BILINGUAL LANGUAGE CONTROL: BOTTOM-UP VERSUS TOP-DOWN

by
Jean P. Bodet III

A thesis submitted to the Psychology Department
College of Liberal Arts and Social Sciences
in partial fulfillment of the requirements for the degree of

Master of Arts

in Psychology

Chair of Committee: Arturo E. Hernandez, Ph.D.

Committee Member: Benjamin Tamber-Rosenau, Ph.D.

Committee Member: J. Leigh Leasure, Ph. D.

University of Houston
May 2021

ACKNOWLEDGEMENTS

Special thanks goes out to Isabelle Parra Valencia, Imaal Ahmed, Ashley Gwananji, Hana Hassan, Andrea Ortiz, and Brendan Traynor for their help with data processing for this project. I must thank also my colleagues Hannah Claussenius-Kalman, My Nguyen, Juliana Ronderos, and Dr. Kelly Vaughn for their edits, input, and help throughout the data processing and writing phases of this process. Thank you to my advisor, Dr. Arturo Hernandez, for his guidance and insight throughout this experiment and its analysis, as well as Dr. Ben Tamber-Rosenau and Dr. David Francis for their help with data analysis. Last, but certainly not least, I would like to thank my family for their constant support of my educational pursuits.

ABSTRACT

Bilinguals' language control mechanisms are well-researched and modelled. An overlooked aspect of their environment, however, is the language context. This research study asks whether context (bottom-up influence) may impact effortful language control mechanisms (top-down control). The present research tests the hypothesis that context can affect top-down language control, and that the extent of these effects are context-dependent, by manipulating auditory language distractors in a picture-naming paradigm with no cued language-switching. This was a departure from the norm of cueing a bilingual to switch languages in order to evidence a proposed language control mechanism. In short, participants heard brief, trial-length audio distractors while engaged in English-only picture-naming. The distractor languages varied per block and were Hungarian, English, Spanish, and Mixed (English and Spanish distractor trials randomly dispersed throughout the block). By removing any cued language-switching (top-down switching) and only changing the context (using auditory distractors), the results yielded can be attributed to the impact of changing bottom-up influence. In the end, the results did not support the hypotheses presented—image-naming response times did not differ significantly between contexts. However, results *may* suggest that while the immediate effects of a change in context are consistent, the lasting effects may differ from context to context. A number of measures of individual differences significantly influenced these results as well, including cognitive control abilities and English proficiency. While these results do support the view that context matters, future studies are needed to better elucidate the way in which it does or does not matter.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

Background

Ask a bilingual about their most frustrating language experiences and they will often recount accidents involving a language switch. These unplanned switches can be an anxious (or humorous) social experience, but when asked why they did this, bilinguals frequently ascribe it to one contextual issue or another ("the guy next to me was speaking my other language", or, "my other-language aunt was there"). Armed with this explanation, science must investigate. How *does* context influence language control? More empirically, this research asks two main questions: 1. Does bottom-up language information influence top-down language control in bilinguals? 2. If so, is this influence dependent on the type of bottom-up language information (context)?

These questions require explanation of *bottom-up* and *top-down*, terms most commonly found in attentional control research. Bottom-up (or *exogenous*) attention describes unconscious processing of environmental inputs while top-down (or *endogenous*) attention describes conscious control over behavioral outputs (Corbetta & Shulman, 2002; Katsuki & Constantinidis, 2014; Petersen & Posner, 2012; Posner & Petersen, 1990). This dichotomy pits internal goals against external stimuli, the two main foci of this research project.

Bottom-Up Information in Bilinguals

For the purposes of this research, bottom-up attention (or bottom-up input, or bottom-up influence) refers to the effect of a language's continued presence in the environment. In other words, the effect of a certain language context. Usually, researchers categorize language contexts through the top-down goals of the bilingual (i.e., to speak only Spanish). For example, Grosjean (1999) proposed that speech "modes" can range from monolingual to bilingual in any given situation based on the bilingual's language goals (i.e., speaking one language versus two). Building on that concept, researchers also distinguished code-switching as a particularly unique extension of the bilingual mode in which the goal is language-switching within individual utterances (Beatty-Martínez, Navarro-Torres, & Dussias, 2020; Green & Abutalebi, 2013). An example of this categorization in practice is Olson's recent research on the influence of context, in which they define three contexts based on the amount of cued language switching in each: two monolingual (one in mostly a first language, L1, one in mostly a second language, L2) and one bilingual (frequent switching between L1 and L2).

For this research, however, experimenters categorize contexts based on the *environment*, not the individual's language goals. Contexts here follow generally the same guidelines, though based on the types of bottom-up language information rather than top-down goals. Thus, in this case, a monolingual context refers to the presence of only the target language (the language desired for speaking) in the environment, a bilingual context refers to the presence of the target language and a non-target language, and a code-switching context refers to the presence of target and non-target languages within individual utterances.

Top-Down Control in Bilinguals

In bilingual research, top-down attention would be the cognitive mechanisms of language control (Grainger & Dijkstra, 1992; Zheng et al., 2018). *Language control,* in this paper, refers to the brain's means of choosing and producing the right words, from the right language, at the right time. The brain selects words based on their activation level at the time of production, so language control ensures the timely activation of desired words through effortful control over whole language networks and individual lexical items (Declerck & Philipp, 2017; Levelt, 1999). This is made only more difficult by the competitive nature of languages, which vie for use

(Bialystok et al., 2009; Dell et al., 1993). Most relevant to this research is evidence that this competition can arise from auditory (bottom-up) inputs (Lagrou et al., 2011). The mechanisms of such a control system are modeled in the Adaptive Control Hypothesis (ACH; Green & Abutalebi, 2013) and the Bilingual Interaction-Activation Model (BIA; Grainger & Dijkstra, 1992).

The Adaptive Control Hypothesis. The ACH (Green & Abutalebi, 2013) proposes that language control exerts its influence through a number of mechanisms, each best-suited to certain goals and situations. In other words, the ACH says that different environments impose different language challenges requiring different tactics to manage. This model posits that a *meta-process* sets goals, both broad and specific (*e.g.*, speaking only L1 versus being prepared to code-switch), which working memory is responsible for maintaining. At the time of speaking, the speech pipeline produces the words selected. The ACH is unique in that it outlines the different mechanisms that might be used in monolingual, bilingual, and code-switching modes. The inclusion of situation-based control strategies shows that this model does consider external context as an influence, though mainly in the sense that the meta-process sets top-down goals based on it.

The Bilingual Interaction-Activation Model. By contrast, the BIA explicitly considers the effects of top-down and bottom-up influences in language production (Grainger et al., 2010; Grainger & Dijkstra, 1992; van Heuven et al., 1998). Essentially, the BIA proposes that any given *lexical item* (word) has both top-down and bottom-up inputs. A word's top-down input is an excitatory connection to a *language node*, which governs the overall activation of the words in a language's network. A word's bottom-up input is excitatory sensory information—this activation can be from direct sensory activation of that word or from the activation of words in its neighborhood (that have similar sound or meaning). Thus, the activation of any word may be affected by influences endogenous (such as the top-down activation of a language node) or exogenous (such as the perception of that word, or similar words, in the environment). So, as a hypothetical, a Spanish-English bilingual wants to speak Spanish. The Spanish language node activates and heightens the activity of all the words connected to it (and may even simultaneously inhibit English, a competing language node). Unfortunately, this bilingual can hear English nearby, which excites words in English despite their being actively inhibited. Now, Spanish faces heightened competition from English because of its presence in the environment. In fact, the BIA predicts that this increased competition could come from written or spoken forms of a non-target language, as bottom-up activation can come from modality (e.g., reading comprehension or auditory comprehension).

Research supports this model: by preceding an image with written target language words and non-target language words rather than with language-switch cues, Peeters and colleagues (2014) tested whether cues alone could account for asymmetrical switch costs (the observed effect that switching into an L1 incurs a greater temporal cost than switching into an L2; Meuter & Allport, 1999). Despite having no switch cues and distracting subjects through only a single written word, the experimenters still saw asymmetrical switch costs on the level of classic cued paradigms. Since a non-target-language comprehension task could cause a lag in target-language production, their findings support the BIA's claim that language nodes and lexical items may be activated through bottom-up information in written form. This means that endogenous inputs (like language goals) and exogenous inputs (like language comprehension) can both affect the efficiency of language production.

Gambi and Hartsuiker's recent research (2016) also support the BIA's predictions, this time using naturalistic auditory distractors. They wondered whether passive comprehension (rather than active comprehension, like in Peeters et al., 2014) could produce the same slowed response times as those seen by Peeters et al. (2014). Experimenters had pairs of participants seated next to each other respond to images, each naming every other image. In certain blocks, one of the participants could switch languages at random while the experimenters measured the other participant's response times. The nonswitch participants' responses were slower in the trials immediately after the other participant exercised a language switch. These results suggest that not only can auditory language information alter language nodes' activity, but it can do so even when *passively* processed (a phenomenon seen in other modalities as well; Grainger & Beauvillain, 1987; Thomas & Allport, 2000). In other words, casually hearing a non-target language utterance can, seemingly, activate the non-target language node and cause lags in subsequent target-language production.

Combining Models. The BIA predicts that bilingual language control can be influenced by intermodal, passive, bottom-up information (Grainger & Dijkstra, 1992). The ACH posits that working memory's multi-faceted and goal-dependent toolbox (including inhibitory control, conflict monitoring, and attention to cues, among others) allows language control to enforce the meta-process' desired goals (Green & Abutalebi, 2013). Together, they outline an variable, goal-dependent system of control driven by endogenous *and* exogenous influences. Such a system would, in theory, be perceptibly influenced both by changes in top-down strategy and bottom-up information.

If, for instance, a monolingual context featured only a bilingual's target language, the BIA says there would be little bottom-up activation of the non-target language node. With the goal of monolingual fluency, the ACH says language control would use interference suppression and conflict monitoring to quiet any spurious non-target activation. By contrast, in a bilingual context where a non-target language can be heard, the BIA says there would be greater unwanted bottom-up activation of the non-target language. To accommodate this interference, ACH predicts that language control would focus on interference suppression and, now, alertness to cues (to switch goals [languages] efficiently if the need should arise). Finally, in a code-switching environment, the BIA says the non-target and target language nodes would be simultaneously and highly activated due to top-down processes (like borrowing words and phrases or switching languages entirely) and bottom-up processes (like comprehending multiple languages simultaneously nearby). With unpredictable cross-linguistic interference and goal-switching, language control would opt for opportunistic planning to keep up.

Note that these predictions are based on contexts defined by language goals; they center on what tools the bilingual would use (and when) to efficiently enforce the goal of code-switching, not switching at all, et cetera. Yet, past research presented above suggests that even if production goals do *not* change (i.e., the bilingual keeps to one monolingual goal throughout a situation, never switching languages), this will not save the brain from accruing costs in dealing with bottom-up language information (Gambi & Hartsuiker, 2016; Peeters et al., 2014). From written word-priming to partners language-switching, non-target-language bottom-up influence seems to be able to cause delays in production just like top-down cues to switch languages.

Hypotheses and Predictions

This idea of bottom-up influence causing production delays leads to the main purpose of this study: investigating whether language environment (context) influences bilingual language control. To answer that question, this experiment had subjects name images while simultaneously hearing short, trial-length audio blips in various languages (language distractors). Each 15-trial block had audio in either Hungarian, English, Spanish, or English-Spanish Mixed (randomly switching between English and Spanish per trial) before moving to the next block of 15 (repeated over 16 blocks). Each distractor had four blocks through the course of the 16 total, totaling 60 trials per language distractor. The participants responded in *only* English—as such, no conscious top-down language control changes (like a top-down cue to switch) would occur. Thus, any difference in response times per block would be attributed to the different contexts (i.e., bottom-up activation of a non-target language). Since recruiting was at the University of Houston in Houston, Texas, the target population was Spanish-English bilinguals and the non-target language distractors were mostly Spanish. The University of Houston was an ideal location for this study, as it is one of the most diverse Universities in the nation and is situated in a city where 38.9% of the population speaks Spanish (University of Houston Department of Institutional Research, 2020; U.S. Census Bureau, 2020). The following were the empirical questions and hypotheses concerning them:

1. ***Does bottom-up information influence top-down control in bilingual language control?*** Experimenters expect this to be the case, as recent evidence suggests external information (i.e., known, non-target languages) can affect language production efficiency – these would appear as longer response times in experimental contexts (those with known non-target language audio, like Spanish) than control contexts (those with target-language audio or unknown-language audio, like English or Hungarian).
2. *Is this influence context-dependent?* Experimenters expect that, yes, top-down control mechanisms and amount of bottom-up activation are context-dependent, so context’s effects

should be as well – different distractors would lead to different costs in image-naming throughout the course of a block.

These hypotheses require elaboration. While previous research investigated the immediate effects of non-target language interference (through written and auditory modalities) on the next trial, this research investigates the effect of a language environment. If different environments cause different amounts of bottom-up, non-target language activation and require different control mechanisms, they should show different levels of production efficiency. So, much like Gambi and Hartsuiker (2016), this research removes cues to keep endogenous goals consistent. Unlike Gambi and Hartsuiker, this is done to observe the general effect on production of each language distractor over the course of a number of trials, rather than the immediate effect of non-target language comprehension.

In other words, across all contexts top-down language control seeks to maintain its primary goal of ‘speak English.’ For Spanish-English bilinguals, bottom-up distractors are an unknown language context (Hungarian), a known target-language context (English), a known non-target-language context (Spanish), and a known target *and* non-target language context (Mixed). Each distractor (in the above order) would cause more activation in undesired words than the last, within and/or beyond the target-language (English) (Declerck & Philipp, 2017). When considered as monolingual (only the target language is present, Hungarian and English contexts), bilingual contexts (target language and other known language, Spanish context), and code-switching environments (frequent target to non-target switching, Mixed context), the ACH suggests each would require different control mechanisms (Green & Abutalebi, 2013). For instance, Hungarian should not activate *any* words, while Mixed would activate both English and Spanish non-target words (other than the stimulus). Together with findings that non-target language input interferes with language production (Gambi & Hartsuiker, 2016; Lagrou et al., 2011), this research predicts that the change in undesired words activated and the change in top-down control strategies would cause different costs (different response times) per context.

