BILINGUAL STROOP EFFECT IN HIGH AND LOW PROFICIENT NEPALI-ENGLISH BILINGUALS

Lekhnath Sharma Pathak & Poshak Pathak

This research article presents experimental psycholinguistic study on high and low proficient Nepali – English bilingual adults. Language comprehension was measured using lexical decision tasks and language production was measured using verbal fluency task. Executive control was measured using bilingual Stroop task in mouse tracking paradigm. The results show high proficient bilinguals show the bilingual advantage in language comprehension, language production and executive control tasks.

Keywords: Stroop task, Nepali-English bilinguals, mouse tracking, bilingual advantage

1. Introduction

Researches in bilingualism have established that juggling between languages confers bilingual advantage (Green & Abutalebi, 2013; Biyalstok, 2017). We explored the bilingual advantage effect (Valin, 2015; Bialystok et al, 2004; Baum & Titone, 2014)) using Stroop task (Stroop, 1935; Incera & McLennan, 2015) in mouse tracking paradigm (Freeman & Ambady, 2010; Freeman, Dale & Farmer, 2011). Colour-naming Stroop task was advanced by J. R. Stroop in 1935 and ever since, this task has become a powerful tool in measuring executive control using various experimental designs to measure the conflict monitoring and resolution across diverse conditions. Recently, it has been adopted in studying bilingualism in bilingual Mouse Tracking Stroop task (Incera et. al, 2013; Incera & McLennam, 2015) and multilingual processing using mouse-tracker (Pathak, 2020). Mouse Tracking paradigm in Psycholinguistic research has been used in many recent studies (see for example, Spivey et.al, 2005; Incera & McLennan, 2016; Viswanathan & Kelty-Stephen, 2017, Erb, 2018).

Bilingual Advantage has been established by many previous studies in which researchers have shown people using more than one language are more capable in managing conflict by directing the cognitive resources to the task relevant stimuli and suppressing the allocation of such resource to task irrelevant stimuli (see for example, Bialystok, et al. 2004; Blom, et.al, 2017; Baum & Titone, 2014; Bialystok, 2016; Valin, 2015). However, a small number of studies also have shown that bilingualism affords no such advantage (Paap & Sawi, 2015). But studies supporting bilingual advantage on cognitive mechanism far outnumber the studies which don't support such claims.

2. Present study

The research question that we are asking in this study is: how does language proficiency affect the performance in bilingual Stroop task among Nepali- English bilinguals? We hypothesize that: high proficient bilinguals perform better in Stroop task compared to low proficient bilinguals. So we predict that: high proficient bilinguals will respond faster and more accurately compared to low proficient bilinguals in incongruent Stroop conditions that require participants to suppress their pre-potent response. The novelty and originality of this study is that this is the first study of bilingual Stroop task among adult participants in this language pair (Nepali and English) in mouse-tracking paradigm (see Rijal, (2020) for the study with similar design in school children with monolingual and bilingual medium of instruction, also Rijal & Pathak, 2020).

3. Method

3.1 Participants

Total twenty-one participants (Female=10) took part in this study. Their mean age was 43.65 (SD = 7.72) years. Two participants were discarded from analysis for exceeding mouse initiation time. So, we analyzed the data for 19 participants. Bilingual proficiency of the participants was established through median split. The median for Verbal Fluency Task (VFT) was 40, for Ghent Vocabulary Test it was 34 and for LexTale it was 60. The participants who had crossed the

threshold (median) in all the three tasks were categorized as high proficient and those below the threshold were categorized as low proficient. All the participants gave written informed consent before the experiment and filled up the language background questionnaire after the experiment. Participation was voluntary. Table 1 presents the demographic profile of the participants.

