Psychology of Language and Communication 2007, Vol. 11, No. 2

HANNA OKUNIEWSKA
University of Warsaw

IMPACT OF SECOND LANGUAGE PROFICIENCY ON THE BILINGUAL POLISH-ENGLISH STROOP TASK

This study addressed the dependence of bilingual Stroop interference effects on L2 proficiency in late, unbalanced bilingualism. Two groups (22 subjects each) of Polish natives, well-educated, hihgly intelligent (Raven's measure) and at age (20-36) of superior inhibitory control (Stroop interference score) participants differed in current English use (self-report) and English proficiency (formally assessed). The effect of proficiency (Mägiste, 1984), the within-language Stroop superiority effect (Goldfarb, Tzelgov, 2007) and the bilingual advantage effect (Craik, Bialystok, 2006) were expected. The two first effects was confirmed which replicates the common findings. The expected advantage in within-language Stroop interference effect among the highly proficient bilinguals was not revealed. The possible reasons for this result are discussed.

Key words: bilingualism, Stroop interference, second language proficiency, inhibitory control

Introduction

In recent years the Stroop paradigm has become one of the best-known procedures in the field of cognitive psychology to study the efficiency of cognitive control mechanisms. In this paradigm, research participants are asked to name ink colors of printed conflicting color words, for example, when the word *blue* is printed in red ink, the correct expected task response is *red*. The delay in color naming evoked in this incongruent condition is called the *Stroop effect* or *interference effect* (Stroop, 1935) and refers to an increase of response time measured in this color-naming condition comparing to the response time in a neutral (control) naming condition (for example: naming the colors of color patches or rectangles or colored Xs). It is universally difficult for anyone who can read to ignore the

Grateful acknowledgment is due to Professor Ida Kurcz who aroused the author's interest in this topic. The study is part of research supported by the grant BSt 1250/41. Address for correspondence: Hanna Okuniewska, University of Warsaw, Faculty of Psychology, Stawki 5/7, 00-183 Warszawa, Poland. E-mail: haniaok@psych.uw.edu.pl

meaning of the word when naming the color it is printed in. The ability to inhibit an automated response (reading the color words) in favor of an unusual one (naming of the ink) was assumed in the Stroop task and discussed as an efficiency of attentional control mechanism by many theoreticians formulating different models of performance of Stroop color-word task (see McLeod, 1991, for review).

Psycholinguistics of bilingualism has frequently used the Stroop paradigm to study the ability to control and regulate bilingual language processing (Tzelgov, Henik, Leiser, 1990; Gerhand, Deręgowski, McAllister, 1995; Lee, Chan, 2000; Roselli, Ardila et al., 2002; Zied et al., 2004; Goldfarb, Tzelgov, 2007). In essence, being bilingual means having more than one active lexical representation to express the same meaning (Desmet, Duyck, 2007). There is common agreement among researchers that in producing a word in a particular language, the conceptual system of a bilingual person activates in parallel the lexical representations of words in both languages (Costa, Santesteban, Ivanova, 2006). The Stroop task is often used to study whether such a parallel activation from semantic representation (color meaning) evokes competition between lexical representations (Roelofs, 2003) or one of the activated representations undergoes further inhibitory mechanisms leading to producing one correct color name in a required language (Green, 1998). In a bilingual person the mappings between these two representations can be distracted through interfering processes of language selection and language switching, experimentally induced. A typical bilingual version of the Stroop task presents as stimuli color terms from bilinguals' first (L1) and bilinguals' second (L2) language and requires color naming answers in either L1 or L2, in congruent and/or incongruent scheme. The reaction time for producing a correct answer is measured and compared between conditions used.

Previous studies using the bilingual Stroop task have demonstrated that both *within-language* and *between-language* interference can be expected (Dalrymple-Alford, 1968; Preston, Lambert, 1969; Mägiste, 1984; Chen, Ho, 1986; Tzelgov, Henik, Leiser, 1990).

Intralingual or within-language Stroop effect takes place when there is a congruency between the language in which color words are written (e.g. Polish) and the language in which answers are requested (Polish). When the language of the written color words (e.g. English) is incongruent with the language of requested answers (Polish), the effect evoked is called between-language interference or interlingual interference (Dalrymple-Alford, 1968).