Method

Participants

This study recruited 18–40-year-old students at the University of Houston in Houston, Texas, a largely bilingual city. The target population was Spanish-English bilinguals, though all students could participate. Exclusionary criteria included psychoactive medications, not speaking English, and knowledge of Hungarian. Students received SONA hours (credited to their ongoing University of Houston psychology courses) for their participation and entered a raffle for *Visa* gift cards of various amounts up to $50. The study took place online.

Given the sample sizes of similar tests (Gambi & Hartsuiker, 2016; Peeters et al., 2014) and the use of online methods for this experiment, the target sample size was 100 participants. Through data collection 168 subjects participated. Following scoring of response time (RT) and response accuracy (RAcc), experimenters excluded 57 participants from final analysis due to corrupted or uninterpretable audio files and 20 more due to missing survey data. Experimenters removed all trials with incorrect or non-English responses and, afterwards, excluded any subject that had less than half of trials correct in any block (two subjects). Finally, experimenters excluded one monolingual. The final sample size was 88.

Descriptive statistics of participants in the final analysis are in *Table 1*. Descriptive statistics concerning languages spoken are in *Table 2*. All participants included in analyses self-rated as highly proficient in English (M = 9.469 out of 10, SD = .780). Fourteen identified as male (19%) and 74 identified as female (81%). Sixty-six were bilingual (75%), 14 were trilingual

(16%), and six spoke more than three languages (7%). The average age of acquisition (AoA) of English was four years old with a wide standard deviation (SD = 3.83). Since the primary non-target-language distractor was Spanish, experimenters separated participants into two language groups: Spanish-speakers (*Hispanophones*, 51 subjects, 58%) and non-Spanish-speakers (*Non-Hispanophones*, 37 subjects, 42%). Descriptive statistics per group are in *Table 3*.

Materials

The onset of the COVID-19 virus and safety-related shutdowns in March of 2020 saw the closing of in-person University research just as data collection was to begin, and conditions did not allow for in-person testing through the data collection period (Office of the Texas Governor, 2020). Experimenters manually adapted the project for online administration using the *jsPsych* library (de Leeuw, 2015) and *JATOS* (Lange et al., 2015), both of which are open-access tools for online experimentation. Research shows the *jsPsych* library of online source code to be highly reliable for response time recording on the scale of in-lab metrics (de Leeuw & Motz, 2016; Hilbig, 2016). For data safety, researchers used the *DigitalOcean* platform to host the experiment on a secured website. Data collection occurred between August and December of 2020. Experimenters prepared data in *R* (R Core Team, 2020) and ran analyses using *JASP* (JASP Team, 2020).

Measures

Each participant took a number of pre-tests aimed at recording individual differences in cognitive control abilities and language experience:

The Flanker Task. This task (Eriksen & Eriksen, 1974) tests cognitive control abilities through the inhibition of extraneous information. Participants responded to a specific stimulus embedded in irrelevant, and at times incongruent, stimuli. The stimuli were arrays of five arrows appearing in a horizontal row (as in the case of Ridderinkhof et al., 1999). The center arrow is the target, the other four are distractors. On congruent trials, all arrows pointed the same direction, on incongruent trials, the center arrow pointed opposite the surrounding arrows. Participants saw two blocks of 50 trials each with equiprobably dispersed congruent and incongruent trials. If the middle arrow pointed left, participants were to press ‘z’ with their left index finger, and if it pointed right, they were to press ‘m’ with their right index finger. Participants received instructions to do this as quickly and accurately as possible. Incongruency forces the effort of inhibiting the surrounding arrows, leading to slowed and, at times, incorrect responses. Participants’ inhibitory control abilities mediate response accuracy and speed, so better inhibitory control abilities would lead to faster and more correct responses on incongruent trials. These results gave an empirical measurement of individuals’ inhibitory control abilities (the difference between congruent and incongruent trial response times) and baseline processing speed (congruent trials).

The Shape-Color Task. This test (Kray & Lindenberger, 2000) focuses on one’s ability to task-switch in dynamic, uncertain contexts. The variant used here had participants seeing circles and squares colored either red or blue (Rodríguez-Pujadas et al., 2013). Before each set of images participants saw a cue to focus on *either* the color or the shape. When the set began with the cue “COLOR,” they pressed “z” with their left hand if the shape was blue and “m” with their right hand if the shape was red. When the block began with the cue “SHAPE,” they pressed “z” with their left hand for a circle and “m” with their right hand for a square. After each trial, a black dollar sign appeared, indicating the participant “switch” rules (if horizontally oriented) or “stay” with the current rule (if vertically oriented) for the next stimulus. “Stay” and “switch” cues ere equiprobably dispersed across all trials. Since trials could be congruent (the correct response was the same in either task) there were four trial types: congruent-stay, incongruent-stay, congruent-switch, and incongruent-switch.

Participants first underwent a number of training trials and practice sets. There were five recorded blocks of 70 trials each, the first being a practice block. Each consisted of 35 right-hand and left-hand responses, 35 shape and color cues. Shape-Color yielded a measure of individuals' task-switching abilities (response time differences between all stay and all switch trials) and baseline processing speed (congruent-stay trials).

The Digit-Span Task. This test, often used in intelligence scales and other psychological batteries (Richardson, 2007; Wechsler, 2003), tested auditory working memory abilities. Participants heard multiple series of numbers between zero and nine in random order, then repeated them back in numerical order. The number of digits increased from three per trial to nine per trial, each trial repeating twice. Participants then underwent the same paradigm, though this time repeating the numbers back in backward order. The ability to internalize, manipulate, and repeat back the numbers reflects specifically auditory working memory, which typically has a capacity of 5 +/-2 and is stretched by the later trials of eight and nine digits (Baddeley, 1992). Accuracy measures gave a record of individuals' auditory working memory ability.

The Bilingual Switching Questionnaire (BSWQ). This survey (Rodriguez-Fornells et al., 2012) is an empirically supported self-assessment of language switching experience and skill. The survey has four subsections: 1) L1 Switching Tendencies; 2) L2 Switching Tendencies; 3) Contextual Switching; 4) Unintended Switching. The first two subsections measure the likelihood of switching into a language based on competency and proficiency. The third measures the frequency of language-switching in based on environments. The fourth measures any other incidental switching. It is comprised of twelve questions, three corresponding to each of the four subsections. The contextual switching and unintentional switching scores, in particular, gave a metric by which to compare bilinguals from diverse language backgrounds on their language-switching habits. Thus, the BSWQ was a key measure of individual differences in switching preferences and experience.

Language and Social Background Questionnaire (LSBQ). This survey (Anderson et al., 2018) yielded scores of bilinguals' language use history as it pertained to their social life. While the BSWQ is a quick means of querying specific language-switching tendencies, the LSBQ covers broad-strokes background information about when and where a bilingual (or multilingual) uses their languages and to what extent. It yields the extent of use, proficiency in, and social tendencies with non-English languages, each output being a composite result of the Likert-scale questions that make up the test. Specifically, this experiment drew scores of individual differences from sections concerning parental education (as a measure of socio-economic status), age of acquisition of (first beginning to learn) English, age of moving to the United States, English proficiency (compiled from self-report measures of speaking, understanding, reading, and writing abilities), mother's first language, father's first language, and the number of languages spoken.

The Babel Paradigm. This paradigm was the experiment of main interest. There were image stimuli and audio distractors. Similar to Gambi and Hartsuiker (2016), this paradigm did not ask subjects to language switch at any point through the experiment. Unlike their research, in this paradigm the distractor no longer preceded the stimulus, it played throughout presentation of the stimulus. This accomplished the goal of removing top-down cues while testing the impact of continuous bottom-up interference with different language distractors.

The image stimuli were black-and-white, two-dimensional, static images from a 520-image database compiled by Bates et al. (2003). Using statistics compiled by Szekely et al. (2005), experimenters removed images with complex answers, with an average correct response of less than 85% in both English and Spanish, and English-Spanish cognates, leaving 178 images total.

The audio distractors were recordings of men and women speaking Spanish, English, and Hungarian. The English and Spanish recordings were narrations of a number of stories, and the Hungarian recordings were excerpts from a news podcast. Experimenters cut the recordings into 2000ms clips at random to neutralize any potential recognition of the story or any flow of narrative. Experimenters removed clips containing recognizable words in Hungarian (such as prominent political figures, country names, and any English cognates or homophones).

First, participants read instructions to put in earphones. Participants then read instructions to name the objects in the images to follow as quickly as possible in English. There were 16 sets of 15 images each. Before each trial, a fixation cross appeared (250-750ms). After each fixation, an image stimulus and audio distractor presented simultaneously (2000ms). See *Figure 1* for visualization of this procedure.

The audio distractors changed between each block and were one of four language contexts: Hungarian, English, Spanish, or Mixed. Hungarian served as a high-level baseline, wherein a non-target language was present but incomprehensible. English was the target language throughout the experiment and its context served as a target-language distractor. Spanish served as a non-target language distractor that was comprehensible to English-Spanish bilinguals (the main population of interest). Mixed played both English and Spanish audio clips at random throughout the set, representing a code-switching context. Each block repeated four times at random (four context distractors, four sets each, 16 sets total). This yielded 60 trials in each context and 240 trials total.

Babel Paradigm Scoring. The outcomes of interest in the Babel paradigm were response times (RT) and response accuracies (RAcc). For every participant, at least two of eight trained research assistants manually scored RT and RAcc for the recorded response of each trial. Scorers used *Audacity*, a free audio manipulation tool, to record both outcomes. RAcc scores were 'correct,' 'incorrect or no response,' or 'wrong language' (to record any accidental language switching). Experimenters adjusted the final length of recordings that did not extend to the full length of 2000ms as set in the experimental design to meet 2000ms, in order to correct for differences in the onset of recordings (unless the recording was greater than 500ms shorter). In other words, since recordings always *ended* at the same point, only the latency of their onset explained differences in recording length. By adding to each RT the length of the recording missing, experimenters corrected for differences in recording onset when scoring RT.

Results

To score the results of this experiment researchers first assigned subjects into two groups, Hispanophone and non-Hispanophone, based on whether or not they spoke Spanish (since distractors of interest were in Spanish). Researchers separated trials into "contexts" based on distractor languages (Hungarian – "HU", English – "EN", Spanish – "SP", and Mixed English-Spanish – "ENSP") and removed all incorrect responses. Both groups had nearly the same average number of incorrect trials per block, averaging between seven and eight incorrect per block for both groups (*Table 4*). 167 subjects participated, 88 had data sufficient to include in final analyses. Of those, 51 were Hispanophone and 37 were non-Hispanophone.

It is important to first note statistical assumptions. Across analyses where subjects split into groups, the number of trials and subjects seemed to quell most statistical issues that could have arisen from the differences in sample sizes. The small issues that did arise are included in discussion of related results below.

Hispanophones

The main paradigm of this experiment was designed to elucidate the effects of different contexts relevant to Spanish-speakers. Thus, experimenters first investigated these effects, the within-subject effects of Hispanophone participants. This was a “whole block” analysis, comparing the average RT of all trials per context to investigate any general effects of the distractors on RT (One-Way ANOVA, four levels of context). This analysis yielded no significant main effects of context (*Table 5*).

Researchers then asked whether the effects of context were limited to the first trials after a change in distractor, as recent similar research would suggest (Gambi & Hartsuiker, 2016). So, experimenters ran “first-thirds” analysis comparing the average RT from the first third of trials in each block of each context (the first five trials from each of the four blocks per context) to investigate whether any contexts had immediate effects that averaged away over later trials (One-Way ANOVA, four levels of context). This, too, did not yield significant main effects (*Table 6*), though the effect size was small-to-moderate ($F = 2.370$, $p = .073$, $\eta^2 = .045$) and prompted researchers to investigate whether RT did indeed differ between the first third of trials and later trials per block.

So, experimenters ran “first-last thirds” analyses comparing the first third of trials and last third of trials (trials 1-5 and trials 11-15) in each block of each context (four One-Way ANOVA; one per context, levels were the first and last third of trials). These revealed that certain contexts did indeed have a significant effect on the difference between first-third and last-third RTs. Specifically, Hungarian ($F = 24.104$, $p < .001$, $\eta^2 = .325$) and English ($F = 7.603$, $p < .01$, $\eta^2 = .132$) contexts had strong main effects on RT differences between the first five and last five trials of their blocks, where the last third were significantly faster (*Table 7*).

All Participants (Grouped)

Researchers then expanded the analysis to compare groups, since so many non-Hispanophones did participate. These analyses focused on within-subject *and* between-subject differences, since they investigated differences between Hispanophones and non-Hispanophones, so they included covariates (discussed later in *Results*). Following the significant effects of time in Hispanophones (differences between the first-third and last-third trials' RT) for certain contexts, researchers began with one large analysis investigating the interactions between context, time, and group (4x3x2 ANCOVA; four levels of context, three levels of time spent in a context [first, middle, and last third of trials; trials 1-5, 6-10, and 11-15], and two group levels). In terms of statistical assumptions, the levels of time violated sphericity, though Huynh-Felt corrections showed no *p*-value errantly showed significance, and the first-thirds and last-thirds levels of Mixed block violated homogeneity of variance (*Table 8*). The 4x3x2 analysis showed a significant interaction between time and context ($F = 2.591$, $p < .05$, $\eta^2 = .032$), a number of significant within-subject covariate interactions with time and context, a significant effect of language group between-subjects ($F = 4.445$, $p < .05$, $\eta^2 = .054$) and, though non-significant, a small-to-moderate effect size of time (*Table 9*).

Exploratory post-hoc comparisons showed a number of significant differences between levels of time within and between contexts (*Table 10*). In short, thirds did not differ significantly *between* any contexts (e.g., first-third RTs were comparable between all contexts, as were middle-third and last-third). *Within* contexts, first-third trials differed significantly from either middle-third or last-third trials (or both) in each context except Spanish. Middle-third and last-third trials did not differ significantly within *any* contexts. No significant language group interaction with time or context arose, likely because the changes in RT over time per context were nearly identical in each group (*Figure 2*). Interestingly though, the Mixed block may have proved particularly difficult for Hispanophones, as their first-third and middle-third RTs were nearly identical. This trend appeared in no other block in either group (*Figure 2*). This was unique, though without a significant interaction involving language group more data would be necessary to tease out its verity.

