

The effects of time duration and bilingualism/trilingualism on second-language production

HSIU-LING HSU* 

Department of Applied Foreign Languages, R.O.C. Naval Academy, Taiwan, Taiwan R.O.C.

Received: March 30, 2022 • Revised manuscript received: August 29, 2022 • Accepted: August 30, 2022

Published online: November 8, 2022

© 2022 Akadémiai Kiadó, Budapest



ABSTRACT

This investigation explored the effects of time duration and bilingualism/trilingualism on speakers' language production. A word-naming task was conducted under three conditions—700 ms, 1,000 ms, and unlimited time. The results showed that the participants incurred fewer errors and successfully corrected errors at 1,000 ms and unlimited time; the bilingual/trilingual advantage was identified in error self-repairs at 1,000 ms; and trilinguals were more strategic in correcting errors than monolinguals and bilinguals. This suggests that unlimited time did not ensure higher accuracy in lexical production and efficient error correction, and that 1,000 ms was the optimal timeframe for processing single monosyllabic Chinese characters.

KEYWORDS

bilingualism, trilingualism, time-driven effect, delay effect, self-repair

1. INTRODUCTION

Over the past few decades, research on the relationship between inhibitory control and bilingual speech production has attracted increasing attention (Abutalebi & Green 2007; Gollan & Acenas 2004; Hsu 2014; Kroll & Gollan 2014; Linck et al. 2012; Martin & Nozari 2021). Inhibitory control is assigned a critical position in successful language production, particularly for bilingual and multilingual language production (Jiao et al. 2020; Kroll et al. 2015). Efficient, enhanced inhibitory control results in a better performance in language production (Green 1998;

* Corresponding author. E-mail: bbshally@gmail.com

Kroll & Gollan 2014; Linck et al. 2012). It is therefore important and necessary to have a better understanding of what constraints affect inhibitory control in regulating bilingual language production. The operation of inhibitory control is closely related to a number of cognitive resources (Linck et al. 2012). Given that these resources can be substantiated by the duration of time, it is important to understand how time duration influences the operation of inhibitory control. Previous studies have focused largely on the relationship between language production and the bilingual disadvantage. The literature thus far has not explicitly bridged the relationship between the bilingual advantage in cognitive control and language production, nor has it provided a direct relationship between inhibitory control and the duration of time or attempted to examine how time-driven inhibitory control shapes bilingual online speech production (except for Hsu 2022).

A large body of prior research has reported that bilinguals perform more poorly than monolinguals on language production tasks, such as picture-naming and verbal fluency tasks (e.g., Gollan et al. 2008; Kroll & Gollan 2014; Martin & Nozari 2021; Pyers et al. 2009). Contrary to those earlier studies, Hsu's (2014, 2022) studies revealed the presence of the bilingual/trilingual advantage in language production. Hsu (2014) found that the bilinguals and trilinguals outperformed the monolinguals, and the trilinguals surpassed the bilinguals, in a read-aloud task with 'no time limitation.' Similarly, Hsu (2022) further conducted a read-aloud task with two time durations—700 ms and 1 s—to examine the differences between monolinguals, bilinguals, and trilinguals and concluded that when given more time to make a response, the bilinguals and trilinguals had better performances compared with the monolinguals. However, interestingly, Hsu's (2022) findings on the verbal tasks were not completely compatible with those of Martin-Rhee & Bialystok's (2008) findings on non-verbal tasks. The latter observed a delay effect in a Simon task: having more time to make a response, namely, a delay in response, resulted in a better performance, whereas a delay in time led to the absence of the bilingual advantage. That is, having more time to respond masked or weakened the bilingual advantage in cognitive control, leading to no differences in reaction times between the monolinguals and the bilinguals. Such inconsistent findings are relatively interesting because they show the importance of time duration when making a response in a task. These previous findings raised the following questions: (a) what will the outcome be after comparing the performances of monolingual, bilingual, and trilingual speakers in a read-aloud task under three time conditions—700 ms, 1,000 ms, and unlimited time; and (b) as argued in Hsu's (2022) research, will having more time to respond result in better performances?

Furthermore, past literature has reported that trilingual speakers are equipped with better language control than monolinguals and bilinguals due to their experience in acquiring three languages (for a detailed review, see Hsu 2014, 359–360). For example, Hoffmann (2000, 90–91) pointed out that because trilinguals have experience in three different languages, they have an "enhanced awareness of the analysis and control components of processing" relative to monolinguals and bilinguals. Andreou (2007, 11) proposed that "...additional demands are placed on their [trilingual] control abilities, in comparison with the demands placed on bilingual or monolingual children," and this gives trilinguals an advantage over bilinguals and monolinguals. Hsu (2014) found that the trilinguals outperformed the bilinguals in lexical accuracy and error correction in a preprogrammed read-aloud task. Therefore, based on these earlier research findings, in the current study it was expected that the trilinguals' performance would surpass that of the bilinguals in the verbal task.



Regarding measurement in the present study, it was important to compare the number of self-repair attempts in the three different timeframes to examine how the time-driven effect and the inhibitory mechanism influenced lexical production. Efficient inhibition of interference from the non-target language ensures more fluent language production as measured by self-repair attempts (Hsu 2014, 2017, 2022; Linck et al. 2012). The number of self-repair attempts provides an opportunity to ‘see’ the efficacy of the inhibitory control mechanism in language production, and a lower number indicates enhanced inhibitory control. Hence, the present study adopted self-repair analysis to provide specific insight into how monolinguals, bilinguals, and trilinguals operate inhibitory control and deploy cognitive resources to process lexical items within different timeframes.

Based on past literature (see the Literature Review section for a detailed review of previous relevant studies), the five research questions in this study are as follows. First, will the time-driven effect affect the speakers’ accuracy in lexical production in the word-naming task? Second, will the time-driven effect influence the speakers’ error self-repairs in the word-naming task? Third, how many attempts will the speakers require to repair each error made in the word-naming task when given 700 ms, 1,000 ms, and unlimited time? Fourth, under which specific time condition(s) will the bilingual/trilingual advantage in inhibitory control be revealed in the word-naming task? In other words, will a longer time, namely, the delay effect, mitigate the bilinguals’ and trilinguals’ disadvantage in lexical production and therefore reduce group differences between the monolinguals, bilinguals, and trilinguals? Finally, will the bilinguals and trilinguals perform differently in the word-naming task when the time given to make a response is different? To answer these research questions, five hypotheses were proposed:

Hypothesis 1: The longer the time provided to name a word is, the fewer errors the speakers will make, because having enough preparation time for speech planning and the slow buildup of inhibition will reduce speech errors.

Hypothesis 2: The longer the time provided to name a word is, the more errors the speakers will correct successfully.

Hypothesis 3: When a longer time is provided to name a word, the speakers will make fewer attempts to correct an error, leading to a lower number of self-repair attempts, because having more time will allow the speakers to preplan how to covertly repair the errors made in their inner speech, leading to less effort in repairing each error.

Hypothesis 4: Due to the delay effect and bilinguals’/trilinguals’ enhanced inhibitory control, when a longer time is provided to name a word, the performance differences between the monolinguals and bilinguals/trilinguals will be reduced or even eliminated in the word-naming task.

Hypothesis 5: Due to trilinguals’ experience in inhibiting interference from the two non-target languages, the trilinguals will outperform the bilinguals in suppressing irrelevant information, thereby possibly leading to better a performance in error correction.

To address the research questions, the present study designed a word-naming task and recruited Mandarin-speaking monolinguals, L1 Minnan-L2 Mandarin bilinguals, and L1 Hakka-L2 Mandarin-L3 Minnan trilinguals.¹ Bilingualism and trilingualism/multilingualism are defined as sequential rather than simultaneous; that is, two or three languages are not acquired

¹Mandarin and Minnan are the two major languages in Taiwan; Hakka, other Chinese dialects, and Formosan languages (such as SaySiyat and Bunun) are minor languages.



at the same age. The previous studies referred to here also include the same kind of bilingualism and multilingualism in the current study (e.g., Gollan & Acenas 2004; Linck et al. 2012). The results of this study provide evidence of direct mapping between the time-driven inhibitory control effect and lexical error-correction ability and the time conditions under which the inhibitory control mechanism facilitates bilingual/trilingual lexical production.

2. LITERATURE REVIEW

2.1. Connection between cognitive advantage and language production

In explaining the emergence of the bilingual/trilingual advantage in a verbal task, it is important to connect the advantage with cognitive control in language. Prior literature has reported that the bilingual advantage in cognitive control emerges in perceptual tasks, such as Simon tasks and flanker tasks (Bialystok & Martin 2004; Bialystok et al. 2004; Costa et al. 2006, 2008; Martin-Rhee & Bialystok 2008), rather than in production tasks, such as picture-naming tasks, language-switching tasks, and verbal fluency tasks (Gollan et al. 2007; Oller & Eilers 2002; Ransdell & Fischler 1989). In other words, bilinguals perform more poorly than monolinguals on language production tasks (Gollan & Acenas 2004; Gollan & Brown 2006; Gollan et al. 2002, 2008; Kroll & Gollan 2014; Martin & Nozari 2021; Oller & Eilers 2002; Pyers et al. 2009; Rosselli et al. 2000). For example, in a picture-naming task, bilinguals were slower in naming the pictures of objects (Gollan et al. 2008), incurred more tip-of-the-tongue states (Gollan & Acenas 2004; Pyers et al. 2009), and produced more errors than monolinguals did (Roberts et al. 2002). The *bilingual disadvantage* is assumed to originate from two possible main sources—decreased frequency of word use in each language relative to monolinguals and competition for selection between two languages or interference from the non-target language (for a detailed review, Kroll & Gollan 2014; Michael & Gollan 2005).

Interestingly, different from the findings of the earlier studies mentioned above, Hsu's (2014, 2017, 2022) studies identified the bilingual/trilingual advantage in read-aloud tasks in which the bilinguals and trilinguals outperformed the monolinguals under certain conditions. However, Hsu (2014, 2017, 2022) failed to explicitly state the connection between cognitive advantage and language production. From this perspective, it could be argued that bilingual language production is inherently more costly and therefore more cognitive resources are required to achieve successful production, although it is still not as efficient as monolingual language production because bilinguals' languages are jointly activated and compete for selection. Therefore, the present study attempted to fill the gap in the explicit relationship between cognitive advantage and language production.

First, it is important to distinguish two theoretically different types of inhibitory control—domain-general and domain-specific. Domain-general inhibitory control is usually utilized to refer to the control involved in perceptual tasks, such as Simon tasks and flanker tasks, and domain-specific inhibitory control is involved in language production tasks, such as picture-naming and word-naming tasks. 'Domain-specific inhibitory control' also refers to 'language-oriented inhibitory control' (henceforth 'language control') (Linck et al. 2012),² revealing that

²The two terms 'language-oriented inhibitory control' and 'domain-specific inhibitory control' were used in Linck et al.'s (2012, 655) study.



inhibitory control plays a role in language production as well. More importantly, [Linck et al.'s \(2012\)](#) study related language control to domain-general inhibitory control in a picture-naming task and found that better inhibitory control facilitated a decrease in the task switch cost (i.e., when switching from the less dominant L2/L3 to the more dominant L1 or from L1 to L2/L3). [Linck et al. \(2012, 651\)](#) emphasized that in a language-switching task, inhibitory control is the most important, and “a domain-general inhibitory control mechanism supports language switching.” Similarly, [Festman et al. \(2010\)](#) manifested that better language control in language production was closely associated with ‘executive functions,’ which are what domain-general inhibitory control is based on. [Festman et al. \(2010\)](#) conducted a picture-naming task and a series of perceptually cognitive control tasks to measure speakers’ language control ability (i.e., the frequency of naming errors) and executive functions, respectively, and found that the two types of inhibitory control were related. The rationale for the close relationship between the two types of control was based on previous relevant research, as will be reviewed in the following.

It has been argued that enhanced language control mainly develops from cross-linguistic competition (e.g., [Jiao et al. 2020](#); [Kroll et al. 2015](#)). Although competition between languages causes a bilingual disadvantage in language production on the one hand, it conduces the enhancement of inhibitory control in bilinguals on the other hand. In other words, effective inhibitory control is required to reduce cross-linguistic interference ([Beatty-Martínez & Dussias 2017](#); [Beatty-Martínez et al. 2020](#); [Hofweber et al. 2019](#); [Hsu 2014, 2017](#); [Kroll et al. 2008](#); [Martin-Rhee & Bialystok 2008](#)). Bilinguals’ enhanced inhibitory control has been hypothesized to come from the continuing management of two languages ([Martin-Rhee & Bialystok 2008](#)). Specifically, due to the parallel activation of underlying linguistic representations in the two languages, bilingual speakers must continuously suppress interference from the non-target language to successfully produce the lexical items in the target language, thereby leading to enhanced inhibitory control ([Jiao et al. 2020](#); [Kroll et al. 2015](#)).

Furthermore, some research has pointed out that inhibitory control has its own place in bilingual language production (e.g., [Costa, Santesteban, & Ivanova 2006](#); [Green 1998](#); [Hsu 2014, 2017](#)). [Green \(1998\)](#) proposed the Inhibitory Control model, which assumes that inhibitory control is an essential component in restraining interference from the non-target language in bilingual language production. As reported by [Green & Abutalebi \(2013, 515\)](#), “[s]peech comprehension and production are governed by control processes.”³ In this view, it is possible that enhanced inhibitory control mitigates the widely manifested bilingual ‘disadvantage’ in certain linguistic aspects under specific conditions during language production, for example, reducing language-switching costs in a task-switching paradigm ([Hartanto & Yang 2016](#); [Linck et al. 2012](#)) and effectively self-repairing speech errors in an unpreprogrammed read-aloud task ([Hsu 2014, 2017](#)).

To the best of our knowledge, the present study is the first attempt to specify the positive effect of the other side of cross-language competition by directly linking competition to the effect of the bilingual ‘advantage’ in inhibitory control on language production. To highlight this effect, this study adopted a word-naming task along with self-repair (i.e., error correction) analysis to address numerous uncertainties, particularly, the frequency of word use and conceptual/lexical

³In their definition of eight control processes, two of them are interference suppression and selective response inhibition.



access in picture-naming, equivalent translation, verbal fluency, and verbal Stroop tasks. In the word-naming task, no mismatches between the stimuli and responses were involved, and the ‘lexical access and retrieval’ factor was washed out to a larger degree.

2.2. Relationship between inhibitory control and resource allocation

Earlier studies have demonstrated that long-term bilingual/trilingual experience increases control abilities and enhances the efficacy of cognitive resource allocation, which in turn benefits bilinguals’ language production to a certain extent. Linck et al. (2012, 660) argued that “better inhibitors may MORE RAPIDLY deploy inhibition to” competing representations and thus reduce the activation of those competitors in language-naming tasks. Linck et al. (2012, 661) claimed that naming in a weaker language requires more inhibition than naming in a more dominant language, but the amount of inhibition required for naming in a weaker language decreases “in a more efficient inhibitory control system.” Furthermore, Jiao et al. (2020, 3) pinpointed that the “[l]ife-long use of these demanding cognitive processes may lead to more efficient resource allocation and enhance more early, automatic processes to prepare the control system for potential conflict.” Jiao et al.’s (2020) study reported that compared with the monolinguals, the bilinguals allocated more resources to process conflicting information ‘early’ during cognitive processing and inhibited irrelevant information between 300 and 400 ms after stimulus onset. This implies that the bilinguals were readier to handle interference from unrelated information relative to the monolinguals.