The standard pattern obtained from a number of bilingual studies is that the Stroop color-naming interference is greater in the within-language condition than in the between-language condition (Dalrymple-Alford, 1968; Preston, Lambert, 1969; Mägiste, 1984; Chen, Ho, 1986). McLeod (1991) concludes his review on Stroop interference in bilinguals with an also recently generally accepted claim of larger Stroop effect for color words appearing in the response language than in another language (the so-called *within language Stroop superiority effect*, WLSSE;

Goldfarb, Tzelgov, 2007, p. 180). A paucity of studies show that the Stroop effect between languages is smaller (reaches approximately 75%) than the within-language Stroop effect (McLeod, 1991).

Three main variables are suggested to influence the pattern of between- and within-language interference in the bilingual Stroop task: Level of L2 proficiency (Mägiste, 1984; Zied et al., 2004), L2 AoA – age of acquisition of L2 (Tzelgov, Henik, Leiser, 1990; Hernandez, Li, 2007) and similarity of L1 and L2 (Chen, Ho, 1986; Lee, Chan, 2000; Sumiya, Healy, 2004). The relative contribution of each of them is still under debate, the variables interact in affecting the bilingual mental lexicon. Language proficiency is considered as the main factor influencing the bilingual processing and can be defined as the degree of control an individual has over a language (Hernandez, Li, 2007). Costa et al. (2006) in a four experiments format explored the impact of language similarity, L2 AoA and L2 proficiency on the bilingual language-switching performance. He finds that neither language similarity nor L2 AoA affects, in a significant manner, the way bilinguals control the language-switching task. Highly and low-proficient subjects differed in pattern of switching cost: For more proficient bilinguals the costs were symmetrical in switching to L1 and to L2 regardless of L2 AoA and language similarity, for low-proficient bilinguals the switching costs were asymmetrical – greater to L1 than to L2.

The role of language proficiency was thoroughly examined also in series of studies applying bilingual Stroop task.

Mägiste (1984) tested German-Swedish highly proficient bilinguals and found for this similar languages (the same alphabetic, similar phonetic and the same German branch) an equivalent amount of Stroop interference, at state of balanced language proficiency among subjects. In contrast, unbalanced German-dominant participants experienced more Stroop interference from German (than Swedish) color terms regardless of German or Swedish response language. Her *language proficiency hypothesis* proposes that bilingual Stroop color-naming task in the dominant language should create a greater interference effect than those in less proficient language and that the interference created by the two languages should become equivalent when subjects are balanced bilinguals.

Chen and Ho (1986) investigated Chinese-English (dissimilar languages) bilinguals from five grade levels (2, 4, 6, 8, 10) and from college. Incongruent color words were printed in Chinese or in English and participants had to name the ink color of each word in such a way that when presented with English (or Chinese in the next condition) word trials, they had to respond in English in half of the trials (within-language condition) and in Chinese on the other half (between-language condition). The authors found some results consistent (when response language was Chinese) and some inconsistent with Mägiste's (1984) findings. When the response was in English, the pattern of interference changed systematically with subjects' experience in English proficiency. It was not a monotonic function of grade level and Stroop's RTs but rather a developmental shift from greater between-language interference (when Chinese was a distractor language and subjects were at first stages of English learning) through the similar amount of both interference effects in grade 4, to greater within-language interference at later stages of English learning. The proposed interpretation is that at the first stages of L2 learning subjects access L2 words via associated lexical representations from L1. As proficiency in L2 increases, dependence on L1 decreases and subjects access the words in new L2 rather through the created in the course of learning more direct links to conceptual representations.

This experimental design poses a question whether using the second language as a response language in subjects at very early stage (grade 2-4) of second language learning was appropriate because of the possibility that latencies obtained depend rather on weak L2 naming skills in children than on the amount of interference they experienced.

Despite the above comment, it is important to stress that the theorizing of Chen and Ho (1986) yielded strong confirmation from results of study with adult Arabic-Hebrew (exp.1) and Hebrew-Arabic (exp. 2) bilinguals of different proficiency and age of acquisition of L2 (Tzelgov, Henik, Leiser, 1990). The pattern of the between- and within-language interference was equal for Arabic-Hebrew bilinguals reflecting the declared equal language proficiency of participants. In exp. 2 Hebrew caused more interference than Arabic in both within- and between-language condition. Hebrew was also the first language of the subjects in exp. 2 and Arabic was learned late in academic studies merely for 4 years. The interference created by Arabic was relatively large only under the within-language condition. The authors confirm the previous interpretation of Chen and Ho (1986) that as proficiency in the second language increases, dependence on the first language decreases and subjects tend to rely on amodal concepts in their lexical selection (Tzelgov, Henik, Leiser, 1990). They obtained two distinct patterns of interference as a function of proficiency: When the second language was of equally high proficiency comparing to L1 and when it was the low-proficient one. In the exp. 1 the interference pattern in L2 did not differ from that revealed in L1, i.e. within-language interference was greater than between-language interference. In contrast, the interference pattern for low-proficient L2 obtained in exp. 2 showed a greater between-language than within-language interference (similar to the results of Chen and Ho's study).