Individual Differences

Grouped Data. Experimenters collected measures of individual differences in three categories: social background, language experience, and cognitive abilities. Social background included socioeconomic status (SES), age, and age of arrival to United States. Language background included parents' first language (Mother and Father L1), English proficiency, number of languages spoken, age of acquisition of English (AoA), and language-switching tendencies divided into two scales (accidental switching and contextual switching). Cognitive abilities included scores of baseline processing speed (Flanker congruent-trial RT and Shape-Color stay-trial RT), inhibitory control abilities (Flanker congruency effect—the average RT difference between congruent and incongruent trials), task-switching abilities (Shape-Color switching effect—the average RT difference between stay and switch trials), and working memory abilities (Digit-Span accuracy scores).

Six of these measures differed significantly between groups: SES ($p < .01$, d = .645, M = 5.648, M_{Hisp} = 4.98, $M_{Non\text{-}Hisp}$ = 6.568), age of entry to US ($p < .01$**, d = .625, M = 2.898, M_{Hisp}

= 1.51, $M_{Non\text{-}Hisp}$ = 4.811), number of languages spoken (*p* < .05, d = .521, M = 2.375, M_{Hisp} = 2.216, $M_{Non\text{-}Hisp}$ = 2.595), mother's L1 (*p* < .001, d = 1.874, M = 2.341, M_{Hisp} = 1.980, $M_{Non\text{-}Hisp}$ = 2.838), father's L1 (*p* < .001, d = 1.9, M = 2.386, M_{Hisp} = 2.039, $M_{Non\text{-}Hisp}$ = 2.865), and contextual language switching tendencies (*p* < .01, M_{Hisp} = 9.784, $M_{Non\text{-}Hisp}$ = 8.135). The group difference between accidental language switching was just below significant (*p* = .055). All group-difference t-tests are in *Table 11*, grouped descriptive statistics are in *Table 3*.

For measures of individual differences in cognitive abilities, Shape-Color captured its target effects, with a 2x2x2 ANOVA (levels of switching, congruency, and language groups) showing a significant effect of congruency (where the correct response in either task was the same; F = 19.820, *p* < .001, η^2 = .187) and switching (whether a trial cued the same task as the preceding trial or not; F = 12.933, *p* < .001, η^2 = .131) but no effect of or interaction with language group (*Table 12, Figure 3*). Flanker also captured its target effects, with a 2x2 ANOVA (levels of congruency and language groups) showing a significant effect of congruency (F = 30.521, *p* < .001, η^2 = .262) but, again, no significant effect of or interactions with language group (*Table 13*).

For the measures that significantly differed between groups, explanation of scores are as follows. For SES, possible scores ranged from 2, being less than high-school-level education for both parents, to 10, being graduate-level education for both parents. For age of entry to US, American-born individuals marked 0. Number of languages spoken is as the name suggests. For contextual language switching tendencies, possible scores ranged from 3 to 15, with 3 representing someone who infrequently chooses to switch languages and 15 representing someone who switches frequently, even within single conversations. For mother's L1 and father's L1, possible scores were 1 (English), 2 (Spanish), and 3 (Other). Frequency statistics for parents' L1 can be summarized as follows: the Hispanophone group was overwhelmingly Spanish L1 for both (86% and 84% for mothers and fathers, respectively), the non-Hispanophone group was overwhelmingly Other L1 for both (92% for both), and each group had a small number of English L1 mothers and fathers (less than 10% for both).

Covariates. The analyses of all participants had six covariates: English proficiency, SES, contextual switching tendencies, accidental switching tendencies, Flanker congruent trial RT, Shape-Color congruent-stay trial RT, Flanker congruency effects, and Shape-Color switching effects. Experimenters chose these as covariates based on their evidenced effect in bilingual paradigms in past research and their relationship to the skills hypothesized to be relevant in this task. In particular, SES significantly differed between groups and seems to interact with bilingualism in other studies (Ardila et al., 2005; Calvo & Bialystok, 2014; Engel de Abreu et al., 2012; Hackman et al., 2010; Jakubów & Corrêa, 2019). English proficiency, though not significantly different between groups, is related to the target language and experimenters controlled for its mastery (also because research suggests that proficiency does impact production skill; Costa & Santesteban, 2004). Contextual and accidental switching tendencies as covariates both correct for the effect of different levels of experience in deliberate language switching (Rodriguez-Fornells et al., 2012). Flanker congruency effects (the difference between Flanker congruent trials RT and incongruent trials RT) serve as an analog for inhibitory control, correcting for the ability to block out distractors (Ridderinkhof et al., 1999). Shape-Color switching effects (the difference between Shape-Color stay trials RT and switch trials RT) corrects for task-switching abilities, a skill relevant to language-switching and potentially relevant when adjusting to switches in context (Green & Abutalebi, 2013; Kray & Lindenberger, 2000). Finally, Flanker congruent and Shape-Color congruent-stay correct for baseline processing speed, as both require response, but neither is an experimental condition.

Correlations with the DV. Measures of individual differences had significant effects across analyses (both between-subjects and as interactions within-subjects), so experimenters investigated their relationships to RT as an attempt to explain significant group differences in RT (*Table 13*). Overall (using every participant's overall average RT) four covariates significantly correlated with RT, including Flanker congruent trial RT (R = .288, $p < .01$), Flanker incongruent trial RT (R = .272, $p = .01$), Shape-Color congruent-switch trial RT (R = .221, $p < .05$), and sex (R = -.213, $p < .05$). No covariates correlated with RT when using only Hispanophone RTs. Non-Hispanophone RTs correlated with Flanker incongruent trial RT (R = .344, $p < .05$), Shape-Color incongruent-stay trial RT (R = .356, $p < .05$), and Shape-Color congruent-switch (R = .412, $p < .01$). An odd result, though non-significant in either group, is the opposite correlations with Flanker congruency effect (the difference between congruent and incongruent trials, an analog for inhibitory control) between groups, where Hispanophone RTs correlated negatively and non-Hispanophone RTs correlated positively.

Familiarity and Target Analyses

Experimenters needed to establish whether familiarity with the distractor language influenced the results. These analyses excluded the Mixed block as it did not fit into the levels defined. First, experimenters re-grouped data based on knowledge of a language rather than block-wise per context. Specifically, experimenters grouped Hispanophone trials into English and Spanish (familiar) versus Hungarian (unfamiliar) and non-Hispanophone trials into English (familiar) versus Spanish and Hungarian (unfamiliar). Experimenters ran a familiarity by time by group ANCOVA (2x3x2) that yielded a significant interaction between familiarity and language group (F = 6.515, $p < .05$, $\eta^2 = .077$), a significant effect of time (F = 7.699, $p < .001$, $\eta^2 = .09$), and a significant between-subjects effect of language group (F = 4.357, $p < .05$, $\eta^2 = .053$) (*Table 14*).

Finally, experimenters also needed to establish whether the distractor language being the target language influenced the results. So, experimenters re-grouped data based on whether the contextual language was a target language or not, excluding Mixed once again, and ran a target by time by group ANCOVA (2x3x2). That analysis returned a significant effect of time (F = 5.104, $p < .01$, $\eta^2 = .061$) and a significant between-subjects effect of language group (F = 4.009, $p < .05$, $\eta^2 = .049$) (Table 15). In both of these analyses time violated sphericity, though again (as with the context by time by group analysis) corrections showed it did not significantly impact any results.

Discussion

From the results, there are a number of key takeaways concerning effects on RT:

1. Context did *not* have a significant effect in Hispanophones alone.
2. Time had a significant or near-significant effect in every all-subjects analysis.
3. Context and Time had a significant interaction effect in an all-subjects analysis.
4. Language familiarity and group had a significant interaction effect in an all-subjects analysis.
5. Language group had a significant between-subjects effect in every all-subjects analysis.
6. Relationships existed between covariates and RT, but their interpretations were unclear.

Effect of Context on RT in Hispanophones. The lack of an effect of context on RT in Hispanophones suggests that there are not context-dependent differences in bilinguals Hispanophones' ability to produce their desired languages (in this case, English). The lack of a main effect in all-subject analyses only further supports the conclusion that there are not context-dependent differences in language control abilities. Actually, the RT averages per context were *remarkably* similar, matching to the third decimal place across contexts in Hispanophones and the second decimal place across contexts in all participants.

The Effect of Time Spent in a Context. The significant effect of time held true in nearly every all-subjects analysis, and the result where it was not significant still had a small-to-moderate effect size worth considering since comparable analyses all returned highly significant p-values. The significant difference between first-third and last-third was true in every context except Spanish and suggests that a change in distractor does have an immediate effect that is lost in later trials. In other words, as bilinguals spend time in a language context, they acclimate to it. Upon visual inspection of the data (*Figure 2*), if there is to be a significant improvement in speed it seems to occur within the first ten seconds of a change in context (the first five trials, each two seconds) and improvement is markedly reduced afterwards. Given that Gambi and Hartsuiker (2016) did see single-trial effects of switches to auditory non-target language distractors, the immediate effects seen in these results seem to support their findings.

Interaction between Context and Time Spent in a Context. Building on the above findings, there was a significant interaction between context and time in an all-subjects analysis. These results suggest that the impact of context differs as the amount of time spent in that context increases. For example, in the Hungarian context the first-third and middle-third differences are highly significant, while middle-third and last-third are not significant at all. By comparison, in the English and Mixed contexts, there is a significant difference between the first-third and last-thirds, but not in-between. Finally, in the Spanish context, there are no significant differences between any of the thirds.

The results of the Hungarian context fit with predictions—as the context causing the least unwanted activation, it poses the least of a problem and the immediate effects are swiftly overcome. The English context, too, fits with predictions by showing improvement though less immediately than the Hungarian context. However, the results of the Spanish and Mixed blocks are complicated by competing averages between groups. The Hispanophone group' Spanish context results showed a decrease in performance (slower RT) in the last-third compared to the middle-third—since the non-Hispanophone group showed the same result in Hungarian and no other context, perhaps this is simply noisy data. Interestingly, the Mixed block is marked by Hispanophones' uniquely poorer performance in acclimating through the thirds of trials, perhaps evidencing that more complex contexts are more difficult for bilinguals. This would be an interesting proof-of-concept, and though the Hispanophone results concerning the interaction between context and time are non-significant (when run on a similar ANCOVA to the all-subjects context by time analysis), they do have a noteworthy low-to-moderate effect size when considering the significance of similar analyses. Taken together this suggests that there may be something unique about the Mixed context's effect on Hispanophones that leads to their difficulty acclimating to it, though more data is necessary to discern the trustworthiness of these results.

Interestingly, other analyses suggest that these results can neither be attributed to knowledge of the distractor language nor the distractor being the target language (since neither analysis showed an interaction between time and that independent variable). In that case, the above explanation of differences in difficulty acclimating (which is essentially based on different familiarity with the distractor languages) may be a moot point. Another interesting finding is that the contexts did not differ in their first-thirds. This suggests that the immediate impact of a switch in context is comparable regardless the language—a context switch to Hungarian, unfamiliar to all participants, had as much effect on RT as a switch to any other language. These results contradict the findings in Gambi and Hartsuiker (2016), where asymmetrical costs suggested the impact of a switch in auditory distractor depended on whether the switch was into or from a non-target language. In the end, once again, more data will be necessary to tease out the origin of this project's findings.

Interaction between Language Familiarity and Language Group. Even though the effect of time spent in a context was not significantly impacted by familiarity with that context's language (no interaction with time), a significant interaction between familiarity and language group did arise in that analysis. The difference between groups in question seems to be that Hispanophones' first-third of trials in a known language is faster than the first-third of trials in an unknown language, whereas non-Hispanophones display the opposite effect. Perhaps this interaction is attributable to the significant difference in Hispanophones' experience with contextual language-switching. Unfortunately, the lack of any interaction with the contextual switching tendencies measure seems to refute this, perhaps it is simply a result of smaller, unbalanced samples. Or, perhaps this reflects the increased load of more familiar languages being in the Hispanophones' environment throughout the task. Again, more data will be necessary to fully understand the results seen here.

Effects of Language Group. The between-subjects significance of language groups in every analysis could be explained by any number of unaccounted-for group differences, as no included covariate explained the difference significantly (or caused it to be non-significant). Since the focus was originally on Hispanophones alone, experimenters did not incorporate a sufficient between-group control. Thus, as a rough analog for a control, experimenters compared any Hungarian blocks that came first across participants (i.e., an unknown-language ‘control’ block that appeared before subjects could be affected by context switches and known other-language distractors). The between-group difference remained significant in this comparison, suggesting the difference to be independent of experimental manipulations (though the sample size was quite small for this analysis). Experimenters also compared RT averages per stimulus from each participant to the RT averages in the image database (Bates et al., 2003; Szekely et al., 2005) to see whether one group may have had comparable RT to the norms established in past research. If one group were comparable and another were not, this could suggest that the non-similar group was affected by experimental conditions. However, these comparisons do not support this explanation, as they showed both groups departed from the observed averages significantly across all images (*Figure 4*). One final consideration is the difference in languages spoken between groups (*Table 2*). The vastly more diverse language backgrounds in the non-Hispanophone group could be masking effects unique to one language or language family that would otherwise help explain the between-group difference. Thus, this group difference in RT will take more research to accurately interpret.

Individual Differences and RT. Some interesting results to arise from this research concern the measures of individual differences and their effects on RT. For instance, the necessity of including English proficiency for the significant interaction effect of context and time in the all-subjects context by time by language group analysis, which suggests that differences in English proficiency play a role in the connection between context and improved performance over the course of a block. Another interesting interaction is between time and Shape-Color switching effects in the context by time by language group analysis. Nearly the same covariates are involved in the target by time by group analysis, where English proficiency and two measures from Shape-Color interact significantly with time spent in a context: the measure of baseline processing speed (congruent-stay trial RT) and switching effects. In the familiarity by time by group analysis, the same covariates arise once again. The continued interactions of English proficiency and task-switching abilities in these analyses strongly suggest that the two are both related to improvement in RT over the course of a block.

Recent research on task-switching supports this conclusion, as Hartanto and Yang (2016) evidenced a strong relationship between language-switching experience and a task similar to the Shape-Color used here. Perhaps Shape-Color serves as a rough analog for language-switching abilities, leading to its significant role in these between-subject analyses, though with no evidenced relationship between Shape-Color and contextual switching experience ($R = -.075$, $p = .486$) this explanation might not be reliable. At the very least, these past results do further support the relationship evidenced here between language control and Shape-Color results, despite unclear interpretation. Gambi and Hartsuiker (2016) saw an effect of proficiency similar to that evidenced here, though theirs was with non-target language proficiency. This supports a role for how balanced a bilingual is (how comparable their proficiencies in spoken languages are) in their ability to manage known non-target language interference.

