Language and Cognition (2025), 17, e15, 1–17
doi:10.1017/langcog.2024.43

ARTICLE

Language control in auditory bilingual comprehension: uncovering novel evidence from the $n - 2$ repetition paradigm

Hong Liu[1] and Adel Chaouch-Orozco[2]

[1]Department of Applied Linguistics, Xi'an Jiaotong-Liverpool University, Suzhou, China and [2]Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hong Kong, China
Corresponding author: Adel Chaouch-Orozco; Email: adel.chaouchorozco@polyu.edu.hk

(Received 21 December 2023; Revised 07 June 2024; Accepted 27 June 2024)

Abstract

In language production, inhibitory control is assumed to be the primary mechanism responsible for successful bilingual processing. To convey messages in one language, bilinguals must inhibit the unintended language. However, it remains unclear whether the same mechanism works in bilingual comprehension. Following up and expanding on Declerck and Philipp (2018, 'Is inhibition implemented during bilingual production and comprehension? n–2 language repetition costs unchained', *Language, Cognition and Neuroscience*, vol. *33*, pp. 608–617), the present study investigates whether inhibition is involved in the linguistic identification system during bilingual comprehension with the $n - 2$ repetition paradigm. This is the second study exploring comprehension with this methodological setup to date. We used an auditory word–picture matching task with Chinese late trilinguals who learned their two non-native languages (L2 English and L3 Spanish) via formal school instruction. Our results indicate that participants responded faster in the $n - 2$ repetition trials (i.e., when the target language in the n and $n - 2$ trials matched). That is, we observed an $n - 2$ repetition *benefit* effect, a novel finding in this literature that goes counter the $n - 2$ repetition cost effect reported in previous studies using production-based tasks. In addition, our results underscore the complex interplay between proficiency and use and the resulting dynamics within the bilingual lexicon. We discuss the results in light of the different bilingual representation and processing models.

Keywords: bilingual comprehension; inhibition; language control; $n - 2$ repetition; proficiency

1. Introduction

Accumulating evidence suggests that bi-/multilinguals' (hereafter 'bilinguals') languages are activated at the early stages of processing and thus enter in competition for selection; evidence of this is found for both language learners and more balanced bilinguals (Kroll et al., 2006). This results in the need to apply a particular mechanism to prevent nontarget languages from interfering (Declerck & Koch, 2023; Declerck & Philipp, 2015; Kroll et al., 2012). In the realm of language production, inhibitory control is assumed to be the primary mechanism responsible for mitigating interference and facilitating successful bilingual processing. Essentially, in order to convey messages in one language, bilinguals must inhibit the unintended language (Green, 1998; Kroll et al., 2008). Note that there are also activation-based accounts suggesting that successful language selection can be achieved by the stronger activation of the selected language (e.g., Blanco-Elorrieta & Caramazza, 2021). Regarding bilingual comprehension, the involvement of the inhibitory mechanism remains an open question. The present study set out to investigate whether inhibition is involved in the linguistic identification system during bilingual comprehension with the $n-2$ repetition paradigm, following up and expanding on Declerck and Philipp (2018), the first study that adopted this setup in bilingual comprehension. Similar to Declerck and Philipp (2018), we also investigated processing in late trilinguals: Chinese native speakers who learned their two non-native languages (L2 English and L3 Spanish) via formal school instruction. In the following, we review how previous studies have approached the involvement of inhibition in comprehension and how the $n-2$ repetition paradigm can provide valuable insights into this area of study.

1.1. The role of inhibition in bilingual comprehension

Influential models of bilingual word recognition differ in their views regarding the involvement of inhibition. For instance, the Bilingual Interaction Model (BIA, Dijkstra & Van Heuven, 1998) and its successor, the BIA+ (Dijkstra & van Heuven, 2002), assume that the two languages are activated to different degrees based on the similarity between the input and the mental representations. However, these two models diverge in their views on how activation is managed.

For the BIA, the activation level of the words in a language results in that language's baseline activation. Factors like stimuli list (previous items in an experiment or previous words in a conversation) may raise or decrease those activation levels holistically. To decrease the activation level of nontarget language representations, the BIA incorporates an inhibitory mechanism *within* the language identification system. Specifically, in this view, inhibition operates at two levels, language-independent lateral inhibition at the word level and top-down inhibition at the language level. First, all words have the ability of inhibiting each other regardless of language membership. At the same time, the target language node, collecting activation from the target input stimuli, exerts inhibition to nontarget representations.

The BIA+ model takes a different stance. In this updated view, the baseline activation of lexical units is independent of language membership. This means that the activation levels of the words are not modulated by previous items. However, their recognition thresholds can be modulated. This modulation is *not* achieved through inhibition, which is absent from the language identification system in the BIA+, but through a task/decision system responsible for regulating the activation output and adapting it to the current task. Notably, this task/decision system is also preserved in the more recent bilingual processing model Multilink (Dijkstra et al., 2019).

To date, empirical evidence for the involvement of inhibition in bilingual comprehension remains inconclusive. A case in point is the asymmetrical switch cost, an empirical effect sometimes observed in production studies, which is either diminished or altogether absent in comprehension (Declerck et al., 2019; Hirsch et al., 2015; Macizo et al., 2012). In language switching tasks, participants switch between languages when responding (in, e.g., picture naming). It is often observed that responding is slower and less accurate when the preceding trial ($n - 1$) is in a different language. In addition, this switch cost is larger when switching from the L2 to the L1 than vice versa. This asymmetry is often attributed to inhibitory control. Using the more robustly represented L1 results in strong activation that requires more inhibition when responding in the L2. Then, when the L1 has to be used again, overcoming that strong inhibition exerted over the L1 is more effortful than in the opposite scenario (L2 to L1).

Importantly, however, not only the reliability of the asymmetrical switch cost has been questioned (see the meta-analysis in Gade et al., 2021), but how to interpret the effect remains open to debate (Bobb & Wodniecka, 2013; Declerck & Koch, 2023). For instance, an alternative explanation that does not resort to inhibition was provided by Philipp et al. (2007). For the authors, the strong activation of the nontarget language used in the $n - 1$ trial can persist into the subsequent trial (n), interfering with the target language and resulting in switch costs. Since the most proficient language has an overall higher baseline activation level, the less proficient language is activated more strongly, leading to more interference when switching from the less to the most proficient language.

