Inhibitory control in Russian-English bilingual adults.

Akhmed Dugrichilov, Aleksandra Cherniaeva, Yana Zakharova, Ali Tepsaev

Introduction

One of the first tools for studying the way languages are stored in memory and the way they can intertwine is the Stroop Test (Stroop, 1935; Jensen AR & Rohwer WD, 1966) that works as follows: the participant's task is to '*identify the colour of a printed word (e.g. "red" printed in green), while ignoring the word itself*' (Šaban & Schmidt, 2022). For interlinguistic Stroop Test Stroop Effect (Stroop Effect occurs when speaker is impaired by incongruent stimuli, for instance, "yellow" in black) was found within and between languages (Altarriba & Mathis, 1997), it was stated that the differential patterns of the effect are dependent on the level of proficiency (Preston & Lambert, 1969) and the language usage (Mägiste, 1985) and that the magnitude of the Stroop effect is normally greater in the intralanguage condition than in the interlanguage condition (Fang et al., 1981; Macleod, 1991 and references).

The present study aimed to examine the role of language order switching in inhibitory control performance through the usage of the Stroop Test. Based on previous studies (Costa & Santesteban, 2004 and references), we hypothesize that there would be an asymmetrical switching cost in inhibitory control performance within the bilingual Stroop Task. Namely, we expect: 1) that response time to stimuli written in Russian will be faster than in English as Russian is L1 for all of our participants; 2) that reaction time to incongruent stimuli will be longer than to congruent stimuli (as was proven in Fang et al., 1981; Sichel & Chandler, 1969); 3) Eng—Rus sequences will be processed slower than Rus—Eng sequences, as switching to L2 is easier because it is easier to inhibit (Meuter & Allport, 1999; Costa & Santesteban, 2004); 4) therefore, there will be no significant difference in processing between sequences in only Russian (Rus—Rus) or only English (Eng—Eng) language as participants have a high level of proficiency in English and colors are basic vocabulary. Ахмед

Methods

Participants. Thirty six Russian-English bilinguals took part in the study (25 women; aged 15-46; M_{age}= 22.1, SD=6.4) The mean education level among participants was 14.4 years, SD=2.5, range 8-22. Based on study design (bilingual adults with English proficiency level being no lower than B2), 14 participants were excluded from further analysis. The final sample consisted, therefore, of 22 participants (18 women; aged 18-46; M_{age} = 22.5, SD = 7.7; mean years of education = 14.2, SD = 1.9, range 11-19). All participants reported being Russian dominant in a language history questionnaire. According to participants' answers, English was mostly used in the following contexts: *Listening to music, Watching videos* and *Reading.*

Materials and design. Prior to the experimental portion, participants filled out a series of questions from the Russian version of the Language Experience and Proficiency Questionnaire (LEAP-Q; Marian et al., 2007). The experiment was run online using PCIbex software (Zehr & Schwarz, 2018). The participants performed a word–word variant of the Stroop task with English and Russian colored words. The target color word appeared on the screen until the participants' response: button *1* served as a response in matching cases whilst button *0* served as a response in mismatching cases. Before the main part of the experiment, participants completed the training session (*n*=16) to get accustomed to the task design. During the main part of the experiment, participants were presented with 33 colored words. Each word served as a distracter for the following target word. Thus, four groups of pairs were examined (*Russian distracter – English target* (*n*=8); *Russian distracter – Russian target* (*n*=8); *English distracter – English target* (*n*=8); *English distracter – Russian target* (*n*=8)).

Data analysis. Statistical analysis was performed in R (R Core Team, 2022). The following libraries and packages were used: *lme4* (Bates et al., 2015), *tidyverse* (Wickham, 2017), *dplyr* (Wickham et al., 2022), *tidyr* (Wickham et al., 2024) and *lmerTest* (Kuznetsova et at., 2017)

There were several model structures. Firstly, we fitted linear models to analyze the relationships between reaction time and target language. The structure of the model was as follows: *reaction time ~ target language + word length + (1 | participant)*. Then, we fitted linear models to analyze the relationships between reaction time and congruence. The structure of the model was as follows: *reaction time ~ congruence + word length + (1 | participant)*. Additionally, we fitted linear models to analyze the relationships between reaction time, target language and congruence. The structure of the model was as follows: *reaction time ~ target language + congruence + word length + (1 | participant)*.