Aside from interactions within analyses, these groups differed in most language experience and social background measures except English AoA, age, accidental switching tendencies, and English proficiency. This makes for a difficult task of teasing out the genesis of group RT differences, especially since English abilities and cognitive abilities are consistent between groups while most other covariates differ widely. It also happened to be that most highly differing measures suffered from restricted range and abnormal clustering, further complicating their interpretation.

In hopes of finding connections between covariates and outcomes to help explain the significant group differences in all analyses, experimenters explored correlations between all of the measures of individual differences and each subject's average overall RT. These were not entirely helpful, either, as only Flanker congruent trial RT, Flanker incongruent trial RT, Shape-Color congruent-switch trial RT, and sex resulted in significant correlations with RT. In short, correlations with RT were mainly limited to cognitive control measures. The Shape-Color correlation seems to show a connection between switching abilities and RT (since this trial type *only* queried switching abilities with no effect of incongruency), though this is not certain due to the absence of an RT correlation with Shape-Color switching effects (the truer representation of switching abilities). The RT correlations with Flanker are also difficult to interpret. While the congruent trials' correlation shows a connection between baseline processing speed and RT, the incongruent trials do not clearly represent inhibitory control (especially in the absence of a correlation with Flanker congruency effect, the truer representation of that ability). Thus, interpretations of the observed correlations are difficult, though imply a relationship between RT and cognitive abilities (task-switching and inhibitory control) at some level.

The large difference between the number of correlations in each grouping of data (four for all participants, none for Hispanophones, three for non-Hispanophones) may simply be a result of differing group sizes (88 data point overall, with 37 for Hispanophones and 51 for non-Hispanophones). However, given their Pearson's coefficients, it seems more likely that the Flanker correlations existed in both groups at less-significant levels until the inclusion of the other, while Shape-Color correlations were largely non-Hispanophone and were only weakened by the combination with Hispanophone data. The RT correlation with sex, which existed in all participants, was present in both groups but was significant in neither alone (each may have lacked a large enough sample to reach significance).

Thus, upon further inspection, it appears that Flanker results interacted with RT in both groups at some level, while Shape-Color effects were limited to non-Hispanophones. This helps narrow down the origins of the interactions seen with these covariates in the main ANCOVA, suggesting that task-switching skill (or, at least, whatever portion of it may be represented by the mixed trial type that interacted in the ANCOVA) played a role in RT for non-Hispanophones, but not Hispanophones, while inhibitory control (or what portion of it is represented by incongruent trials but not congruency effects) played a role in RT for both groups.

Finally, the lack of a significant difference between language groups in Flanker congruent trials as well as between Shape-Color congruent-stay trials suggest the two groups had similar baseline processing speeds. The Shape-Color and Flanker ANOVA suggest their inhibitory and task-switching skills were similar as well, given no interactions with language group. Thus, one can conclude that the difference in impact of those skills between groups (seen in their correlations with RT) did not relate to their expertise in them, but in their use of them in this paradigm. This may suggest that Hispanophones make use of different cognitive strategies in managing contextual interlingual conflict—a result perhaps connected to or resulting from the observed significant difference between the groups in contextual switching tendencies. The ACH predicts results such as these and recent research does, indeed, support these conclusions (Beatty-Martínez, Navarro-Torres, Dussias, et al., 2020; Green & Abutalebi, 2013). These results, more generally, substantiate recent calls for greater reporting of participant background, as its role in bilingual experimentation deserves greater attention than researchers afford it (de Bruin, 2019).

Summary. Taken together, while these results do support the view that context matters, future studies are needed to better elucidate the way in which it does or does not matter. It does not suggest that, generally, context affects RT. It does suggest that there are immediate effects of a change in an auditory stimulus' language, and it *may* suggest that long-term effects of such a stimulus differ depending on the language heard, though this interpretation of the data is not certain. In other words, this data points to a change in heard languages causing slower production that *may* abate at different rates depending on the language heard. More data is necessary in order to parcel out the impact of this paradigm on different bilinguals and the best interpretations of this data. Thus, whether or not bottom-up contextual language information affects top-down language control in bilinguals, and to what degree, will require much more research to discern.

Limitations. This research has a number of limitations. Given that it occurred during the COVID-19 pandemic, it took place entirely online. Experimenters wrote the code for this experiment with little training so issues with data collection did occur that, in part, led to high attrition. There was also high attrition due to participants losing interest in the study, even though it only lasted around 30 minutes. The original paradigm was 90 minutes long, so much of its efficacy was lost in its shortening for online administration. Thus, attrition and a time limit hurt data collection and led to losses in power and interpretation.

Online recording and RT-scoring presented its own challenges. As no readily available system existed for online recordings, the experimenters coded their own. The accuracy of this system, based on this research, seems reliable given the similarity in scores across groups and participants, though it is indeed not empirically evaluated. Subjects being able to participate under any given circumstance complicates matters—some subjects even carried conversations in the middle of the experiment or laughed with friends (and were not part of the final analysis as a result). As such, there is no true way to ensure that all participants paid their fullest attention (or even a reasonable amount throughout) to the task at hand. Also, recording quality and length differed between participants depending on their laptop, browser, microphone, and internet speed (some did not even use an external microphone, despite instructions to do so). This issue alone could have led to group differences—the difference in SES between groups could suggest that Hispanophones were working with older equipment, leading to lags in recording. Though this is far-fetched, since experimenters ensured only recordings of similar length stayed in the analysis and corrected for recording length differences, it could still prove to play a role in final results. In short, the hardships of online experimentation were on full display in this project and could have had large effects on its results.

Given that this experiment was not originally designed for between-group comparisons, the lack of a true baseline makes interpretation of between-group differences difficult. It also makes difficult the interpretation of any effect of context within-subjects. Future research would benefit from a clearly defined control block (i.e., one taking place before any experimental manipulations and featuring some kind of non-language noise distractor, to establish individual differences in image-naming speed). Further, the lack of monolingual participation makes it difficult to parcel out the effects of speaking multiple languages and other related confounds. On the note of speaking multiple languages, the non-Hispanophone group constituted individuals from a *number* of language backgrounds and language families, perhaps muddying any effects of specific languages. Further, only 28 of the final 88 participants identified English as their L1. This may explain the significance of English proficiency across analyses, since the diversity of experience in the language would be quite large. However, those 29 Anglophones did split nearly evenly between Hispanophones (17) and non-Hispanophones (12) (though there was a proportionally higher amount in the Hispanophones group), potentially dispersing any effects that may have had between groups. This, combined with the comparable English proficiency levels between groups, suggests the small number of English-L1 speakers alone does not explain the group difference. Either way, to help simplify explanations, future research would be benefited by limiting other-language groups to avoid effects that may be unique to one group of speakers being lost in the noise of other groups, and by encouraging more English-L1 participation (or, if that is not possible, ensuring comparable levels of proficiency and more comparable numbers of English-L1 speakers in each group).

Future research would also benefit from including short-length and long-length blocks in different participants to investigate the effects of more frequent context changes versus the longer-term effects of static contexts (i.e., whether RT do indeed reach a floor after a certain number of trials, and whether that floor would be in the ten-to-twenty-second window this research suggests). Greater numbers of trials and blocks would indeed be invaluable to isolating the remarkably small effects that this research suggests may exist in terms of the interaction between context and time spent in a context.

Concerning proficiency (a measure of individual differences that significantly affected outcomes in all-subject analyses), this research originally used a more empirical measure of English and Spanish proficiency (pre-online version) which would have been invaluable in understanding their effects on RT, so the inclusion of a more empirical measure in the future would be invaluable. Gambi and Hartsuiker (2016) found that non-target language proficiency modulated the effects of a switch to a non-target language auditory distractor. Even though this research suggests familiarity does not play a role in the impact of a distractor, future inclusion of reliable non-target language proficiency measures as a covariate would significantly aid in supporting or refuting these effects.

References

Anderson, J. A. E., Mak, L., Keyvani Chahi, A., & Bialystok, E. (2018). The language and social background questionnaire: Assessing degree of bilingualism in a diverse population. *Behavior Research Methods*, *50*(1), 250–263. https://doi.org/10.3758/s13428-017-0867-9

Ardila, A., Rosselli, M., Matute, E., & Guajardo, S. (2005). The Influence of the Parents' Educational Level on the Development of Executive Functions. *Developmental Neuropsychology*, *28*(1), 539–560. https://doi.org/10.1207/s15326942dn2801_5

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556–559. https://doi.org/10.1126/science.1736359

Bates, E., D'Amico, S., Jacobsen, T., Székely, A., Andonova, E., Devescovi, A., Herron, D., Ching Lu, C., Pechmann, T., Pléh, C., Wicha, N., Federmeier, K., Gerdjikova, I., Gutierrez, G., Hung, D., Hsu, J., Iyer, G., Kohnert, K., Mehotcheva, T., … Tzeng, O. (2003). Timed picture naming in seven languages. *Psychonomic Bulletin & Review*, *10*(2), 344–380. https://doi.org/10.3758/BF03196494

Beatty-Martínez, A. L., Navarro-Torres, C. A., & Dussias, P. E. (2020). Codeswitching: A Bilingual Toolkit for Opportunistic Speech Planning. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.01699

Beatty-Martínez, A. L., Navarro-Torres, C. A., Dussias, P. E., Bajo, M. T., Guzzardo Tamargo, R. E., & Kroll, J. F. (2020). Interactional context mediates the consequences of bilingualism for language and cognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(6), 1022–1047. https://doi.org/10.1037/xlm0000770

Bialystok, E., Craik, F. I. M., Green, D. W., & Gollan, T. H. (2009). Bilingual Minds. *Psychological Science in the Public Interest*, *10*(3), 89–129. https://doi.org/10.1177/1529100610387084

Calvo, A., & Bialystok, E. (2014). Independent effects of bilingualism and socioeconomic status on language ability and executive functioning. *Cognition*, *130*(3), 278–288. https://doi.org/10.1016/j.cognition.2013.11.015

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(3), 201–215. https://doi.org/10.1038/nrn755

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

de Bruin, A. (2019). Not All Bilinguals Are the Same: A Call for More Detailed Assessments and Descriptions of Bilingual Experiences. *Behavioral Sciences*, *9*(3), 33. https://doi.org/10.3390/bs9030033

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

de Leeuw, J. R., & Motz, B. A. (2016). Psychophysics in a Web browser? Comparing response times collected with JavaScript and Psychophysics Toolbox in a visual search task. *Behavior Research Methods*, *48*(1), 1–12. https://doi.org/10.3758/s13428-015-0567-2

Declerck, M., & Philipp, A. M. (2017). Is there lemma-based language control? The influence of language practice and language-specific item practice on asymmetrical switch costs.

Language, Cognition and Neuroscience, *32*(4), 488–493. https://doi.org/10.1080/23273798.2016.1250928

Dell, G. S., Juliano, C., & Govindjee, A. (1993). Structure and content in language production: A theory of frame constraints in phonological speech errors. *Cognitive Science*, *17*(2), 149–195. https://doi.org/10.1016/0364-0213(93)90010-6

Engel de Abreu, P. M. J., Cruz-Santos, A., Tourinho, C. J., Martin, R., & Bialystok, E. (2012). Bilingualism Enriches the Poor: Enhanced Cognitive Control in Low-Income Minority Children. *Psychological Science*, *23*(11), 1364–1371. https://doi.org/10.1177/0956797612443836

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149. https://doi.org/10.3758/BF03203267

Gambi, C., & Hartsuiker, R. J. (2016). If you stay, it might be easier: Switch costs from comprehension to production in a joint switching task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(4), 608–626. https://doi.org/10.1037/xlm0000190

Grainger, J., & Dijkstra, T. (1992). On the Representation and Use of Language Information in Bilinguals. In R. J. Harris (Ed.), *Advances in Psychology* (Vol. 83, pp. 207–220). North-Holland. https://doi.org/10.1016/S0166-4115(08)61496-X

Grainger, J., Midgley, K., & Holcomb, P. J. (2010). Chapter 14. Re-thinking the bilingual interactive-activation model from a developmental perspective (BIA-d). In *Language Acquisition across Linguistic and Cognitive Systems*. John Benjamins. https://www.jbe-platform.com/content/books/9789027287564-lald.52.18gra

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, *25*(5), 515–530. https://doi.org/10.1080/20445911.2013.796377

Grosjean, F. (1999). *The bilingual's language modes*. 27.