Also Roselli et al. (2000) study of Spanish-English bilinguals and monolinguals supports the claim that the pattern of Stroop interference can change in accordance with language experience of L2. The results did not confirm the relationship between the age of L2 acquisition and Stroop interference.

In the experiment by Zied et al. (2004) the functioning of inhibitory mechanisms (as measured in Stroop color-word interference test) in younger and older French-Arabic bilinguals were examined. Aging was found to be accompanied with a decline in Stroop inhibition in bilinguals with pattern depending on their level of proficiency. The results documented that balanced bilinguals (young and old) were always responding faster to all Stroop conditions than their unbalanced counterparts and that pattern of performance observed was different for both groups. It seems as if being a balanced bilingual determines the improvement of inhibitory abilities in Stroop interference in comparison to an unbalanced status. Unbalanced older Arabic-dominant bilinguals obtained more between-language interference with French responses and older French-dominant bilinguals presented more between-language interference with Arabic responses. Unfortunately, in their article the authors did not present the whole set of results obtained in this interesting examination and concentrated on the most salient ones with regard to predictions formulated.

In sum, research using the bilingual Stroop task has demonstrated interference in low and high proficient bilinguals in both intra and interlingual conditions. Neither the switching language paradigm nor the Stroop paradigm appeared sensitive to L2 AoA influence and language similarity can at best modify the size but not the pattern of interference. The relationship between language proficiency and the pattern of interference was established although the particular ratios may vary between bilinguals groups and examination designs. Moreover, as the Stroop effect seems a good measure for frontal lobe function and inhibitory processes (Lezak, 1995), the possible influence of lifelong bilingualism on efficiency of attentional inhibitory control has also been documented with the Stroop paradigm.

An issue that has not been adequately addressed in previous studies is how the second language learning starting relatively later in ontogenesis but leading to proved high proficiency in the L2 (unbalanced, L1-dominant; the so-called late bilinguals) influences the pattern of interference effects in comparison with low-proficient learners of L2 (unbalanced, L1-dominant learners). It is fair to ask this question taking into account that certain cognitive mechanisms enable us to learn a second language (or more languages) through our life span. Although the level of performance achieved by an adult learner might be less than optimal, fluent late bilinguals are able to select a word in a target language without too much interference. Moreover, as Mechelli et al. (2004) have shown, second language learning increases the density of grey matter in specific region of the brain (left inferior parietal cortex) and causes its structural reorganization to degree which is modulated by the proficiency attended and the age of second language acquisition. The increased grey-matter density was common to both early and late bilinguals although greater in the early bilinguals. The plastic changes in the brain caused by second language learning together with previous conclusions based mostly on early bilinguals studies could be taken as an essential support for the following expectations: (1) late highly proficient bilinguals would show a greater between-language interference effect than low-proficient learners, with L1 responses; (2) both, late high proficient bilinguals and learners would show more interference in the within-language condition than in the between-language condition; with L1 responses (3) late highly proficient bilinguals would show less within-language interference than learners; with L1 responses.

The goal of the current study was to analyze the impact of late acquired second language proficiency on the performance of the bilingual Polish-English Stroop task and to examine the pattern of within- and between-language interference looking for evidence of bilingual processing advantage as postulated recently by Bialystok and coworkers (Craik, Bialystok, 2006; Bialystok at al., 2005;). The growing body of research supports the claim that bilingualism enhances cognitive control in children and delays the onset of dementia by 4 years in old age (Bialystok, Craik, Freedman, 2007). Bilinguals must exercise central inhibitory control on account of the constant need to monitor and properly use two language systems (Green, 1998). Bilingualism also boosts performance in various executive function tasks for adult participants. The cognitive processes underlying this advantage are not clear and a possible biological background mentioned above did not discriminate the second language learning as a way of gaining the same profits. Therefore, it seems reasonable to ask: what about late bilinguals who also have massive practice throughout their lives in exercising control processes, have a balanced daily use of both languages and are high in their L2 proficiency? Do they demonstrate enhanced attentional control and smaller Stroop effect as the early bilinguals do? If they do, there should be observed a greater advantage in interference (smaller interference effect) in the more trained, more proficient group than in the less trained group of learners. This is stated in hypothesis 3.