These earlier studies noted the positive interaction between the high demand for inhibitory control and the allocation of cognitive resources. Accordingly, bilinguals’ better inhibitory control and effective resource allocation benefit their language production at certain loci of online speech production to some extent. However, to what extent does inhibitory control facilitate bilinguals’ speech production, and how many cognitive resources impact their performance? Examining the execution of error correction and the time-driven effect may illuminate these questions by comparing the performances of monolinguals, bilinguals, and trilinguals. This is one major contribution of the present study.

2.3. Cognitive resources and the time-driven effect

The operation of inhibitory control is strongly associated with cognitive resources, and how resources are deployed is substantially affected by the duration of time (Eriksen & James 1986; LaBerge et al. 1991; Linck et al. 2012). From this perspective, it was of interest to manipulate time duration to create different demands on language control, which would allow us to understand how inhibitory control exerts its influence on bilingual/trilingual language production.

Time duration has been empirically demonstrated to be a relatively critical factor in determining the outcome of experiments involving language tasks (e.g., Diamond et al. 2002; Huguet et al. 1999; McFall et al. 2009; Schremm et al. 2016; Sharma et al. 2010). For instance, McFall et al. (2009) carried out an experiment involving three time conditions in a verbal Stroop color-word task—750 ms, 1 s, and 2 s—to explore how the different times would affect the effects of social presence on the participants’ performance. They found that only after the participants were given enough time, namely, 2 s, to respond to the stimuli, the effect of social presence would reduce Stroop interference and error rates as well as improve error correction.



In other words, 2 s was long enough to operate inhibitory control to suppress interference (McFall et al. 2009).

Interestingly, the findings of past studies have highlighted two conflicting tendencies with respect to how the duration of time influences participants' performances on cognitive tasks. One finding revealed that the more time that was given, the worse the participants performed (Eriksen & James 1986; LaBerge et al. 1991). This suggests that the longer durations of stimuli presentation caused the spatial scope of attentional focus to widen, which resulted in sparse available attentional resources and less efficient processing of the stimuli. LaBerge et al. (1991) conducted three experiments to examine how the gradually narrowing attentional range of a letter target affected the participants' response times by varying the durations of stimuli presentation, for example, the durations in Experiment 1 including 600 ms, 350 ms, and 83 ms. They found that the shorter durations narrowed the subjects' extent of attentional focus and therefore debilitated interference from the flanking non-target letters.⁴ Furthermore, Eriksen & James (1986) connected the size of the attentional focus area with the density of attentional resources within the area. They carried out two experiments in which the subjects were instructed to search for one of the two target letters from eight letters displayed in a circle. Eriksen & James (1986) found that when the spatial scope of attentional focus increased, the subjects' response latencies were longer. For this result, they explained that when the size of attentional focus was enlarged, the density of processing resources was reduced within the focus, causing less efficient processing of the stimuli in the attended area (also see Balz & Hock 1997; Egeth 1977). This implies that the shorter the time was, the denser the resources available for processing stimuli were; thus, the denser the resources were, the more effectively the stimuli were processed.

Contrasting with the research mentioned above, other findings showed that more time facilitated individuals to perform better (Houghton & Tipper 1994; Hsu 2014, 2022; Sharma et al. 2010). For example, Sharma et al. (2010) conducted two blocks of Stroop color-word tasks with two different time conditions—1000-ms RSI (response-to-stimulus interval) and 32-ms RSI. They found that the longer the RSI was, the stronger inhibition was, thereby leading to a decrease in interference from distractors. Sharma et al. (2010, 53) argued that “[o]ne important feature of inhibition is that it builds relatively slowly, peaking sometime after stimulus onset”; in other words, more time given resulted in stronger inhibition, which ensured the successful processing of the target items (also see Houghton & Tipper 1994). This finding implies that because the time was long enough to operate inhibitory control, the participants processed the targets better. The rationale is that with time, the intensity of inhibition will be higher and thus more interference will be effectively suppressed, leading to participants' better performances. These two interesting but contradictory findings lead to the question of how much time is needed to make a difference in the performance of different language groups on a verbal task.

More interestingly, for the latter finding, which suggested that having more to make a response would result in a better performance, there were two somewhat inconsistent trends in

⁴For this result, LaBerge et al. (1991) posited that a short duration of the first target would induce the subjects to narrow their attentional focus on the location of the target, such that the letters located on either side of the target fell outside the attended area, and therefore it was less likely that the subjects would process information at the location of the flankers, which in turn reduced distractor interference from the second target.



the emergence of the bilingual advantage (Hsu 2014, 2022; Martin-Rhee & Bialystok 2008). Hsu (2014, 2022) reported that when there was more time given to make a response, the bilingual advantage was present, whereas Martin-Rhee & Bialystok (2008) reported that when more time was given, the bilingual advantage was absent. Hsu (2014) found that the bilinguals and trilinguals outperformed the monolinguals in a read-aloud task with unlimited time given and the exclusion of the lexical retrieval and access factor, which is assumed to be a major source of the bilingual disadvantage. Similarly, Hsu (2022) administered 700-ms and 1-s read-aloud tasks and reached the conclusion that under the condition of more time, the bilinguals and trilinguals made fewer errors and repaired errors more effectively relative to the monolinguals. Conversely, Martin-Rhee & Bialystok (2008) reported that when given more time to make a response, the monolinguals and bilinguals performed equally. They conducted a Simon task with three time conditions—immediate, short delay, and long delay—and found that the bilingual advantage emerged only in the immediate condition, not in the short (500 ms) and long (1,000 ms) delays inserted between the stimuli and responses. This suggests that the delay in time allowed all the participants to have enough time to resolve perceptual conflicts and thus reduce the positive bilingual effect, leading to equal performances between the bilinguals and the monolinguals.

It is worth noting that Hsu's (2014, 2022) studies did not use the notion of 'the delay effect' to explicate the findings, while the findings of Hsu (2014, 2022) and Martin-Rhee & Bialystok (2008) came to the same conclusion: the delay effect reduced the 'lingual disadvantage' in certain tasks, namely, it mitigated the bilingual/trilingual disadvantage in the verbal tasks and the monolingual disadvantage in the non-verbal tasks. Regarding the delay effect, both the slow buildup of inhibition and the bilingual/trilingual cognitive advantage raised the question of whether a longer time would produce a delay effect to facilitate bilinguals and trilinguals to perform better in a word-naming task by effectively inhibiting irrelevant information. Specifically, given that having a longer time to respond provides bilinguals and trilinguals with more time to inhibit interference from the non-target language(s) during lexical production, would the bilingual/trilingual disadvantage in language production decrease, which would in turn reduce the different performances between the bilinguals/trilinguals and the monolinguals?

The present study adopted the term 'delay effect' and three timeframes—700 ms, 1,000 ms, and unlimited time—to examine the time-driven effect on the performance of language groups in a word-naming task to produce results that would be comparable to those in past research. In this study, 'delay' is defined as the condition in which a single Chinese character is delayed in disappearing from the computer screen relative to another character; that is, a character remains on the screen longer than another does. The aim was to find out whether 700 ms, 1,000 ms, and unlimited time would yield a short to long delay, respectively (see the Methods section for a detailed explanation of the determination of the three time conditions).

Taken together, the delay effect, the cross-linguistic competition theory, and the bilingual/trilingual cognitive advantage provided the theoretical motivation and framework for the present study to explore the interplay between the effects of time duration and bilingualism/trilingualism on language production. It is important to specify the role of the delay effect and delineate the dynamic interplay between cognitive control and time duration for a substantial understanding of the differences between monolinguals, bilinguals, and trilinguals in lexical production and to avoid blurring the implications of bilinguals' and trilinguals' enhanced inhibitory control in lexical processing. The findings of the present research will in turn contribute to the theoretical development of language production models.



3. METHODS

The current research attempted to explore how the effects of time duration and bilingualism/trilingualism interacted to shape speakers' verbal performances in a read-aloud task by measuring error rates and self-repair attempts. Three time conditions in the read-aloud task were conducted to gain a more substantial understanding of how differently monolinguals, bilinguals, and trilinguals allocated attentional resources and operated cognitive control in processing Chinese characters in different timeframes.

The present study replicated but further amended Hsu's (2014, 2017, 2022) experimental methods, design, and procedure. For example, some confounding Chinese characters (e.g., 囊 /naŋ2/ 'a bag' and 纏 /tʂʰan2/ 'to wind around') were removed from the test items and the skipped items were viewed as an error. In this study, none of the data collected in Hsu's (2014, 2017, 2022) studies were used, and new data were collected by conducting a new experiment with the modification of the experimental design and the recruitment of a new group of participants.

3.1. Determination of the three time conditions: 700 ms, 1,000 ms, and unlimited time

In the current research, the amount of time provided to name a word were theoretically and empirically determined in accordance with Levelt's (1989) perceptual loop theory, Indefrey & Levelt's (2004) speech production model, and Hsu's (2014, 2022) studies. In Levelt's (1989) model, the parallel processing of different components, the monitoring of inner speech, the detection of erroneous outputs, and covert repairs all occur via three loops that check the output of each process involved. The first loop inspects the preverbal plan before it is formulated. The second loop (i.e., the internal loop) covertly (or pre-articulatorily) monitors inner speech or the linguistic message before it is articulated. Finally, the third loop (i.e., the external loop of monitoring, also called overt or post-articulatory monitoring) inspects the produced utterance after articulation. Levelt (1989) further assumed that the internal loop and the Articulator would simultaneously begin to operate on a phonological word sent by the Phonological Encoder. The Articulator took about 200–250 ms to unpack the phonetic plan and to initiate articulation; that is, the time between the delivery of the phonetic plan and the initiation of articulation was 200–250 ms (Klapp et al. 1973; Klapp & Erwin 1976). The recognition of a syllable or a monosyllabic word takes approximately 200 ms after word onset (Marslen-Wilson 1987; Marslen-Wilson & Tyler 1980), which includes the time taken in the process of listening to overt or external speech. Based on this recognition assumption, Levelt (1989) inferred that the time for inner speech would be faster, and the latency of recognizing inner speech was estimated to be 150–200 ms. This would therefore allow the internal loop to have 0 ms (200 ms–200 ms) to 100 ms (250 ms–150 ms) to send out an interrupt signal to the Articulator when a trouble item was detected. Levelt (1989) argued that in numerous cases, the internal loop was short enough to stop articulation before the erroneous lexical item was spoken, whereas it was possible that the speaker was too late in interrupting the troublesome item, leading to the possibility of the trouble item being stopped shortly after it was spoken so that the partially interrupted error could be repaired.

More recently, a refined version of estimated timeframes to encode spoken words was put forth by Indefrey & Levelt (2004, 108) (for a review, see Hsu 2022, 9). This model shows that the



Table 1. Estimated time durations for encoding spoken words adopted from Indefrey & Levelt (2004, 108)

Operation	Duration (ms)
Conceptual preparation (from picture onset to selecting the target concept)	175
Lemma retrieval	75
Word-form encoding:	
Phonological code retrieval	80
Syllabification	125
Phonetic encoding (up to initiation of articulation)	145
Total	600

routine of encoding a word starts at the conceptual level (or lemma level) and ends at the phonetic level. The model assumes that a speaker spends, at most, 175 ms from a picture’s onset to the selection of the target concept, 75 ms to retrieve a target lexical item, 80 ms to retrieve a phonological code, 125 ms to syllabify the retrieved forms, and 145 ms for phonetic encoding. Therefore, according to the model, a speaker needs, at most, 600 ms to produce a word after seeing an image and, at most, 425 ms (600 ms–175 ms) after seeing a printed word (cf. Table 1).

According to Indefrey & Levelt’s (2004) model, 600 ms is long enough to read a word aloud.⁵ However, in a preliminary investigation of the delay effect on language production, Hsu (2022) conducted two small-scale pilot studies with the manipulation of 600 ms, 2 s, and 3 s and found that there were no statistically significant differences between monolinguals, bilinguals, and trilinguals when a Chinese character was presented on the screen for 600 ms, 2 s, and 3 s. They were either so short or so long that all three language groups produced a relatively high percentage of errors (70%) or made far fewer errors. Consequently, Hsu (2022) lengthened the duration of the presentation of a character from 600 to 700 ms and shortened the time from 2 to 3 s to 1 s.

Based on the literature mentioned above, the present study adopted three time conditions—700 ms, 1,000 ms, and unlimited time—under which the participants would identify and name individual Chinese characters. Additionally, the use of the three time conditions produced results that were comparable to those in previous studies to a certain extent. According to prior studies, the 700-ms duration was long enough, by far, for speakers to identify a Chinese character and call it out. What is important to note is that even though 700 ms is a relatively

⁵Based on Levelt’s (1989) theory, Indefrey & Levelt’s (2004) speech model, and the WEAVER++ model, speakers produce a monosyllabic lexical item at 425 ms (600 ms–175 ms), at most, because they do not have to start the processing of the target character from the conceptual tier but directly from the lemma tier. Based on the WEAVER++ model, a lexical processing model (Levelt et al. 1999; Roelofs 2003, 2007), lexical items can be directly processed and articulated upon seeing its visual forms (e.g., the visual Chinese character 風 /fēn1/ ‘wind’) without going through its conceptual tier. Before conducting the present research, 100 target one-syllable Chinese characters were read by 10 monolingual Chinese speakers (five males and five females), and the mean utterance duration of each character was 269.80 ms (Minimum: 182 ms; Maximum: 420 ms; SD: 59 ms).



brief time, the execution of repairing errors is still possible in light of Levelt's (1989) model, in which it is assumed that the internal loop and the Articulator operate in a parallel manner. Three conditions of time available to make a response were therefore designed. In a certain sense, the delay effect was replicated in the present study by having the participants read aloud the monosyllabic characters visually presented one at a time on the screen, each of which lasted 700 ms, 1,000 ms, and unlimited time. These three timeframes yielded a short to long delay, respectively.

3.2. Participants

Seventy-three young adults were initially recruited for the experiment: 21 Mandarin monolingual speakers, 26 Minnan-Mandarin bilinguals, and 26 Hakka-Mandarin-Minnan trilinguals. All were either undergraduates or postgraduates between 18 and 24 years of age, and all were receiving their formal education in Mandarin Chinese in Taiwan.⁶ The Minnan-Mandarin bilinguals' dominant language was their L1 Minnan and for the Hakka-Mandarin-Minnan trilinguals' it was L2 Mandarin. These speakers had been exposed to a bilingual or trilingual environment in which they always spoke L1 at home and L2 or L3 in public areas or the workplace. A one-way ANOVA with one independent variable (language groups: monolinguals, bilinguals, and trilinguals) was run and indicated that the differences in ages in the language groups were not significant, $F(2, 60) = 1.042$, $P = 0.359$.

Of the 73 young adults initially recruited, 10 were excluded from the experiment because they were unable to pass or complete the language assessment tests that were carried out prior to the experiment, as presented in the following. For the language assessment tests, six credentialed and experienced language teachers, who were also the research assistants and experimenters in this study, were invited to create the two tests—'Mandarin-Reading-Aloud' and 'Listening and Speaking Minnan/Hakka'—for the bilingual and trilingual groups. Of the six teachers, two each were either Mandarin, Minnan, or Hakka native speakers teaching Chinese, Minnan, or Hakka in an elementary school in southern Taiwan. They had received the MOE Certificate of Qualification to Teach Mandarin Chinese as a Second/Foreign Language, the Hakka Language Proficiency Certificate, and the Minnan Language Proficiency Certificate, respectively. These language tests were designed by referring to the assessments of listening, speaking, and reading skills furnished by the Interagency Language Roundtable (ILR).⁷ The two language tests took the bilingual speakers approximately 40 min to complete and the trilinguals one hour or so.