Hackman, D. A., Farah, M. J., & Meaney, M. J. (2010). Socioeconomic status and the brain: Mechanistic insights from human and animal research. *Nature Reviews Neuroscience*, *11*(9), 651–659. https://doi.org/10.1038/nrn2897

Hartanto, A., & Yang, H. (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. https://doi.org/10.1016/j.cognition.2016.01.016

Hilbig, B. E. (2016). Reaction time effects in lab- versus Web-based research: Experimental evidence. *Behavior Research Methods*, *48*(4), 1718–1724. https://doi.org/10.3758/s13428-015-0678-9

Jakubów, A. P. D. S. P., & Corrêa, L. M. S. (2019). Variation under a psycholinguistic perspective: Number agreement in Brazilian Portuguese and its impact for the schooling. *Ilha Do Desterro A Journal of English Language, Literatures in English and Cultural Studies*, *72*(3), 101–122. https://doi.org/10.5007/2175-8026.2019v72n3p101

JASP Team. (2020). *JASP (Version 0.14.1)[Computer software]*. https://jasp-stats.org/

Katsuki, F., & Constantinidis, C. (2014). Bottom-Up and Top-Down Attention: Different Processes and Overlapping Neural Systems. *The Neuroscientist*, *20*(5), 509–521. https://doi.org/10.1177/1073858413514136

Kray, J., & Lindenberger, U. (2000). Adult age differences in task switching. *Psychology and Aging*, *15*(1), 126–147. https://doi.org/10.1037//0882-7974.15.1.126

Lagrou, E., Hartsuiker, R., & Duyck, W. (2011). Knowledge of a Second Language Influences Auditory Word Recognition in the Native Language. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *37*, 952–965. https://doi.org/10.1037/a0023217

Lange, K., Kühn, S., & Filevich, E. (2015). "Just Another Tool for Online Studies" (JATOS): An Easy Solution for Setup and Management of Web Servers Supporting Online Studies. *PLOS ONE*, *10*(6), e0130834. https://doi.org/10.1371/journal.pone.0130834

Levelt, W. J. M. (1999). Models of word production. *Trends in Cognitive Sciences*, *3*(6), 223–232. https://doi.org/10.1016/S1364-6613(99)01319-4

Meuter, R. F. I., & Allport, A. (1999). Bilingual Language Switching in Naming: Asymmetrical Costs of Language Selection. *Journal of Memory and Language*, *40*(1), 25–40. https://doi.org/10.1006/jmla.1998.2602

Office of the Texas Governor. (2020). *Governor Abbott's Proactive Response to the Coronavirus Threat: A high-level snapshot of actions taken by Governor Abbott; not inclusive of all briefings, hearings, proclamations, calls, actions or related events in Texas.* https://open.texas.gov/uploads/files/organization/opentexas/OpenTexas-Report.pdf

Peeters, D., Runnqvist, E., Bertrand, D., & Grainger, J. (2014). Asymmetrical switch costs in bilingual language production induced by reading words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(1), 284–292. https://doi.org/10.1037/a0034060

Petersen, S. E., & Posner, M. I. (2012). The Attention System of the Human Brain: 20 Years After. *Annual Review of Neuroscience*, *35*(1), 73–89. https://doi.org/10.1146/annurev-neuro-062111-150525

Posner, M. I., & Petersen, S. E. (1990). The Attention System of the Human Brain. *Annual Review of Neuroscience*, *13*(1), 25–42. https://doi.org/10.1146/annurev.ne.13.030190.000325

R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Richardson, J. T. E. (2007). Measures of short-term memory: A historical review. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *43*(5), 635–650. https://doi.org/10.1016/s0010-9452(08)70493-3

Ridderinkhof, K. R., Band, G. P. H., & Logan, G. D. (1999). A study of adaptive behavior: E□ects of age and irrelevant information on the ability to inhibit one's actions. *Acta Psychologica*, 23.

Rodriguez-Fornells, A., Krämer, U. M., Lorenzo-Seva, U., Festman, J., & Münte, T. F. (2012). Self-Assessment of Individual Differences in Language Switching. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00388

Rodríguez-Pujadas, A., Sanjuán, A., Ventura-Campos, N., Román, P., Martin, C., Barceló, F., Costa, A., & Ávila, C. (2013). Bilinguals Use Language-Control Brain Areas More Than Monolinguals to Perform Non-Linguistic Switching Tasks. *PLoS ONE*, *8*(9), e73028. https://doi.org/10.1371/journal.pone.0073028

Szekely, A., Damico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., Jacobsen, T., Arevalo, A., Vargha, A., & Bates, E. (2005). Timed Action and Object Naming. *Cortex*, *41*(1), 7–25. https://doi.org/10.1016/S0010-9452(08)70174-6

University of Houston Department of Institutional Research. (2020). *University of Houston Statistical Handbook*. https://uh.edu/ir/reports/new-statistical-handbook/

U.S. Census Bureau. (2020). *American Community Survey 2020*. https://www.census.gov/acs/www/data/data-tables-and-tools/data-profiles/

van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic Neighborhood Effects in Bilingual Word Recognition. *Journal of Memory and Language*, *39*(3), 458–483. https://doi.org/10.1006/jmla.1998.2584

Wechsler, D. (2003). *WISC-IV: Administration and scoring manual*. Psychological Corporation.

Zheng, X., Roelofs, A., & Lemhöfer, K. (2018). Language selection errors in switching: Language priming or cognitive control? *Language, Cognition and Neuroscience*, *33*(2), 139–147. https://doi.org/10.1080/23273798.2017.1363401

Figures and Tables

Table 1. All-Subjects Descriptive Statistics. *Information collected on potential confounds relevant to this research.*

	n	Mean	Std. Dev.	Minimum	Maximum
flAcc	88	92.659	15.132	12.667	100
flCongRT	88	699.042	226.119	174.556	1653.151
flDiffRT	88	58.434	99.911	-206.623	407.76
scAcc	88	66.92	9.94	33	80
scConStayRT	88	1010.524	323.519	231.083	2029.286
scDiffRT	88	67.049	172.35	-439.464	689.575
dsAcc	88	60.248	25.638	0	94.643
AvRT	88	1.217	0.136	0.882	1.544
EnProf	88	9.469	0.78	6.75	10
SES	88	5.648	2.569	2	10
MomL1	88	2.341	0.623	1	3
DadL1	88	2.386	0.596	1	3
Age	86	21.977	2.735	19	36
AgeUS	88	2.898	5.5	0	18
Sex	88	1.841	0.368	1	2
numLangs	88	2.375	0.748	2	5
EnAoA	88	4	3.836	0	16
bsqCtxS	88	9.091	2.966	3	14
bsqAccS	88	7.352	1.924	3	14

Note. **Ds** = Digit Span; **fl** = Flanker; **sc** = Shape-Color. **Acc** (ds, fl, and sc) are accuracy scores. **flCongRT** is the RT of congruent Flanker trials. **flDiffRT** is the Flanker congruency effect. **scConStayRT** is the congruent-stay trial-type RT from Shape-Color. **scDiffRT** is the Shape-Color switching effect. **AvRT** is the overall average RT across all trials. **EnProf** is English Proficiency. **SES** is socioeconomic status. **MomL1** and **DadL1** are parents' first languages. **Age** is subject age. **AgeUS** is Age of Arrival to United States. **Sex** is self-identified sex of subject. **numLangs** is the Number of Languages Spoken. **EnAoA** is English Age of Acquisition. **bsqCtxS** is Contextual Switching Tendencies from the BSWQ, **bsqAccS** is Accidental Switching Tendencies from the BSWQ. **langGroup** is Language Group.

Table 2. Frequency Statistics of Language Spoken. *For all subjects, Hispanophones, and non-Hispanophones.*

All Subjects

L1		L2		L3		L4		L5	
Language	*n*	Language	*n*	Language	*n*	Language	*n*	Language	*n*
vietnamese	11	english	58	none	69	none	82	none	87
chinese	1	spanish	17	spanish	6	korean	1	telugu	1
spanish	31	vietnamese	4	italian	1	french	1		0
english	29	hindi	3	farsi	1	marwadi	1		0
polish	1	mandarin	1	english	1	hindi	1		0
urdu	6	tagalog	1	chinese	1	arabic	1		0
cantonese	1	malayalam	1	magahi	1	tamil	1		0
russian	1	arabic	2	french	2		0		0
malayalam	3	urdu	1	marathi	1		0		0
hindi	1		0	sindhi	1		0		0
filipino	1		0	memoni	1		0		0
tagalog	1		0	punjabi	1		0		0
igbo	1		0	bangla	1		0		0
	0		0	hindi	1		0		0

Hispanophones

L1		L2		L3		L4		L5	
Language	*n*	Language	*n*	Language	*n*	Language	*n*	Language	*n*
spanish	31	english	34	none	44	none	50	none	51
english	17	spanish	16	spanish	4	korean	1		0
vietnamese	1	vietnamese	1	italian	1		0		0
malayalam	2		0	farsi	1		0		0
	0		0	french	1		0		0

Non-Hispanophones

L1		L2		L3		L4		L5	L1
Language	*n*	Language	*n*	Language	*n*	Language	*n*	Language	*n*
vietnamese	10	english	24	none	27	none	32	none	36
chinese	1	vietnamese	3	english	1	french	1	telugu	1
polish	1	hindi	3	chinese	1	marwadi	1		0
urdu	6	mandarin	2	magahi	1	hindi	1		0
english	12	tagalog	1	french	1	arabic	1		0

cantonese	1	malayalam	1	marathi	1	tamil	1	0
russian	1	arabic	2	marathi	1		0	0
hindi	1	urdu	1	sindhi	1		0	0
filipino	1		0	memoni	1		0	0
tagalog	1		0	punjabi	1		0	0
igbo	1		0	bangla	1		0	0
malayalam	1		0	hindi	1		0	0

Table 3. All-Subjects Grouped by Languages Spoken Descriptive Statistics. *Information on collected information relevant to this research.*

	Group	*n*	Mean	Std. Dev.	Minimum	Maximum
flAcc	**1**	51	93.216	16.499	12.667	100
	2	37	91.892	13.201	46	100
flCongRT	**1**	51	734.004	250.782	388.324	1653.151
	2	37	650.852	179.153	174.556	1173.557
flDiffRT	**1**	51	51.515	110.33	-206.623	407.76
	2	37	67.973	84	-134.613	268.42
scAcc	**1**	51	68.255	9.796	33	80
	2	37	65.081	9.976	36	76
scConStayRT	**1**	51	1039.686	372.766	397.917	2029.286
	2	37	970.327	238.916	231.083	1743.727
scDiffRT	**1**	51	72.26	186.255	-439.464	689.575
	2	37	59.867	153.337	-285.085	454.1
dsAcc	**1**	51	58.894	25.1	0	92.857
	2	37	62.114	26.595	0	94.643
AvRT	**1**	51	1.246	0.14	0.985	1.544
	2	37	1.176	0.121	0.882	1.538
EnProf	**1**	51	9.431	0.778	6.75	10
	2	37	9.52	0.791	7	10
SES	**1**	51	4.98	2.445	2	10
	2	37	6.568	2.478	2	10
MomL1	**1**	51	1.98	0.374	1	3
	2	37	2.838	0.553	1	3
DadL1	**1**	51	2.039	0.398	1	3
	2	37	2.865	0.481	1	3
Age	**1**	49	22.286	3.109	19	36
	2	37	21.568	2.115	19	28
AgeUS	**1**	51	1.51	3.706	0	16
	2	37	4.811	6.895	0	18
Sex	**1**	51	1.843	0.367	1	2
	2	37	1.838	0.374	1	2
numLangs	**1**	51	2.216	0.577	2	5
	2	37	2.595	0.896	2	5
EnAoA	**1**	51	4.431	3.754	0	16
	2	37	3.405	3.919	0	16
bsqCtxS	**1**	51	9.784	2.7	3	14
	2	37	8.135	3.084	3	14

bsqAccS	**1**	51	7.686	1.703	3	13
	2	37	6.892	2.132	3	14

Note. See Table 1 for descriptions of names.
Note. Hispanophones (1) and non-Hispanophones (2).

Table 4. Correctness Descriptive Statistics. *Number of incorrect trials per context and group with relevant statistics.*

	HU		**EN**		**SP**		**ENSP**	
	1	**2**	**1**	**2**	**1**	**2**	**1**	**2**
Valid	51	37	51	37	51	37	51	37
Missing	0	0	0	0	0	0	0	0
Mean	7.863	7.811	6.804	8.000	7.843	8.270	7.333	7.730
Std. Deviation	5.095	5.929	5.091	4.601	5.640	6.623	5.256	5.269
Variance	25.961	35.158	25.921	21.167	31.815	43.869	27.627	27.758
Minimum	1.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000
Maximum	22.000	27.000	19.000	21.000	22.000	27.000	20.000	22.000

Note. Hispanophones (1) and non-Hispanophones (2).

Table 5. Hispanophone Whole-Block Results. *Within and between-subject effects from Hispanophone Whole-Block analysis, as well as descriptive statistics.*

Within Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p	η^2_p
Context	0.002	3	6.195e -4	0.378	0.769	0.008
Residuals	0.246	150	0.002			

Note. Type III Sum of Squares

Between Subjects Effects

Cases	Sum of Squares	df	Mean Square
Residuals	3.919	50	0.078

Note. Type III Sum of Squares

Descriptives

Context	Mean	SD	N
EN	1.250	0.146	51
ENSP	1.245	0.152	51
HU	1.249	0.142	51
SP	1.242	0.136	51

Table 6. Hispanophone First-Thirds Results. *Within and between-subject effects from Hispanophone First-Thirds analysis showing a near-significant effect of Context on RT.*

Within Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p	η^2_p
Context	0.033	3	0.011	2.370	0.073	0.045
Residuals	0.698	150	0.005			

Note. Type III Sum of Squares

Between Subjects Effects

Cases	Sum of Squares	df	Mean Square
Residuals	4.022	50	0.080

Note. Type III Sum of Squares

Descriptives

Context	Mean	SD	N
EN	1.273	0.149	51
ENSP	1.254	0.162	51
HU	1.288	0.150	51
SP	1.262	0.153	51

Table 7. Hispanophone First-Last Thirds Results. *Within and between-subject effects from Hispanophone First-Last Thirds analyses showing a significant effect of Thirds in Hungarian and English.*

HU Within Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p	η^2_p
Third	0.090	1	0.090	24.104	< .001	0.325
Residuals	0.187	50	0.004			

Note. Type III Sum of Squares

HU Between Subjects Effects

Cases	Sum of Squares	df	Mean Square
Residuals	1.968	50	0.039

Note. Type III Sum of Squares

HU Descriptives

Third	Mean	SD	N
First	1.288	0.150	51
Last	1.228	0.144	51

EN Within Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p	η^2_p
Third	0.037	1	0.037	7.603	0.008	0.132
Residuals	0.245	50	0.005			

Note. Type III Sum of Squares

EN Between Subjects Effects

Cases	Sum of Squares	df	Mean Square
Residuals	2.187	50	0.044

Note. Type III Sum of Squares

EN Descriptives

Third	Mean	SD	N
First	1.273	0.149	51
Last	1.235	0.163	51

SP Within Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p	η^2_p
Third	0.017	1	0.017	3.472	0.068	0.065
Residuals	0.241	50	0.005			

Note. Type III Sum of Squares

SP Between Subjects Effects

Cases	Sum of Squares	df	Mean Square
Residuals	2.029	50	0.041

Note. Type III Sum of Squares

SP Descriptives

Third	Mean	SD	N
First	1.262	0.153	51
Last	1.237	0.148	51

ENSP Within Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p	η^2_p
Third	0.011	1	0.011	2.666	0.109	0.051
Residuals	0.213	50	0.004			

Note. Type III Sum of Squares

ENSP Between Subjects Effects

Cases	Sum of Squares	df	Mean Square
Residuals	2.308	50	0.046

Note. Type III Sum of Squares

ENSP Descriptives

Third	Mean	SD	N
First	1.254	0.162	51
Last	1.233	0.155	51

Table 8. All-Subjects Context by Time by Group Assumption Checks. *Ensuring homogeneity of variance and sphericity across all levels of the 4x3x2 ANCOVA. ENSP1 and ENSP3 violate homogeneity of variance, Time violates sphericity. Upon reviewing corrections, it was evident neither of these significantly impacted results.*

Test for Equality of Variances (Levene's)

	F	df1	df2	p
HU1	1.441	1	86	0.233
HU2	3.234	1	86	0.076
HU3	1.106	1	86	0.296
EN1	0.878	1	86	0.351
EN2	3.074	1	86	0.083
EN3	1.080	1	86	0.302
SP1	2.069	1	86	0.154
SP2	0.004	1	86	0.949
SP3	1.955	1	86	0.166
ENSP1	4.274	1	86	0.042
ENSP2	1.368	1	86	0.245
ENSP3	4.203	1	86	0.043

Note. 1, 2, and 3 refer to First, Middle, and Last Thirds.