Table 1: Demographic profile of the participants

Measure	Full Sample	Language Proficiency	
		High	Low
Sample size	21	10	11
Demographics			
Gender	12M, 9F	9M, 1F	3M, 8F
Age	42 (6.38)	36 (5.92)	47 (7.21)
Parents education (Median)	4	4	4
Primary Language Measures			
Age of L2 acquisition	12 (2.81)	11(3.22)	14 (3.53)
Non-L1 language usage (0-10 scale)	5.21 (1.40)	6.63 (2.43)	4.97 (1.29)
Other Language Measures			
Proficiency in L2 (0-10) scale	7.02 (2.18)	7.23 (1.75)	6.71 (2.54)
Learnt L2 at school (age)	7.42 (1.03)	8.82 (2.67)	7.21 (2.15)
Learnt L2 at home (age)	3.87 (1.29)	4.29 (1.53)	3.35 (1.18)

Note: Note: Standard deviations are given in brackets wherever applicable. L1 = First Language (Nepali), L2 = Second Language (English)

3.2 Procedure

All participants were seated in a quiet room. Task instructions were given both in English and Nepali. Baseline task was created as a practice for mouse movement and also to compare mouse trajectories without experimental manipulation. Proficiency establishment tasks were performed after the main experiment.

3.3 Material and Stimuli

All the computer based experimental tasks were run in DELL 15.6" Inspiron 15 3000 Series laptop with screen resolution of 1320 x 720 pixel. Optical mouse was used to make response on the computer screen. I ball microphone was used to record the language production task in verbal fluency test.

3.3.1 Stimuli for language comprehension task

We used two standard lexical decision tasks: LexTale and Ghent Vocabulary. These two tasks measure the proficiency of participants in their L2 (English) as these two tests measure their degree of familiarity and proficiency in English. See figures 1 – 4 for sample trials of lexical decision task.

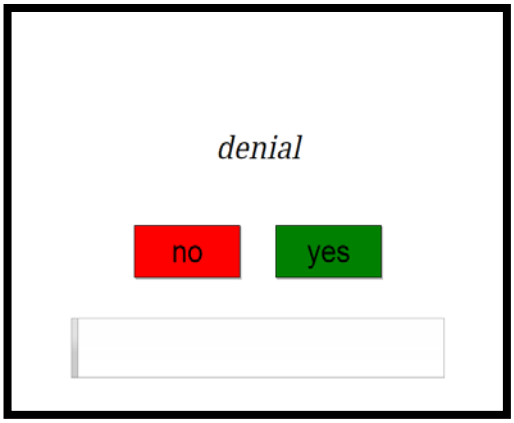

Figure 1: Lexical decision task in LexTale (Word: denial)

Figure 1 shows a sample trial in LexTale in which participants see the word "denial" and have to press or click on the response "yes" on the screen since this is a legitimate word in English.

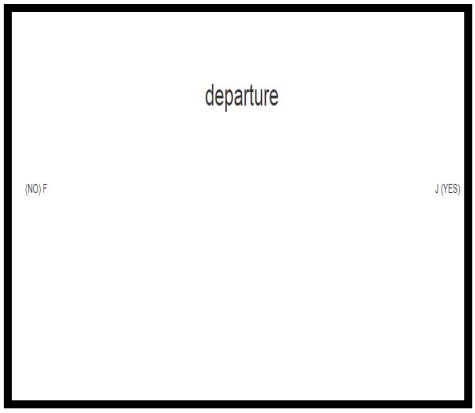

Figure 2: Lexical decision task in Ghent Vocabulary Test (Word: departure)

Figure 2 shows lexical decision task in Ghent Vocabulary Test in which participants see the word "departure" and press the button "J" on the keyboard indicating that it is a legitimate word in English.