1.2. The $n - 2$ repetition paradigm

Given the above limitations, the $n - 2$ repetition paradigm has been regarded as more precise for tapping into the role of inhibition (Koch et al., 2010; Mayr & Keele, 2000; Philipp et al., 2007; see review in Declerck & Philipp, 2015). In $n - 2$ repetition experiments, subjects switch between three languages while performing a task. The analysis entails comparing sequences like AB*A* and CB*A*, each letter referring to a language to be employed in each subsequent trial (e.g., Chinese–English–Chinese vs. Spanish–English–Chinese trials). Thus, in the CB*A* sequence, there are more intervening trials (at least two) before language A is used again. This larger number of intervening trials allows the inhibition applied earlier upon language A to dissipate further compared to what occurs in the AB*A* sequence, where there is only one trial in between (i.e., $n - 2$ repetition). In other words, there is a larger amount of residual inhibition in the AB*A* than in the CB*A* sequence. Consequently, less effort is required to overcome the residual inhibition exerted on that language when responding to A in the CB*A* sequence. This leads to better performance (i.e., shorter RTs and/or fewer errors) compared to the AB*A* sequence (i.e., $n - 2$ repetition cost).

To date, the studies employing this methodology have consistently observed significant $n - 2$ repetition costs (e.g., Declerck et al., 2015; Philipp et al., 2007; see Guo et al., 2013 effects appearing in the event-related potential data only). Crucially, unlike asymmetrical switch costs, $n - 2$ repetition costs have been claimed to be only attributable to inhibitory processes (Declerck & Philipp, 2015). For instance, note that the persisting activation account discussed above would predict an $n - 2$ repetition *benefit* (Koch et al., 2010). Therefore, the $n - 2$ repetition cost is considered a more reliable, less ambiguous index of inhibitory control in bilingual processing. However, note that a lack of $n - 2$ repetition cost does not necessarily imply that inhibition is not involved. This is because this cost is a relative measure of inhibition, resulting from more inhibition in the $n - 2$ repeat trials than in the non-repeat trials. A lack of $n - 2$ repetition cost could be modulated by other factors, such as language proficiency. Consequently, the claim that inhibitory control in bilingual comprehension is not necessary is best supported with empirical evidence showing a $n - 2$ repetition *benefit*.

To the best of our knowledge, only Declerck and Philipp (2018) explored this question, using the $n - 2$ repetition paradigm in comprehension. In their study, two production tasks (picture naming and reading aloud) and two comprehension tasks (picture and word categorization) were administered among German–English–French trilinguals, who were dominant in German, relatively proficient in English and less proficient in French. While an $n - 2$ repetition cost was observed in the two production tasks, the effect only appeared in one of the comprehension experiments (picture categorization) and, importantly, only when inspecting the error rates of the least dominant language (French). The authors explained the cost in the least dominant language in the comprehension task by resorting to the BIA model. In their view, the more proficient – and overall more activated – language applied more inhibition to the weaker language than vice versa. As a result, the least proficient language received the largest amount of inhibitory feedback. Thus, the authors concluded that inhibition might occur in comprehension, but its involvement depends on various factors, including task demands (i.e., whether the task involves lexical-semantic processing) and language proficiency (i.e., inhibition may be applied to a less proficient language).

Declerck and Philipp's (2018) study represents a necessary step in a holistic approach to the investigation of bilingual language control with the $n - 2$ repetition paradigm, while also prompting further inquiry. Their investigation posits that inhibition's role may be contingent on specific task requirements, particularly those involving lexical-semantic processing. However, tasks can vary along other dimensions too, such as potential conflicts at the response level.

For instance, in comparing a general lexical decision task versus a language-specific lexical decision task, the demands on recognizing homographs differ. In the general task, where participants make 'yes' responses to words from either language, homographs typically elicit faster responses than control words (e.g., Experiment 3, Dijkstra et al., 1998). This lack of conflict at the response level, where either interpretation of a homograph facilitates decision-making, can explain the quicker response. Conversely, in a language-specific task, where the 'yes' response is reserved for words from only one language, homographs (presented in a mixed-language list) tend to slow down response times (RTs; e.g., Experiment 2, Dijkstra et al., 1998). This slowing down arises from conflict at the response level, as one must disregard the reading in the nontarget language.

Bilingual language processing in everyday contexts presents a range of tasks that often differ from those typically used in experimental settings. Real-world bilingual comprehension frequently centers on meaning extraction rather than on resolving response-level conflicts. In such contexts, the activation of nontarget languages may not necessitate suppression since both languages converge on a shared meaning. This may be particularly true for language learners, who often depend on their dominant language to access meaning while processing in their nondominant language (Kroll et al., 2010). Consequently, it is of great significance to investigate language control within bilingual comprehension through tasks that closely mirror this aspect of bilingual processing (i.e., the absence of overt conflict at the response level).

The way such task demands interact with inhibition remains unspecified in the BIA model – an account focused on word recognition. Nonetheless, the model's assumption of lateral inhibition and language-wide inhibition in the language identification system, particularly during lexical-semantic processing (Declerck et al., 2015; Green, 1998; Schwieter & Sunderman, 2008), would imply that response-level task demands should not substantially impact the inhibitory mechanism or alter the $n-2$ repetition cost. In other words, according to the BIA, the $n-2$ repetition cost should manifest consistently regardless of the task demands at the response level – as observed by Declerck and Philipp (2018). However, as noted above, it is still possible to observe a null effect under the inhibition account, as this cost is a relative measure and can be modulated by language proficiency, as shown in Declerck and Philipp (2018). The more proficient language, receiving less inhibition from other less proficient languages overall, may not show an $n-2$ repetition cost.

Conversely, in the BIA+, inhibition is not a component of the identification system but the task/decision system is responsible for regulating the system's output. Thus, the activation of nontarget representations may be more susceptible to task demands at the response level. This could, in turn, influence the $n-2$ repetition effect, resulting in more nuanced outcomes depending on the task under examination. Therefore, examining the interplay between task demands and language control mechanisms is crucial in shedding light on the complex reality of bilingual communication.