In the current study, we will examine the magnitude of the Stroop color-word interference effect in two language conditions. One, in which the Polish words (color names) were printed in incompatible colors of ink and participants were required to name color of the ink while suppressing more habitual tendency to read the color word, and in the second one, when the same color names were written in English and concurred as a distractor with the same ("name the color of ink in Polish") response mode. Evidence from previous studies enables us to confine the experimental tasks to L1 responses. A complex relationship between L1 and L2 proficiency and Stroop interference pattern has been highlighted earlier in a number of studies using dissimilar languages. Reaction times are usually greater in L2 responses than in L1. One response language, L1 (the native one) allows us to have a greater confidence in attributing any changes in performance to interference effects and not to compound it with other linguistic variables coming from using L2 which are not controlled, such as problems with foreign articulation, processing speed, differences in daily use of English etc.

Method

Both age and intellectual level may contribute to performance on the Stroop task (McLeod, 1991). Effort was made to recruit persons from not distant age range and with above average intellectual abilities.

Table 1. Demographic data

Subject group	N	Gender		Age		Years of education		Intelligence (Raven)	
		Male	Female	Mean	SD	Mean	SD	Mean	SD
Group L	22	12	10	23,68	2,75	16,36	2,19	24,68	4,91
Group T	22	9	13	26,91	4,78	17,27	1,88	23,64	4,95
Total	44	21	23	25,30	4,19	16,82	2,07	24,16	4,90

Participants

The L2 high-proficient group consisted of 22 graduates of English studies aged 22-36 years (mean 26.9) employed as English teachers (15 ss.) or translators (7 ss.). We will refer to it as to “teacher and translator group”, in short “Gr.T”. All of them had acquired English after the age of 10, had taken English formal courses for at least 10 years and used English daily at work. It should be stressed that based on self-report of the bilingual participants they were highly proficient in English reading, writing, understanding and communicating and they used these functions in their daily occupational activity.

The L2 low-proficient group consisted of 22 students of different faculties from the University of Warsaw aged 20-28 years (mean 23.7) learning English at the university level A2-B1 (according to CEF). All of them began to learn English late in elementary school (after the age of 10) and continued it in a non-systematic way. They report using only Polish for everyday communication. We will refer to it further as to “learner group”, in short “Gr.L”.

Participants were volunteers recruited in Warsaw via an announcement at the university and contacts with Warsaw high schools. All of them were native speakers of Polish and resided in Poland at the time of the testing. All participants were right-handed and had normal or corrected-to-normal vision and correct colors recognition. The two groups did not differ in the number of years of education and in reading and color naming span as measured by Stroop subtasks I and II. The detail demographic data are shown in Table 1.

Procedure

The participants were tested individually in one session that lasted about 60-80 min. Care was taken to ensure a comfortable and free-from-distraction testing conditions: a separate room, silence, good daytime lighting. The participants were informed that they were taking part in a test study of attention in various visual and linguistic interference conditions. Directly before testing they were shortly interviewed to gather information about English proficiency, age of English acquisition, current use of English as well as common demographic facts. The health condition (exclusion criteria: dyslexia, color blindness, speech or hearing disturbances) was a separate part of the interview. The subjects' ability to identify and accurately name the four colors used in the test was proved by asking them to name the colors of eight rectangular patches which were pointed to by the experimenter (only all correct answers were approved). Subjects received verbal instructions in Polish prior to each task and subtask. The tests were administered by a trained experimenter.

Materials

The Raven procedure was applied prior to the main Stroop testing with the goal to assess nonverbal learning and reasoning abilities on visual material in those well-educated participants with assumed high degree of verbal intelligence (Andreou, Karapetsas, 2004).

Polish version of *Raven's Advanced Progressive Matrices* (Jaworowska, Szustrowa, 1991) is a multiple-choice paper and pencil test designed for very capable subjects and normalized on polish good students population. It was chosen as a test that measures abstract nonverbal reasoning ability basing on visuospatial modality. Applied in its timed version (30 min.) enabled a suitable assessment of perceptual analytic reasoning rate.