In the Mandarin-Reading-Aloud test, which was conducted by two Mandarin monolinguals who created the test, the potential participants were asked to read aloud six pieces of news, about 350 words each, with different topics (i.e., social news, weather news, sports news, entertainment news, medical and health news, and technology news) twice in Mandarin. The test was created to inspect bilinguals' and trilinguals' L2 Mandarin pronunciations, namely, to evaluate the degree of their accent influenced by their L1. For example, for the Minnan native speakers, they occasionally slipped [nan] in Mandarin as [lan] since Minnan phonotactic rules do not allow two nasals to occur in the same syllable (Chung 1996). If the Minnan native speakers could not

⁶Mandarin Chinese is used as a medium of instruction in Taiwan.

⁷The related assessment data are available at <https://www.govtilr.org/>.



utter the sequences of sounds in Mandarin that were banned in Minnan, or had relatively great difficulty in pronouncing them, they were excluded from the experiment. As a result, out of the 73 potential participants, three Minnan native speakers were excluded.

After the Mandarin-Reading-Aloud test, the Listening and Speaking Minnan/Hakka test, which was designed to evaluate bilingual and trilingual speakers' listening and speaking proficiency in L1 and L3, was administered. The test included two parts: Conversation and Picture Description. The four language teachers who created the test were also invited to conduct it. In the Conversation part, the teachers and the recruited speakers made conversation on a variety of topics, ranging from their experiences in language acquisition to daily life (e.g., school life, shopping, food, and hobbies). After completing the Conversation test, the speakers took the Picture Description test in which they were shown three picture books, which consisted of 11–15 pictures each, for example, a group of kids playing in a playground or a couple sitting at the table, in which no words were printed. For the Minnan-Mandarin bilingual speakers, all the topics and pictures had to be talked about or described in Minnan (L1); for the Hakka-Mandarin-Minnan trilingual speakers, they had to use their L1 and L3 asynchronously, Hakka (L1) first and then Minnan (L2).

A 5-point Likert scale, from 1 = 'the least fluent' to 5 = 'the most fluent', was utilized to evaluate the speakers' performance on each test. In this study, 3 points denoted a level between B1 and B2, 4 points level C1, and 5 points level C2 (refer to <https://blgts.moe.edu.tw/tmt/view.php?page=questionBase>). Only speakers who had scored above 4 points on each test, except for the trilinguals' Minnan Speaking test, were invited to take part in this experiment.⁸ For the trilingual participants, if they scored higher than 3 points on the Speaking Minnan (L3) test, they were invited to participate in the read-aloud experiment because in Taiwan, it is relatively difficult to recruit L1 Hakka-L2 Mandarin young adults (i.e., native Hakka speakers) who can speak Minnan (L3) as well as Hakka (L1) and Mandarin (L2). Consequently, out of the 21 trilinguals, 16 scored 3 points and five scored 4 points. However, for those who earned 3 points on the Minnan test, they conversed smoothly, talked fluently about their life experiences and familiar topics without any difficulty, and described the pictures without many problems. They understood the key points of more abstract and complex texts as well. More importantly, their language behavioral pattern was different from that of bilinguals; that is, they were more aware of or more sensitive to the differences between the languages they acquired relative to bilinguals. From this perspective, although they earned only 3 points, they were considered 'trilingual' speakers rather than bilinguals in a certain sense. After the Listening and Speaking test, out of the 70 speakers, two from the Minnan group and five from the Hakka group were excluded. Finally, 63 young adults participated in the experiment in this study, and they were divided into three language groups: monolingual (L1 Mandarin), bilingual (L1 Minnan-L2 Mandarin), and trilingual (L1 Hakka-L2 Mandarin-L3 Minnan), 21 participants in each group. The participants' demographic characteristics and details on their language background are depicted in Table 2 below:

⁸For the Mandarin test, all of the Minnan-Mandarin bilinguals scored either 4 (23.80%) or 5 (76.20%) points, and all the trilinguals also scored 4 (9.52%) or 5 (90.48%) points. For the Minnan test, all of the Minnan-Mandarin bilinguals scored either 4 (47.62%) or 5 (52.38%) points and the trilinguals scored either 3 (76.20%) or 4 (23.80%) points. For the Hakka test, the trilinguals scored either 4 (71.43%) or 5 (28.57%) points.



Table 2. Demographic characteristics of the monolingual, bilingual, and trilingual participants

	Monolinguals (<i>n</i> = 21)	Bilinguals (<i>n</i> = 21)	Trilinguals (<i>n</i> = 21)
Male-female ratio	6M:15F	8M:13F	9M:12F
Age range in years	18.00–23.00	18.00–24.00	18.00–24.00
Mean age in years(SD)	21.14 (1.45)	21.24 (1.30)	21.57 (1.86)
L2 AOA in years(SD)		4.05 (2.29)	2.71 (2.00)
L3 AOA in years(SD)			6.86 (2.15)
Scores on Mandarin test	4.86 (0.32)	4.76 (0.44)	4.90 (0.30)
Scores on Minnan test		4.52 (0.51)	3.24 (0.44)
Scores on Hakka test			4.29 (0.46)

Note: SD = standard deviation; *n* = number of participants; M = male; F = female; L2/L3 AOA = age of L2/L3 acquisition; Mandarin test = Mandarin-Reading-Aloud test; Minnan test = Listening and Speaking Minnan test; Hakka test = Listening and Speaking Hakka test.

3.3. Stimuli and design

In this experiment, 192 Chinese characters were adopted. Based on Hsu’s (2014, 2017, 2022) test materials, the current experiment further refined the experimental materials by removing some characters after two pilot studies. Originally, there were 210 characters taken into consideration in accordance with Chinese character frequency⁹ and Hsu’s (2009, 2011, 2014, 2017, 2022) studies on speech errors. However, after two pilot studies in which 24 monolingual, bilingual, and trilingual undergraduates participated (12 in each pilot study and four in each language group), 18 out of the 210 characters were eliminated from the current research because they incurred considerably higher error rates (above 75% in each language group). The 18 characters involved more complicated glyphs (e.g., 囊 /naŋ2/ ‘a bag’, 攬 /tɕʰan1/ ‘to support with hand’, and 纏 /tɕʰan2/ ‘to wind around’), easily confounding glyphs (e.g., 兵 /pin1/ ‘arms’ versus 𠂔 /pʰan1/ ‘the word used just for its sound’), or heteronyms (e.g., 藏 pronounced as /tsaŋ4/ ‘Tibet’ or /tɕʰan2/ ‘to hide’ with different meanings). Consequently, out of the 210 characters, 192 were singled out in the present study (see Appendix).

Only monosyllabic Chinese characters were utilized as the stimuli in the read-aloud experiment. Minimal pairs or groups of target characters were designed; to be precise, the phonological structure and suprasegmentals of the characters in each pair or group were the same in every way, except for the onset sound, rhyme, or tone. Owing to such similar phonetic features in each pair or group, the possibility of yielding speech errors increased. For example, a group of four

⁹The Chinese character frequency list is available at the following two academic websites: https://language.moe.gov.tw/001/Upload/files/SITE_CONTENT/M0001/PIN/yyu7.htm?open; <http://spokentaiwanmandarin.nccu.edu.tw/character-frequency.html> (NCCU Corpus of Spoken Taiwan Mandarin).



characters—/fan1/ ‘square’, /fan2/ ‘a house’, /fan3/ ‘to visit’, and /fan4/ ‘to release’¹⁰—had the identical onset /f/ and rhyme /an/ but different tones, while the pair /k^han4/ ‘to see’ and /k^han4/ ‘to resist’ carried the same onset and tone but different coda nasals.

The present study selected two coda nasals /n, ŋ/ and four initial consonants /f, p^h, h, xw/ as the target materials. For the coda nasals, the current study only selected the combination of /a+n/ and /a+ŋ/ as the target materials and excluded the /m/-/ɱ/ and /ən/-/əŋ/ pairs since the former (/an/ vs. /aŋ/) can be distinguished from each other in Taiwan Mandarin, while the latter cannot (Hsu 2014; Li 1992). As pointed out by Li (1992), /m/ and /ɱ/ as well as /ən/ and /əŋ/ have merged in Taiwan Mandarin; that is, /m/ cannot be distinguished from /ɱ/, and vice versa, and /ən/ cannot be differentiated from /əŋ/, and vice versa. Based on Li’s (1992) study, Hsu (2014) only counted the mispronounced /an/ and /aŋ/ as speech errors and did not take /m/-/ɱ/ and /ən/-/əŋ/ into consideration when designing the test materials. Accordingly, for the coda nasals, the present research only considered the two combinations of /a+n/ and /a+ŋ/ and ruled out the four combinations of /m/, /ɱ/, /ən/, and /əŋ/ when designing the test materials.

All the target characters were presented for 700 ms, 1,000 ms, or without a time limit on the computer screen, thereby creating three blocks. In each block, every lexical item constituted a trial and was repeated twice randomly. Thus, a total of 1,152 trials were formed for this experiment. Before the experimental block proceeded, 10 characters selected from the target characters, such as 幫 /paŋ1/ ‘to help’, 房 /fan2/ ‘a house’, 晚 /wan3/ ‘night’, and 看 /k^han4/ ‘to see’, were used for practice. Each participant practiced the same set of characters. The order of the presentation of the blocks in the experiment was counterbalanced across the participants.

3.4. Procedure

The participants completed the experiment individually in a soundproof room and were paid NT\$1,500 after they finished the experiment, which took approximately two hours with three five-minute breaks. Prior to the experiment, the participants were given an informed consent form to sign after they agreed to take part in the experiment; they were also informed of their rights based on the human research ethics of National Cheng Kung University in Taiwan.¹¹

The E-prime 2.0 Professional program was run to present the experimental and practice Chinese characters, all of which were shown centrally against a white background on a 15.6-inch computer (ACERTMP852-MG) monitor. The test characters appeared on the computer screen

¹⁰The numbers 1, 2, 3, and 4 refer to Tone Number, and their phonetic tone values correspond to 55, 35, 21, and 51 in Taiwan Mandarin. Regarding Tone 3, its citation tone value is either 214 or 213. However, it tends to be pronounced with tone value 21 in Taiwan, particularly in southern Taiwan. According to the two criteria of speech errors defined in the present study, tone value 21 cannot be classified as an error because the participants in this study always pronounced third-tone syllables with tone value 21, not 214 or 213. In other words, the participants in this study never produced Tone 3 with tone values 214 or 213 throughout the experiment because they believed that pronouncing Tone 3 with tone value 21 was correct. Therefore, third-tone syllables with tone value 21 were not treated as errors in such cases because they were habitual (see the Definition of Speech Errors section for a more detailed definition).

¹¹The present study recruited participants through a bulletin board, Facebook, leaflets, and other media on which the research assistant’s e-mail address was provided. If minors were interested in this study and their legal guardian agreed, then any minors and their legal guardian would be invited to visit the lab or a place where they felt comfortable to sign the informed consent.



one at a time and disappeared after 700 ms, 1,000 ms, or after the participant made a response by pressing the left mouse button, whereby the next trial appeared.

Before the experiment, the participants had to familiarize themselves with each test character, all of which were printed on A4-size paper with phonetic notations and tones. Next, the experimenter gave instructions to the participants: “Read each Chinese character aloud as quickly and as accurately as you can; do not worry about your errors, but correct any error that you make as often as possible.” The participants wore a Shure SM10A headset with an Audio-Technica AT9901 noise-filtering microphone, which was connected to the computer with the Audio-Technica monaural power module, so that their speech could be recorded immediately into the computer through Speech Analyzer 3.0.1. After that, they practiced 10 trials to be well familiar with the experimental procedure. Finally, the experiment started.

The experiment was conducted in the following order: (1) a cross (+) for fixation appeared on the center of the computer screen for 2000 ms; (2) next, 10 training trials showed up in a row, with each character shown on the screen for 700 ms, 1,000 ms, or until the participant made a response, and the next character appeared thereafter; (3) following the practice trials, the + appeared again for 2000 ms; (4) after that, the experimental stimuli appeared and remained on the screen for 700 ms, 1,000 ms, or until the participant made a response; (5) after presenting the last test character, an asterisk (*) appeared for 1,000 ms, signaling the termination of the block. [Figure 1](#) below depicts the schematic layout of the stimuli presentation in each block in the three time conditions.

As soon as the experiment was completed, the participants were invited to answer a post-experimental questionnaire about their willingness to correct errors. They were asked four main questions: (1) Did you make any errors in the experiment?; (2) Did you try to repair the error(s) that you made?; (3) Did you ignore the error(s) that you made without attempting to correct the error(s)?; and (4) If the answer to Question (3) is ‘Yes,’ why did you not try to correct the error(s) that you made?

3.5. Data analysis

In the current study, error rates, self-repair rates, successful self-repair rates, and the number of self-repair attempts for each error were calculated and analyzed to examine performance differences between the monolingual, bilingual, and trilingual participants. First, the outputs viewed as errors and self-repairs in this research will be defined. Next, a description of how the errors and self-repairs were measured will be given (for a more detailed description of self-repairs and the classification of errors, see [Hsu 2009, 2011](#)).

3.5.1. Definition of speech errors. In this research, speech errors are defined in accordance with the definition by [Dell \(1986, 284\)](#): “A slip of the tongue can be identified as an unintended, nonhabitual deviation from a speech plan.” A wrong output was treated as a speech error according to two criteria: (1) an erroneous item was considered an error only when it was pronounced accidentally, not intentionally and normally; and (2) the speaker who produced the error corrected it (i.e., the error was caused by a performance defect such as external distractions rather than a lack of competence). The two criteria in question were utilized to see whether the speakers corrected the errors that they made. The most convincing criterion was that as soon as



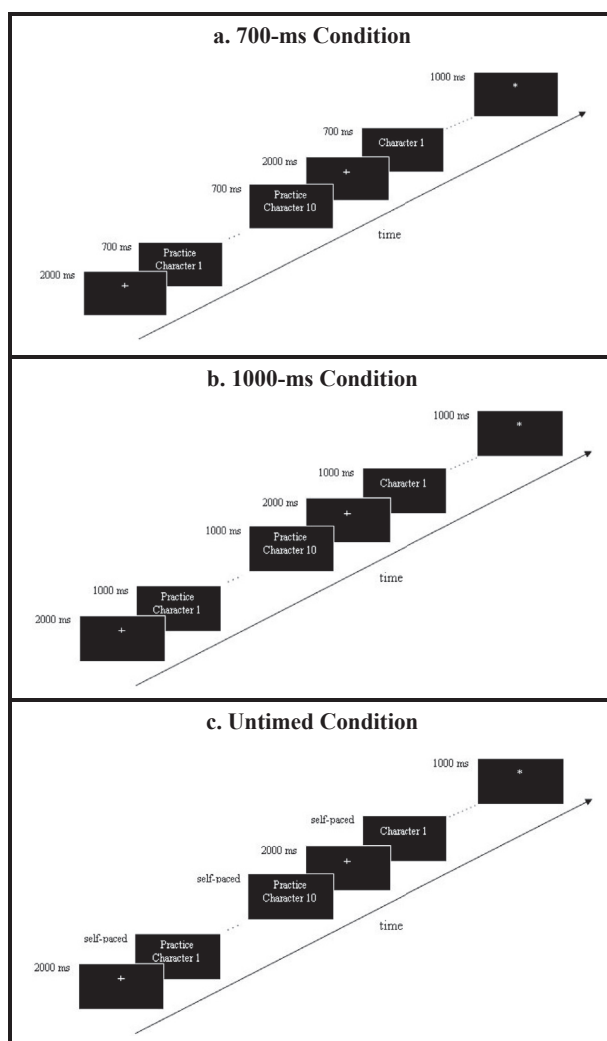


Figure 1. Schematic illustration of the three time conditions in the read-aloud task

a speaker produced an error, he or she corrected it. For the other criterion, the occurrence of the corrected forms of errors in the data suggested that the speaker knew and was capable of pronouncing the target sounds, meaning that the error was not habitual.