1.3. The present study

The present study follows up and expands on Declerck and Philipp (2018). We investigated the involvement of inhibition in bilingual comprehension by adopting the $n-2$ repetition paradigm, this being the second study exploring comprehension with this methodological setup to date. First, to extend the finding of Declerck and Philipp (2018), we used a picture–word matching task based on Jiao et al. (2020, 2022), which also employed it to investigate switching costs and language processing in bilingual comprehension. In this task, participants decide whether a picture and a spoken word in one of the three languages match. Following Declerck and Philipp (2018), we tested Chinese late trilinguals who learned their two non-native languages (L2 English and L3 Spanish) via formal school instruction. On average, they had spent 13.92 years (*SD* = 2.80) learning English and 3.67 years (*SD* = 2.13) learning Spanish.

In addition, we used a relatively large stimuli list, containing 27 items (more details are provided in the 'Methods' section). Previous studies using the $n-2$ paradigm exclusively used relatively small lists. For example, Philipp et al. (2007) employed nine digits, whereas six pictures were used in Declerck and Philipp (2018). A limited number of items within a dataset necessarily results in increased item repetitions, potentially constraining the ecological validity of insights derived from such data (Winter & Grice, 2021). In the context of language control studies, the repetition of items may exert a significant influence on the activation levels specific to those items, which, in turn, could have implications for the underlying control processes

(Kleinman & Gollan, 2018; Shen & Chen, 2023). Analogously, repetition priming effects would entail the overall higher activation of all the items within the set, which would interact with the language activation levels, leading to some unintended impact on the $n - 2$ repetition effect. Therefore, given the prevalent use of limited stimulus lists in the field and the potential influence item repetition may have on the $n - 2$ repetition effect, we decided to increase the number of stimulus items. Such efforts are imperative to uphold the validity of conclusions drawn from bilingual language control studies.

2. Methods

2.1. Participants

Thirty-seven Chinese–English–Spanish trilinguals (mean age = 21.52, *SD* = 2.37) from several universities in mainland China participated. Mandarin Chinese was their dominant language and the language of daily communication. They were undergraduate students in Spanish, and they had learned English as a compulsory subject at school before entering university. The medium of instruction was Chinese in these universities, but Spanish lessons were mainly delivered in Spanish. During the first 2 years of university, they had 2 h of English class each week. After that, English classes became optional.

Before the experiment, we measured the participants' language proficiency in each language with the Multilingual Naming Test (MINT; Gollan et al., 2012). The MINT is a standardized naming test where participants name 68 pictures of varying word frequencies. It has been validated as a reliable measure for capturing variance in bilinguals' language proficiency and language dominance in English, Spanish and Mandarin (Gollan et al., 2012; Ivanova et al., 2013; Sheng et al., 2014). The MINT has also been shown to correlate with the global proficiency of second language learners at both high and low proficiency levels (Liu & Chaouch-Orozco, 2024). Participants completed the MINT first in Chinese, then in English and lastly in Spanish. Their MINT scores, shown in Table 1, significantly differed across the three languages (L1 vs. L2, $t = 14.22$, $p < 0.001$; L1 vs. L3, $t = 15.96$, $p < 0.001$; L2 vs. L3, $t = 2.48$, $p = 0.01$), indicating that the participants' proficiency in the three languages correlated with their order of acquisition. Their self-reported proficiency in the three languages mirrored the MINT scores, except that the difference in perceived proficiency between L2 English and L3 Spanish only approached significance (L1 vs. L2, $t = 15.61$, $p < 0.001$; L1 vs. L3, $t = 11.21$, $p < 0.001$; L2 vs. L3, $t = 1.49$, $p = 0.07$). The subjects also completed the Language History Questionnaire 3 (LHQ3, Li et al., 2020).

Table 1. Participants' language use and proficiency information

	L1 Chinese	L2 English	L3 Spanish
Total use in years	21.17 (2.45)	13.92 (2.80)	3.67 (2.13)
Language use (in %)	63.27 (13.01)	15.80 (9.12)	20.92 (9.83)
MINT scores	60.84 (2.96)	40.45 (5.51)	28.14 (8.68)
Self-reported proficiency (on a scale of 7)	6.64 (0.49)	4.83 (0.66)	4.55 (1.05)
Listening	6.80 (0.44)	4.74 (0.85)	4.24 (1.14)
Speaking	6.62 (0.60)	4.30 (0.88)	4.10 (1.04)
Reading	6.72 (0.46)	5.42 (0.71)	5.12 (1.14)
Writing	6.40 (0.82)	4.86 (0.93)	4.74 (1.20)

Their language use also significantly differed across the three languages (L1 vs. L2, $t = 13.90$, $p < 0.001$; L1 vs. L3, $t = 11.82$, $p < 0.001$; L2 vs. L3, $t = -2.19$, $p = 0.02$).

Concerns regarding the sample size of the present study are acknowledged, given the inherent logistical challenges associated with recruiting trilingual participants. The rarity of individuals proficient in L1 Chinese, L2 English and L3 Spanish significantly constrains the pool of potential participants. This difficulty is exacerbated by the need to carefully match the proficiency levels among the languages – specifically, ensuring higher proficiency in English than in Spanish, yet maintaining sufficient proficiency in Spanish to participate in the experiment, in alignment with the participant background in Declerck and Philipp (2018).

Given the complexities of conducting power analysis for mixed-effects models, we draw upon comparable studies within the literature for guidance. Notably, a power analysis in a recent study by Koch et al. (2024), which also investigated $n - 2$ repetition costs, indicated that a sample of 28 participants would be adequate to discern differences in $n - 2$ language repetition costs with a power of 0.80. Consequently, we consider that our sample size is sufficiently robust to detect significant effects within an $n - 2$ repetition design framework, while acknowledging the challenges in participant recruitment.

2.2. Tasks and materials

The participants completed a picture–word matching task. They listened to a word and decided whether a picture on the screen matched the heard stimuli. They were asked to respond as accurately and quickly as possible. Twenty-seven pictures were taken from MultiPic (Duñabeitia et al., 2018). The nouns labeling each of these entities in the three languages were comparable in word frequency and the number of syllables. The frequencies for the Chinese, English and Spanish names for the pictures were based on SUBTLEX-CH (Cai & Brysbaert, 2010), SUBTLEX-UK (van Heuven et al., 2014) and SUTBLEX-ESP (Cuetos et al., 2011), respectively. See Appendix A for the complete list of stimuli. The pronunciation of the stimuli words was created on the website https://soundoftext.com/, using Mandarin Chinese, British English and Peninsular Spanish. The same pictures and stimuli words were used to construe match and non-match trial.