Secondly, linear models were used for above-mentioned pairs of stimuli. Thus, we fitted linear models to analyze the relationships between reaction time and type of stimuli pair. The structure of the model was as follows: *reaction time ~ language pair + word length + (1 | participant)*. Likewise, we fitted linear models to analyze the relationships between reaction time and congruence within the pairs. The structure of the model was as follows: *reaction time ~ pair congruence + word length + (1 | participant)*.

Lastly, the final model was fitted to analyze the relationships between reaction time, type of stimuli pair and congruence within the pairs. The structure of the model was as follows: *reaction time ~ language pair + pair congruence + word length + (1 | participant)*.
Саша

Results

After separating the data into two datasets regarding the language of the stimuli: English or Russian, the measures of central tendency were calculated: for Russian — M_{time} = 1158.663, SD = 473.5574, range – 416-2477, for English — M_{time} = 1120.608, SD = 417.6278, range – 436-2347, Wilcoxon's test – p-value = 0.5494 (not statistically significant). Division of the obtained data into the two groups of congruence and incongruence of the stimuli's color and the word shown and the subsequent processing of the obtained subsets gave the following measures of central tendency: for congruent — M_{time} = 1084.634, SD = 436.7456, range – 416-2238, for incongruent — M_{time} = 1179.023, SD = 451.3169, range – 431-2477, Wilcoxon's test – p-value = 0.009331 (statistically significant). The following two separations were carried out, resulting in Russian and English congruent and incongruent stimuli groups with the following measures:Russian congruent — M_{time} = 1061.064, SD =

433.2354, range – 416-2238; Russian incongruent — M_{time} = 1213.439, SD = 487.2942, range – 431-2477; English congruent — M_{time} = 1104.128, SD = 440.3041, range – 472-2223; English incongruent — M_{time} = 1134.935, SD = 397.7704, range – 436-2347. Both the language and congruence factors were statistically significant (p-value = 0.00801 and p-value < 0.0001, respectively). The next four divisions were the pairs of predictor and target words: *rus-rus* — M_{time} = 1105.738, SD = 467.7309, range – 416-2307, *rus-eng* — M_{time} = 1081.688, SD = 415.9198, range – 436-2222, *eng-rus* — M_{time} = 1080.864, SD = 417.0268, range – 416-2238, *eng-eng* — M_{time} = 1110.065, SD = 417.4064, range – 487-2347. No factor had any statistically significant impact on the reaction time (including the length of the target). Subsequent divisions of the data were performed regarding the congruence of the target words: *cong-cong* — M_{time} = 1013.61, SD = 421.5605, range – 416-2238, *cong-incong* — M_{time} = 1133.415, SD = 439.8479, range – 431-2347, *incong-cong* — M_{time} = 1133.24, SD = 452.6939, range – 472-2223, *incong-cong* — M_{time} = 1129.838, SD = 406.0899, range – 510-2307. All the predictors were statistically significant (*cong-incong* – p-value = 0.0002, *incong-cong* – p-value = 0.000152, *incong-incong* – p-value < 0.0001), except for the length of the word – p-value = 0.99954.

Furthermore, having set the aforementioned models, we have the following results:

model_lang_x <- lmer(RT ~ Language_stim + length_word + (1 | Hash), data = joined_times_filtered) — reaction time dependence on the language of the target (either English or Russian) and the length of the target's word with participants' number as a random effect showed that the language of the target does not affect the reaction time in a statistically significant way, 1174.458 ms being the average reaction time of English target and 1237.359 ms of Russian target (p = 0.0355). There was no statistical difference in the RT depending on the length of the target as well (1166.479 ms – p-value = 0.3022)

model_congr_x <- lmer(RT ~ Congruence + length_word + (1| Hash), data = joined_times_filtered) — reaction time dependence on the congruency of the color and the presented word and the length of the target's word with participants' number as a random effect displays that the reaction time of congruent target is 1130.258 ms, whereas for incongruent target it is 1226.5 ms (statistically significant difference — p < 0.001).

model_lang_congr1_x <- lmer(RT ~ Language_stim + Congruence + length_word + (1|Hash), data = joined_times_filtered) — reaction time dependence on the target's language and the congruency of the color and the word shown and the length of the target's word with participants' number as a random effect shows the statistically significant difference between English and Russian target's word, its incongruence and the length of the word (1170.388 ms, 1249.19 ms – p-value = 0.00801, 1274.479 ms – p-value < 0.001 and 1151.75 ms – p-value = 0.02004, respectively).