3.5.2. Definitions of error correction and self-repair attempts. Error correction in the current study is defined as a direct measurement of the efficacy of the inhibition of competing lexical and phonological presentations activated in the preceding trial. The efficacy of inhibiting interference is mainly disclosed by measuring the percentage of successful error corrections and



the number of self-repair attempts (Hsu 2014, 2017, 2022), which are indicative of speakers’ language control ability to suppress immediately irrelevant information that is being activated to successfully correct errors they have just made.

The term ‘the number of self-repair attempts’ used in this research refers to how many times the speakers strived to repair each error made before producing the correct form or sound(s), which denoted how much effort they required to self-repair an error. A self-repair attempt is defined as a self-induced error followed by a self-produced repair (Kormos 1999, 2000; Levelt 1983). The speakers who participated in this research did not manage to successfully correct all their self-produced errors the first time. In other words, upon making an error, the speakers struggled to correct the error more than once before uttering its correct form, as displayed in (1):

- (1) Correcting one error twice by a female monolingual speaker¹²
 Speaker: {kwan4}* {kwan4}* [kwan4]¹³
 kwan kwan stroll
 Intended Utterance: [kwan4]
 ‘to stroll’

In (1), the monolingual speaker repeated two phonologically identical errors (i.e., /kwan4/) before producing the target /kwan4/ ‘逛 (to stroll)’ to pronounce its correct output. In this case, the speaker attempted to correct an error that she had just made because she was not capable of efficiently inhibiting interference from the preceding error, thereby causing the occurrence of an identical error. Consequently, in order to yield the target form correctly, the speakers had to operate the inhibitory control mechanism effectively to prevent the prior verbal response(s) from affecting the following self-repair attempts, and then redirect their attention to the target item after amending the error that they had just made.

3.5.3. Recording and measurement of errors, self-repairs, and durations. In the current research, the percentage of errors, overall self-repairs, successful self-repairs, and the number of self-repair attempts. The mean error rates were calculated as the aggregate number of errors divided by the total number of trials. Overall, the mean self-repair rates were measured by the number of unsuccessful and successful self-repairs divided by the total number of errors; the mean successful self-repair rates were measured by the number of successful error repairs divided by the number of overall self-repairs. Moreover, the mean number of self-repair attempts was calculated as the aggregate attempts in the successful error repairs divided by the number of successful self-repair cases.

The participants’ verbal responses were saved as WAV files, and the recorded characters were analyzed using Speech Analyzer 3.0.1. For the untimed task, apart from the recording of the speakers’ verbal responses, their response durations were recorded from the first appearance of a stimulus on the computer screen to the pressing of the left mouse button. Regarding the

¹²In (1), the asterisk (*) shows the point at which a self-correction was induced, the form in braces ({} represents an error that was corrected, and the corrected form is underlined. The first line shows the mispronounced forms, followed by its English gloss, then the intended character, and, lastly, the character translation from Chinese to English.

¹³The phone [a], which is a variant of the phoneme /a/, appears before the velar [ŋ] in Mandarin.



accuracy of the responses, each target sound was first analyzed by two monolingual native speakers of Mandarin, who were experimenters in this study. The native Mandarin speakers helped with error identification to find out which characters were mistakenly pronounced and which errors were corrected; during the identification process, they manually recorded the verbal responses as correct, incorrect, unpronounced (or skipped), or self-corrected. After that, the researcher checked each error and self-repair identified by the two native Mandarin speakers. Both the incorrect and skipped items were included in the subcategory ‘Error’ and self-repair cases were further subcategorized as ‘Unsuccessful Self-correction’ and ‘Successful Self-correction.’ Finally, the number/percentage of the errors and self-repair attempts were counted/calculated.

4. RESULTS

This study used a 3 (language group: monolingual, bilingual, or trilingual) by 3 (time condition: 700 ms, 1,000 ms, or untimed) factorial design to measure the performance differences in lexical processes among the three language groups across three time conditions. Due to the non-normal distribution and heterogeneity of the data, IBM SPSS Version 20.0 was utilized to administer two nonparametric tests—the Kruskal-Wallis test (K-W test) and the Friedman test. Before running the two tests, the original continuous or scaled data were transformed into ordinal data. A Pearson correlation coefficient was further run for the data obtained from the task in the unlimited time (henceforth UT) condition, showing a quite weak linear relationship between the response latencies and the other measurements (i.e., error rates, self-repair rates, successful self-repair rates, and the number of self-repair attempts) (all $r < 0.15$), eliminating the tradeoff between speed and accuracy and between speed and attempt to repair errors as a factor in the performance of the three language groups.¹⁴ The statistical results for the error rates and error-correction data are depicted below in order. In Tables 3 through 18, N, M, B, and T stand for number, monolingual, bilingual, and trilingual, respectively; UT represents unlimited time; SD is shown in parentheses, > represents statistically significantly more than, and * indicates a P -value less than 0.05.

4.1. Error rates

Figure 2 shows the mean percentages of the words that incurred errors in the word-naming task by the three language groups across the three conditions of time given to name a word. The continuous data were transformed into ordinal data, as shown in Tables 3 through 6.

The K-W test results revealed that there were no statistically significant differences between the median error ratings across the three groups under all three time conditions (700 ms: $H(2) = 0.228$, $P = 0.892$; 1,000 ms: $H(2) = 5.503$, $P = 0.064$; UT: $H(2) = 1.554$, $P = 0.460$) (see Table 3). The Mann-Whitney test for pairwise comparisons further showed that a significant difference was found only between the monolinguals and the trilinguals in the 1000-ms condition ($z = -2.155$, $P = 0.031$), while no significant differences were observed between the

¹⁴For the unlimited time condition in the present study, there were no significant differences in response latencies among the monolinguals, bilinguals, and trilinguals, $F(2, 60) = 0.848$, $MSE = 17205.444$, $P = 0.433$.



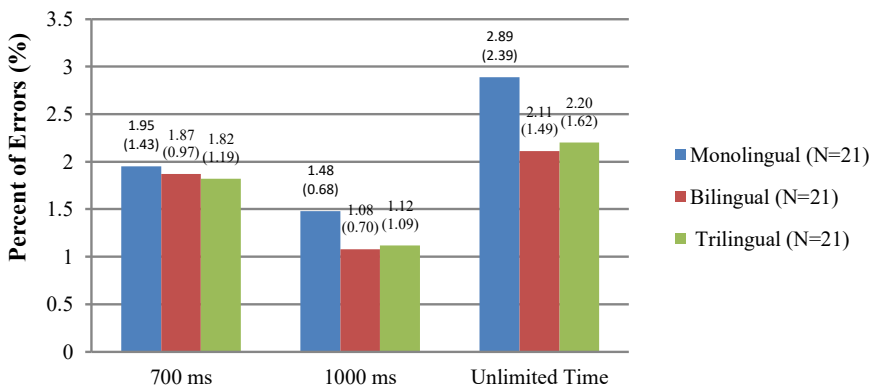


Figure 2. Mean percentages of erroneous words (SD in parentheses)

Table 3. K-W test: mean rank of error rates

Language Group	700 ms		1,000 ms		Untimed	
	N	Mean Rank (SD)	N	Mean Rank (SD)	N	Mean Rank (SD)
M	21	31.38	21	39.52	21	36.02
B	21	33.55	21	29.45	21	29.48
T	21	31.07	21	27.02	21	30.50
Sum	63	32.00 (18.32)	63	32.00 (18.31)	63	32.00 (18.31)

other language groups across the three time conditions ($P > 0.05$) (see Table 4). This suggests that the trilinguals significantly incurred fewer errors relative to the monolinguals ($T: 27.02 < M: 39.52$) when the duration for the character presentation was 1,000 ms. The statistical results suggest that the participants’ performance on lexical accuracy varied depending on different time conditions.

Table 4. Mann-Whitney test for error rates

Language Group	Statistics	700 ms	1,000 ms	Untimed
M vs. B	<i>z</i>	−0.315	−1.825	−1.310
	<i>P</i>	0.753	0.068	0.190
M vs. T	<i>z</i>	−0.013	−2.155	−0.819
	<i>P</i>	0.990	0.031*	0.413
B vs. T	<i>z</i>	−0.504	−0.479	−0.025
	<i>P</i>	0.615	0.632	0.980



Table 5. Friedman test: error rates

Language Group	N	700 ms		1,000 ms		Untimed		Chi-Squared	df	p
		Mean (SD)	Mean Rank	Mean (SD)	Mean Rank	Mean (SD)	Mean Rank			
M	21	30.36 (16.87)	1.90	20.07 (15.92)	1.57	40.57 (19.28)	2.52	9.810	2	0.007*
B	21	30.95 (17.01)	2.29	22.38 (16.36)	1.38	36.67 (18.40)	2.33	12.095	2	0.002*
T	21	35.60 (15.58)	2.26	24.19 (15.67)	1.52	36.21 (21.33)	2.21	7.253	2	0.027*

The Friedman test was conducted to examine the time-driven effect on the accuracy of lexical production, as shown in Table 5. For each language group, there was a significant difference across the three time conditions ($P < 0.05$). A follow-up Wilcoxon test showed that for all three groups, there were significant differences between the UT and 1000-ms conditions (M: $P = 0.001$; B: $P = 0.011$; T: $P = 0.038$); for the bilinguals and trilinguals, there were significant differences between the 1000-ms and 700-ms conditions (B: $P = 0.004$; T: $P = 0.005$); for the monolinguals, there were significant differences between the UT and 700-ms conditions ($P = 0.025$) (see Table 6). These results indicate that for the bilinguals and trilinguals, they incurred fewer errors in the 1000-ms condition than in the other two time conditions, and for the monolinguals, the 700-ms and 1000-ms conditions resulted in fewer errors than the UT condition, highlighting the time-driven effect on the accuracy of the participants’ lexical production.

4.2. Self-repair rates

Figure 3 shows the mean percentages of the overall self-repairs (including both successful and unsuccessful repairs) made by the three language groups across the three time conditions. The continuous data were transformed into ordinal data, as illustrated in Tables 7 through 10.

The K-W test results shown in Table 7 above indicate that there were statistically significant differences across the three language groups in the 1000-ms and UT conditions (1,000 ms: $H(2) = 8.705$, $P = 0.013$; UT: $H(2) = 10.697$, $P = 0.005$), whereas there was no significant difference in the 700-ms condition (700 ms: $H(2) = 0.633$, $P = 0.729$). As shown in Table 8,

Table 6. Wilcoxon single-rank test: error rates

Language Group	N	1,000 ms–700 ms		UT–700 ms		UT–1000 ms		Pairwise Comparisons
		z	p	z	p	z	p	
M	21	−1.512	0.130	−2.243	0.025*	−3.320	0.001*	UT > 700, 1,000
B	21	−2.868	0.004*	0.000	1.000	−2.538	0.011*	UT, 700 > 1,000
T	21	−2.801	0.005*	−0.056	0.955	−2.070	0.038*	UT, 700 > 1,000



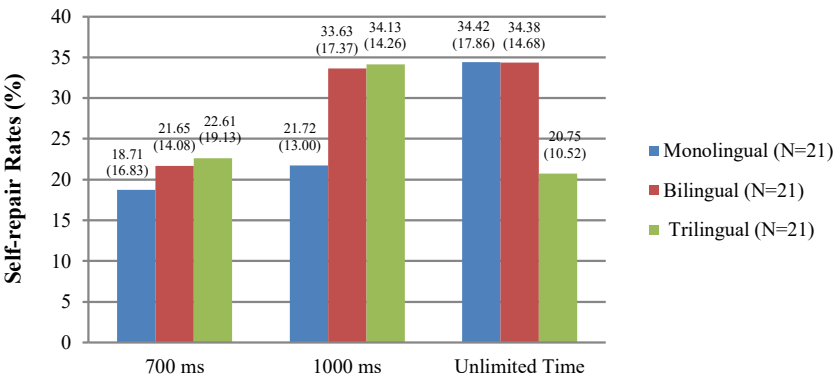


Figure 3. Mean percentages of self-repairs (SD in parentheses)

Table 7. K-W test: mean rank of self-repair rates

Language Group	700 ms		1,000 ms		Untimed	
	N	Mean Rank (SD)	N	Mean Rank (SD)	N	Mean Rank (SD)
M	21	29.76	21	22.43	21	36.38
B	21	34.26	21	35.86	21	38.24
T	21	31.98	21	37.71	21	21.38
Sum	63	32.00 (18.32)	63	32.00 (18.32)	63	32.00 (18.31)

the Mann-Whitney test for pairwise comparisons shows that there were significant differences between the monolinguals and the bilinguals ($z = -2.491$, $P = 0.013$) and between the monolinguals and the trilinguals ($z = -2.568$, $P = 0.010$) in the 1000-ms condition, and the bilinguals and trilinguals corrected more errors than the monolinguals (B: 35.86, T: 37.71 > M: 22.43). The test also showed that there were significant differences between the monolinguals and the trilinguals ($z = -2.709$, $P = 0.007$) and between the bilinguals and the trilinguals

Table 8. Mann-Whitney test for self-repair rates

Language Group	Statistics	700 ms	1,000 ms	Untimed
M vs. B	z	−0.944	−2.491	−0.390
	P	0.345	0.013*	0.696
M vs. T	z	−0.239	−2.568	−2.709
	P	0.811	0.010*	0.007*
B vs. T	z	−0.252	−0.453	−2.910
	P	0.801	0.650	0.004*



Table 9. Friedman test: self-repair rates

Language Group	N	700 ms		1,000 ms		Untimed		Chi-Squared	df	p
		Mean (SD)	Mean Rank	Mean (SD)	Mean Rank	Mean (SD)	Mean Rank			
M	21	23.69 (18.84)	1.57	29.86 (16.41)	1.90	42.45 (14.98)	2.52	9.810	2	0.007*
B	21	22.00 (18.96)	1.38	35.71 (15.52)	2.33	38.29 (16.73)	2.29	12.095	2	0.002*
T	21	27.52 (21.62)	1.71	41.40 (14.40)	2.52	27.07 (15.03)	1.76	8.667	2	0.013*

($z = -2.910$, $P = 0.004$) in the UT condition, and the monolinguals and bilinguals corrected more errors compared with the trilinguals (M: 36.38, B: 38.24 > T: 21.38). These statistical results suggest that language background took effect in the error-correction rates when the time given to make a response was 1,000 ms or unlimited.