There were four blocks of 40–41 trials. Each word was repeated three times in each language as audio stimuli, and each picture was repeated nine times as visual stimuli. The sequence of trials was pseudorandomized across languages, stimulus type, language sequence (e.g., AB*A* vs. CB*A*) and answer type (match vs. non-match). Immediate repetition of a language and immediate repetition of a stimulus (both visual and audio) either in the subsequent trial (either match or non-match) or the one following it was not allowed.

2.3. Procedure

The online experiment was created and presented on *Gorilla Experiment Builder* (Anwyl-Irvine et al., 2020). After providing their consent and reading the instructions, the participants performed an audio check by listening to a few words to check the sound volume. Then, they were presented with 12 practice trials and were allowed to repeat the practice as many times as needed. Right-handed subjects had to manually press ‘0’ on the keyboard to indicate ‘match’ and ‘1’ for ‘non-match’. The order was inverted for left-handed participants. Following previous procedures, we presented the stimuli with a cue indicating the relevant language. Each trial started with a fixation point (600 ms), followed by the stimulus and the cue. The stimulus and the cue remained on the screen until a response (key press) was given. The participants were allowed to rest between blocks.

2.4. Data analysis

Data and analysis code can be found in the corresponding author’s OSF repository (https://osf.io/3qmy4/). As the task was completed online, some measures were taken to ensure the quality of the data. First, the average and standard deviation time taken to complete the task were calculated. Participants who took more than the average plus two standard deviations to complete were excluded from the analysis. This was done because excessive time could indicate that participants were distracted while completing the experiment. Second, each participant’s data were examined carefully to detect random responses. Our experiment did not permit more than five consecutive identical responses (either match or mismatch), thus, instances where a participant provided the same answer for more than 10 consecutive trials were flagged as random. As there were no more than five consecutive runs for the same answer (match or mismatch), if a participant provided the same answer for more than 10 consecutive trials, we considered such answers random. Furthermore, considering the high-frequency nature of the words used and the familiarity of the depicted objects, we anticipated a high level of accuracy in participant responses. Therefore, an accuracy rate below 80% was interpreted as an indication of potential distraction among participants. Following these exclusion criteria, two participant’s data were moved. Thus, the experiment contained data from 35 participants.

For each participant, the first two trials of each block were removed. RTs three standard deviations below and above the mean were also removed. Responses below 200 ms that hardly reflected conscious responses were further removed. In total, 2.88% of the observations were removed.

The exploration of the RT distributions indicated that the Inverse Gaussian transformation provided the best correction to the typical skewness in the distribution of the raw RTs. Thus, in the RT analysis, these transformed RTs were used and incorrect responses were removed. RTs and error rates were analyzed with (generalized) linear mixed-effects models (Baayen et al., 2008) using *R* (version 3.6.1; R Core Team, 2021) with the *lme4* package (Bates et al., 2015). Following Scandola and Tidoni (2021), complex random intercepts (CRIs; i.e., using random intercepts for each grouping factor instead of random slopes) were employed. For each analysis, a maximal model was fitted. In the case of non-convergence, the CRI that explained the least variance was removed until the model converged.

The fixed effects in the models included main effects and interactions of interest (Brauer & Curtin, 2018). The grouping factors were trial type (i.e., repetition vs. non-repetition), language (i.e., Chinese, English and Spanish) and their interactions. Full-CRI structures with random intercepts for subjects and each grouping factor were specified. Sum contrasts for the language variable were employed. For the accuracy analysis, the variable was dummy-coded (1 for ‘correct’ and 0 for ‘incorrect’) and employed generalized linear mixed-effects models with a binomial family fit.

Because R only allows for $n - 1$ levels contrasts to be specified, two maximal models differing in their contrasts for the factor language were built in each experiment's analysis. If a maximal model did not converge, the CRI explaining the least variance was removed and the resulting model was run. This procedure was repeated until the model converged. Further, we checked model assumptions (e.g., normality of residuals' distribution, homoscedasticity) and removed observations with absolute standardized residuals above 2.5 *SD* (Baayen & Milin, 2010). Model 1 contrasted Chinese versus English (and Chinese vs. Spanish), and Model 2 contrasted Chinese versus Spanish (and English vs. Spanish). When the effects are identical in all models, we report them only once. When the effects differ between the two models, we report the results of both models (see Appendix B for all maximal and convergent models and their results).

3. Results

Table 2 displays the RTs and error rates. The RT analysis showed that the main effect of trial type was significant ($\beta = 0.02$, $t = 2.01$, $p < 0.05$), indicating that responses to repetition trials were faster. In other words, we observed an $n - 2$ repetition *benefit* effect. Moreover, the *post hoc* pairwise comparisons revealed that this effect was significant in English ($\beta = 0.03$, $z = 2.41$, $p < 0.05$). The effect of trial type was also significant in Spanish ($\beta = 0.03$, $z = 2.56$, $p < 0.05$).

However, the error rate analysis only partially replicated these patterns, with a significant effect found for English ($\beta = 1.46$, $z = 4.70$, $p < 0.001$) but not for Spanish ($\beta = 0.03$, $z = 1.27$, $p = .90$). Interestingly, this effect reflects an $n - 2$ repetition cost, instead of the benefit we observed in the rest of results. Note, however, that accuracy effects have been traditionally regarded as less reliable difference of processing mechanisms. Ultimately, the results of interest are RTs, as the standard assumption is that processing difficulties are reflected in those RTs.

4. Discussion

The present study set out to investigate whether inhibition is involved in the linguistic identification system during bilingual comprehension with the $n - 2$ repetition paradigm.