Test consists of two parts: an exercise set of 12 items and the main set of 36 items. Each item contains a pattern problem with one part removed and from six to eight pictured inserts of which one contains the correct pattern. Subjects select a correct pattern piece and write its number on an answer sheet. Both the sets and the items between the sets are arranged from the relative simple to the very complex and abstract. Subjects are given a score for each correct answer, the sum of scores for each participant was further analyzed in this study. Converting the raw scores into standardized ranks through tables based on similar age range placed the performance of all subjects on a more than average level.

The Polish bilingual version of the *Stroop* task were applied. The Stroop task material consisted of four A4 (297 x 210 mm) cards. Each card comprised 120 items (color rectangles or color names) presented in 5 columns of 24 items each. Words were printed in Helvetica capitals 12/30 point bold, left justified, with 15 mm space between columns. Four colors were represented: red, green, blue, and brown. Each color was used six times in each column. The Polish names of colors consisted of three syllables and the English names of one syllable. The response language was Polish in each test trial.

On Card I the Polish color words (CZERWONY 'red', ZIELONY 'green', NIEBIESKI 'blue', and BRĄZOWY 'brown') were printed in black and arranged in a quasi-random order. Card II consisted of 120 color rectangles. On Card III the Polish color words (CZERWONY 'red', ZIELONY 'green', NIEBIESKI 'blue', and BRĄZOWY 'brown') were also arranged in a quasi-random order, so that no name was printed in its matching color (e.g. the name for 'red' was never printed in red ink), and no color or name were repeated one after the other in the column.

On Card IV the English color words (RED, GREEN, BLUE, and BROWN) were arranged in the same way as the Polish color words on Card III.

To avoid language order effect, the order of applying Card III and Card IV was counterbalanced among subjects.

The first subtask (word reading) was to read aloud the words on Card I as quickly and as accurately as possible. The second (color naming) subtask was to name aloud the color of series of rectangles. In the third and fourth subtask the subjects were required to name aloud as fast and as accurately as possible the color of ink of the words written on Card III (within-language interference subtask) and Card IV (between-language interference subtask) respectively. The participants were required to complete each card starting in the upper left corner and proceed from the left to the right side of the card and from top to the bottom of each column. Errors should be immediately corrected by participants themselves. The experimenter timed performance at the end of each card with a stopwatch. Performance times in seconds and sum of errors were indicated on the scoring sheet for each part of the test separately. For all further interference analyses an index of the Stroop effect size was used. The index was calculated for each participant as a difference between the baseline score (RT on color naming subtask) minus the interference score (RT on appropriate interference condition, within- or between-language).

Results

The groups differed significantly in age, $F(1,42) = 7.535$; $p < 0.01$, Gr.T was on average older than Gr.L. The groups did not differ in intelligence measured with the Raven Test, $F(1,42) = 0.494$; n.s. No significant effects of age, gender and intelligence on Stroop interference were found in a stepwise linear regression analysis.

An analysis of errors showed no significant differences. The mean percentages of errors were very small in both groups of participants (smaller than 0.5%): in the learners group – 0.38%, 0.15%, and 0.30%, and in the teachers group – 0.15%, 0.34%, and 0.49%, for color naming, within-language interference, and between-language interference condition, respectively.

The mean RT Stroop task raw scores, standard deviation, and mean interference effects are presented in Table 2.

The scores were analyzed in a three-way ANOVA, with level of L2 proficiency (teachers, learners) and order of Stroop interference tasks (English first, Polish first) as between-subjects variables and stimulus language of Stroop interference task (English, Polish) as a within-subjects variable. In all cases below, the significance level was 0.05, and an interference index was taken as a input score for further analyses.

The main effect of order of tasks did not reach statistical significance and order of tasks did not interact with any other variable. In between-subject com-

Table 2. Mean RT scores in seconds

StroopTasks and Effects	Subject groups			
	Group L N = 22		Group T N = 22	
	Mean	SD	Mean	SD
Word reading	57,94	16,93	55,99	18,09
Color naming	118,76	19,89	111,57	18,50
Interference (Polish)	177,97	37,59	175,61	37,68
Interference (English)	151,80	32,70	161,34	34,26
Within-language Interference Effect	59,21	31,18	64,05	32,06
Between-language Interference Effect	33,03	22,80	55,54	29,46

parison the analysis revealed a main effect of level of L2 proficiency. The teacher group appeared significantly slower than learner group, showed greater between-language interference effect, $F(1,40) = 7.828$. In contrast, within-language interference effect did not differentiate both groups.

