Original Papers

Polish Psychological Bulletin
2016, vol. 47(2) 228–232
DOI - 10.1515/ppb-2016-0026

Anna Turula[*]

Thinking in a foreign language, fast and slow

Abstract*: Several studies (Keysar et al., 2012; Lazar et al., 2014) suggest that decisions made in a foreign language are more rational. The authors imply that when thinking in a language which is not our native tongue, analytical, slow, deep-thinking is activated. The question that underlies the present article is whether this is a characteristic of every mental operation in the foreign medium. Studies carried out by Costa et al. (2014), Geipel et al. (2015) and Hadjichristidis et al. (2015) suggest the issue is much more complex than it may seem.*
The answer to the question above was sought through a study in which 84 Polish advanced users of English as a foreign language were asked to solve mathematical problems from the Cognitive Reflection Test (Frederick, 2005). Initially the subjects were randomly assigned to two groups, who subsequently solved the problems in Polish (native tongue) and in English (foreign language). The article presents the results, discusses them and arrives at a number of conclusions as well as implications for further research.

Key words*: FL effect, CRT, working memory, affect, cognition*

1. Thinking fast and slow: the two systems

Human beings think in two different ways: automatically, based on intuition, as well as more reflectively, processing incoming data with deliberation and care. Kahneman (2011) calls these two types of thinking fast and slow, respectively, and ascribes them to the operation of two different systems, System 1 and System 2.

The two systems collaborate and reinforce each other in ways that are advantageous and lie at the core of our survival and evolutionary success as a species. However, there are also drawbacks to this collaboration. System 1 is fast and emotionally loaded, whereas System 2 is "lazy" and, as a result, prone to trusting its intuitive counterpart too much. This results in the generation and propagation of different kind of biases, prejudices and stereotypes. The conclusion Kahneman (2011) proposes is as simple as it is disquieting: we are decidedly less rational than we think we are. And even if our rational thinking has a chance to step in, the first impressions created by System 1 are very difficult to override.

An example of this failure to reconsider erroneous intuitive solutions was demonstrated in a study by Frederick (2005), whose aim was to show the relationship between certain cognitive abilities and decision making[1]. The tool applied in the study, the Cognitive Reflection Test (CRT), consisted of problems whose solution required deep, reflective thinking typical of System 2. At the same time, however, each task was constructed in such a way as to provoke System 1 into offering an intuitive (and incorrect) answer. What Frederick (2005) discovered was that for all three items on the CRT the intuitive answers were very frequent.

A similar tendency to opt for thinking fast, even if against what seems logical, was demonstrated in the much earlier *Asian disease* experiment by Tversky and Kahneman (1981). Based on their results, the authors observed that "people systematically violate the requirements of consistency and coherence", and they traced these violations to "the psychological principles that govern the perception of decision problems and the evaluation of options." (1981, p. 453).

2. Turning on the rational system: the power of foreign language

However, as discovered in various studies, the rationality of thinking can be secured in a number of ways, which include: any kind of strain, perceptual or cognitive, a direct comparison of options or a lack of time constraints. Additionally, sufficient motivation and awareness raising are beneficial factors (Kahneman, 2011).

Increased rationality in the choices made has been noted in one other type of research setting: when the problems are

[*] **Pedagogical University, Krakow, Poland**

[1] Time and risk preference; for details and the discussion of the correlation of these two measures and the CRT scores – cf. Frederick (2005).

considered in a foreign language (FL). One such study (Keysar et al., 2012), based on a modified version of the *Asian disease* experiment, showed that while the results for the subjects processing the problems in their native tongue (NT) were similar to those obtained by Tversky and Kahneman (1981), in the FL respondents the effect of the gain / loss frames did not apply, proving that "a robust asymmetry of risk preferences disappears when a decision takes place in a foreign language" (Keysar et al., 2012, p. 664). Keysar et al. ascribe this effect to a reduction of the emotional resonance which normally occurs in NT processing. Their interpretation is corroborated by a later study (Lazar et al., 2014), which, based on three physiological measures (Electrocardiogram, Galvanic Skin Response and Elecroencephalogram), showed that its 69 subjects scored lower on the positive affect arousal scale when using L2 in decision making.

These results, however, were only partly confirmed in Costa et al.'s (2014) replica of Keysar et al.'s (2012) study. The new research corroborated the previous finding that decisions involving loss aversion made in a foreign language showed less irrationality. However, Costa et al. extended the study on the FL effect in decision making by adding a CRT test, in English and Spanish, where the tasks require decisions which are emotionally neutral. The results of this part of the study showed no FL effect, leading to a conclusion that there are certain boundaries to this phenomenon.