Test of Sphericity

	Mauchly's W	Approx. X^2	dfSphericity	p-value	Greenhouse-Geisser ε	Huynh-Feldt ε	Lower Bound ε
Context	0.969	2.412	5	0.790	0.980	1.000	0.333
Time	0.906	7.622	2	0.022	0.914	0.935	0.500
Context ✻ Time	0.778	19.009	20	0.522	0.931	1.000	0.167

Table 9. All-Subjects Context by Time by Group Results. *Showing the significant effect of the Context by Time interaction, an interaction between the two and English proficiency, and a significant between-subjects effect of language group. No significant effect of Time arose, though its effect size was small-to-moderate, and it significantly interacted with Shape-Color switch effects.*

Within Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p	η^2_p
Context	0.022	3	0.007	1.553	0.202	0.020
Context ✻ langGroup	0.010	3	0.003	0.717	0.543	0.009
Context ✻ flDiffRT	0.015	3	0.005	1.081	0.358	0.014
Context ✻ scConStayRT	0.007	3	0.002	0.485	0.693	0.006
Context ✻ EnProf	0.019	3	0.006	1.309	0.272	0.017
Context ✻ SES	0.013	3	0.004	0.951	0.416	0.012
Context ✻ bsqCtxS	0.007	3	0.002	0.466	0.707	0.006
Context ✻ bsqAccS	0.009	3	0.003	0.614	0.607	0.008
Context ✻ flCongRT	0.023	3	0.008	1.636	0.182	0.021
Context ✻ scDiffRT	0.017	3	0.006	1.223	0.302	0.015
Residuals	1.104	234	0.005			
Time	0.026 [a]	2 [a]	0.013 [a]	2.603 [a]	0.077 [a]	0.032
Time ✻ langGroup	0.004 [a]	2 [a]	0.002 [a]	0.454 [a]	0.636 [a]	0.006
Time ✻ flDiffRT	0.010 [a]	2 [a]	0.005 [a]	1.032 [a]	0.359 [a]	0.013
Time ✻ scConStayRT	0.021 [a]	2 [a]	0.011 [a]	2.178 [a]	0.117 [a]	0.027
Time ✻ EnProf	0.016 [a]	2 [a]	0.008 [a]	1.580 [a]	0.209 [a]	0.020
Time ✻ SES	0.027 [a]	2 [a]	0.014 [a]	2.779 [a]	0.065 [a]	0.034
Time ✻ bsqCtxS	0.004 [a]	2 [a]	0.002 [a]	0.367 [a]	0.694 [a]	0.005
Time ✻ bsqAccS	0.005 [a]	2 [a]	0.003 [a]	0.532 [a]	0.588 [a]	0.007
Time ✻ flCongRT	0.014 [a]	2 [a]	0.007 [a]	1.391 [a]	0.252 [a]	0.018
Time ✻ scDiffRT	0.037 [a]	2 [a]	0.018 [a]	3.751 [a]	0.026 [a]	0.046
Residuals	0.767	156	0.005			
Context ✻ Time	0.056	6	0.009	2.591	0.018	0.032
Context ✻ Time ✻ langGroup	0.021	6	0.004	0.992	0.430	0.013
Context ✻ Time ✻ flDiffRT	0.026	6	0.004	1.227	0.291	0.015
Context ✻ Time ✻ scConStayRT	0.032	6	0.005	1.468	0.187	0.018
Context ✻ Time ✻ EnProf	0.050	6	0.008	2.301	0.034	0.029
Context ✻ Time ✻ SES	0.015	6	0.002	0.674	0.670	0.009
Context ✻ Time ✻ bsqCtxS	0.021	6	0.003	0.965	0.449	0.012
Context ✻ Time ✻ bsqAccS	0.019	6	0.003	0.871	0.516	0.011
Context ✻ Time ✻ flCongRT	0.021	6	0.003	0.970	0.445	0.012

Within Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p	η^2_p
Context ✻ Time ✻ scDiffRT	0.017	6	0.003	0.803	0.568	0.010
Residuals	1.678	468	0.004			

Note. Type III Sum of Squares
Note. Explanation of names in *Table 1* note.
[a] Mauchly's test of sphericity indicates that the assumption of sphericity is violated (p < .05).

Between Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p	η^2_p
langGroup	0.932	1	0.932	4.446	0.038	0.054
flDiffRT	0.047	1	0.047	0.225	0.636	0.003
scConStayRT	0.021	1	0.021	0.102	0.751	0.001
EnProf	0.403	1	0.403	1.923	0.169	0.024
SES	0.129	1	0.129	0.615	0.435	0.008
bsqCtxS	0.065	1	0.065	0.312	0.578	0.004
bsqAccS	0.235	1	0.235	1.124	0.292	0.014
flCongRT	0.465	1	0.465	2.218	0.140	0.028
scDiffRT	0.020	1	0.020	0.098	0.756	0.001
Residuals	16.345	78	0.210			

Note. Type III Sum of Squares
Note. Explanation of names in *Table 1* note.

Table 10. All-Subjects Context by Time by Group Post-Hoc Comparisons. *For Time, explored because of the significance of Time in similar analyses and a small-to-moderate effect size in this analysis, a significant difference arose between First-Thirds and Middle-Thirds of trials. For Context by Time, results are more complex, though comparing within-context Thirds comparisons adds clarity (as discussed in paper).*

Post Hoc Comparisons - Time

			95% CI for Mean Difference				
		Mean Difference	**Lower**	**Upper**	**SE**	**t**	**p holm**
X1	X2	0.031	0.018	0.044	0.005	5.722	< .001 ***
	X3	0.037	0.024	0.050	0.005	6.846	< .001 ***
X2	X3	0.006	-0.007	0.019	0.005	1.124	0.263

*** p < .001

Note. P-value and confidence intervals adjusted for comparing a family of 3 estimates (confidence intervals corrected using the bonferroni method).

Note. Results are averaged over the levels of: langGroup, Context

Post Hoc Comparisons - Context ✻ Time

			95% CI for Mean Difference				
		Mean Difference	**Lower**	**Upper**	**SE**	**t**	**p holm**
HU, X1	EN, X1	0.006	-0.027	0.038	0.010	0.595	1.000
	SP, X1	0.018	-0.015	0.050	0.010	1.851	1.000
	ENSP, X1	0.014	-0.019	0.046	0.010	1.411	1.000
	HU, X2	0.043	0.010	0.075	0.010	4.469	< .001 ***
	EN, X2	0.031	-0.003	0.065	0.010	3.102	0.090
	SP, X2	0.048	0.014	0.082	0.010	4.786	< .001 ***
	ENSP, X2	0.038	0.004	0.072	0.010	3.790	0.009 **
	HU, X3	0.046	0.013	0.078	0.010	4.751	< .001 ***
	EN, X3	0.039	0.005	0.073	0.010	3.843	0.007 **
	SP, X3	0.048	0.014	0.082	0.010	4.825	< .001 ***
	ENSP, X3	0.052	0.018	0.086	0.010	5.145	< .001 ***
EN, X1	SP, X1	0.012	-0.021	0.045	0.010	1.256	1.000
	ENSP, X1	0.008	-0.025	0.040	0.010	0.816	1.000
	HU, X2	0.037	0.003	0.071	0.010	3.697	0.013 *
	EN, X2	0.025	-0.007	0.058	0.010	2.650	0.347
	SP, X2	0.042	0.008	0.076	0.010	4.215	0.002 **
	ENSP, X2	0.032	-0.002	0.066	0.010	3.218	0.066
	HU, X3	0.040	0.006	0.074	0.010	3.966	0.005 **
	EN, X3	0.033	3.861e -4	0.065	0.010	3.425	0.033 *
	SP, X3	0.043	0.009	0.077	0.010	4.254	0.001 **
	ENSP, X3	0.046	0.012	0.080	0.010	4.574	< .001 ***

Post Hoc Comparisons - Context ✻ Time

		Mean Difference	95% CI for Mean Difference		SE	t	p holm	
			Lower	Upper				
SP, X1	ENSP, X1	-0.004	-0.037	0.028	0.010	-0.440	1.000	
	HU, X2	0.025	-0.009	0.059	0.010	2.491	0.519	
	EN, X2	0.013	-0.021	0.047	0.010	1.325	1.000	
	SP, X2	0.030	-0.002	0.063	0.010	3.150	0.079	
	ENSP, X2	0.020	-0.014	0.054	0.010	2.012	1.000	
	HU, X3	0.028	-0.006	0.062	0.010	2.760	0.255	
	EN, X3	0.021	-0.013	0.055	0.010	2.065	1.000	
	SP, X3	0.031	-0.002	0.063	0.010	3.191	0.072	
	ENSP, X3	0.034	-1.452e -4	0.068	0.010	3.368	0.040	*
ENSP, X1	HU, X2	0.029	-0.005	0.063	0.010	2.913	0.162	
	EN, X2	0.018	-0.016	0.051	0.010	1.747	1.000	
	SP, X2	0.034	4.907e -4	0.068	0.010	3.431	0.033	*
	ENSP, X2	0.024	-0.008	0.057	0.010	2.549	0.453	
	HU, X3	0.032	-0.002	0.066	0.010	3.182	0.072	
	EN, X3	0.025	-0.009	0.059	0.010	2.487	0.519	
	SP, X3	0.035	8.826e -4	0.069	0.010	3.470	0.029	*
	ENSP, X3	0.038	0.006	0.070	0.010	3.968	0.005	**
HU, X2	EN, X2	-0.012	-0.044	0.021	0.010	-1.214	1.000	
	SP, X2	0.005	-0.027	0.038	0.010	0.539	1.000	
	ENSP, X2	-0.005	-0.037	0.028	0.010	-0.499	1.000	
	HU, X3	0.003	-0.030	0.035	0.010	0.282	1.000	
	EN, X3	-0.004	-0.038	0.030	0.010	-0.425	1.000	
	SP, X3	0.006	-0.028	0.040	0.010	0.557	1.000	
	ENSP, X3	0.009	-0.025	0.043	0.010	0.877	1.000	
EN, X2	SP, X2	0.017	-0.016	0.050	0.010	1.753	1.000	
	ENSP, X2	0.007	-0.026	0.040	0.010	0.715	1.000	
	HU, X3	0.014	-0.020	0.048	0.010	1.435	1.000	
	EN, X3	0.007	-0.025	0.040	0.010	0.775	1.000	
	SP, X3	0.017	-0.017	0.051	0.010	1.723	1.000	
	ENSP, X3	0.021	-0.013	0.054	0.010	2.043	1.000	
SP, X2	ENSP, X2	-0.010	-0.043	0.023	0.010	-1.038	1.000	
	HU, X3	-0.002	-0.036	0.031	0.010	-0.249	1.000	
	EN, X3	-0.009	-0.043	0.024	0.010	-0.943	1.000	
	SP, X3	3.920e -4	-0.032	0.033	0.010	0.041	1.000	
	ENSP, X3	0.004	-0.030	0.038	0.010	0.359	1.000	
ENSP, X2	HU, X3	0.008	-0.026	0.041	0.010	0.748	1.000	
	EN, X3	5.356e -4	-0.033	0.034	0.010	0.053	1.000	

Post Hoc Comparisons - Context ✻ Time

		Mean Difference	95% CI for Mean Difference		SE	t	p_{holm}
			Lower	Upper			
	SP, X3	0.010	-0.024	0.044	0.010	1.036	1.000
	ENSP, X3	0.014	-0.019	0.046	0.010	1.419	1.000
HU, X3	EN, X3	-0.007	-0.040	0.026	0.010	-0.723	1.000
	SP, X3	0.003	-0.030	0.036	0.010	0.300	1.000
	ENSP, X3	0.006	-0.027	0.039	0.010	0.633	1.000
EN, X3	SP, X3	0.010	-0.023	0.042	0.010	1.023	1.000
	ENSP, X3	0.013	-0.020	0.046	0.010	1.356	1.000
SP, X3	ENSP, X3	0.003	-0.029	0.036	0.010	0.333	1.000

Note. Results are averaged over the levels of: langGroup

Note. P-value and confidence intervals adjusted for comparing a family of 66 estimates (confidence intervals corrected using the bonferroni method).

* $p < .05$, ** $p < .01$, *** $p < .001$

Table 11. Measures of Individual Differences Comparisons and Descriptive Statistics. *T-Tests evaluating group differences in individual differences and the averages of each measure per groups.*

Independent Samples T-Test

	t	df	p
flAcc	0.403	86	0.688
flCongRT	1.722	86	0.089
flDiffRT	-0.761	86	0.449
scAcc	1.489	86	0.140
scStayRT	0.581	86	0.563
scDiffRT	0.331	86	0.741
dsAcc	-0.579	86	0.564
EnProf	-0.525	86	0.601
SES	-2.989	86	0.004
MomL1	-8.677	86	< .001
DadL1	-8.797	86	< .001
AgeUS	-2.895	86	0.005 [a]
numLangs	-2.411	86	0.018 [a]
EnAoA	1.242	86	0.217
Age	1.209	84	0.230
bsqCtxS	2.664	86	0.009
bsqAccS	1.942	86	0.055

Note. Student's t-test.

Note. Explanation of names in *Table 1* note.