There was a main effect of stimulus language, $F(1,40) = 20.876$, and the interaction of stimulus language and level of proficiency, $F(1,40) = 5.415$. The stimulus language effect revealed a significant difference between the reaction times for the within-language condition and the between-language condition in the Gr.L ($p < 0.001$) but no significant differences for within- and between-language condition in the performance of subjects from Gr.T.

The mean within- and between-language interference effects are shown in Figure 1.

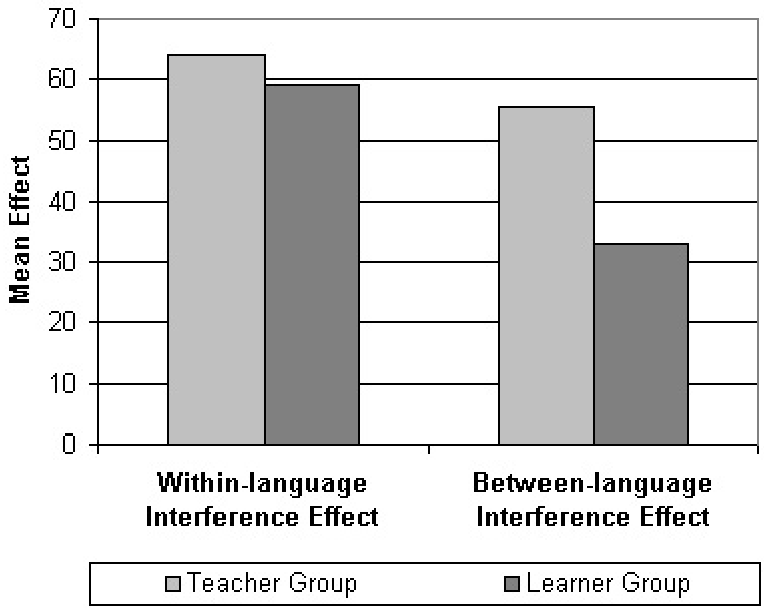
Discussion

This study addressed two main questions. We will now consider them successively in a more complete way. The first one is as follows: What can the performance pattern of bilingual Stroop interference task tell us about control mechanisms which are necessary for processing two languages?

We compared the Stroop task performance in two groups of well-educated subjects, native speakers of Polish, differing in level of English proficiency. We expected to replicate the common effect of proficiency and the within-language Stroop superiority effect as mentioned in the introduction. In our experimental scheme the effect of proficiency could be observed in group comparison (Gr.T vs. Gr.L) on between-language condition, and in condition/subject comparison (within-subject comparison).

The higher level of proficiency in the teacher group accounts for a significantly greater amount of interference than that seen in the learner group when the English

Figure 1. Mean within- and between-language interference effect in both subject groups



incongruent color words were used as distractors. This expected finding fits well in the so-called inhibitory model of bilinguals word production (Green, 1998). For participants more proficient in English (Gr.T), a stronger level of activation of the lexical representations of the nonresponse language required more inhibition to prevent it from being produced than weaker activation caused by the distractor words in participants with low proficiency. It is assumed that the baseline level of activation of L2 lexical items depends on acquired proficiency in L2 and is lower for the less proficient language. It should be stressed that the interview, all test instructions, successive Stroop subtask (reading, colors naming) strongly primed Polish as a response language during the whole examination. According to Grosjean (1998), a state of L1 or L2 activation can influence the task performance. Despite that, the interference from English color words was strong enough to influence the interlingual subtask performance in an asymmetric way in both groups with magnitude parallel to its level of proficiency. It appears that this result supports the inhibitory model and Costa's et al. (2006) suggestion that the amount of inhibition applied to the nonresponse language depends on the difference between the proficiency level of the two languages of a bilingual: the greater the difference, the weaker the interference effect, at least in the case of our young unbalanced bilinguals, similar (late) age of acquisition in subjects, and dissimilar languages.

Thus, the first expectation of the current study was that the between-language interference should be greater in Gr.T than in Gr.L. This was suggested by various studies (e.g. Preston & Lambert, 1969; Mägiste 1984, 1985;) which showed that the between-language interference increased with increasing L2 proficiency. The results of the experiment confirmed this expectation.