Moreover, the Friedman test was run to examine the time-driven effect (see Table 9), showing that for each language group, the differences in self-repair rates were statistically significant across the three time conditions ($P < 0.05$). A sequence of Wilcoxon tests showed that for the monolinguals, there were significant differences between the UT and 700-ms conditions ($P = 0.004$) and between the UT and 1000-ms conditions ($P = 0.016$), and more errors were corrected in the UT condition (M: 42.45) than in the 700-ms (M: 23.69) and 1000-ms (M: 29.86) conditions. For the bilinguals, the differences between the 1000-ms and 700-ms conditions ($P = 0.029$) and between the UT and 700-ms conditions ($P = 0.003$) were significant, suggesting that more errors were corrected in the 1,000 ms (B: 35.71) and UT (B: 38.29) conditions than in the 700 ms (B: 22.00) condition. For the trilinguals, there were significant differences between the 1000-ms and 700-ms conditions ($P = 0.019$) and between the UT and 1000-ms conditions ($P = 0.004$), and more errors were corrected in the 1000-ms condition (T: 41.40) than in the 700-ms (T: 27.52) and UT (T: 27.07) conditions (see Table 10). These results suggest that the time durations for making a response affected the three language groups’ performance on error-correction rates.

Table 10. Wilcoxon single-rank test: self-repair rates

Language Group	N	1,000 ms–700 ms		UT–700 ms		UT–1,000 ms		Pairwise Comparisons
		z	P	z	P	z	P	
M	21	−0.939	0.348	−2.885	0.004*	−2.399	0.016*	UT > 700, 1,000
B	21	−2.190	0.029*	−2.990	0.003*	−0.643	0.520	1,000, UT > 700
T	21	−2.347	0.019*	−0.191	0.848	−2.870	0.004*	1,000 > 700, UT



4.3. Successful self-repair rates

Figure 4 shows the mean percentages of the successful self-repairs made by the three language groups across the three time conditions. The continuous data were transformed into ordinal data, as shown in Tables 11 through 14.

The K-W test showed that there were significant differences across the three language groups under the 1000-ms condition (1,000 ms : $H(2) = 20.367$, $P = 0.000$), whereas there were no significant differences under the 700-ms and UT conditions (700 ms : $H(2) = 1.925$, $P = 0.382$; UT : $H(2) = 4.869$, $P = 0.088$) (see Table 11). This suggests that the differences in successful self-repair rates among the monolinguals, bilinguals, and trilinguals changed at the 1000-ms time point. Furthermore, the follow-up Mann-Whitney test results shown in Table 12 indicate that there were significant differences in successful self-repair rates between the three language groups in the 1000-ms condition (M vs. B : $z = -2.565$, $P = 0.010$; M vs. T : $z = -4.143$, $P = 0.000$; B vs. T : $z = -2.879$, $P = 0.004$) as well as between the monolinguals and the trilinguals in the UT condition ($z = -2.164$, $P = 0.030$). In the 1000-ms condition, the trilinguals corrected the most errors successfully and the monolinguals repaired the fewest errors, with the bilinguals somewhere in the middle (T : 43.64 > B : 32.14 > M : 20.21). In the UT condition, the trilinguals corrected significantly more errors successfully compared with the monolinguals. These statistical results show that language background exerted its effect on the speakers' performance on successful error-correction rates only when the time given to make responses was 1,000 ms or unlimited.

Additionally, through the Friedman test, the time-driven effect was measured (see Table 13). The results showed that for each language group, there were statistically significant differences in successful self-repair rates across the three time conditions ($P < 0.05$). The Wilcoxon test indicated that for the monolinguals, there were significant differences between the UT and 700 ms ($P = 0.009$) conditions and between the UT and 1,000 ms ($P = 0.007$) conditions; the monolinguals successfully corrected more errors in the UT condition (M : 42.62) than in the 700-ms (M : 28.02) and 1000-ms (M : 25.36) conditions. For both the bilinguals and the

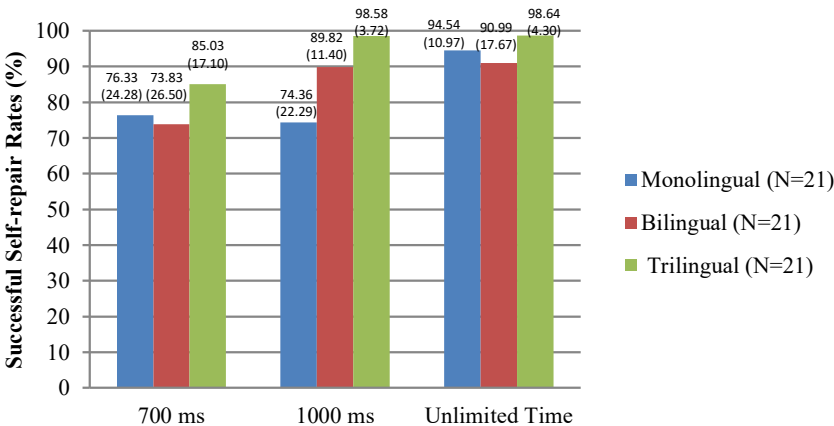


Figure 4. Mean percentages of successful self-repairs (SD in parentheses)



Table 11. K-W test: mean rank of successful self-repair rates

Language Group	700 ms		1,000 ms		Untimed	
	N	Mean Rank (SD)	N	Mean Rank (SD)	N	Mean Rank (SD)
M	21	30.67	21	20.21	21	28.36
B	21	29.00	21	32.14	21	29.98
T	21	36.33	21	43.64	21	37.67
Sum	63	32.00 (17.96)	63	32.00 (16.82)	63	32.00 (14.61)

Table 12. Mann-Whitney test for successful self-repair rates

Language Group	Statistics	700 ms	1,000 ms	Untimed
M vs. B	<i>z</i>	−0.306	−2.565	−0.288
	<i>P</i>	0.760	0.010*	0.773
M vs. T	<i>z</i>	−1.039	−4.143	−2.164
	<i>P</i>	0.299	0.000*	0.030*
B vs. T	<i>z</i>	−1.311	−2.879	−1.841
	<i>P</i>	0.190	0.004*	0.066

Table 13. Friedman test: successful self-repair rates

Language Group	N	700 ms		1,000 ms		Untimed		Chi-Squared	df	<i>p</i>
		Mean (SD)	Mean Rank	Mean (SD)	Mean Rank	Mean (SD)	Mean Rank			
M	21	28.02 (18.48)	1.83	25.36 (17.84)	1.64	42.62 (11.57)	2.52	9.844	2	0.007*
B	21	23.19 (17.23)	1.57	33.79 (15.74)	2.10	39.02 (16.07)	2.33	8.000	2	0.018*
T	21	21.60 (17.12)	1.48	36.90 (9.05)	2.24	37.50 (9.47)	2.29	16.178	2	0.000*

trilinguals, there were significant differences between the 1000-ms and 700-ms conditions (B : $P = 0.047$; T : $P = 0.003$) and between the UT and 700-ms conditions (B : $P = 0.008$; T : $P = 0.006$); the bilinguals and trilinguals repaired more errors at the 1000-ms (B : 33.79; T : 36.90) and UT (B : 39.02; T : 37.50) time points than at the 700-ms time point (B : 23.19; T : 21.60) (see [Tables 13](#) and [14](#)). These results suggest that the time durations to make a response influenced the performance of the language groups in successful error-correction rates.



Table 14. Wilcoxon single-rank test: successful self-repair rates

Language Group	N	1,000 ms–700 ms		UT-700 ms		UT-1000 ms		Pairwise Comparisons
		<i>z</i>	<i>P</i>	<i>z</i>	<i>P</i>	<i>z</i>	<i>P</i>	
M	21	−0.429	0.668	−2.597	0.009*	−2.676	0.007*	UT > 700, 1,000
B	21	−1.982	0.047*	−2.657	0.008*	−1.334	0.182	1,000, UT > 700
T	21	−2.989	0.003*	−2.739	0.006*	−0.406	0.684	1,000, UT > 700

4.4. Self-repair attempts

Figure 5 displays the mean times for repairing an error made by the three language groups across the three time conditions. The continuous data were transformed into ordinal data, as displayed in Tables 15 through 18.

The K-W test indicated that there were no statistically significant differences in self-repair attempts across the three groups under all three time conditions (700 ms : $H(2) = 0.002$, $P = 0.999$; 1,000 ms : $H(2) = 3.571$, $P = 0.168$; UT : $H(2) = 4.686$, $P = 0.096$) (see Table 15). The Mann-Whitney test further showed that there was a significant difference identified only between the monolinguals and the trilinguals under the UT condition ($z = -2.151$, $P = 0.032$) (see Table 16). In the UT condition, the monolinguals made more attempts to correct each error made compared with the trilinguals (M : 35.83 > T : 27.76). These results show that the significant differences in the attempts to repair an error were observed between the language groups, namely, the monolinguals and the trilinguals, in the UT condition.

Furthermore, the Friedman test was conducted to examine the time-driven effect (see Table 17). The results indicated that for the monolinguals, the differences in self-repair attempts for each error were statistically significant across the three time conditions ($P = 0.022$). A series of Wilcoxon tests showed that for the monolinguals, there were significant discrepancies in the attempts to repair each error among the three time conditions, whereas for the bilinguals and

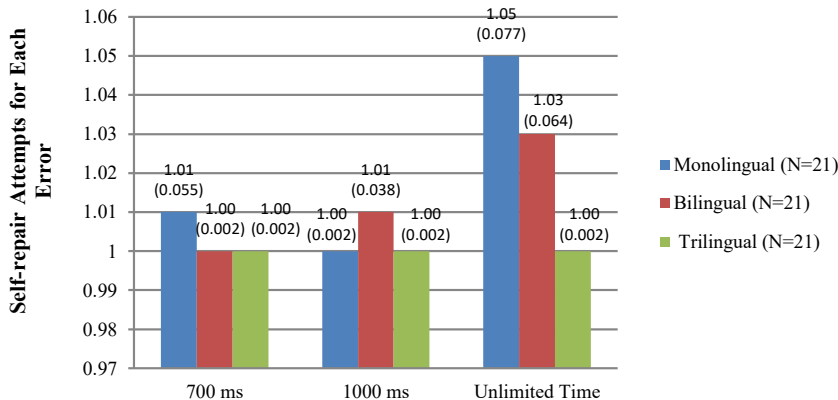


Figure 5. Mean number of self-repair attempts for each error (SD in parentheses)



Table 15. K-W test: mean rank of the number of self-repair attempts for each error

Language Group	700 ms		1,000 ms		Untimed	
	N	Mean Rank (SD)	N	Mean Rank (SD)	N	Mean Rank (SD)
M	21	32.05	21	30.43	21	35.83
B	21	31.98	21	35.14	21	32.40
T	21	31.98	21	30.43	21	27.76
Sum	63	32.00 (6.76)	63	32.00 (9.33)	63	32.00 (12.13)

Table 16. Mann-Whitney test: self-repair attempts for each error

Language Group	Statistics	700 ms	1,000 ms	Untimed
M vs. B	<i>z</i>	−0.034	−1.476	−0.842
	<i>P</i>	0.973	0.140	0.400
M vs. T	<i>z</i>	−0.034	0.000	−2.151
	<i>P</i>	0.973	1.000	0.032*
B vs. T	<i>z</i>	0.000	−1.476	−1.498
	<i>P</i>	1.000	0.140	0.134

Table 17. Friedman test: self-repair attempts for each error

Language Group	N	700 ms		1,000 ms		Untimed		Chi-Squared	df	<i>p</i>
		Mean (SD)	Mean Rank	Mean (SD)	Mean Rank	Mean (SD)	Mean Rank			
M	21	29.67 (7.64)	1.90	29.33 (6.11)	1.86	37.00 (14.61)	2.24	7.600	2	0.022*
B	21	28.83 (6.11)	1.83	33.40 (12.52)	2.05	33.76 (13.26)	2.12	3.714	2	0.156
T	21	32 (6.87)	2.00	32 (6.87)	2.00	32 (6.87)	2.00	0.000	2	1.000

trilinguals, there were no significant differences between the three time conditions. For the monolinguals, the differences in self-repair attempts between the UT and 700-ms conditions ($P = 0.046$) and between the UT and 1000-ms conditions ($P = 0.028$) were significant. The monolinguals made more attempts to correct each error in the UT condition ($M : 37.00$) than in the 700-ms ($M : 29.67$) and the 1000-ms ($M : 29.33$) conditions, meaning that for the monolinguals, their attempts to correct errors varied across the different time conditions (see Table 18).



Table 18. Wilcoxon single-rank test: self-repair attempts for each error

Language Group	N	1,000 ms–700 ms		UT-700 ms		UT-1,000 ms		Pairwise Comparisons
		z	p	z	p	z	p	
M	21	−0.447	0.655	−1.992	0.046*	−2.201	0.028*	UT > 700, 1,000
B	21	−1.604	0.109	−1.826	0.068	−0.423	0.672	700 = 1,000 = UT
T	21	0.000	1.000	0.000	1.000	0.000	1.000	700 = 1,000 = UT

Table 19. Time effect: summary of differences between the three time conditions for each group

Language Group	Error Rates	SR Rates	Successful SR Rates	SR Attempts
M	UT > 700, 1,000	UT > 700, 1,000	UT > 700, 1,000	UT > 700, 1,000
B	UT, 700 > 1,000	1,000, UT > 700	1,000, UT > 700	700 = 1,000 = UT
T	UT, 700 > 1,000	1,000 > 700, UT	1,000, UT > 700	700 = 1,000 = UT
Delay Effect	Partial Support	Partial Support	Partial Support	Partial Support

Note: M = monolingual; B = bilingual; T = trilingual; SR = self-repair; UT = unlimited time; (>) = statistically significantly more than.

5. DISCUSSION

The goal of the present investigation was to explore how inhibitory control and attentional resources modulated bilingual/trilingual lexical production under different time conditions. To address these questions, a word-naming task with three time conditions was conducted, errors and self-repairs were measured, and the differences between monolinguals, bilinguals, and trilinguals were compared in terms of the accuracy of lexical production and the execution of error correction. The findings relevant to the time-driven effect and the bilingual/trilingual effect are outlined in Tables 19 and 20 respectively, to give a clear overall picture of the results of the present study.

Table 20. Bilingual/trilingual effect: summary of differences between the three language groups under each time condition

Measurements	700 ms	1,000 ms	UT
Error Rates	M = B = T (ns.)	M > T (T Advantage)	M = B = T (ns.)
SR Rates	M = B = T (ns.)	B > M; T > M (B/T Advantage)	M > T; B > T
Successful SR Rates	M = B = T (ns.)	B > M; T > M; T > B (B/T Advantage)	T > M (T Advantage)
SR Attempts	M = B = T (ns.)	M = B = T (ns.)	M > T (T Advantage)

Note: SR = self-repair; UT = unlimited time; M = monolingual; B = bilingual; T = trilingual; (>) = statistically significantly more than; ns. = non-significant.