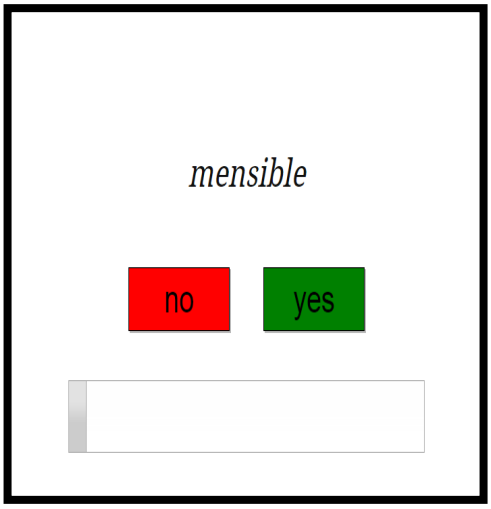

Figure 3: Lexical decision task in LexTale (Non-word: mensible)

Figure 3 shows a sample trial in LexTale in which participants see the non-word "mensible" and have to press or click on the response "no" on the screen since this is not a legitimate word in English.

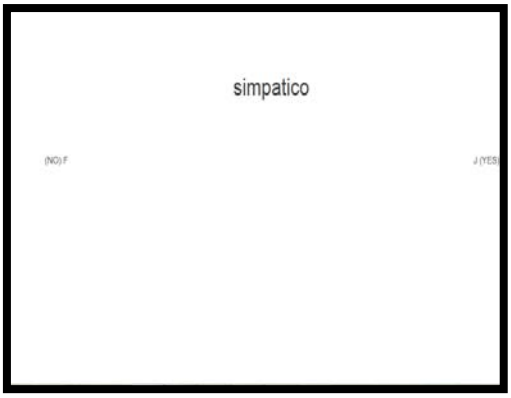

Figure 4: Lexical decision task in Ghent Vocabulary Test (Non-word: simpatico)

Figure 4 shows lexical decision task in Ghent Vocabulary Test in which participants see the word "simpatico" and press the button "F" on the keyboard indicating that it is not a legitimate word in English.

3.3.2 Stimuli for language production task

In order to measure the bilingual proficiency in language production, we used verbal fluency task (VFT), semantic and phonological fluency tasks in both L1 (Nepali) and L2 (English). (A) Phonological Verbal Fluency Task in L1 consisted of F, A, S and in L2 consisted of फ, अ, स. In order to test proficiency in semantic fluency we used Semantic Verbal Fluency Task asking participants to make words from the categories of animals, fruits, vegetables, clothes and flowers in L1 and L2. In verbal fluency task, a participant is asked to make as many words as possible from the given letter and category within a minute (60 seconds). For L2 phonetic fluency we adopted the standard letters used in such tests. However, for Nepali because of lack of such standard task we merely converted the English letters to Nepali phonological counterparts.

3.3.3 Stimuli and design for executive control task

We used four colour words; RED, BLUE, YELLOW, GREEN following Klein (1964) in both English and Nepali. Total number of trials were 72 with 36 trials in each language. All trials were counterbalanced and randomized.

Stroop task measure executive control by testing the ability of the participant to suppress pre-potent response and maintain goal directed behavior by monitoring and resolving the conflict presented by the distractor that can easily impact the response. We created a bilingual stroop task in Nepali and English. When the participants clicked on the START button, a color word either in Nepali or English on an x would appear on the screen just above the START button and on the top right and left response buttons would appear with all the four words distributed I such a way that the button on the left would have two words of which one would be the displayed word and another would be an irrelevant word and on the opposite screen there would be the word naming the color of the ink in which the word is written and another word would be an irrelevant word Participants had to click on the word that named the color of the ink in which word was written. They had to ignore the actual word displayed. In congruent trials, the ink and color word would match and in congruent trials it wouldn't match. In neutral trials the displayed characters would be four 'X's in all four colors. Figures 5 and 6 present congruent trials in L1 and L2.

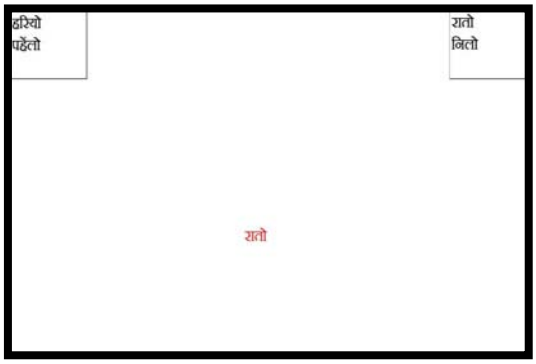

Figure 5: Congruent trial in L1 (word रातो 'red' written in red ink)

In this trial, participants look at the word रातो 'red' and click on the matching word on the top right response button.