The effect of proficiency is noticeable also in another kind of comparison. When we compare the differences in performance on within- and between-language condition among learner group with the same difference for teachers, we realize only one significant WLSS effect. There is a clear WLSSE among learner group but not so evident among teacher group. The latter group shows the same tendency ($p = 0.121$) as well but not at a significant level. According to terms of inhibitory mechanisms, the different ratio between inter- and intralingual interference in both groups can be explained through the higher level of L2 proficiency in Gr.T than in Gr.L, which leads to a higher level of activation of L2 lexical representations and, consequently, requires more inhibition than analogous processing in the lower-proficiency group. Thus, the expected general stronger within-language than between-language interference was confirmed although in Gr.T did not reach statistical significance.

This paper deals also with a second question, i.e. whether late bilinguals of high proficiency in their L2 develop and train through the frequent and regular use of two languages a cognitive inhibitory mechanism to such a degree that enables them to achieve an essential advantage in dealing with Stroop interference. If they do, a greater advantage in interference (smaller interference effect) should be observed in a more trained, more proficient group (Gr.T) than in less trained group of learners (Gr.L) on both interference conditions. It should especially be present in the within-language condition where the interference effect is not influenced by group differences in language proficiency. The hypothesis that bilinguals from Gr.T would outperform the learner group in a Stroop task because of their acquired enhanced inhibitory control has not been confirmed. There was no significant difference between Gr.T and Gr.L performance on within-language condition (learners were even slightly faster than teachers). The here not hypothezied reason could be that the benefiting influence of bilingualism on cognitive control was observed mainly in early, relatively balanced bilinguals (Bialystok et.al., 2004, 2006; Zied et.al., 2004) and all subjects in the current study were late unbalanced bilinguals, i.e. they acquired L2 after the age of 10. There is still the question of the role that the age of acquisition plays for developing cognitive control and bilingual advantage.

However, when interpreting the latter finding we should also consider some facts that could cover the eventually existing enhanced inhibitory efficiency in participants from Gr.T.

Firstly, theoreticians of bilingual advantage have suggested that this expected effect increases with age. Perhaps our bilinguals from Gr.T will experience their advantage in the last age range only.

Secondly, for the guided second language acquisition in another country in scholarly settings a special importance of written communication is typical (Klein, 1986). The role of experience in written communication for Stroop interference effects is demonstrated in a very interesting study of Gerhand, Deręgowski, and McAllister (1995). The authors reported that in Gaelic-English bilinguals living on the Isle of Lewis in Scotland the within-language interference in their dominant language (Gaelic) was virtually absent because of impact of expertise in English literacy. It seems plausible that a long-lasting, intensive experience in written communication in subjects from Gr.T may have affected the pattern of interference we found. The suggestion is that the distracting strength of written L2 may be greater in this circumstances than spoken L1. This issue requires further investigation.

Thirdly, all our subjects are young, at age when attentional control and coping with semantic interference reach the lifespan peak (Craik, Bialystok, 2006). When analyzing changes in the performance on Polish Stroop task across lifespan Okuniewska (2001, 2006) found the 20-29 age range to be optimal for inhibition efficiency. Thus, one can assume that the presumed maximum level of attentional inhibitory abilities in our young bilinguals at least in part coincides with advantage in cognitive control making a separate assessment difficult without further precise examination.

The three arguments against rejecting the bilinguals advantage hypothesis as an explanation of results received are not exhausting enough and it seems that the Stroop paradigm alone in its bilingual version may be not sufficient to resolve issues that remain open.

References

Andreu, G. & Karapetsas, A. (2004). Verbal abilities in low and highly proficient bilinguals. *Journal of Psycholinguistic Research*, 33, 5, 357-364.

Bialystok, E., Craik, F.I.M., & Freedman, M. (2007). Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia*, 45, 459-464.

Bialystok, E., Craik, F.I.M., Grady, C., Chau, W., Ishii, R., Gunji, A., & Pantev, C. (2005). Effects of bilingualism on cognitive control in the Simon task: Evidence from MEG. *Neuroimage*, 24, 40-49.

Bialystok, E., Craik, F.I.M., Klein, R. & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging*, 19, 290-303.