The study by Costa et al. (2014) is not the only research to show that the question of the FL effect may be much more complex. In two very recently published studies (Geipel et al., 2015 and Hadjichristidis et al., 2015) the authors argue that the FL effect may actually amount to a reduction in the intuitive, emotional thinking, rather than activating the rational processing mode.

3. The question

In the light of the considerations above, it may be debated whether reasoning in a foreign language is actually the catalyst for BOTH – as Keysar et al. (2012) claim – emotional distance and deliberation. In this context it seems interesting to consider once again the limitations of the phenomenon pointed by Costa et al. (2014), as well as the argument for the FL-induced disactivation of the intuitive processing mode – rather than the reinforcement of the rational mode – put forward by Geipel et al. (2015) and Hadjichristidis et al. (2015). As a result, the question to be asked addresses the character and scope of the FL effect in problem processing.

An important issue to consider in this context is the potential detrimental effect of processing mathematical – as opposed to non-mathematical – tasks in a foreign language. It seems that such an effect may result from a certain cognitive overload. This is because the worktable on which the problems are accommodated – the working memory (WM) – is in charge of foreign language (Robinson, 2001, 2005) as well mathematical (Ashcraft and Krause, 2007; Raghubar et al., 2010) processing. If we add to this the negative affective arousal typical of FL contexts (Lazar et al., 2014), as well as the tensions accompanying solving mathematical problems (Ashcraft and Krause, 2007), we may be looking at serious emotional strain. Such a strain is likely to result in ego depletion, which Kahneman (2011) sees as responsible for falling back on System 1 in input processing.

4. The study

The answer to the question of the role and scope of the FL effect on CRT-like tasks, including the working memory issue, was sought in a study carried out in 2014 at the Pedagogical University in Krakow, Poland. The study departed from the Null Hypothesis:

H#0: There will be no FL effect, as demonstrated by Costa et al. (2014).

as well as two alternative, one-tailed hypotheses:

H#1: Solving CRT tasks in a foreign language will lead to greater consistency and coherence of thinking and give better results in the test, as proposed by Keysar et al. (2012).

H#2: solving CRT tasks in a foreign language will give inferior results in the test.

In order to verify the hypotheses, 84 students from the MA programme in English Studies were asked to write a modified CRT[2]. The respondent group consisted of 42 first-year (Y1) and 42 second-year (Y2) students. The composition of both groups was similar in that the Y1 group contained 30 full-time and 12 part-time students, with the Y2 group comprising 32 full-time and 10 part-time students. Both Y1 and Y2 students were randomly assigned to the native tongue (NT) and foreign language (FL) subgroups (NT=42; FL=42), and asked to solve the test tasks in Polish and English, respectively. The level of English of the participants in the study can be roughly estimated as B2-C1 for the Y1 group and C2 for Y2 based on two facts: Y1 were BA English Studies graduates which, at least by Polish standards, means they have – generally – reached level C1; the level of Y2 was known because all the respondents had passed their end-of-year-1 examination achieving 60% or more which is C2.

The tool used was a four-item test including the three original CRT tasks (problems 1–3), plus an additional ROSES task (problem 4), borrowed from Kahneman (2011). Task 4, which tests logical but not mathematical thinking, was included to check the extent to which English and mathematics actually competed for WM resources. As for the NT version of the CRT, the problems were translated into Polish with slight cultural (problem 1) and technical (problem 3) adaptations, both of which were meant to ease the processing load by increasing the familiarity of the task content. Another simplification – intended, like the previous adaptation, to help unblock System 2 (cf. Kahneman, 2011) – involved turning the original gap-fill test into a choice task offering a direct comparison of options (two

[2] Both language versions of the CRT used are included in Appendix 1.

answers, the correct and the intuitive were given). As for the procedure, all the respondents received the test, in Polish (NT) or English (FL), and were given 4 minutes to complete it. They were instructed to mark the answer which seemed correct without any written calculations.

The results of the modified version of the CRT are presented in Table 1 and Figures 1–2. As can be seen, the native tongue (NT) correct answers outnumber those from the FL version of the modified CRT, both overall as well as for Y1 and Y2 when analysed separately. Consequently, the accuracy ratios, not very high in any case, are lower if the tasks were processed in the foreign language (overall .49/.36). The scores with regard to the year of study show that Y2 outperformed Y1 on both measures (NT: .52/.46; FL: .38/.34). However, the NT/FL difference between the two groups is statistically non-significant ($\chi^2 = 0.001$; $p = 0.97$).