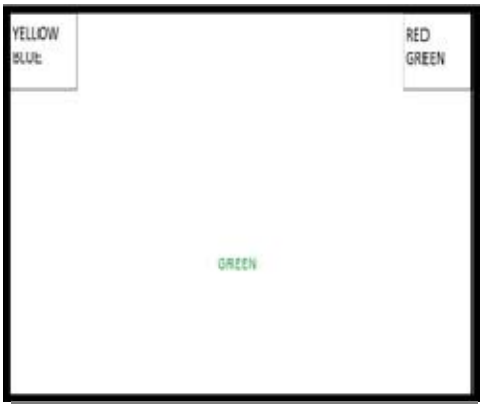

Figure 6: Congruent trial in L2 (word 'green' written in green ink)

In this trial, the participants look at the word GREEN and click on the matching word in the response button.

4. Results

Data analysis was done by conducting t-test, ANOVA. Mastersheets were created in excel and t- test was performed in R and ANOVA in SPSS, plots for data visualization have been done in R.

4.1 Median split for categorizing participants

Table 2: Median score of participants.

Measurement	Median
Verbal Fluency Task	40
Ghent Vocabulary Test	34
LexTale	60

The table presents the median score of the participants in language comprehension and language production task. The median was the threshold for categorizing the participants as high proficient and low proficient. The participants who scored above the threshold were categorized as high proficient and those who scored below the threshold were categorized as low proficient.

4.2 Language comprehension

Language comprehension was measure by the performance of participants on lexical decision task in their L2. This required the participants to look at the string of characters displayed on the screen and to judge whether those characters made a legitimate word or not in English. There

were strings of letters that made words of which some were more frequent and some less. The strings of letters that did not make real word were either pronounceable or not. Table 3 presents the score obtained by the participants in two different lexical decision tasks. Both the tasks were conducted online.

Table 3: Mean score difference in L2 high and low proficient participants. Standard deviation (SD) and p-value are given in brackets.

	LexTale	Ghent
High Proficient	74.3 (10.4)	49.9 (11.7)
Low Proficient	53.6 (6.70)	28.9 (14.5)
Whole Sample	63.5 (13.5)	37.9 (17.1)
Paired t-test value	5.47 (p <.001)	6.02 (p < .001)

The table shows that the score difference in both the tests of language comprehension measure in L2 was highly significant which mean the participants in low proficiency category performed significantly poorer than the participants in high proficiency category.

4.3 Language production

Verbal fluency task (VFT) was used to test the proficiency in language production. Participants were given three letters and five categories each in Nepali and English and they had make as many words as they could from the given item within 60 seconds. They couldn't repeat the words, use proper nouns, inflected and derivational forms. Table 4 presents the results of verbal fluency task.

In table 4, words for semantic fluency in L1 are: ज-जनावर, लु-लुगा, फ-फल, त-तरकारी, फू-फूल; and, words for semantic fluency in L2 are: An-animal, cl-clothes, fr-fruits, ve-veggies and fl-flowers.

The results of verbal fluency task show that high proficient participants performed better in all the measures of the tasks. They were significantly better in L2 production tasks. They also performed better in L1 production tasks.

Table 4: Phonetic and semantic fluency difference in verbal fluency task in high and low proficient participants. Standard deviations are given in brackets.