Costa, A., Santesteban, M., & Ivanova, I. (2006). How do bilinguals control their lexicalization process? Inhibitory and language-specific selection mechanisms are both functional. *Journal of Experimental Psychology*: *Learning, Memory and Cognition*, 32, 5, 1057-1074.

CEF – *Common European framework of reference for languages: learning, teaching, assessment (CEF)*. Strasbourg: Council of Europe Publishing 2000.

Craik, F.I.M. & Bialystok, E. (2006). Cognition through the lifespan: mechanisms of change. *Trends in Cognitive Sciences*, 10, 3, 131-138.

Dalrymple-Alford, E.C. (l968). Interlingual interference in a color-naming task. *Psychonomic Science*, 10, 215-216.

Desmet, T. & Duyck, W. (2007). Bilingual Language Processing. *Language and Linguistics Compass*, 1, 3, 168-194.

Gerhand, S.J., Deręgowski, J.B., & McAllister, H. (1995). Stroop phenomenon as a measure of cognitive functioning in bilingual (Gaelic/English) subjects. *British Journal of Psychology*, 86, 89-92.

Goldfarb, L. & Tzelgov, J. (2007). The cause of the within-language Stroop superiority effect and implications. *Quarterly Journal of Experimental Psychology*, 60, 2, 179-185.

Green, D.W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, 1, 67-81.

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition*, 1, 131-149.

Hernandez, A.E. & Li, P. (2007). Age of acquisition: Its neural and computational mechanisms. *Psychological Bulletin*, 133, 4, 638-650.

Jaworowska, A. & Szustrowa, T. (1991). *Podręcznik do Testu Matryc Ravena. Wersja dla Zaawansowanych. Polska standaryzacja* (Manual for Raven's Advanced Progressive Matrices. Polish standardization). Warszawa: Pracownia Testów Psychologicznych PTP.

Klein, W. (1986). *Second language acquisition*. Cambridge: Cambridge University Press.

Lee, T.M.C., & Chan, C.C. (2000). Stroop interference in Chinese and English. *Journal of Clinical and Experimental Psychology*, 22, 465-471.

Lezak, M.D. (1995). *Neuropsychological assessment*. Oxford: Oxford University Press.

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

Mägiste, E. (1984). Stroop tasks and dichotic translation: the development of interference patterns in bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*,10, 2, 304-315.

Mägiste, E. (1985). Development of intra- and inter-lingual interference in bilinguals. *Journal of Psycholinguistic Research*, 14, 2, 137-154.

Mechelli, A., Crinion, J., Noppeney, U., O’Doherty, J., Ashburner, J., Frackowiak, R.S., & Price, K. (2004). Structural plasticity in the bilingual brain. *Nature*, 431, 757.

Okuniewska, H. (2001). Age differences in the performance of the Polish version of Stroop Interference Test. *Polish Psychological Bulletin*, 32, 4 (URL: http://www.insight-library.net).

Okuniewska, H. (2006). Pokonywanie interferencji semantycznej – wpływ wieku (Coping with semantic interference – the effect of age). *Kolokwia Psychologiczne*, 15, 85-103.

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

Rodriguez-Fornells, A., de Diego Balaguer, R., & Münte, T.F. (2005). Executive functions in bilingual language processing. In M. Gullberg & P. Indefrey (Eds.), *The cognitive neuroscience of second language acquisition* (pp. 133-190). Malden, MA: Blackwell.

Roelofs, A. (2003). Goal-referenced selection of verbal action: Modelling attentional control in the Stroop task. *Psychological Review*, 110, 88-125.

Roselli, M., Ardila, A. Santisi, M.N., Arecco, M., Salvatierra, J., Conde, A., & Lenis, B. (2002). Stroop effect in Spanish-English bilinguals. *Journal of International Neuropsychological Society*, 8, 819-827.

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

Sumiya, H., & Healy, H. (2004). Phonology in the bilingual Stroop effect. *Memory and Cognition*, 32, 752–758.

Tzelgov, J., Henik, A., & Leiser, D. (1990). Controlling the Stroop interference: Evidence from a bilingual task. *Journal of Experimental Psychology*: *Learning, Memory and Cognition*, 16, 5, 760-771.

Zied, K.M., Philippe, A., Pinon, K., Havet-Thomassin, V., Ghislaine, A., Roy, A., & Le Gall, D. (2004). Bilingualism and adult differences in inhibitory mechanisms: Evidence from a bilingual Stroop task. *Brain and Cognition*, 54, 254-256.