[a] Levene's test is significant ($p < .05$), suggesting a violation of the equal variance assumption

Table 12. Shape-Color ANOVA Results. *Showing the significant effects of congruency and switching on RT with no significant effect of or interaction with language group.*

Within Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p	η^2_p
Congruency	559867.351	1	559867.351	19.820	< .001	0.187
Congruency ✻ langGroup	743.014	1	743.014	0.026	0.872	3.058e -4
Residuals	2.429e +6	86	28247.409			
Switching	395762.803	1	395762.803	12.933	< .001	0.131
Switching ✻ langGroup	1637.445	1	1637.445	0.054	0.818	6.218e -4
Residuals	2.632e +6	86	30600.898			
Congruency ✻ Switching	2021.794	1	2021.794	0.081	0.776	9.426e -4
Congruency ✻ Switching ✻ langGroup	61845.933	1	61845.933	2.482	0.119	0.028
Residuals	2.143e +6	86	24917.277			

Note. Type III Sum of Squares
Note. Explanation of names in *Table 1* note.

Between Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p	η^2_p
langGroup	165550.288	1	165550.288	0.396	0.531	0.005
Residuals	3.598e +7	86	418316.060			

Note. Type III Sum of Squares
Note. Explanation of names in *Table 1* note.

Descriptives

Congruency	Switching	langGroup	Mean	SD	N
Congruent	Stay	1	1039.686	372.766	51
		2	970.327	238.916	37
	Switch	1	1080.275	362.187	51
		2	1055.882	317.203	37
Incongruent	Stay	1	1085.828	364.600	51
		2	1076.060	351.748	37
	Switch	1	1189.831	404.033	51
		2	1117.620	362.560	37

Table 13. Flanker ANOVA Results. *Showing the significant effects of congruency on RT with no significant effect of or interaction with language group.*

Within Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p	η^2_p
Congruency	153073.881	1	153073.881	30.521	< .001	0.262
Congruency ✻ langGroup	2904.118	1	2904.118	0.579	0.449	0.007
Residuals	431324.388	86	5015.400			

Note. Type III Sum of Squares
Note. Explanation of names in *Table 1* note.

Between Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p	η^2_p
langGroup	240740.794	1	240740.794	2.636	0.108	0.030
Residuals	7.855e +6	86	91338.885			

Note. Type III Sum of Squares
Note. Explanation of names in *Table 1* note.

Descriptives

Congruency	langGroup	Mean	SD	N
Congruent	1	734.004	250.782	51
	2	650.852	179.153	37
Incongruent	1	785.518	234.962	51
	2	718.824	184.547	37

Table 14. Correlations between Measures of Individual Differences and RT. *Correlations (Rho, p-values, and number of subjects) between measures of individual differences and block-wise average RTs across all subjects (overall), Hispanophones, and Non-Hispanophones.*

		Overall RT	Hisp RT	non-Hisp RT
flAcc	Pearson's r	0.008	-0.104	0.199
	p-value	0.939	0.467	0.237
flCongRT	Pearson's r	0.288	0.257	0.249
	p-value	0.007**	0.069	0.137
flIncongRT	Pearson's r	0.272	0.195	0.344
	p-value	0.01*	0.17	0.037*
flDiffRT	Pearson's r	-0.061	-0.169	0.224
	p-value	0.57	0.236	0.182
scAcc	Pearson's r	0.127	0.132	0.027
	p-value	0.237	0.356	0.876
scConStayRT	Pearson's r	0.151	0.063	0.297
	p-value	0.161	0.663	0.074
scInconStayRT	Pearson's r	0.197	0.107	0.356
	p-value	0.066	0.456	0.031*
scConSwitchRT	Pearson's r	0.221	0.114	0.412
	p-value	0.038*	0.424	0.011*
scInconSwitchRT	Pearson's r	0.179	0.102	0.267
	p-value	0.095	0.476	0.11
scDiffRT	Pearson's r	0.105	0.086	0.126
	p-value	0.329	0.55	0.456
dsAcc	Pearson's r	0.087	0.074	0.157
	p-value	0.422	0.607	0.353
EnProf	Pearson's r	-0.09	-0.028	-0.157
	p-value	0.407	0.848	0.353
SES	Pearson's r	-0.033	-0.031	0.177
	p-value	0.758	0.828	0.296
MomL1	Pearson's r	-0.181	-0.147	0.141
	p-value	0.092	0.305	0.405
DadL1	Pearson's r	-0.203	-0.187	0.157
	p-value	0.057	0.189	0.354
Age	Pearson's r	0.165	0.115	0.189
	p-value	0.129	0.43	0.262
AgeUS	Pearson's r	0.03	0.126	0.121
	p-value	0.784	0.378	0.475

Sex	Pearson's r	-0.213	-0.214	-0.236
	p-value	0.046*	0.131	0.159
numLangs	Pearson's r	0.065	0.275	0.009
	p-value	0.55	0.051	0.957
EnAoA	Pearson's r	0.099	0.03	0.125
	p-value	0.36	0.832	0.462
bsqCtxS	Pearson's r	0.144	0.132	0.005
	p-value	0.182	0.355	0.978
bsqAccS	Pearson's r	-0.016	-0.02	-0.142
	p-value	0.881	0.887	0.402

Note. Pearson’s Partial Correlations.
Note. Explanation of names in *Table 1* note.

Table 15. All-Subjects Familiarity by Time by Group. *Showing a significant effect of time spent in context on RT, a time by English proficiency interaction effect on RT, and a time by Shape-Color switching effect interaction effect on RT. Also showing an interaction between the effects of familiarity and language group on RT.*

Within Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p	η^2_p
Known	0.001	1	0.001	0.394	0.532	0.005
Known ✻ langGroup	0.021	1	0.021	6.545	0.012	0.077
Known ✻ flDiffRT	0.002	1	0.002	0.719	0.399	0.009
Known ✻ scConStayRT	0.004	1	0.004	1.227	0.271	0.015
Known ✻ EnProf	0.005	1	0.005	1.451	0.232	0.018
Known ✻ SES	0.006	1	0.006	1.926	0.169	0.024
Known ✻ bsqCtxS	0.005	1	0.005	1.473	0.229	0.019
Known ✻ bsqAccS	7.517e -5	1	7.517e -5	0.023	0.880	2.940e -4
Known ✻ flCongRT	8.208e -4	1	8.208e -4	0.251	0.618	0.003
Known ✻ scDiffRT	0.003	1	0.003	0.986	0.324	0.012
Residuals	0.256	78	0.003			
Time	0.048 [a]	2 [a]	0.024 [a]	7.699 [a]	< .001 [a]	0.090
Time ✻ langGroup	0.001 [a]	2 [a]	5.209e -4 [a]	0.166 [a]	0.847 [a]	0.002
Time ✻ flDiffRT	0.007 [a]	2 [a]	0.004 [a]	1.139 [a]	0.323 [a]	0.014
Time ✻ scConStayRT	0.007 [a]	2 [a]	0.003 [a]	1.116 [a]	0.330 [a]	0.014
Time ✻ EnProf	0.039 [a]	2 [a]	0.020 [a]	6.310 [a]	0.002 [a]	0.075
Time ✻ SES	0.014 [a]	2 [a]	0.007 [a]	2.163 [a]	0.118 [a]	0.027
Time ✻ bsqCtxS	0.006 [a]	2 [a]	0.003 [a]	0.908 [a]	0.406 [a]	0.012
Time ✻ bsqAccS	0.002 [a]	2 [a]	0.001 [a]	0.340 [a]	0.712 [a]	0.004
Time ✻ flCongRT	0.011 [a]	2 [a]	0.006 [a]	1.821 [a]	0.165 [a]	0.023
Time ✻ scDiffRT	0.023 [a]	2 [a]	0.011 [a]	3.670 [a]	0.028 [a]	0.045
Residuals	0.488	156	0.003			
Time ✻ Known	0.004	2	0.002	0.683	0.507	0.009
Time ✻ Known ✻ langGroup	0.002	2	8.997e -4	0.335	0.716	0.004
Time ✻ Known ✻ flDiffRT	0.008	2	0.004	1.561	0.213	0.020
Time ✻ Known ✻ scConStayRT	0.015	2	0.008	2.826	0.062	0.035
Time ✻ Known ✻ EnProf	0.003	2	0.002	0.642	0.528	0.008
Time ✻ Known ✻ SES	0.005	2	0.003	0.977	0.379	0.012
Time ✻ Known ✻ bsqCtxS	0.005	2	0.002	0.877	0.418	0.011
Time ✻ Known ✻ bsqAccS	0.007	2	0.003	1.293	0.277	0.016
Time ✻ Known ✻ flCongRT	0.007	2	0.004	1.395	0.251	0.018

Within Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p	η^2_p
Time ✻ Known ✻ scDiffRT	0.004	2	0.002	0.786	0.458	0.010
Residuals	0.488	156	0.003			

Note. Type III Sum of Squares
Note. Explanation of names in *Table 1* note.
[a] Mauchly's test of sphericity indicates that the assumption of sphericity is violated (p < .05).

Between Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p	η^2_p
langGroup	0.453	1	0.453	4.357	0.040	0.053
flDiffRT	0.016	1	0.016	0.157	0.693	0.002
scConStayRT	0.006	1	0.006	0.059	0.808	7.585e -4
EnProf	0.253	1	0.253	2.432	0.123	0.030
SES	0.062	1	0.062	0.592	0.444	0.008
bsqCtxS	0.036	1	0.036	0.351	0.556	0.004
bsqAccS	0.152	1	0.152	1.460	0.231	0.018
flCongRT	0.207	1	0.207	1.992	0.162	0.025
scDiffRT	0.010	1	0.010	0.093	0.761	0.001
Residuals	8.116	78	0.104			

Note. Type III Sum of Squares
Note. Explanation of names in *Table 1* note.

Descriptives

Known	Time	langGroup	Mean	SD	N
Known	1	1	1.267	0.142	51
		2	1.202	0.146	37
	2	1	1.237	0.140	51
		2	1.181	0.137	37
	3	1	1.236	0.148	51
		2	1.176	0.142	37
Unknown	1	1	1.288	0.150	51
		2	1.194	0.118	37
	2	1	1.236	0.157	51
		2	1.163	0.131	37
	3	1	1.228	0.144	51
		2	1.162	0.129	37

Table 16. All-Subjects Target by Time by Group. *Showing, again, a significant effect of time spent in context on RT, a time by English proficiency interaction effect on RT, a time by Shape-Color switching effect interaction effect on RT, and now a time by Shape-Color baseline trial RT interaction effect on RT. Also showing an interaction between the time, target, and Shape-Color baseline trial RT on RT.*

Within Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p	η^2_p
Target	0.001	1	0.001	0.472	0.494	0.006
Target ✻ langGroup	0.007	1	0.007	2.447	0.122	0.030
Target ✻ flDiffRT	0.006	1	0.006	2.105	0.151	0.026
Target ✻ scConStayRT	0.005	1	0.005	1.893	0.173	0.024
Target ✻ EnProf	0.007	1	0.007	2.285	0.135	0.028
Target ✻ SES	0.002	1	0.002	0.584	0.447	0.007
Target ✻ bsqCtxS	2.002e -4	1	2.002e -4	0.069	0.793	8.890e -4
Target ✻ bsqAccS	1.162e -4	1	1.162e -4	0.040	0.841	5.160e -4
Target ✻ flCongRT	0.005	1	0.005	1.648	0.203	0.021
Target ✻ scDiffRT	0.010	1	0.010	3.524	0.064	0.043
Residuals	0.225	78	0.003			
Time	0.034 [a]	2 [a]	0.017 [a]	5.104 [a]	0.007 [a]	0.061
Time ✻ langGroup	5.839e -5 [a]	2 [a]	2.920e -5 [a]	0.009 [a]	0.991 [a]	1.112e -4
Time ✻ flDiffRT	0.009 [a]	2 [a]	0.004 [a]	1.291 [a]	0.278 [a]	0.016
Time ✻ scConStayRT	0.021 [a]	2 [a]	0.010 [a]	3.071 [a]	0.049 [a]	0.038
Time ✻ EnProf	0.029 [a]	2 [a]	0.014 [a]	4.240 [a]	0.016 [a]	0.052
Time ✻ SES	0.009 [a]	2 [a]	0.005 [a]	1.406 [a]	0.248 [a]	0.018
Time ✻ bsqCtxS	0.005 [a]	2 [a]	0.003 [a]	0.807 [a]	0.448 [a]	0.010
Time ✻ bsqAccS	0.006 [a]	2 [a]	0.003 [a]	0.892 [a]	0.412 [a]	0.011
Time ✻ flCongRT	0.013 [a]	2 [a]	0.006 [a]	1.862 [a]	0.159 [a]	0.023
Time ✻ scDiffRT	0.027 [a]	2 [a]	0.013 [a]	3.977 [a]	0.021 [a]	0.049
Residuals	0.525	156	0.003			
Time ✻ Target	0.009	2	0.004	1.547	0.216	0.019
Time ✻ Target ✻ langGroup	3.163e -4	2	1.581e -4	0.057	0.945	7.293e -4
Time ✻ Target ✻ flDiffRT	0.013	2	0.007	2.416	0.093	0.030
Time ✻ Target ✻ scConStayRT	0.024	2	0.012	4.296	0.015	0.052
Time ✻ Target ✻ EnProf	0.004	2	0.002	0.654	0.521	0.008
Time ✻ Target ✻ SES	0.006	2	0.003	1.106	0.334	0.014
Time ✻ Target ✻ bsqCtxS	0.004	2	0.002	0.709	0.494	0.009
Time ✻ Target ✻ bsqAccS	0.006	2	0.003	1.050	0.352	0.013
Time ✻ Target ✻ flCongRT	0.015	2	0.008	2.723	0.069	0.034

Within Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p	η^2_p
Time ✻ Target ✻ scDiffRT	0.008	2	0.004	1.426	0.243	0.018
Residuals	0.525	156	0.003			

Note. Type III Sum of Squares
Note. Explanation of names in *Table 1* note.
[a] Mauchly's test of sphericity indicates that the assumption of sphericity is violated ($p < .05$).