5.1. Hypothesis 1

Table 19 shows that for the monolinguals, the UT condition resulted in more errors than the 700-ms and 1000-ms conditions, and for the bilinguals and trilinguals, the 700-ms and UT conditions incurred more errors relative to the 1000-ms condition, implying that having more time to make a response did not entail higher accuracy of lexical production. This finding does not fully support Hypothesis 1: the more time given to make a response, the fewer errors the speakers will make. This result partially lends support to the delay effect (the longer the time, the better the performance) and the slow buildup of inhibition effects (the longer the time, the stronger the inhibition of interference). As predicted, the bilinguals and trilinguals produced fewer errors in the 1000-ms condition than in the 700-ms condition. However, unexpectedly, all three language groups performed better at the 1000-ms time point than at the UT time point (the longest time), revealing that having more time to respond did not mean performing better. This confounding finding can be explicated from the perspective of the “optimal attention allocation principle” proposed by Eriksen & James (1986, 228). Based on their findings, they assumed that the subjects would not utilize all possibly available attentional resources to complete a task and more resources would not ensure that their performance would improve. Instead, the demands of a task would decide the optimal allocation of resources, namely, how many attentional resources would be deployed depended on task variables and processes.

The participants in the present study performed the best in the accuracy of lexical production when they were given 1,000 ms to make a response, suggesting that among the three time conditions, 1,000 ms was the optimal timeframe for the allocation of attentional resources to process individual Chinese characters; neither the shortest time (e.g., 700 ms) nor the longest time (i.e., unlimited time) was an optimal time condition to identify and name characters. Time durations decide the size of attentional focus, and the focus size affects the density of attentional resources, which in turn determines the effectiveness degree of processing resources, as claimed in Eriksen & James' (1986) and LaBerge et al.'s (1991) studies. From this perspective, compared with that in the 700-ms and unlimited time conditions, the density of processing resources in the 1000-ms timeframe was optimal for producing monosyllabic lexical items in the read-aloud task. Put another way, among the three time conditions, the participants were inclined to process individual Chinese characters most efficiently and accurately in the 1000-ms condition.

The finding of a higher percentage of errors made in the longer time condition in the current study challenges McFall et al.'s (2009) finding in which the occurrence of errors was lower when the time given to make a response was longer (i.e., 2 s) relative to when the time was shorter (i.e., 700 ms and 1 s). According to McFall et al.'s (2009) explanation, when the time to make a response is longer, speakers will have more attentional resources available for processing lexical items, possibly leading to fewer errors. In other words, when speakers have enough time and resources to well preplan the target lexeme, monitor its inner speech processes, detect covert errors, and repair the covert errors before the target form is articulated, fewer speech errors will be made (Indefrey & Levelt 2004; Levelt 1989; Levelt et al. 1999; McFall et al. 2009).

Two possible accounts for the incompatible findings of the present study and McFall et al.'s (2009) study may be related to task types and recruitment of participants. First, in terms of task types, McFall et al. (2009) adopted a Stroop color-word task (involving mismatches between stimuli and responses) in which the participants had to name the color-word's ink color rather than its meaning, while the present study utilized a word-naming task (with no mismatches



between stimuli and responses) in which the participants merely needed to identify and call out Chinese characters. In comparison with the task demand for inhibitory control in the present research, the task demand in McFall et al.'s (2009) study was more difficult for the participants to complete. Furthermore, the current research recruited participants from different language backgrounds (i.e., monolinguals, bilinguals, and trilinguals), while McFall et al. (2009) categorized the participants into the same language group. Therefore, it is unclear whether the bilingual and/or trilingual effect exerted an influence on the outcome of McFall et al.'s (2009) study.

5.2. Hypothesis 2

In the present study, the results concerning successful self-repair rates partially support Hypothesis 2: speakers will repair more errors successfully when they are provided with more time (i.e., $UT > 1,000\text{ ms} > 700\text{ ms}$). Table 19 indicates that the monolinguals successfully corrected more errors when the time given to make a response was unlimited, whereas the bilinguals and trilinguals performed better in successful error correction when the time was 1,000 ms or unlimited ($M : UT > 700\text{ ms}, 1,000\text{ ms}$; $B \text{ and } T : UT, 1,000\text{ ms} > 700\text{ ms}$), which partially supports the delay effect.

In much the same vein, the statistical results of the overall self-repairs (including successful and unsuccessful repairs) did not completely meet the expectations of Hypothesis 2. It was originally expected that when the time given to make a response was longer, upon making an error, the participants would be more willing to repair errors because they had no time pressure, leading to the highest overall self-repair rate in the UT condition, followed by that in the 1000-ms condition and the lowest rate in the 700-ms condition (i.e., $UT > 1,000\text{ ms} > 700\text{ ms}$). However, the monolinguals repaired more errors in the UT timeframe than in the other two timeframes ($M : UT > 700, 1,000$), the bilinguals corrected more errors in the 1000-ms and UT timeframe than in the 700-ms timeframe ($B : 1,000, UT > 700$), and the trilinguals corrected more errors in the 1000-ms timeframe than in the other two timeframes ($T : 1,000 > 700, UT$). Although there were no significant differences between certain time conditions, the overall self-repair rate data and the participants' responses in the post-experimental questionnaire about their willingness to correct errors revealed that among the three timeframes, the participants were more willing to correct self-induced errors when the time to make a response was at least 1,000 ms or more.

Interestingly, compared with the other two groups, the trilinguals significantly corrected more errors only in the 1000-ms condition rather than in the UT condition. Further inspection of the feedback on the willingness to repair errors suggests that the trilinguals were 'more strategic' in correcting errors relative to the monolinguals and bilinguals. They explicitly pointed out that they were afraid of failure under certain time conditions, particularly in the UT condition, under which they had made too many errors. Thus, even though they had enough time to repair errors, they strategically withdrew their effort to repair self-induced errors (Elliot et al. 2005); in so doing, they had more attentional resources left and prepared for the processing of subsequent lexical items. Accordingly, 'concern about failure' might have been the cause of the statistical results.

5.3. Hypothesis 3

Regarding error-repair attempts, the findings of this research reject Hypothesis 3, which assumed that if the participants were given more time, they would make fewer attempts to repair each



error because they were able to well preplan how to correct self-induced errors covertly before the wrong form was spoken. The monolinguals made more attempts to repair each error in the UT condition (the longest time in this study) than in the 700-ms and 1000-ms conditions, whereas the bilinguals and trilinguals performed equally across the three time conditions, as displayed in Table 19. This finding does not support the argument that the inhibition effect strengthens slowly (Houghton & Tipper 1994; Sharma et al. 2010). Based on Sharma et al.'s (2010) study, it has been manifested that the longer the time is, the stronger the inhibition effect is. Thus, it should have been predicted that the more time given to make a response, the better the participants would perform in the execution of error self-repairs, which rested heavily on inhibitory control to suppress interference from the previous wrong response and irrelevant information. However, the results in the current study do not support Sharma et al.'s (2010) finding.

One possible explanation for the poor performance in the UT condition is that instead of efficiently using sufficient time to repair errors, having unlimited time to make a response allowed the speakers to inspect other parts of the target Chinese character or to remain focused on certain parts, causing the activation of more related but non-target linguistic nodes and therefore more interference. This augmented interference led to greater difficulty in processing the target lexical items to a certain degree. For the Mandarin monolinguals, more non-target linguistic nodes relevant to the target Chinese characters were activated; for the bilinguals, the linguistic representations not only in Mandarin (L2) but also in their L1 (Minnan) were activated; and for the trilinguals, the linguistic forms related to the target in their three languages were co-activated (Bialystok & Craik 2010; Shook & Marian 2013).¹⁵ As argued by Zhang & Wang (2007, 152), “the less the amount of activation is, the more efficient the [neural] system is.” Accordingly, such increased interference from more irrelevant linguistic nodes activated hindered the speakers’ ability to repair errors efficiently and to correct errors immediately or for the first time.

5.4. Hypothesis 4: performance in error rates

The results obtained in this research do not completely support Hypothesis 4: when more time is available for naming a word, the difference between the monolinguals and bilinguals/trilinguals will be diminished or eliminated. More precisely, the bilingual and trilingual disadvantage in language production was expected to decrease in the 1000-ms condition and be removed in the UT condition.

Table 20 shows that as expected, there were no differences in error rates across the three language groups in the UT condition, suggesting that the group differences were eliminated in the UT condition, which supports Hypothesis 4 (see Table 20). However, the three groups unexpectedly produced similar error rates in the 700-ms condition, whereas their performance was significantly different in the 1000-ms condition, rejecting Hypothesis 4 (see Table 20). The similarity in error rates between the groups in the UT condition might have been attributed to the fact that the bilingual/trilingual disadvantage in language production was offset by the

¹⁵Numerous studies have manifested that bilinguals’ two acquired languages are still activated together even though they are exposed to the conversational context in which only one language is required for communication, thereby leading to ineluctable interference from linguistic representations, particularly phonological and lexical representations, in the non-target language (e.g., Bialystok 2007; Costa & Santesteban 2004; Libben & Titone 2009).



Table 21. Summary of error-rate results of the current study and those in Hsu's (2014, 2022) studies

	700 ms	1,000 ms	UT
Hsu's (2014) study			B, T > M
Hsu's (2022) study	M > B, T	B, T > M	
Current study	M = B = T	T > M; M = B; B = T	M = B = T

Note: M = monolingual; B = bilingual; T = trilingual; UT = unlimited time; (>) = outperformed; (=) = performed equally.

advantage in cognitive control. More interestingly, these results show that the bilingual/trilingual advantage emerged during lexical production when the time condition was 1,000 ms or unlimited. These findings of error rates, however, are not fully compatible with Hsu's (2014, 2022) findings. To clearly see the differences in error rates among these studies, a brief summary of their results is provided in Table 21.

In Hsu's (2014) study, the bilinguals and trilinguals produced fewer errors than the monolinguals in the UT condition, whereas in the current study, the three groups produced similar error rates. Two potential reasons for the conflicting results between the two studies might be the age of L2/L3 acquisition (AOA) and the participants' ages.

First, regarding AOA, the participants in Hsu's (2014) study were early balanced bilinguals and trilinguals, whereas not all the participants in the present research were early balanced bilinguals and trilinguals. Hsu (2014, 358) defined "an early balanced bilingual or trilingual as a speaker who has learned a second or third language before the age of eight and can switch back and forth between the two or three languages easily in different contexts...." In Hsu's (2014, 362) study, "the bilinguals and trilinguals were exposed to their L2 before the age of three and the trilinguals were exposed to their L3 before the age of eight." Compared with those who participated in Hsu's study, some bilinguals and trilinguals in the current study acquired their L2 or/and L3 later. Twelve out of the 21 bilinguals were exposed to their L2 after the age of 3 but not as late as age 7; out of the 21 trilinguals, five acquired their L2 after age 3 but not as late as age 7 and seven learned their L3 after age 8 but not as late as age 10. AOA possibly resulted in the participants' different degrees of enhanced cognitive control in the two studies. That is, the enhanced cognitive control of the bilinguals and trilinguals in this study might not have been as strong as that of the bilinguals and trilinguals in Hsu's (2014) study, and enhanced cognitive control was not strong enough to allow the bilingual/trilingual advantage in cognitive control to emerge, possibly leading to the non-significant differences between the three language groups in the current research.

Second, the participants in this study were younger than those in Hsu's (2014) study. The ages of the former ranged from 18 to 24 years old (monolingual M = 21.14, SD = 1.45; bilingual M = 21.24, SD = 1.30; trilingual M = 21.57, SD = 1.86) and those of the latter were from 22 to 30 years old.¹⁶ Given that the participants in the present research were younger than those in

¹⁶Hsu (2014) did not specify the mean age of each language group. Consequently, it is not clear whether the participants who took part in Hsu's study were older than those in the current study.



Hsu's (2014) study, the young adults' peak capacities may have obscured potential performance differences between the monolinguals, bilinguals, and trilinguals. When young individuals' cognitive capabilities are at a peak, young monolinguals have more attentional resources to operate cognitive control, thereby yielding performances at or near the optimal or maximum state relative to other age groups (Costa et al. 2008); this may have reduced the differences in the strength of cognitive control between the monolinguals and the bilinguals/trilinguals. However, a wider range of ages does not necessarily imply a higher mean age, and therefore whether the age factor determined the contradictory results of this study and Hsu's (2014) study still needs further investigation.

In Hsu's (2022) preliminary study, the bilinguals and trilinguals made more errors in the 700-ms condition but produced fewer errors in the 1000-ms condition relative to the monolinguals. Based on this finding, Hsu (2022, 46) concluded that "when more time was given, bilinguals and trilinguals incurred fewer errors relative to monolinguals." However, in the present study, the three groups had equal performances in the 700-ms condition, while only the trilinguals outperformed the monolinguals in the 1000-ms condition. There are two possible explanations for the inconsistent findings. One might have been the participants' ages, as mentioned above; the other may have been the sample size. First, in Hsu's (2022) study, the participants' ages ranged from 21 to 28 years old, while there was no detailed description of the participants' ages, nor was the age of L2/L3 acquisition furnished. Therefore, this discrepancy in the results awaits further examination in the future. Second, the sample size was small in Hsu's (2022) study: there were only 10 participants in each language group; comparatively, there were 21 in each group in the current study.

5.5. Hypothesis 4: performance in error correction

Table 20 shows that when the time given to make a response was 1,000 ms, the bilinguals and trilinguals significantly repaired more errors than the monolinguals did, and among those errors, the former corrected more errors successfully than the latter. These findings show that the bilingual/trilingual advantage in inhibitory control in error-repair execution was clearly 'seen' when 1,000 ms was given to name a word. In other words, provided that error-correction failure entails difficulty in suppressing unintended information and rapidly accessing/retrieving the intended linguistic representation, the higher rate of successful error correction implies that the bilinguals and trilinguals inhibited erroneous information in the 1000-ms condition more efficiently relative to the monolinguals. Furthermore, Table 20 indicates that all three language groups made similar self-repair attempts in both the 700-ms and the 1000-ms condition, while the trilinguals outperformed the monolinguals in the UT condition. These findings disclose that the bilingual/trilingual cognitive advantage facilitated error correction in the 1000-ms condition, and the trilinguals' inhibitory control was found to be considerably advanced in the 1000-ms and UT conditions.

The results of the self-repair attempts are partially in line with those in Hsu's (2014, 2022) studies, as illustrated in Table 22. The current study reported that there was a similarity in self-repair attempts across the language groups in the 700-ms condition, as in Hsu's (2022) study, and the trilinguals performed better than the monolinguals in the UT condition, as in Hsu's (2014) study. However, there are some differences between the results of these studies. The potential causes may be AOA, the participants' age, and the sample size, as discussed above.



Table 22. Comparison between the current study and Hsu’s (2014, 2022) studies: self-repair attempts

	700 ms	1,000 ms	UT
Hsu’s (2014) study			B, T > M
Hsu’s (2022) study	M = B = T	B, T > M	
Current study	M = B = T	M = B = T	T > M; M = B; B = T

Note: M = monolingual; B = bilingual; T = trilingual; UT = unlimited time; (>) = outperformed; (=) = performed equally.

Although the results of these studies are not all the same, these results reached convergence on the fact that the bilingual/trilingual disadvantage was not observed in any aspects concerning error correction.