model_stim_lang1_x <- lmer(RT_stim ~ Language_pair + length_stim + (1|Hash), data = final_dataset_no_NA) — reaction time dependence on the prime-target pair, the length of the target's word with participants' number as a random effect demonstrates that there's no significance between *eng-eng*, *eng-rus*, *rus-eng*, *rus-rus* prime-target pairs and the length of the target's word (p-value = 0.0732, 0.2255, 0.5854 and 0.1021, respectively).

model_stim_congr1_x <- lmer(RT_stim ~ Congruency_pair + length_stim + (1|Hash), data = final_dataset_no_NA) — reaction time dependence on the congruency of the color and the word and the length of the target's word with participants' number as a random effect shows that congruence pairs *congr-incongr* (1159.7885 ms – p-value = 0.0002)*, incongr-congr* (1159.8394 ms – p-value = 0.0002)*, incongr-incongr* (1168.9025 ms – p-value < 0.0001) differ from the pair *congr-congr* (1054.7677 ms) in a statistically significant manner. The length of the target does not, in turn, have a statistically significant impact on the reaction time (1054.7646 ms – p-value = 0.9995).

model_stim_lang_congr1_x <- lmer(RT_stim ~ Lang_congr_pair + length_stim + (1|Hash), data = final_dataset_no_NA) — reaction time dependence on the language and congruency of the prime and the target and the length of the target's word with participants' number as a random effect could be described in the following table:

	English congruent (ms)	English incongruent (ms)	Russian congruent (ms)	Russian incongruent (ms)
English congruent (ms)	1163.94	1255.55 ($p = 0.13187$)	1174.96 ($p = 0.84523$)	1382.78 ($p = 0.00323$)
English incongruent (ms)	1449.46 ($p < 0.0001$)	1249.24 ($p = 0.05265$)	1233.7 ($p = 0.17048$)	1358.56 ($p = 0.00314$)
Russian congruent (ms)	1132.21 ($p = 0.51893$)	1279.98 ($p = 0.01987$)	1225.64 ($p = 0.18966$)	1268.02 ($p = 0.10115$)
Russian incongruent (ms)	1240.53 ($p = 0.11546$)	1251.48 ($p = 0.07702$)	1308.69 ($p = 0.02249$)	1404.55 ($p = 0.00019$)

The length of the word also has a statistically significant impact on the reaction time — 1141.06 ms (p-value = 0.03598). Яна

Discussion

As expected, incongruent stimuli are processed for a longer period of time than congruent ones (by 94.39 ms – statistically significant). Also, no statistically significant difference in processing time for the Rus-Rus and Eng-Eng sequences was found.

The Stroop Effect was also larger for Russian than for English language – items in Russian also influenced the increase in reaction time for the incongruent stimuli as they could be distracting the participants from the right answer, and the effect for the native language was larger (for the information on similar results see Schmidt et al. 2018, Leiser & Tselgov, 1990 and Macleod, 1991): the response for the congruent Russian stimuli is faster as that for English one and the response for the incongruent one is slower – that being said, our expectations were proven to be right only partially. Furthermore, the average processing time for the Eng-Rus sequence turned out not to be greater than the time for the Rus-Eng pair as the language wasn't a significant variable: we could ascribe it to the small sample size (22 participants) and to the fact that our respondents weren't qualified as equal bilinguals even though all of them were proficient in English.

As for the statistical significance of the congr-congr sequence (response time of congr-congr is lower than of others), we could ascribe it to the smaller RT of congruent trials overall. At the same time, we could also say that, in some way, *contingency* played a role in our results: contingency is the increase of the magnitude of the Stroop Effect that arises when proportion of the congruent stimuli is enlarged, i.e. the respondents start spending less time reacting to the congruent stimuli (Schmidt et al., 2007; Schmidt & Besner, 2008). Therefore, we could hypothesize that the sequence of 2 stimuli had an effect on the rapidity of reaction and the speed of the second response grew. Али

References

Altarriba & Mathis, 1997 — Altarriba J., Mathis K. M. (1997) Conceptual and lexical development in second language acquisition. J Mem Lang 36(4):550–568.

Bates et al., 2015 — Bates, D., Mächler, M., Bolker, B. M. & Walker, S. C. Fitting linear mixed-efects models using lme4. J. Stat. Sofw. 67, 2015. P. 1–48.

Costa & Santesteban, 2004 — Costa, Albert & Santesteban, Mikel. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners.

Fang et al., 1981 — Fang, sheng-ping & Tzeng, Ovid & Alva, Liz. (1981). Intralanguage vs. interlanguage Stroop effects in two types of writing systems. Memory & cognition. 9. 609-17.