Table 2. Mean response times (RTs, in milliseconds; standard deviations), error rates (%) and $n - 2$ repetition effects (in milliseconds)

	$n - 2$ repetition		$n - 2$ non-repetition			
Language	RT	Error rates	RT	Error rates	$n - 2$ repetition effect in RT	$n - 2$ repetition effect in error rates
Chinese	1,124 (512)	0.1	1,127 (588)	0.0	−3	0.1
English	1,274 (778)	0.6	1,324 (899)	0.1	−50*	0.5*
Spanish	1,539 (1,263)	0.4	1,594 (1,196)	0.4	−55*	0

*$p < 0.05$.

Declerck and Philipp (2018), the only study that employed this methodology in comprehension, provided some limited evidence to support that inhibition was involved in bilingual comprehension, but that it was also contingent on task demands and language proficiency. To further examine this question, we followed up on their study by adopting the same $n - 2$ repetition paradigm but with a different comprehension-based task. Chinese–English–Spanish late trilinguals performed a picture-matching task with a relatively large set of items. The increased size of the stimuli list compared to what was used in previous $n - 2$ studies aimed to address the ecological concern with generalizing findings drawn with a small number of items (Winter & Grice, 2021).

Our data indicate that participants responded faster in the $n - 2$ repetition trials (i.e., when the target language in the n and $n - 2$ trials matched). That is, we observed an $n - 2$ repetition *benefit* effect in response latencies, going counter the $n - 2$ repetition cost effect reported in previous studies using production-based tasks that is indicative of the presence of inhibitory control mechanisms and representing a novel finding in this literature. In addition, the $n - 2$ benefit effect was observed in both English and Spanish, but not in Chinese. There was also a speed–accuracy trade-off in English (i.e., an $n - 2$ repetition benefit in RTs but a cost in accuracy). Next, we interpret the results in light of the different bilingual processing models.

In bilingual language production, studies using the $n - 2$ repetition paradigm have consistently observed an $n - 2$ repetition cost (Babcock & Vallesi, 2015; Declerck et al., 2015; Philipp & Koch, 2009; Philipp et al., 2007; Timmer et al., 2018). The explanation for this cost relies on persisting inhibition. In the sequence AB*A*, after language A is used in the first trial, it is inhibited in the next trial (language B). This inhibition persists into the third trial of the sequence, when A is required again. This results in worse performance than in the CB*A* sequence. There, when language A is required in the third trial, the initial inhibition exerted on A in the first trial has had more time to dissipate.

Here, we predicted that if inhibition operates within the language identification system as posited in the BIA model, we should expect an $n - 2$ repetition cost, which would be further modulated by language proficiency, similar to what was found in Declerck and Philipp (2018). In contrast, we observed an $n - 2$ repetition *benefit*, an effect not predicted by the inhibition account but that can be explained by the persisting activation account discussed in the introduction (Declerck & Philipp, 2015; Koch et al., 2010). Responding in the $n - 2$ trial of the AB*A* sequence activated language A above its baseline level. This extra activation – not being suppressed at the subsequent switch to language B – needed some time to dissipate. As such, by the time the participant responds in the n trial (third trial) to language A again, the remaining residual activation facilitates the performance compared to the situation where no persisting activation was to be found (i.e., the CB*A* sequence).

Our findings are, in principle, at odds with those reported in Declerck and Philipp (2018). Using the same $n - 2$ repetition paradigm, Declerck and Philipp (2018) observed costs, although only in the least proficient language, whereas the present study found benefits. Noteworthy, however, the tasks employed in the two studies differed. In the present task, participants needed to recognize the heard word, retrieve the conceptual representation, and judge whether that representation matched the concept in the picture. Notably, the audio stimulus also sent activation to the translation equivalents in other languages. However, this would not create conflict at the response level, as all representations lead to a conceptual representation that is, at least to some extent (Chaouch-Orozco et al., 2023), shared among the translation equivalents.

Conversely, the task in Declerck and Philipp (2018) was picture categorization, in which subjects judged whether a specific letter was present in the word describing the picture in the input language. There, the concept depicted by the picture activated the orthographic representations in all three languages. This can lead to conflict at the response level with regards to the presence or absence of the specific letter in the corresponding word of the target language.

Reconciling the different findings observed in the two studies poses a challenge within the framework of the BIA model, as this account does not explicitly address how differences in task demands at the response level may affect the engagement of inhibitory processes. As such, the BIA model predicts an $n - 2$ repetition cost, irrespective of the task at hand, provided that it involves lexical-semantic involvement processing, a prediction at odds with our findings.

Conversely, the BIA+ model offers a potential explanation for the findings observed in both studies. This model posits that in the absence of inhibition within the language identification system, the activation outputs are regulated through the task/decision system, which can respond to differences in task demands by implementing dynamic adaptations. In the picture–word matching task employed in the present study, the primary challenge lies in understanding the meaning, rather than managing multiple linguistic systems simultaneously. The converging paths to the shared meaning from the different language representations might enhance task performance, as opposed to hindering it. Consequently, the task/decision system adjusts to these demands by allowing the persisting activation of items from the same language to influence the recognition threshold for that language when it is reused two trials later.

In a picture categorization task such as the one used in Declerck and Philipp (2018), because of the conflict at the response level, the task/decision system may manage the persisting activation in a different manner to prevent interference. Such interference prevention measure applied to a language may lead to a performance cost in this language if it is required immediately again.

Drawing on these observations, we propose that inhibition does not constitute an intrinsic mechanism within the language identification system for modulating language activation levels during bilingual comprehension, in line with the BIA+. Rather, the task/decision system responds to the particular demands of a given task, implementing dynamic adjustments to regulate the output from the language identification system. These adjustments give rise to a task-dependent $n - 2$ repetition effect within the context of bilingual comprehension, highlighting the adaptability of the cognitive processes involved in managing language activation and control.

Similar to Declerck and Philipp (2018), the present findings also show that the $n - 2$ repetition effect is modulated by proficiency (i.e., the effect is not equally manifest in all three languages). The benefit was observed in both nondominant languages, English and Spanish, but not in the dominant Chinese. According to the persisting activation account mentioned in the introduction, proficiency plays a critical role in influencing activation dynamics (Philipp et al., 2007). In this framework, a less proficient language experiences stronger activation, which persists into the following trials. Consequently, one would expect a more pronounced $n - 2$ repetition benefit in the less proficient language, given the greater potential for facilitation. In contrast, the already high baseline activation in a more developed language, like L1 Chinese, is likely to result in a ceiling effect, where responses in the n trial receive negligible additional facilitation from the previous $n - 2$ trial. This aligns with the null effect in L1 Chinese but does not explain the similar magnitude of benefit in L2 English and L3 Spanish, as one would have expected to observe the largest $n - 2$ repetition benefit in L3 Spanish – the less proficient language.

