
RT (ms [SE])			
N-1 L1 (German)	1110 (34)	1116 (37)	-6
N-1 L2 (English)	1113 (33)	1110 (34)	3
Error (% [SE])			
N-1 L1 (German)	3.1 (0.6)	3.6 (0.7)	-0.5
N-1 L2 (English)	3.9 (0.6)	3.5 (0.6)	0.4

Note. The data for Experiment 1 are aggregated over RCI.

For the L3 French trials in Experiment 1, the ANOVA on RT did not show significant main effects of n-2 language repetition and of n-1 language, both *F*s < 1, and the interaction was not significant, too, $F(1, 27) = 3.285$, $p = .081$, $\eta_p^2 = .108$. The same pattern was found in the error rates, with no main effects of n-2 language repetition, $F < 1$, and of n-1 language, $F(1, 27) = 2.269$, $p = .144$, $\eta_p^2 = .078$, as well as no interaction, $F(1, 27) = 1.516$, $p = .229$, $\eta_p^2 = .053$.

For the L3 French trials in Experiment 2, the ANOVA on RT did not show any significant effect, all *F*s < 1, just as the ANOVA on error rates, all *F*s < 1.