	Mean (SD)		95% C. I.	Paired t-test p value
	Low Prof.	High Prof.		
L1 Phonetic Fluency				
फ	5.21 (2.26)	8.53 (3.05)	[-5.84, -0.80]	0.013
अ	4.15 (1.18)	7.29 (1.30)	[-4.30, -1.97]	<.001
म	6.74 (1.53)	7.11 (2.21)	[-2.16, 1.42]	0.668
L2 Phonetic Fluency				
F	6.24 (2.71)	10.28 (3.76)	[-7.12, -0.96]	0.013
A	6.14 (2.65)	9.46 (2.92)	[-5.94, -0.70]	0.016
S	5.42 (2.89)	9.28 (2.01)	[-4.71, 1.23]	0.003
L1 Semantic Fluency				
ज	8.14 (1.98)	13.64 (2.38)	[-7.56, -3.44]	<.001
लु	6.45(2.56)	9.29 (2.21)	[-5.09, -0.59]	0.020
फ	8.29 (2.11)	10.22 (3.06)	[-1.49, 0.46]	0.290
त	9.89 (2.56)	12.52 (3.24)	[-5.37, 0.11]	0.050
फू	3.21 (1.43)	4.20 (1.31)	[-2.28, 0.30]	0.124
L2 Semantic Fluency				
An	9.58 (3.56)	15.15 (2.58)	[-8.49, -2.65]	<.001
Cl	9.21 (2.15)	13.69 (2.87)	[-6.86, -2.10]	<.001
Fr	7.95 (2.70)	10.89 (3.02)	[-5.63, -0.25]	0.034
Ve	4.51 (2.06)	9.39 (2.42)	[-6.99, -2.77]	<.001
Fl	3.79 (2.39)	6.44 (2.17)	[-4.79, -0.51]	0.018

4.4 Executive control

Executive control task was measured by the mouse trajectories in a bilingual Stroop task as the

participants made mouse movement to make the response. Temporal measurements of initiation and response time in milliseconds captured the cognitive decision-making processes. Figure 7 shows the sample of mouse trajectories produce by hand movement of the participant as they make response.

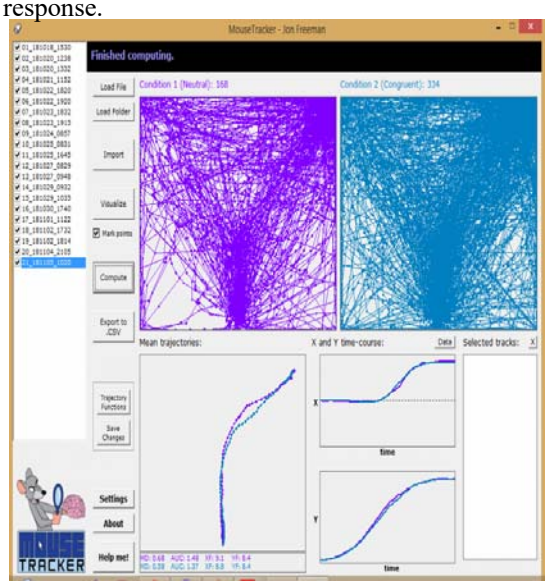

Figure 7: Sample of mouse trajectories

In Figure 7, the trajectories in the figure in the purple color on the left panel are the trajectories in neutral condition and the on the right panel in blue are the trajectories in congruent condition.

4.4.1 Baseline

A t-test was performed to check the significant effect between L1 & L2. Area Under the Curve (AUC), p=0.187 (p>0.05) was not significant. Maximum Deviation (MD), p=0.265 (p>0.05) was not significant. Initiation Time, High Proficient- p=0.186 (p>0.05) was not significant, Low Proficient- p=0.266(p>0.05) was not significant. Response Time, High Proficient- p=0.119 (p>0.05) was not significant, Low Proficient- p=0.138 (p>0.05) was not significant. The t-test shows proficiency and language experience did not affect the mouse movement in baseline condition.