Francis, 1999 — Francis W. S. Cognitive integration of language and memory in bilinguals: semantic representation. Psychological bulletin, 125(2), 1999. P. 193–222.

Green, 1998 — Green D. W. Mental control of the bilingual lexicosemantic system. Bilingualism: Language and Cognition, 1, 1998. P. 67–81.

Jensen AR & Rohwer WD, 1966 — Jensen AR, Rohwer WD (1966). "The Stroop color-word test: a review". Acta Psychologica. 25 (1): 36–93.

Kuznetsova et at., 2017 — Kuznetsova A., Brockhoff P. B., & Christensen R. H. B. lmerTest package: tests in linear mixed effects models. Journal of statistical software 82(13), 2017.

Leiser & Tselgov, 1990 — Leiser, D., & Tzelgov, J. (1990). Controlling Stroop interference: Evidence from a bilingual task. Journal of Experimental Psychology-Learning Memory and Cognition.

Mägiste, 1985 — Mägiste, E. Development of intra- and interlingual interference in bilinguals. J Psycholinguist Res 14, 137–154 (1985).

Marian et al. 2007 — Marian V., Blumenfeld H. K., and Kaushanskaya M. (2007). The language experience and proficiency questionnaire (LEAP-Q): assessing language pro-files in bilinguals and multilinguals. J. Speech Lang. Hear. Res. 50, 2007. P. 940–967.

Macleod, 1991 — MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. Psychological Bulletin, 109, 163-203.

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

Naylor, 2012 — Naylor L. J., Stanley E. M., & Wicha N. Y. Cognitive and electrophysiological correlates of the bilingual Stroop effect. Frontiers in Psychology, 3, 2012.

New et al., 2006 — New, B., ferrand, L., pallier, C. et al. Reexamining the word length effect in visual word recognition: New evidence from the English Lexicon Project. Psychonomic Bulletin & Review 13, 45–52 (2006).

Preston & Lambert, 1969 — Preston, M. S., & Lambert, W. E. (1969). Interlingual interference in a bilingual version of the Stroop color-word task. Journal of Verbal Learning and VerbalBehavior, 8, 295-301

R Core Team, 2022 — R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, 2022.

Roelofs, 2003 — Roelofs A. Goal-referenced selection of verbal action: modeling attentional control in the Stroop task. Psychological review, 110(1), 2003. P. 88–125.

Šaban & Schmidt, 2022 — Šaban I. & Schmidt J. R. Interlinguistic conflict: Word–word Stroop with first and second language colour words. Cognitive Processing, 23(4), 2022. P. 619-636.

Schmidt et al., 2007 — Schmidt, James & Crump, Matthew & Cheesman, Jim & Besner, Derek. (2007). Contingency learning without awareness: Evidence for implicit control. Consciousness and cognition. 16. 421-35.

Schmidt & Besner, 2008 — Schmidt, J. R., & Besner, D. (2008). The Stroop effect: Why proportion congruent has nothing to do with congruency and everything to do with contingency. Journal of Experimental Psychology: Learning, Memory, and Cognition, 34(3), 514–523.

Schmidt et al. 2018 — Schmidt JR, Hartsuiker RJ, De Houwer J. Interference in Dutch-French Bilinguals. Exp Psychol. 2018 Jan;65(1):13-22.

Sichel & Chandler, 1969 — Sichel, J. L., & Chandler, K. A. (1969). The color-word interference test: The effects of varied color-word combinations upon verbal response latency. Journal of Psychology, 72, 219-231.

Stroop, 1935 — Stroop J. R. Studies of interference in serial verbal reactions. Journal of experimental psychology, 18(6), 1935.

Wickham, 2016 — H. Wickham. ggplot 2: Elegant Graphics for Data Analysis. Springer, 2016.

Wickham, 2017 — H. Wickham. tidyverse: Easily Install and Load the 'Tidyverse'. R package version 1(1), 2017.

Wickham H. et al., 2022 — H. Wickham, R. Francois, L. Henry, K. Müller, & D. Vaughan. dplyr: A grammar of data manipulation. R package version 1(2), 2022.

Wickham et al., 2024 — Wickham H., Vaughan D., Girlich M. tidyr: Tidy Messy Data. R package version 1.3.1, 2024.

Zehr & Schwarz, 2018 — Zehr J. & Schwarz F. PennController for Internet Based Experiments (IBEX), 2018.