Table 1. The number of correct answers and accuracy ratios for the test

	Total	Y1	Y2
Number of correct NT answers	83	39	44
Accuracy ratio for NT answers	0.49	0.46	0.52
Number of correct FL answers	61	29	32
Accuracy ratio for FL answers	0.36	0.34	0.38

Figure 1. The overall results of the modified version of the CRT for the native tongue (NT) and foreign language (FL)

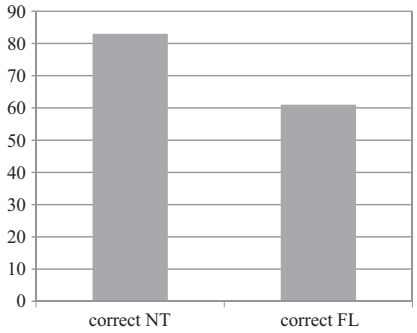
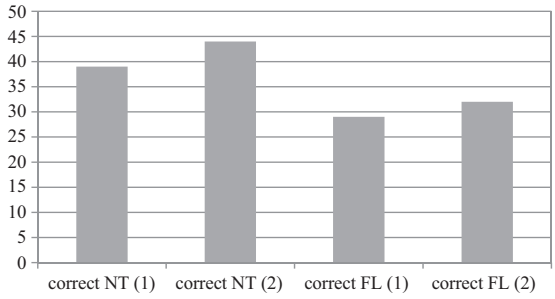


Figure 2. The overall results of the modified version of the CRT for the native tongue and foreign language in the first (1) and second (2) year of studies

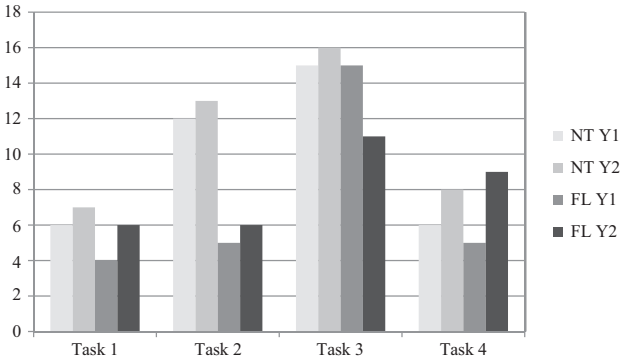


The answers to the individual tasks are presented in Table 2 and Figure 3.

Table 2. Number of correct answers and accuracy ratios for individual tasks

	Task 1	Task 2	Task 3	Task 4
NT Y1	6 0.28	**12** **0.57**	15 0.71	6 0.28
NT Y2	7 0.33	**13** **0.62**	16 0.76	8 0.38
FL Y1	4 0.19	**5** **0.24**	15 0.71	5 0.24
FL Y2	6 0.28	**6** **0.28**	11 0.52	9 0.43

Figure 3. The overall results of the modified version of the CRT for the native tongue and foreign language for Tasks 1, 3 and 4 as opposed to Task 2



The number of correct answers for both languages at the two proficiency levels in the tasks other than Task 2 are comparable: the respondents found questions 1 and 4 rather strenuous cognitively, and 3 relatively easy, regardless of the language (Table 3). This, however, is not the case for Task 2, whose difficulty increases more than two-fold when it has to be processed in the foreign language (Table 2 and Figure 3). Admittedly though, the difference, similarly to Tasks 1, 3 and 4, is statistically non-significant (Table 3)[3].

5. Discussion

For tasks 1, 3 and 4 the results of the study corroborate the Null Hypothesis: while the use of the foreign language in the processing of mathematical and logical tasks has a slightly detrimental effect on the results, the FL effect on processing is statistically non-significant. In this way the study corroborates Costa et al.'s (2014) findings that there are some boundaries to the FL effect. This refers to both types of tasks: mathematical (m: 1 and 2) and non-

[3] Calculations done with the use of *Calculation for the Chi-Square test: An interactive calculation tool for chi-square tests of goodness of fit and independence* (Preacher, 2001).

Table 3. The difference between the NT and FL scores for individual tasks vis à vis overall scores

	Task 1		Task 2		Task 3		Task 4	
	χ^2	p	χ^2	p	χ^2	p	χ^2	p
Y1	0.05	0.82	2.9	0.08	1.15	0.28	0.17	0.68
Y2	0.2	0.65	2.4	0.12	0.06	0.81	2.01	0.15
all	0.59	0.9	5.29	0.15	2.45	0.48	4.13	0.25

mathematical (nm: 4). A given task is either difficult (tasks 1 [m] and 4 [nm]) or relatively easy (task 3 [m]), and this distinction is not affected by the language of the test.

In the light of the above, it seems necessary to reconsider the word RATIONAL which Keysar et al. (2012) use when discussing their results. It appears that in making claims about the role of the FL in decision making and other operations involving interpretation, reasoning and judgement, we need to differentiate between rational understood as (i) "unemotional" and (ii) "reflective", a point also made by Costa et al. (2014, p. 249), who ascribe the lack of the FL effect in the CRT to the "unemotional" character of the tasks. As a result, Keysar et al.'s (2012) claim about the potential of the FL to increase emotional distance AND deliberation needs revision. While decisions made in the foreign language tend to be more rational in the sense that they become less emotional (Costa et al.'s 2014), it is doubtful whether foreign language processing of tasks actually increases deliberation or reinforces cognitive abilities such as interpretation and reasoning. All this, observed in the present study, is very much in line with the results reported recently by Geipel et al. (2015) and Hadjichristidis et al. (2015).