4.4.2 Main experiment: Area Under the Curve (AUC)

A subject wise 2 x 2 x 3 repeated measures ANOVA was performed with language (L1 and L2) and congruency (neutral, congruent and incongruent) as within subject factors and proficiency (high and low proficiency) as between subject factors. The main effect for language was significant, $F(1,18)= 9.303$, $p<0.05$, $\eta_p^2 = 0.354$ showing overall deviation of the mouse trajectory toward unselected response was higher for L2 (mean= 1.475, SE= 0.155) than L1 (mean= 1.160, SE= 0.124). The main effect for congruency was not significant, $F(1,18)= 1.471$, $p>0.05$, $\eta_p^2 = 0.080$. The main effect for proficiency was not significant, $F(1,18)= 1.465$, $p>0.05$, $\eta_p^2= 0.079$.

4.4.3 Main experiment: Maximum Deviation (MD)

A subject wise 2 x 2 x 3 repeated measures ANOVA was performed with language (L1 and L2) and congruency (neutral, congruent and incongruent) as within subject factors and proficiency (high and low proficiency) as between subject factors. The main effect for language was significant, $F(1,18)= 11.280$, $p<0.05$, $\eta_p^2 = 0.118$ showing overall deviation of the mouse trajectory toward unselected response was higher for L2 (mean= 0.764, SE= 0.053) than L1 (mean= 0.684, SE= 0.060). The main effect for congruency was not significant, $F(1,18)= 0.503$, $p>0.05$, $\eta_p^2 = 0.059$. The main effect for proficiency was not significant, $F(1,18)= 0.142$, $p>0.05$, $\eta_p^2= 0.008$.

4.4.4 Main experiment: Initiation Time (IT)

A subject wise 2 x 2 x 3 repeated measures ANOVA was performed with language (L1 and L2) and congruency (neutral, congruent and incongruent) as within subject factors and proficiency (high and low proficiency) as between subject factors. The main effect for language was significant, $F(1,18)= 9.492$, $p<0.05$, $\eta_p^2 = 0.018$ showing the participants took more time to initiate the mouse movement for L1(mean= 455.076 ms, SE= 36.019 ms) than L2 (mean= 448.624 ms, SE= 34.583 ms). The main effect for congruency was not significant, $F(1,18)= 0.739$, $p>0.05$, $\eta_p^2= 0.068$. The main effect for proficiency was significant, $F(1,18)= 7.395$, $p<0.05$, $\eta_p^2 = 0.085$

showing initiation time of the mouse was higher for L2 high proficient participants (mean= 462.307, SE= 49.940) than L1(mean= 449.637, SE= 47.364.

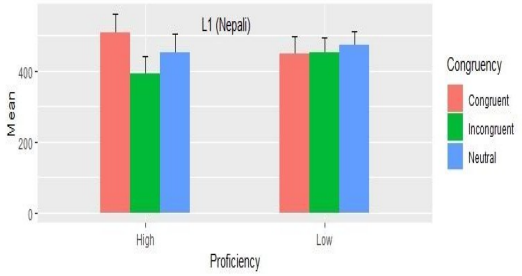

Figure 8: Initiation Time (IT) when the input language was L1

Figure 8 shows the high proficient participants initiated the mouse movement earlier in incongruent condition compared to congruent and neutral condition when the input language was

L1.

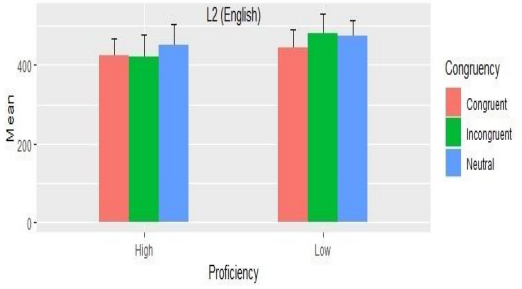

Figure 9: Initiation Time (IT) when the input language was L2

Figure 9 shows the high proficient participants initiated the mouse movement earlier in incongruent condition compared to congruent and neutral condition when the input language was L2. The temporal measurement is in milliseconds.