Between Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p	η^2_p
langGroup	0.420	1	0.420	4.009	0.049	0.049
flDiffRT	0.024	1	0.024	0.233	0.631	0.003
scConStayRT	0.017	1	0.017	0.159	0.691	0.002
EnProf	0.264	1	0.264	2.519	0.117	0.031
SES	0.052	1	0.052	0.495	0.484	0.006
bsqCtxS	0.046	1	0.046	0.437	0.510	0.006
bsqAccS	0.142	1	0.142	1.352	0.248	0.017
flCongRT	0.211	1	0.211	2.014	0.160	0.025
scDiffRT	0.008	1	0.008	0.080	0.778	0.001
Residuals	8.167	78	0.105			

Note. Type III Sum of Squares
Note. Explanation of names in *Table 1* note.

Descriptives

Target	Time	langGroup	Mean	SD	N
Non-Target	1	1	1.274	0.142	51
		2	1.194	0.118	37
	2	1	1.233	0.141	51
		2	1.163	0.131	37
	3	1	1.232	0.139	51
		2	1.162	0.129	37
Target	1	1	1.273	0.149	51
		2	1.202	0.146	37
	2	1	1.244	0.157	51
		2	1.181	0.137	37
	3	1	1.235	0.163	51
		2	1.176	0.142	37

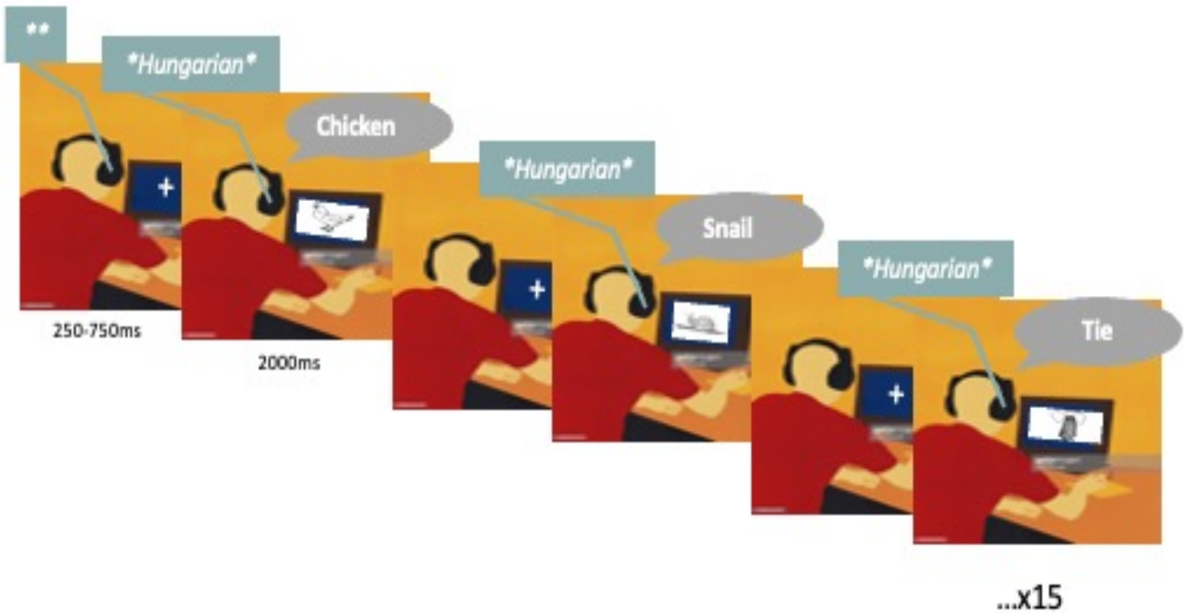

Figure 1. The Babel Paradigm procedure. Distractor audio plays through length of presentation of stimulus. Fixations (**; cross on-screen) are silent. Each block of 15 trials repeats 4 times randomly over 16 blocks (60 trials per distractor).

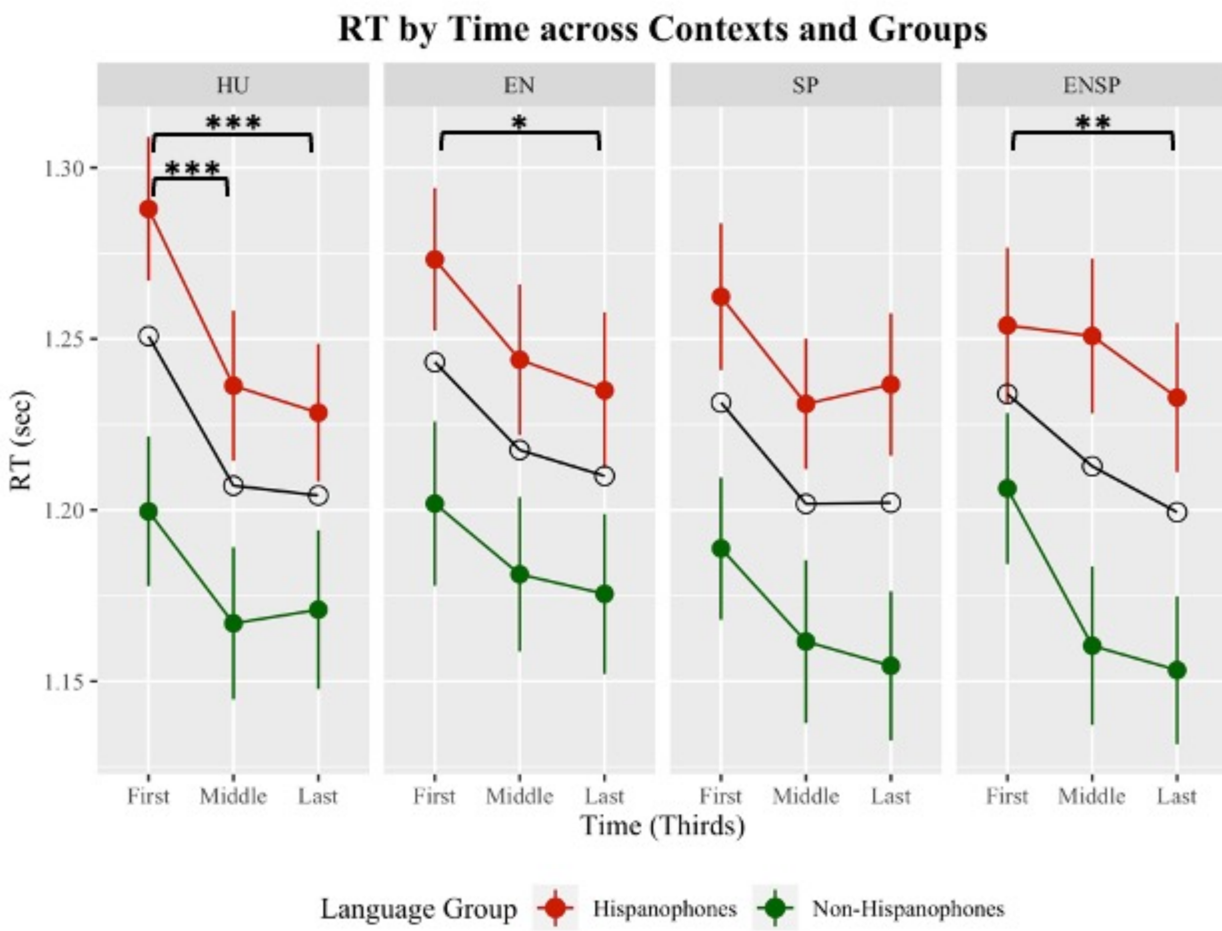

Figure 2. RT by Time Spent in Context across Context and Language Groups. Hispanophones (upper bar) and non-Hispanophones (lower bar). Black (middle bar) is the average used in analysis and post-hoc comparisons. Significant differences from post-hoc t-tests between First-Third, Middle-Third, and Last-Third shown in asterisks. Note the comparatively abnormal trend of ENSP Hispanophone RTs over the course of the block.

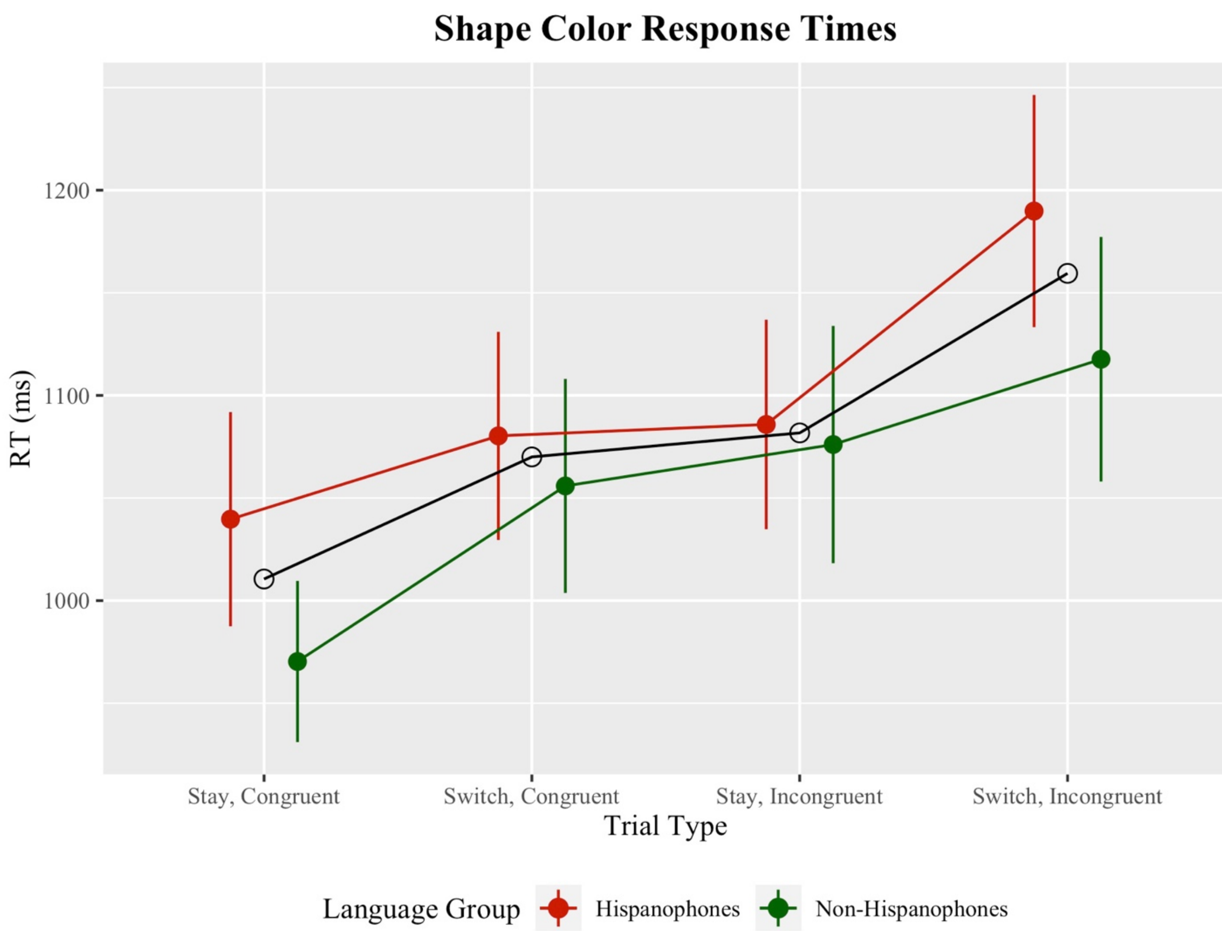

Figure 3. Shape-Color RT by Trial Type. Hispanophones (upper bar) and non-Hispanophones (lower bar). Visualization of significant differences between switch and stay trials (t-tests), generally, as well as significant main effects of congruency and switching (ANOVA). Increasing RTs from baseline trial type (congruent-stay) to mixed trial types (congruent-switch and incongruent-stay) to experimental trial type (incongruent-switch) suggest Shape-Color captured its target effect of task-switching ability.

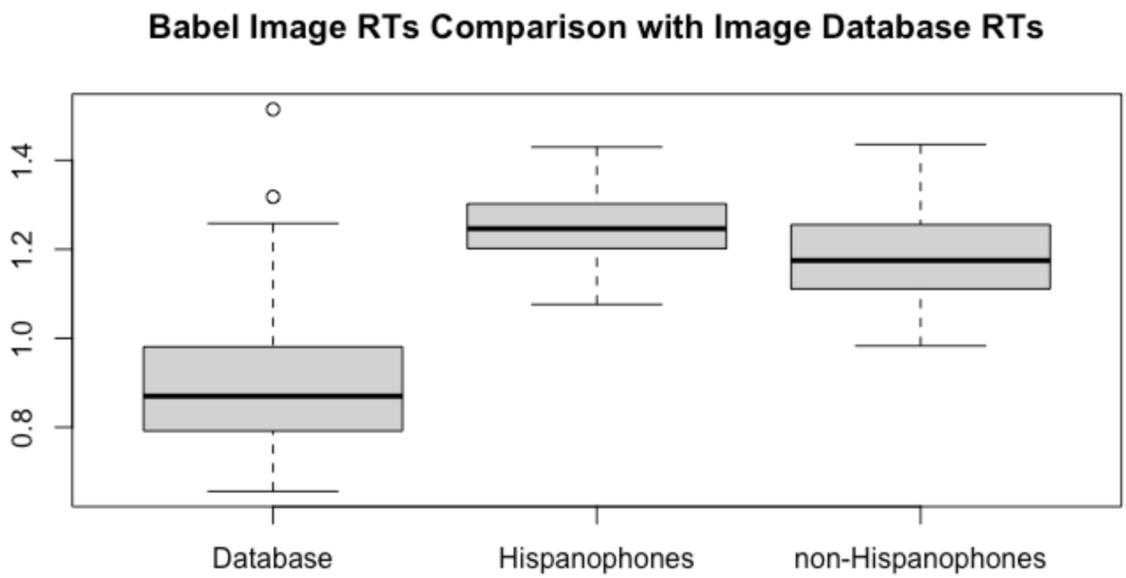

Figure 4. Babel Image Response Times Compared to Database Image Response Times. Hispanophones and non-Hispanophones both differed significantly from database images and differed significantly from eachother. ($M_{database}$ = 0.899, M_{Hisp} = 1.255, $M_{nonHisp}$ = 1.18; $t_{database\text{-}Hisp}$ = -27.648, $p < .0001$; $t_{database\text{-}nonHisp}$ = -20.682, $p < .0001$; $t_{Hisp\text{-}nonHisp}$ = 7.346, $p < .0001$).