Theoretically speaking, bilinguals and trilinguals encounter greater interference from target and non-target languages than monolinguals do. Stated differently, the former are confronted with a higher demand for inhibitory control to suppress cross-linguistic interference than the latter. From this perspective, bilinguals and trilinguals should perform worse in lexical production compared with monolinguals. However, as can be seen in Table 20, the bilinguals and trilinguals performed equivalently to the monolinguals in accuracy of lexical production (reflected by the error rates) in the 700-ms and UT conditions, and the trilinguals even performed better than the monolinguals in the 1000-ms condition. The trilinguals surpassed the monolinguals in attempting to repair errors in the UT condition as well. This suggests that the trilinguals allocated attentional resources to process the target stimuli and operate the inhibitory mechanism more efficiently relative to their monolingual counterparts, and such an advantage in language control was clearly ‘observed’ when the time given to name a word was 1,000 ms or unlimited.

5.6. Hypothesis 5

The results obtained in this study do not fully support Hypothesis 5: trilinguals are predicted to outperform bilinguals in inhibiting interference due to the trilingual experience of inhibitory control, thereby leading to a better performance, particularly in error correction. In the present study, the bilinguals and trilinguals performed similarly in nearly all of the measures across the three time conditions, except for the self-repair rates in the UT condition and successful self-repair rates in the 1000-ms condition (see Table 20). For the self-repair rates in the UT condition, the trilinguals were less willing than the bilinguals to correct errors in the UT condition. That is, they ‘strategically’ repaired fewer errors compared with the bilinguals when there was no time limitation for naming a word. According to their post-experimental questionnaire responses about their willingness to correct errors, the trilinguals ‘tactically’ made their speech as fluent as possible at the expense of error correction, namely, they placed speech fluency as the priority, which in turn led to the lower percentage of naming errors. It should be noted, however, that the overall self-repairs were composed of successful and unsuccessful error repairs, so the rates of the overall self-repairs were not directly relevant to the effectiveness of cognitive control, particularly inhibitory control. Therefore, this finding in question does not



imply that the bilinguals had more advanced cognitive control than the trilinguals did, but instead it suggests that the bilinguals were more willing to repair the errors that they made no matter whether they repaired the errors successfully or not.

For the successful self-repair rates, the bilinguals and trilinguals had similar overall rates in the 1000-ms condition, whereas the trilinguals had significantly higher successful rates of correcting errors than the bilinguals did. Successfully correcting errors depends heavily on inhibitory control, as claimed by [Hsu \(2014, 2017, 2022\)](#). This suggests that the trilinguals had better control over the use of inhibition in error correction than the bilinguals in the 1000-ms condition.

The present study provides a new direction for future research on the bilingual/trilingual disadvantage in language production by taking into account self-repair analysis and the speakers' language control. This work has made a critical step toward directly connecting language control to the effect of time-driven inhibitory control in lexical production to depict a more nuanced picture of the relationship between these factors.

6. CONCLUSION AND IMPLICATIONS

The findings of the current research have shed some light on how the two effects of time and bilingualism/trilingualism on inhibitory control dynamically interacted in a read-aloud task under certain time conditions. The bilingual/trilingual disadvantage in the word-naming task was reduced or eliminated, leading to a smaller difference or no difference between the bilinguals/trilinguals and the monolinguals. For the time-driven effect, the bilinguals and trilinguals performed better in the accuracy of lexical production in the 1000-ms condition than in the 700-ms and 1000-ms conditions, whereas the monolinguals had a better performance in the 700-ms and 1000-ms conditions than in the UT condition. This suggests that more time did not ensure more accurate lexical production, and for the bilinguals and trilinguals, 1,000 ms provided an optimal timeframe for the allocation and usage of available attentional resources in the word-naming task. Furthermore, the bilingual advantage was evidently observed in error correction in the 1000-ms condition and the trilingual advantage was observed in the 1000-ms and UT conditions. This suggests that compared with the monolinguals, the bilinguals and trilinguals had a more efficient inhibitory control mechanism whose effect could be seen under certain time conditions, particularly in the 1000-ms condition. Due to their enhanced inhibitory control in the efficient inhibition of interference, the bilingual/trilingual disadvantage in language production was decreased. Finally, in having more time to correct errors, it was apparent that the trilinguals were more strategic and efficient in operating the inhibitory mechanism and allocating attentional resources for the execution of repairing errors in comparison with the monolinguals.

The present study has illustrated the detailed error-correction measurements and the specific time conditions under which the bilingual/trilingual disadvantage decreased or was removed in lexical production. This will provide future research with an empirical and theoretical foundation for self-repair analysis and the determination of the amount of time to allocate in language production experiments. Furthermore, the findings of the present study are not in line with those of [McFall et al.'s \(2009\)](#) and [Hsu's \(2014, 2022\)](#) studies. Future studies should replicate the methods of [McFall et al.'s \(2009\)](#) and [Hsu's \(2014, 2022\)](#) studies, particularly the



task types, the selection of participants, and the age of the participants, to further examine and analyze the discrepancies and similarities among these studies.

Conflict of interest: The author declares that she and the institute supporting this work have no conflict of interest.

ACKNOWLEDGEMENTS

This work was supported by the Ministry of Science and Technology (MOST) under Grant MOST 107-2410-H-346-002.

REFERENCES

- Abutalebi, Jubin and David W. Green. 2007. Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics* 20. 242–275.
- Andreou, G., 2007. Phonological awareness in bilingual and trilingual school children. *The Linguistics Journal* 3(3). 8–15.
- Balz, Gunther W. and Howard S. Hock. 1997. The effect of attentional spread on spatial resolution. *Vision Research* 37. 1499–1510.
- Beatty-Martínez, Anne L. and Paola E. Dussias. 2017. Bilingual experience shapes language processing: Evidence from codeswitching. *Journal of Memory and Language* 95. 173–189.
- Beatty-Martínez, Anne L., Christian A. Navarro-Torres, Paola E. Dussias, María T. Bajo, Rosa E. G. Tamargo and Judith F. Kroll. 2020. Interactional context mediates the consequences of bilingualism for language and cognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 46(6). 1022–1047.
- Bialystok, Ellen. 2007. Cognitive effects of bilingualism: How linguistic experience leads to cognitive change. *The International Journal of Bilingual Education and Bilingualism* 10. 210–223.
- Bialystok, Ellen and Fergus I. Craik. 2010. Cognitive and linguistic processing in the bilingual mind. *Current Directions in Psychological Science* 19. 19–23.
- Bialystok, Ellen, Fergus I. Craik, Raymond Klein and Mythili Viswanathan. 2004. Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging* 19. 290–303.
- Bialystok, Ellen and Michelle M. Martin. 2004. Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental Science* 7. 325–339.
- Chung, Raung-Fu. 1996. *The segmental phonology of Southern Min in Taiwan*. Taipei, Taiwan: Crane.
- Costa, Albert, Wido La Heij and Eduardo Navarrete. 2006. The dynamics of bilingual lexical access. *Bilingualism: Language and Cognition* 9. 137–151.
- Costa, Albert, Mireia Hernández and Núria Sebastián-Gallés. 2008. Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition* 106. 59–86.
- Costa, Albert and Mikel Santesteban. 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.
- Costa, Albert, Mikel Santesteban and Iva Ivanova. 2006. How do highly proficient bilinguals control their lexicalization process? Inhibitory and language-specific selection mechanisms are both functional. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 32. 1057–1074.



- Dell, Gary S. 1986. A spreading activation theory of retrieval in sentence production. *Psychological Review* 93. 283–321.
- Diamond, Adele, Natasha Kirkham and Dima Amso. 2002. Conditions under which young children can hold two rules in mind and inhibit a prepotent response. *Developmental Psychology* 38(3). 352–362.
- Elliot, Andrew J., Mandy M. Shell, Kelly B. Henry and Markus Maier. 2005. Achievement goals, performance contingencies, and performance attainment: An experimental test. *Journal of Educational Psychology* 97. 630–640.
- Egeth, Howard. 1977. Attention and preattention. In G.H. Bower (ed.) *The psychology of learning and motivation*, Vol. 11 (pp. 277–320). New York, NY: Academic Press.
- Eriksen, Charles W. and Jame D. St. James. 1986. Visual attention within and around the field of focal attention: A zoom lens model. *Perception and Psychophysics* 40. 225–240.
- Festman, Julia, Antoni Rodríguez-Fornells and Thomas F. Münte. 2010. Individual differences in control of language interference in late bilinguals are mainly related to general executive abilities. *Behavioral and Brain Functions* 6, Article 5.
- Gollan, Tamar H. and Lori A. Acenas. 2004. What is a TOT? Cognate and translation effects on tip-of-the-tongue states in Spanish-English and Tagalog-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 30. 246–269.
- Gollan, Tamar H. and Alan S. Brown. 2006. From tip-of-the-tongue (TOT) data to theoretical implications in two steps: When more TOTs means better retrieval. *Journal of Experimental Psychology: General* 135(3). 462–483.
- Gollan, Tamar H., Christine Fennema-Notestine, Rosa I. Montoya and Terry L. Jernigan. 2007. The bilingual effect on Boston naming test performance. *Journal of the International Neuropsychological Society* 13. 197–208.
- Gollan, Tamar H., Rosa I. Montoya, Cynthia Cera and Tiffany C. Sandoval. 2008. More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language* 58. 787–814.
- Gollan, Tamar H., Rosa I. Montoya and Grace A. Werner. 2002. Semantic and letter fluency in Spanish–English bilinguals. *Neuropsychology* 16. 276–562.
- Green, David W. 1998. Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition* 1(2). 67–81.
- Green, David W. and Jubin Abutalebi. 2013. Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology* 25(5). 515–530.
- Hartanto, Andree and Hwajin Yang. 2016. Disparate bilingual experiences modulate task-switching advantages: A diffusion-model analysis of the effects of interactional context on switch costs. *Cognition* 150. 10–19.
- Hoffmann, Charlotte. 2000. Bilingual and trilingual competence: Problems of description and differentiation. *Estu dios de Sociolinguística* 1. 83–92.
- Hofweber, Julia, Theodoros Marinis and Jeanine Treffers-Daller. 2019. Effects of dense code-switching on executive control. In I.A. Sekerina, L. Spradlin and V. Valian (eds.) *Bilingualism, executive function, and beyond: Questions and insights* (pp. 161–180). Amsterdam, Philadelphia, PA: John Benjamins.
- Houghton, George and Steven P. Tipper. 1994. A model of inhibitory mechanisms in selective attention. In D. Dale and T.H. Carr (eds.) *Inhibitory processes in attention, memory, and language* (pp. 53–112). San Diego, CA: Academic Press.



- Huguet, Pascal, Marie P. Galvaing, Jean M. Monteil and Florence Dumas. 1999. Social presence effects in the Stroop task: Further evidence for an attentional view of social facilitation. *Journal of Personality and Social Psychology* 77. 1011–1023.
- Hsu, Hsiu-ling, Jr. 2009. A study of nasals in coda positions: Mandarin slips and Taiwanese L2 English. Doctoral dissertation. National Kaohsiung Normal University, Taiwan.
- Hsu, Hsiu-ling. 2011. The emergence of an unmarkedness effect in Mandarin speech errors: Nasals in a coda position. *Language and Speech* 54(3). 307–340.
- Hsu, Hsiu-ling. 2014. Effects of bilingualism and trilingualism in L2 production: Evidence from errors and self-repairs in early balanced bilingual and trilingual adults. *Journal of Psycholinguistic Research* 43(4). 357–379.
- Hsu, Hsiu-ling. 2017. The interaction between the effects of bilingualism and cross-language similarity in balanced and imbalanced bilingual adults' L2 Mandarin word-reading production. *Journal of Psycholinguistic Research* 46(4). 935–962.
- Hsu, Hsiu-ling. 2022. The effects of bilingualism and trilingualism on language production: Evidence from error-correction and switch cost (in Chinese). *Journal of Taiwanese Languages and Literature* 17(1). 1–46.
- Indefrey, Peter and Willem Levelt. 2004. The spatial and temporal signatures of word production components. *Cognition* 92. 101–144.
- Jiao, Lu, John G. Grundy, Cong Liu and Baoguo Chen. 2020. Language context modulates executive control in bilinguals: Evidence from language production. *Neuropsychologia* 142. 107441.
- Klapp, Stuart T., Wallace G. Anderson and Raymond W. Berrian. 1973. Implicit speech in reading, reconsidered. *Journal of Experimental Psychology* 100. 368–374.
- Klapp, Stuart T. and Claudia I. Erwin. 1976. Relation between programming time and duration of the response being programmed. *Journal of Experimental Psychology: Human Perception and Performance* 2. 591–598.
- Kormos, Judit. 1999. Monitoring and self-repair in L2. *Language Learning* 49(2). 303–342.
- Kormos, Judit. 2000. The role of attention in monitoring second language speech production. *Language Learning* 50(2). 343–384.
- Kroll, Judith F., Susan C. Bobb, Maya Misra and Taomei Guo. 2008. Language selection in bilingual speech: Evidence for inhibitory processes. *Acta Psychologica* 128. 416–430.
- Kroll, Judith F., Paola E. Dussias, Kinsey Bice and Lauren Perrotti. 2015. Bilingualism, mind, and brain. *Annual Review of Linguistics* 1(1). 377–394.
- Kroll, Judith F. and Tamar H. Gollan. 2014. Speech planning in two languages: What bilinguals tell us about language production. In M. Goldrick, V.S. Ferreira and M. Miozzo (eds.) *The Oxford handbook of language production* (pp. 165–181). Oxford: Oxford University Press.
- LaBerge, David, Vincent Brown, Marc Carter, David Bash and Alan Hartley. 1991. Reducing the effects of adjacent distractors by narrowing attention. *Journal of Experimental Psychology: Human Perception and Performance* 17. 65–76.
- Levelt, Willem J. M. 1983. Monitoring and self-repair in speech. *Cognition* 33. 41–103.
- Levelt, Willem J. M. 1989. *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Levelt, Willem J. M., Ardi Roelofs and Antje S. Meyer. 1999. A theory of lexical access in speech production. *The Behavioral and Brain Sciences* 22. 1–38.
- Li, Wen-C., Jr. 1992. Four mergers in the Mandarin finals of the speech of Taipei. MA thesis. Oxford University, Oxford.