It is important to note that our participants used Spanish more frequently, as they were majoring in Spanish, and it was the primary medium of instruction in their classes. In contrast, English lessons were limited to only 2 h per week. Their answers to the language background questionnaire confirmed that their use of Spanish was significantly more frequent than English. This disparity in language use may have lessened the differences in activation levels, resulting in similar levels in both L2 English and L3 Spanish. This interpretation underscores the complex interplay between proficiency and use and the resulting dynamics within the bilingual lexicon (Chaouch-Orozco et al., 2021, 2024) and the need for future studies to disentangle their independent contributions in carefully designed experimental setups.

Moreover, a speed–accuracy trade-off was observed in English. The repetition of a language two trials back resulted in faster RTs but more errors in English. Such speed–accuracy trade-off effects have been documented in previous research, often linked to participants adjusting their response criteria, leading to faster yet more error-prone responses(Philipp et al., 2007). In this study, it appears participants may have similarly lowered their response criteria for English. However, it is not quite clear why this adjustment was exclusively done to English.

Lastly, it should also be noted that a color cue was used for each language in the present experiment design. Previous research showed that there were cue-repetition benefits independent of language-repetition effects (Philipp & Koch, 2009), so there might be a possibility of priming at the level of visual cue encoding contributing to the observed language-repetition benefits. Further investigation, particularly through replication studies, is necessary to elucidate the mechanisms underlying these observations.

5. Conclusion

The present study set out to investigate whether inhibition is involved in the linguistic identification system during bilingual comprehension with the $n - 2$ repetition paradigm, investigating Chinese late trilinguals who learned their two non-native languages (L2 English and L3 Spanish) via formal school instruction. Our results indicate that participants responded faster in the $n - 2$ repetition trials (i.e., when the target language in the n and $n - 2$ trials matched), but more errors were also found in these trials, at least in L2 English. That is, we observed an $n - 2$ repetition benefit effect in response latencies but not in error rates, a novel finding in this literature. This effect also goes counter the findings in Declerck and Philipp (2018), the first study adopting this paradigm in bilingual comprehension. We propose that these different findings can be reconciled under the lens of the BIA+ framework, for which adjustments by the task/decision system would give rise to a task-dependent $n - 2$ repetition effect within the context of bilingual comprehension. Furthermore, the present $n - 2$ repetition benefit was only found in the two non-native languages but not in the dominant language. Such results underscore the complex interplay between proficiency and use and the resulting dynamics within the bilingual lexicon.

Data availability statement. The data and scripts are stored in the Open Science Framework Repository at https://osf.io/3qmy4/.

Funding statement. The work was funded by RDF from Xi'an Jiaotong-Liverpool University (Fund No.: RDF-22-01-026) awarded to both authors.

Competing interest. The authors declare no competing interests.

Ethics statement. The studies involving human participants were reviewed and approved by the University Ethics Committee at Xi'an Jiaotong-Liverpool University (ER-HSS-0010000127020221202133703). All participants signed consent form to provide informed consent to participate in this study.

References

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, 52, 388–407. https://doi.org/10.3758/s13428-019-01237-x

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research*, 3 (2), 12–28. https://doi.org/10.21500/20112084.807

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. https://doi.org/10.3389/fpsyg.2015.01622

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

Blanco-Elorrieta, E., & Caramazza, A. (2021). A common selection mechanism at each linguistic level in bilingual and monolingual language production. *Cognition*, 213, 104625. https://doi.org/10.1016/j.cognition.2021.104625

Bobb, S., & 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 Psychology*, 25(5), 568–585. https://doi.org/10.1080/20445911.2013.792822

Brauer, M., & Curtin, J. J. (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods*, 23(3), 389–411. https://doi.org/10.1037/met0000159

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PloS one*, 5(6), e10729. https://doi.org/10.1371/journal.pone.0010729

Chaouch-Orozco, A., Alonso, J. G., Duñabeitia, J. A., & Rothman, J. (2023). The elusive impact of L2 immersion on translation priming. *Studies in Second Language Acquisition*, 45(2), 393–415. https://doi.org/10.1017/S0272263122000249

Chaouch-Orozco, A., Alonso, J. G., & Rothman, J. (2021). Individual differences in bilingual word recognition: The role of experiential factors and word frequency in cross-language lexical priming. *Applied Psycholinguistics*, 42(2), 447–474. https://doi.org/10.1017/S014271642000082X

Chaouch-Orozco, A., González Alonso, J., Duñabeitia, J. A., & Rothman, J. (2024). Are translation equivalents really equivalent? Evidence from concreteness effects in translation priming. *International Journal of Bilingualism*, 28(2), 149–162. https://doi.org/10.1177/13670069221146641

Cuetos, F., Glez-Nosti, M., Barbón, A., & Brysbaert, M. (2011). SUBTLEX-ESP: Spanish word frequencies based on film subtitles. *Psicológica*, 33(2), 133–143.

Declerck, M., & Koch, I. (2023). The concept of inhibition in bilingual control. *Psychological Review*, 130(4), 953–976. https://doi.org/10.1037/rev0000367

Declerck, M., Koch, I., Duñabeitia, J. A., Grainger, J., & Stephan, D. N. (2019). What absent switch costs and mixing costs during bilingual language comprehension can tell us about language control. *Journal of Experimental Psychology: Human Perception and Performance*, 45, 771–789. https://doi.org/10.1037/xhp0000627

Declerck, M., & Philipp, A. M. (2015). A review of control processes and their locus in language switching. *Psychonomic Bulletin & Review*, 22, 1630–1645. https://doi.org/10.3758/s13423-015-0836-1

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. https://doi.org/10.1080/23273798.2017.1398828

Declerck, M., Thoma, A. M., Koch, I., & Philipp, A. M. (2015). Highly proficient bilinguals implement inhibition – Evidence from n-2 language repetition costs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(6), 1911–1916. https://doi.org/10.1037/xlm0000138

Dijkstra, T., & van Heuven, W. J. B. (1998). The BIA model and bilingual word recognition. In J. Grainger & A. M. Jacobs (Eds.), *Localist connectionist approaches to human cognition* (pp. 189–226). Lawrence Erlbaum.