4.4.5 Main experiment: Response Time

A subject wise 2 x 2 x 3 repeated measures ANOVA was performed with language (L1 and L2) and congruency (neutral, congruent and incongruent) as within subject factors and proficiency (high and low proficiency) as between subject factors. The main effect for language was not significant, $F(1,18)= 1.248$, $p>0.05$, $\eta_p^2 =.0.068$ The main effect for congruency was significant, $F(1,18)= 4.513$, $p<0.05$, $\eta_p^2= 0.361$ showing response time for incongruent condition was higher (mean=2409.304, SE=46.678) than congruent condition (mean=2343.850, SE=54.853,) and neutral (mean=2329.222, SE= 61.695). The main effect for proficiency was not significant, $F(1,18)= 0.026$, $p>0.05$, $\eta_p^2 = 0.002$.

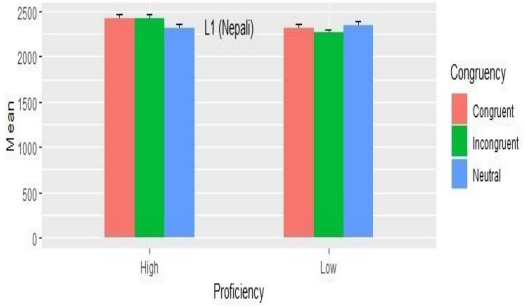

Figure 10: Response Time (RT) when the input language was L1

Figure 10 shows the participants with higher proficiency in L2 were slower in responding in incongruent condition compared to the low proficient participants when the input language was L1.

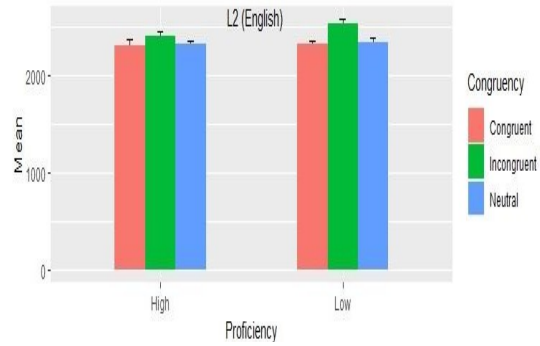

Figure 11: Response Time (RT) when the input was L2

Figure 11 shows the participants with higher proficiency in L2 were faster in responding in incongruent condition compared to the low proficient participants when the input language was L2.

5. Discussion and conclusion

Our study replicates and conforms to the previous findings on cognitive advantage of bilingual

experience in conflict resolution and extends the bilingual advantage discourse further. Incera and McLennan (2015) had reported that bilinguals are slower in initiating the mouse movement but they are faster in responding which was consistent with expertise theory that bilinguals behave as experts by taking time in initiating the action but are faster in making the response. Theirs was a comparison between monolinguals and bilinguals whereas our study compared high proficient and low proficient bilinguals. In our study the high proficient bilinguals were faster initiating the mouse movement and also faster in making response in incongruent condition compared to low proficients.

Currently, there doesn't exist experimental psycholinguistics studies in Nepal. The only detailed psycholinguistic study so far is a longitudinal case study on child language acquisition by Pathak (2004, 2005, 2007). Recently, after the initiation of Cognitive Science and Psycholinguistics Lab in the Central Department of Linguistics at Tribhuvan University, some experimental studies in psycholinguistics have started to emerge from Nepal (for example, Rijal and Pathak, 2020; Rijal, 2020; Pathak, 2020).

6. Further work

The study can be extended further in terms of analysis and increasing the statistical power. Analysis of X and Y Coordinates gives the time course of activation or deviation of the mouse trajectories and shown clearly the Area Under the Curve and Maximum Deviation. The present study has used normalized time analysis. Raw time analysis will capture the mouse movement pattern as it evolves in raw time. This study has rather small sample size, so increasing the sample size will increase the statistical power and higher generalizability.

References

Baum, Shari, & Titone, Debra. 2014. Moving toward a neuroplasticity view of bilingualism, executive control, and aging. Applied Psycholinguistics, *35*(5), 857-894.

Bialystok, Ellen. 2017 The bilingual adaptation: How minds accommodate experience, Psychological Bulletin, 143:3, 233-262.