- Libben, Maya R. and Debra A. Titone. 2009. Bilingual lexical access in context: Evidence from eye movements during reading. *Journal of Experimental Psychology: Learning, Memory and Cognition* 35. 381–390.
- Linck, Jared A., John W. Schwieter and Gretchen Sunderman. 2012. Inhibitory control predicts language switching performance in trilingual speech production. *Bilingualism: Language and Cognition* 15(3). 651–662.
- Marslen-Wilson, William D. 1987. Functional parallelism in spoken word-recognition. *Cognition* 25. 71–102.
- Marslen-Wilson, William and Lorraine K. Tyler. 1980. The temporal structure of spoken language understanding. *Cognition* 8. 1–71.
- Martin, Clara D. and Nazbanou Nozari. 2021. Language control in bilingual production: Insights from error rate and error type in sentence production. *Bilingualism: Language and Cognition* 24(2). 374–388.
- Martin-Rhee, Michelle and Ellen Bialystok. 2008. The development of two types of inhibitory control in monolingual and bilingual children. *Bilingualism: Language and Cognition* 11(1). 81–93.
- McFall, Sametria R., Jeremy P. Jamieson and Stephen G. Harkins. 2009. Testing the mere effort account of the evaluation–performance relationship. *Journal of Personality and Social Psychology* 96(1). 135–154.
- Michael, Erica B. and Tamar H. Gollan. 2005. Being and becoming bilingual: Individual differences and consequences for language production. In J.F. Kroll and A.M.B. de Groot (eds.) *Handbook of bilingualism: Psycholinguistic approaches* (pp. 389–407). New York, NY: Oxford University Press.
- Oller, Kimbrough and Rebecca Eilers. 2002. *Language and literacy in bilingual children*. Buffalo, NY: Multilingual Matters.
- Pyers, Jennie E., Tamar H. Gollan and Karen Emmorey. 2009. Bimodal bilinguals reveal the source of tip-of-the-tongue states. *Cognition* 112. 323–329.
- Ransdell, Sarah E. and Ira Fischler. 1989. Effects of concreteness and task context on recall of prose among bilingual and monolingual speakers. *Journal of Memory and Language* 28. 278–291.
- Roberts, Patricia M., Linda J. Garcia, Alain Desrochers and Denise Hernandez. 2002. English performance of proficient bilingual adults on the Boston Naming Test. *Aphasiology* 16. 635–645.
- Roelofs, Ardi. 2003. Goal-referenced selection of verbal action: Modeling attentional control in the Stroop task. *Psychological Review* 110. 88–125.
- Roelofs, Ardi. 2007. Attention and gaze control in picture naming, word reading, and word categorizing. *Journal of Memory and Language* 5. 232–251.
- Rosselli, Mónica, Alfredo Ardila, Katia Arujo, Viviana A. Weekes, Virginia Caracciolo, Mabel Padilla and Feggy Ostrosky-Solís. 2000. Verbal fluency and repetition skills in healthy older Spanish–English bilinguals. *Applied Neuropsychology* 7. 17–24.
- Sharma, Dinkar, Rob Booth, Rupert Brown and Pascal Huguet. 2010. Exploring the temporal dynamics of social facilitation in the Stroop task. *Psychonomic Bulletin and Review* 17(1). 52–58.
- Schremm, Andrea, Merle Horne and Mikael Roll. 2016. Time-driven effects on processing relative clauses. *Journal of Psycholinguist Research* 45(5). 1033–1044.
- Shook, Anthony and Viorica Marian. 2013. The bilingual language interaction network for comprehension of speech. *Bilingualism: Language and Cognition* 16. 304–324.
- Zhang, Yang and Yue Wang. 2007. Neural plasticity in speech acquisition and learning. *Bilingualism: Language and Cognition* 10(2). 147–160.



Appendix

Table A1. 192 target characters utilized in this study with relevant frequency statistics (for frequency, refer to <https://sites.google.com/site/tmdshare/home/99nian-xia-jiao-xue-gong-zuo-fang-zhi-biao-zu-zi-liao/zi-pin-biao>)

High Frequency (48.96%)			
No.	Target Characters	Rank of Frequency	Frequency
1	上 /ʃan4/ 'on top of'	17	8,668
2	會 /xwei4/ 'be good at'	28	7,044
3	然 /zan2/ 'but'	36	6,264
4	發 /fa1/ 'to send out'	46	5,489
5	方 /faŋ1/ 'square'	61	4,896
6	看 /kʰan4/ 'to see'	65	4,757
7	法 /fa4/ 'France'	86	3,962
8	山 /ʃan1/ 'a mountain'	121	3,245
9	回 /xwei2/ 'to go back to'	126	3,161
10	灣 /wan1/ 'a bay'	143	2,859
11	常 /tʃʰan2/ 'constant'	146	2,818
12	花 /xwa1/ 'a flower'	153	2,763
13	場 /tʃʰan3/ 'a level open space'	155	2,754
14	還 /xwan2/ 'to return to'	162	2,583
15	感 /kan3/ 'to feel'	164	2,509
16	光 /kwaŋ1/ 'light'	167	2,490
17	黑 /xei1/ 'black'	168	2,489
18	產 /tʃʰan3/ 'to produce'	185	2,306
19	觀 /kwan1/ 'to observe'	187	2,250
20	展 /tʃan3/ 'to unfold'	207	2,109
21	話 /xwa4/ 'speeches'	217	2,061
22	廣 /kwaŋ3/ 'extensive'	244	1,892
23	放 /faŋ4/ 'to put'	258	1,806
24	三 /san1/ 'three'	265	1,771
25	管 /kwan3/ 'a tube'	281	1,654
26	滿 /man3/ 'full'	293	1,576

(continued)



Table A1. Continued

High Frequency (48.96%)			
No.	Target Characters	Rank of Frequency	Frequency
27	反 /fan3/ 'opposite'	305	1,523
28	父 /fu4/ 'father'	308	1,496
29	轉 /tʃwan3/ 'to turn'	316	1,469
30	非 /fer1/ 'not'	327	1,436
31	歡 /xwan1/ 'cheerful'	329	1,434
32	南 /nan2/ 'south' 1,415	333	1,415
33	服 /fu2/ 'clothes'	335	1,413
34	單 /tan1/ 'single'	362	1,318
35	房 /fan2/ 'a house'	366	1,300
36	專 /tʃwan1/ 'to be engrossed in'	368	1,289
37	環 /xwan2/ 'a ring'	385	1,251
38	費 /fer4/ 'expenses'	401	1,211
39	王 /wan2/ 'a king'	417	1,161
40	黃 /xwan2/ 'yellow'	428	1,123
41	黨 /tan3/ 'a faction'	437	1,097
42	飛 /fer1/ 'to fly'	441	1,091
43	半 /pan4/ 'half'	451	1,069
44	府 /fu3/ 'the seat of government'	455	1,066
45	夫 /fu1/ 'a husband'	457	1,056
46	案 /an4/ 'a case'	503	951
47	萬 /wan4/ 'ten thousand'	513	934
48	晚 /wan3/ 'night'	518	925
49	班 /pan1/ 'a class'	560	841
50	裝 /tʃwan1/ 'to dress up'	566	831
51	港 /kan3/ 'a port'	575	820
52	護 /xu4/ 'to defend'	576	816
53	配 /p ^h ei4/ 'to match'	613	762
54	防 /fan2/ 'to guard against'	618	758
55	福 /fu2/ 'good fortune'	624	750

(continued)



Table A1. Continued

High Frequency (48.96%)			
No.	Target Characters	Rank of Frequency	Frequency
56	玩 /wan2/ 'to play'	631	745
57	傷 /ɕaŋ1/ 'to hurt'	636	737
58	談 /tʰan2/ 'to talk'	643	731
59	雙 /ɕwan1/ 'two'	652	718
60	否 /foʊ3/ 'to negate'	657	710
61	站 /tɕan4/ 'to stand'	660	706
62	唱 /tɕʰan4/ 'to sing'	686	666
63	康 /kʰan1/ 'health'	698	657
64	富 /fu4/ 'wealthy'	732	623
65	飯 /fan4/ 'cooked rice'	736	622
66	散 /san4/ 'to idle' 610	743	610
67	擔 /tan1/ 'to undertake'	758	596
68	幫 /paŋ1/ 'to help'	759	596
69	船 /tɕʰwan2/ 'a boat'	760	595
70	婦 /fu4/ 'a woman'	765	592
71	板 /pan3/ 'a board'	768	588
72	換 /xwan4/ 'to change'	769	588
73	穿 /tɕʰwan1/ 'to wear'	797	552
74	忙 /maŋ2/ 'busy'	800	550
75	盤 /pʰan2/ 'a tray'	807	540
76	忽 /xu1/ 'to neglect'	810	537
77	寒 /xan2/ 'cold'	825	527
78	窗 /tɕʰwan1/ 'a window'	835	518
79	呼 /xu1/ 'to exhale'	839	512
80	忘 /waŋ4/ 'to forget'	843	509
81	訪 /faŋ3/ 'to visit'	866	495
82	浪 /laŋ4/ 'billows'	873	486
83	普 /pʰu3/ 'general'	874	486
84	剛 /kaŋ1/ 'hard'	876	485

(continued)



Table A1. Continued

High Frequency (48.96%)			
No.	Target Characters	Rank of Frequency	Frequency
85	掌 /tʂaŋ3/ 'a plam'	893	473
86	賞 /ʂaŋ3/ 'to bestow'	905	468
87	皇 /xwaŋ2/ 'an emperor'	908	460
88	湖 /xu2/ 'a lake'	909	460
89	抗 /k ^h aŋ4/ 'to resist'	921	449
90	旁 /p ^h aŋ2/ 'side'	936	435
91	翻 /fan1/ 'to turn over'	965	410
92	範 /fan4/ 'a model'	971	406
93	棒 /paŋ4/ 'a bat'	976	404
Medium Frequency (44.79%)			
No.	Target Characters	Rank of Frequency	Frequency
94	患 /xwaŋ4/ 'trouble'	1,003	386
95	章 /tʂaŋ1/ 'a chapter'	1,014	381
96	哈 /xa1/ 'to breathe on or out'	1,025	372
97	床 /tɕ ^h waŋ2/ 'a bed'	1,030	369
98	煩 /fan2/ 'to be annoyed'	1,034	368
99	殘 /ts ^h an2/ 'incomplete'	1,042	363
100	漢 /xan4/ 'the Han dynasty'	1,048	358
101	航 /xaŋ2/ 'to navigate'	1,077	341
102	互 /xu4/ 'mutual'	1,081	339
103	刊 /k ^h an1/ 'to publish'	1,124	316
104	閃 /ʂan3/ 'to flash'	1,144	308
105	寬 /k ^h wan1/ 'wide'	1,153	304
106	含 /xan2/ 'to contain'	1,156	302
107	爬 /p ^h a2/ 'to crawl'	1,172	294
108	蛋 /tan3/ 'an egg'	1,187	289
109	灰 /xwer1/ 'grey'	1,195	284
110	判 /p ^h an4/ 'to judge'	1,201	282
111	幹 /kan4/ 'the main part'	1,203	282

(continued)



Table A1. Continued

Medium Frequency (44.79%)			
No.	Target Characters	Rank of Frequency	Frequency
112	乏 /fa2/ 'lack'	1,223	272
113	犯 /fan4/ 'to violate' 270	1,225	270
114	凡 /fan2/ 'ordinary'	1,229	267
115	繁 /fan2/ 'numerous' 267	1,231	267
116	朗 /laŋ3/ 'bright'	1,232	265
117	肥 /fer2/ 'fertile'	1,236	261
118	輔 /fu3/ 'to assist'	1,242	260
119	糖 /t ^h aŋ2/ 'sugar'	1,247	258
120	甘 /kan1/ 'the sweet after-taste'	1,248	257
121	網 /waŋ3/ 'a net'	1,259	254
122	滑 /xwa2/ 'slippery'	1,290	242
123	緩 /xwan3/ 'slow'	1,300	239
124	喊 /xan3/ 'to shout'	1,310	237
125	貪 /t ^h an1/ 'greedy'	1,330	232
126	膚 /fu1/ 'the skin'	1,332	232
127	虎 /xu3/ 'a tiger'	1,346	225
128	撞 /tʂwaŋ4/ 'to collide with'	1,365	218
129	荒 /xwaŋ1/ 'barren'	1,382	213
130	番 /fan1/ 'to take turns'	1,409	203
131	膽 /tan3/ 'gall'	1,466	188
132	喪 /saŋ4/ 'to lose' 181	1,493	181
133	讚 /tsan4/ 'to praise'	1,514	177
134	湯 /t ^h aŋ1/ 'soup'	1,558	168
135	胖 /p ^h aŋ4/ 'fat'	1,572	165
136	攀 /p ^h an1/ 'to climb'	1,591	161
137	昌 /tʂ ^h aŋ1/ 'prosperous'	1,602	158
138	爛 /lan4/ 'rotten'	1,629	155
139	葡 /phu2/ 'a grape'	1,688	142
140	鋪 /p ^h u1/ 'to pave'	1,694	141

(continued)



Table A1. Continued

Medium Frequency (44.79%)			
No.	Target Characters	Rank of Frequency	Frequency
141	籃 /lan2/ 'a basket'	1,718	138
142	罰 /fa2/ 'to punish'	1,779	129
143	陪 /p ^h ei2/ 'to accompany'	1,781	128
144	廊 /lan2/ 'a (usually long) covered porch or veranda'	1,782	128
145	販 /fan4/ 'to sell'	1,820	122
146	佩 /p ^h ei4/ 'to wear'	1,893	110
147	攤 /t ^h an1/ 'to spread out'	1,907	109
148	慌 /xwaŋ1/ 'in a hurry'	1,910	108
149	芳 /faŋ1/ 'fragrant'	1,932	105
150	汪 /waŋ1/ 'a pool of water'	1,961	102
151	肺 /fei4/ 'a lung'	1,977	100
152	傘 /san3/ 'an umbrella'	1,992	98
153	泛 /fan4/ 'to float'	2,002	97
154	葬 /tsaŋ4/ 'to inter'	2,049	91
155	扇 /ʂan4/ 'a fan'	2,053	90
156	晃 /xwaŋ4/ 'to hang around'	2,069	88
157	桑 /saŋ1/ 'mulberry'	2,070	88
158	罐 /kwan4/ 'a jar'	2,088	86
159	仿 /faŋ3/ 'to copy'	2,089	85
160	悔 /xwei3/ 'to regret'	2,094	85
161	蕩 /taŋ4/ 'to wander'	2,106	84
162	盼 /p ^h an4/ 'to hope'	2,138	80
163	懶 /lan3/ 'lazy'	2,144	80
164	喘 /tʂ ^h wan3/ 'to pant'	2,153	78
165	禪 /tʂ ^h an2/ 'meditation'	2,169	77
166	瀑 /p ^h u4/ 'a waterfall'	2,170	77
167	盲 /maŋ2/ 'blind'	2,235	71
168	濫 /lan4/ 'to flood'	2,250	70
169	伐 /fa2/ 'to fell'	2,269	68

(continued)



Table A1. Continued

Medium Frequency (44.79%)			
No.	Target Characters	Rank of Frequency	Frequency
170	肪 /faŋ2/ 'fat'	2,280	67
171	匪 /fei3/ 'a robber'	2,333	63
172	沸 /fei4/ 'to boil'	2,370	59
173	闖 /tɕ ^h waŋ3/ 'to charge'	2,381	59
174	狼 /lan2/ 'a wolf'	2,403	57
175	噓 /taŋ1/ 'a loud'	2,408	57
176	逛 /kwaŋ4/ 'to stroll'	2,430	55
177	坊 /faŋ1/ 'a workshop'	2,435	54
178	仗 /tɕaŋ4/ 'to battle'	2,411	56
179	杭 /xaŋ2/ 'short for Hangzhou'	2,477	51
180	綁 /paŋ3/ 'to tie'	2,487	51
Low Frequency (6.25%)			
No.	Target Characters	Rank of Frequency	Frequency
181	栓 /ɕwan1/ 'a plug'	2,533	48
182	叛 /p ^h an4/ 'to betray'	2,548	47
183	剖 /p ^h ou3/ 'to cut open'	2,550	47
184	帆 /fan2/ 'a sail'	2,613	44
185	吼 /xou3/ 'to roar'	2,637	43
186	框 /k ^h waŋ1/ 'a frame'	2,657	42
187	妃 /fei1/ 'imperial concubine'	2,688	40
188	詹 /tɕan1/ 'a family name'	2,696	40
189	龐 /p ^h aŋ2/ 'enormous'	2,753	38
190	蠻 /man2/ 'barbarous'	2,792	36
191	嗓 /saŋ3/ 'voice'	2,809	35
192	謊 /xwaŋ3/ 'a lie'	2,845	34