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, 5, 175–197. https://doi.org/10.1017/S1366728902003012

Dijkstra, T., Van Jaarsveld, H., & Ten Brinke, S. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language and Cognition*, 1, 51–66. https://doi.org/10.1017/S1366728998000121

Dijkstra, T., Wahl, A., Buytenhuijs, F., Halem, N. V., Al-Jibouri, Z., De Korte, M., & Rekke, S. (2019). Multilink: A computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition*, 22(4), 657–679. https://doi.org/10.1017/s1366728918000287

Duñabeitia, J. A., Crepaldi, D., Meyer, A. S., New, B., Pliatsikas, C., Smolka, E., & Brysbaert, M. (2018). MultiPic: A standardized set of 750 drawings with norms for six European languages. *Quarterly Journal of Experimental Psychology*, 71, 808–816. https://doi.org/10.1080/17470218.2017.1310261

Gade, M., Declerck, M., Philipp, A. M., Rey-Mermet, A., & Koch, I. (2021). On the existence of asymmetrical switch costs and reversed language dominance – A meta-analysis. *Journal of Cognition*, 4, 55. http://doi.org/10.5334/joc.186

Gollan, T. H., Weissberger, G. H., Runnqvist, E., Montoya, R. I., & Cera, C. M. (2012). Self-ratings of spoken language dominance: A Multilingual Naming Test (MINT) and preliminary norms for young and aging Spanish–English bilinguals. *Bilingualism: Language and Cognition*, 15, 594–615. https://doi.org/10.1017/S1366728911000332

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, 1, 67–81. https://doi.org/10.1017/S1366728998000133

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. https://doi.org/10.1080/20445911.2013.796377

Hirsch, P., Declerck, M., & Koch, I. (2015). Exploring the functional locus of language switching: Evidence from a PRP paradigm. *Acta Psychologica*, 161, 1–6. https://doi.org/10.1016/j.actpsy.2015.07.010

Ivanova, I., Salmon, D. P., & Gollan, T. H. (2013). The multilingual naming test in Alzheimer's disease: Clues to the origin of naming impairments. *Journal of the International Neuropsychological Society*, 19(3), 272–283. https://doi.org/10.1017/S1355617712001282

Jiao, L., Duan, X., Liu, C., & Chen, B. (2022). Comprehension-based language switching between newly learned languages: The role of individual differences. *Journal of Neurolinguistics*, 61, 101036. https://doi.org/10.1016/j.jneuroling.2021.101036

Jiao, L., Liu, C., de Bruin, A., & Chen, B. (2020). Effects of language context on executive control in unbalanced bilinguals: An ERPs study. *Psychophysiology*, 57(11), e13653. https://doi.org/10.1111/psyp.13653

Kleinman, D., & Gollan, T. H. (2018). Inhibition accumulates over time at multiple processing levels in bilingual language control. *Cognition*, 173, 115–132. https://doi.org/10.1016/j.cognition.2018.01.009

Koch, I., Declerck, M., Petersen, G., Rister, D., Scharke, W., & Philipp, A. M. (2024). Reassessing the role of language dominance in n-2 language repetition costs as a marker of inhibition in multilingual language switching. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 50(9), 1516–1528. https://doi.org/10.1037/xlm0001333

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. https://doi.org/10.3758/PBR.17.1.1

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. https://doi.org/10.1016/j.actpsy.2008.02.001

Kroll, J. F., Bobb, S. C., & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and Cognition*, 9(2), 119–135. https://doi.org/10.1017/S1366728906002483

Kroll, J. F., Dussias, P. E., Bogulski, C. A., & Kroff, J. R. V. (2012). Juggling two languages in one mind: What bilinguals tell us about language processing and its consequences for cognition. In B. H. Ross (Ed.), *Psychology of learning and motivation* (pp. 229–262). Academic Press.

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

Li, P., Zhang, F., Yu, A., & Zhao, X. (2020). Language History Questionnaire (LHQ3): An enhanced tool for assessing multilingual experience. *Bilingualism: Language and Cognition*, 23(5), 938–944. https://doi.org/10.1017/S1366728918001153

Liu, H., & Chaouch-Orozco, A. (2024). Evaluation of the Multilingual Naming Test (MINT) as a quick and practical proxy for language proficiency. *Linguistic Approaches to Bilingualism*, 14(5), 759–773. https://doi.org/10.1075/lab.23066.liu

Macizo, P., Bajo, T., & Paolieri, D. (2012). Language switching and language competition. *Second Language Research*, 28, 131–149. https://doi.org/10.1177/0267658311434893

Mayr, U., & Keele, S. W. (2000). Changing internal constraints on action: The role of back-ward inhibition. *Journal of Experimental Psychology: General*, 129(1), 4–26. https://doi.org/10.1037//0096-3445.129.1.4

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. https://doi.org/10.1080/09541440600758812

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. https://doi.org/10.1037/a0016376

R Core Team. (2021). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. https://www.R-project.org/.

Scandola, M., & Tidoni, E. (2021, February 8). Reliability and feasibility of linear mixed models in psychology and neuroscience. https://doi.org/10.31234/osf.io/kfhgv

Schwieter, J. W., & Sunderman, G. (2008). Language switching in bilingual speech production: In search of the language-specific selection mechanism. *The Mental Lexicon*, 3, 214–238. https://doi.org/10.1075/ML.3.2.06SCH

Shen, Q., & Chen, Y. (2023). Investigating the modulation of stimulus types on language switching costs: Do semantic and repetition priming effect matter? *Frontiers in Psychology*, 14, 1090744. https://doi.org/10.3389/fpsyg.2023.1090744

Sheng, L., Lu, Y., & Gollan, T. H. (2014). Assessing language dominance in Mandarin-English bilinguals: Convergence and divergence between subjective and objective measures. *Bilingualism: Language and Cognition*, 17(02), 364–383. https://doi.org/10.1017/S1366728913000424