Bialystok, Ellen, Craik, F. I. M., Klein, R., & Viswanathan, M. 2004. Bilingualism, Aging, and Cognitive Control: Evidence from the Simon Task. Psychology and Aging, *19*(2), 290-303.

Blom, Elma, Boerma, Tessel., Bosma, Evelyn., Cornips, L. Eoni, and Everaert, Emma. 2017. Cognitive advantages of bilingual children in different sociolinguistic contexts. Frontiers in Psychology, *8*, Article ID 552.

Christopher D. Erb. 2018. The developing mind in action: Measuring manual dynamics in childhood, Journal of Cognition and Development

Freeman, John B. & Ambady, N. 2010. MouseTracker: Software for studying real-time mental processing using a computer mouse-tracking method. Behavior Research Methods, 42, 226-241

Freeman, John B, Dale, Richardson, & Farmer, Thomas. A. 2011. Hands in motion reveals mind in motion. Frontiers in Psychology, 2:59

Green, David W., & Abutalebi, Jubin. 2013. Language control in bilinguals: The adaptive control hypothesis, Journal of Cognitive Psychology, 25:5, 515-530

Incera, Sara; Markis, Theresa A. & McLennan, Conor T. 2013. Mouse-Tracking Reveals When the Stroop Effect Happens.Psychology Faculty Publications, Paper 17.

Incera, Sara, McLennan, C.T. 2015. Mouse tracking reveals that bilinguals behave like experts, Bilingualism: Language and Cognition, 1: 1-11.

Klein, George. S. 1964. Semantic power measured through the interference of words with color-naming. American Journal of Psychology, *77,* 576-588.

Paap, Kenneh R., Johnson, H.A., & Sawi, O. 2015. Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. Cortex; a journal devoted to the study of the nervous system and behavior, *69*, 265- 78.

Pathak, Lekhnath S. 2004. Child Language Acquisition: A Case Study Nepalese Linguistics, Vol. 21, November, 2004, pp. 87-93

Pathak, Lekhnath. S. 2005. Thus Spake Poshak: Linguistic Development of a 20 Month Old Child Contemporary Issues in Nepalese Linguistics, Edited by Yadava et al., Linguistic Society of Nepal, 2005, pp. 513-522

Pathak, Lekhnath S. 2007. Poshak Revisited: Further Investigation on Nepali Child Language Acquisition Recent Studies in Nepalese Linguistics, Edited by Rai, N.K, Y.P. Yadava et al., Linguistics Society of Nepal, 2007, pp. 423-437

Pathak, Lekhnath S. 2020. Parallel language activation in Nepali-English-Sanskrit: A psycholinguistic mousetracking study. A mini research submitted to the Research Directorte, Rector's Office, Tribhuvan University, Kathamndu.

Rijal, Sabita. 2020. Effect of first and second language mediated instruction in cognitive control: A psycholinguistic study, MA Thesis submitted to the Central Department of Linguistics, Tribhuvan University.

Rijal, Sabita & Pathak, Lekhnath S. 2020. Bilingual literacy effect on executive control. A poster presented at the 3rd International Workshop of Society for Cognitive Science of Culture at Birla Institute of Technology and Science University, Pillani Campus, Goa, February 3 – 7, 2020

Rosselli, Monica, Adrila, Alfredo, Lalwani, Laxmi. N., and Velez-uribe, Idaly. 2016. The effect of language proficiency on executive functions in balanced and unbalanced Spanish–English bilinguals. Bilingualism: Language and Cognition, *19*(3), 489-503.

Spivey, Michael J., Grosjean, Marc & Knoblich, Gunther. 2005. Continuous attraction towards phonological competitors. Proceedings of the National Academy of Sciences, *102,* 10393-10398.

Stroop, J. Ridley. 1935. Studies of interference in serial verbal reactions. Journal of Experimental Psychology, *18,* 643-662.

Valian, Virginia. 2015. Bilingualism and cognition. Bilingualism: Language and Cognition, *18*(1), 3-24.