Timmer, K., Calabria, M., Branzi, F. M., Baus, C., & Costa, A. (2018). On the reliability of switching costs across time and domains. *Frontiers in Psychology*, 9, 1032. https://doi.org/10.3389/fpsyg.2018.01032

van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, 67(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521

Winter, B., & Grice, M. (2021). Independence and generalizability in linguistics. *Linguistics*, 59(5), 1251–1277. https://doi.org/10.1515/ling-2019-0049

Appendix A The list of stimuli used in the experiment

Chinese	Frequency (Zipf values)	English	Frequency (Zipf values)	Spanish	Frequency (Zipf values)
ji /tɕi/	4.94	Chicken /ˈtʃik.ɪn/	4.82	pollo /’poʎo/	4.67
tou fa /thou, fa/	5.07	hair /heəʳ/	5.03	pelo /’pelo/	4.97
nv ren /nɥ, ʐən/	5.63	woman /ˈwʊm.ən/	5.22	mujer /mu’xeɾ/	5.72
nan ren /nan, ʐən /	5.54	man /mæn/	5.86	hombre /’ombɾe/	6.02
yan jing /jan, tɕiŋ/	5.23	eye /aɪ/	5.13	ojo /’oxo/	4.81
nan hai /nan, xaɪ̯/	5.15	boy /bɔɪ/	5.28	niño /’niɲa/	5.43
zhu /tʂu/	4.78	pig /pɪg/	4.51	cerdo /’θeɾða/	4.60
gou /kɤʊ̯/	5.55	dog /dɒg/	4.67	perro /’pera/	5.22
shou /ʂɤʊ̯/	5.65	hand /hænd/	5.44	mano /’mano/	5.41
lian /liɛn/	5.18	face /feɪs/	5.44	cara /’kaɾa/	5.38
sen lin /sən, lin/	4.40	forest /ˈfɒr.ɪst/	4.67	bosque /’boske/	4.73
rou /ʐɤʊ̯/	4.86	meat /miːt/	4.80	carne /’kaɾne/	4.92
biao /pi̯ɑʊ̯/	4.65	watch /wɒtʃ/	5.30	reloj /re’lox/	4.71
nai lao /naɪ̯, lɑʊ̯/	4.41	cheese /tʃiːz/	4.81	queso /’keso/	4.67
shu /ʂu/	5.33	book /bʊk/	5.21	libro /’liβɾo/	5.20
niu nai /ni̯ɤʊ̯, naɪ̯/	4.42	milk /mɪlk/	4.73	leche /’leʧe/	4.68
chuang /tʂʰu̯ɑŋ/	5.29	bed /bed/	5.12	cama /’kama/	5.25
qian /tɕʰiæn/	5.94	money /ˈmʌn.i/	5.84	dinero /di’neɾo/	5.86
yi zi /i, fu/	4.55	chair /tʃeəʳ/	4.66	silla /’siʎa/	4.74
xiang zi /ɕi̯ɑŋ, tsɿ/	4.65	box /bɒks/	5.12	caja /’kaxa/	5.05
men /mən/	5.42	door /dɔːʳ/	5.26	puerta /’pweɾta/	5.52
fang zi /fɑŋ, tsɿ/	5.39	house /haʊs/	5.83	casa /’kasa/	6.14
chuang hu /tɕu̯ɑŋ, xu/	4.61	window /ˈwɪn.dəʊ/	4.84	ventana /ben’tana/	4.86
chu fang /tɕu, faŋ/	4.68	kitchen /ˈkɪtʃ.ən/	5.20	cocina /ko’θina/	4.92
huo /xu̯ɔ/	5.02	fire /faɪəʳ/	5.18	fuego /’fweγo/	5.12
qian /tɕʰiæn/	5.94	money /ˈmʌn.i/	5.84	dinero /di’neɾo/	5.86
yu mi /ɥy, mi/	4.29	corn /kɔːn/	3.94	maíz /ma’iθ/	4.17
qun zi /tɕʰyn, tsɿ/	4.54	dress /dres/	4.85	vestido /bes’tiðo/	4.91

Appendix B Maximal and convergent models for the RTs analysis, and their outcomes

Maximal Model 1: Chinese vs. English and English vs. Spanish:

invRT ~ Trial type * Language + (1 | Participant) + (1 | Participant: Trial type) + (1 | Participant: Language) + (1 | Participant: Trial type: Language)

Convergent Model 1: Chinese vs. English and English vs. Spanish:

invRT ~ Trial type * Language + (1 | Participant) + (1 | Participant: Language)

Table B1. Summary of Model 1 for the analysis of RTs, including intercept and factors and their coefficients, standard errors, *t*-values and *p*-values

	Coefficient	SE	*t*-Value	*p*-Value
Intercept	−0.91	0.03	−40.24	<0.001
Trial type	0.02	0.01	2.20	0.03
Language 1	−0.18	0.02	−10.00	<0.001
Language 2	−0.20	0.02	−10.62	<0.001
Trial type by language 1	−0.06	0.02	−2.95	<0.5
Trial type by language 2	−0.03	0.02	−1.57	0.12

Maximal Model 2: Chinese vs. Spanish and English vs. Spanish:

invRT ~ Trial type * Language + (1 | Participant) + (1 | Participant: Trial type) + (1 | Participant: Language) + (1 | Participant: Trial type: Language)

Convergent Model 2: Chinese vs. Spanish and English vs. Spanish:

invRT ~ Trial type * Language + (1 | Participant) + (1 | Participant: Language)

Table B2. Summary of Model 2 for the analysis of RTs, including intercept and factors and their coefficients, standard errors, *t*-values and *p*-values

	Coefficient	SE	*t*-Value	*p*-Value
Intercept	−0.91	0.03	−40.24	<0.001
Trial type	0.02	0.01	2.20	0.03
Language 1	−0.18	0.02	−10.00	<0.001
Language 2	−0.01	0.02	−0.62	0.54
Trial type by language 1	−0.06	0.02	−2.95	<0.5
Trial type by language 2	0.03	0.02	1.39	0.17

Cite this article: Liu, H., & Chaouch-Orozco, A. (2025). Language control in auditory bilingual comprehension: uncovering novel evidence from the $n - 2$ repetition paradigm, *Language and Cognition*, 17, e15, 1–17. https://doi.org/10.1017/langcog.2024.43