The argument above does not fully hold for Task 2, whose results incline towards corroborating H#2 rather than H#0, showing that the claim made by Costa et al. (2014), that is there is no FL effect related to the CRT, may have its boundaries as well. This is said extremely cautiously, as the uniqueness of Task 2 proved non-significant upon statistical analysis (Table 3). However, the FL / NT difference is notable (more than two-fold; cf. Figure 3), and as such potentially worth considering and in need of further research.

The detrimental FL effect in Task 2 may have resulted from working memory overload connected with the very high cognitive strain of the task itself. It required counting backwards, an extremely intense mental operation, which demands a lot of – if not full – attention (unlike counting forwards or simple calculations – tasks 1 and 3 – which become automatic with age [NT] or with growing language proficiency [FL]). Related to this could have been an additional factor: mutually reinforcing mathematical and language anxieties (Ashcraft and Krause, 2007; Kahneman, 2011; Lazar et al., 2014). This, however, is another tentative comment, in need of further investigation.

What is also interesting is that there was a visible – though again statistically non-significant – difference between the Y1 and Y2 students, with the latter performing better on the FL version of Task 2. This observation leads to another issue worth considering in any future research: the role of proficiency in the foreign language used in the task processing. Keysar et al.'s (2012) respondents for the modified *Asian disease* experiment declared FL levels which can be described as approximately intermediate[4]. This means that when talking about FOREIGN LANGUAGE EFFECT (ibid.), the authors are actually referring to how rational thought was affected at FL INTERMEDIATE proficiency levels. We do not know – as Keysar et al. (2012) did not investigate – what FL effect would have been noted with more advanced learners.

The proficiency problem may actually be very difficult to solve. On the one hand, the higher the FL level the more automatic the processing and, consequently, the lower the COGNITIVE interference between the foreign language and the requirements of the task at hand, especially those of the CRT-type. On the other hand, more advanced FL users are likely to be as EMOTIONAL in their decision-making as those processing information in their native tongue. As a result, we are facing a dilemma: thinking in the foreign language will be both faster (System 1) and slower (System 2) than native tongue processing. This will be the case at any level of FL proficiency, but in different ways. Processing input in a foreign language will be more emotional and less of a cognitive strain for advanced language users; less emotional but debilitating for intermediate users. Consequently, it is clear that in thinking fast and slow there is no such thing as THE FOREIGN LANGUAGE EFFECT PER SE. We can only talk about the different effects FL processing can have on rational thought, always bearing in mind the two, and not necessarily identical, meanings of the word RATIONAL: "unemotional" and "reflective".

All in all – a conclusion which is made cautiously due to the small sample size of this research, as well as the fact that the differences between the task results reported are notable yet statistically non-significant – it can be said that the present study corroborates the findings of Costa et al. (2014), as well as Geipel et al. (2015) and Hadjichristidis et al. (2015). However, more research seems to be necessary

[4] When asked to assess their FL proficiency on a 1–10 scale, Keysar et al.'s (2012) respondents in experiments (described earlier in this article) declared: 4.2 for Japanese; 4.4 for English; 3.8 for French.

into the role of language proficiency in the FL effect as well as the constraints, both cognitive and affective, resulting from the combined mathematical / FL clutter in the working memory.

References

Ashcraft, M. H., & Krause, J. A. (2007). Working memory, math performance, and math anxiety. *Psychonomic Bulletin & Review, 14*(2), 243–248.

Costa, A., Foucart, A., Arnon, I., Aparici, M., & Apesteguia, J. (2014). "Piensa" twice: On the foreign language effect in decision making. *Cognition*, *130*(2), 236–254.

Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives, 19*(4), 25–42.

Geipel, J., Hadjichristidis, C., Surian, L. (2015). How foreign language shapes moral judgment. *Journal of Experimental Social Psychology, 59*, 8–17.

Hadjichristidis, C., Geipel, J., & Savadori, L. (2015). The effect of foreign language in judgments of risk and benefit: The role of affect. *Journal of Experimental Psychology: Applied, 21*(2), 117–129.

Kahneman, D. (2011). *Thinking, Fast and Slow.* London: Penguin Books.

Keysar, B., Hayakawa, S. L., & An, S. G. (2012). The Foreign-Language Effect: Thinking in a Foreign Tongue Reduces Decision Biases. *Psychological Science, 23*, 661–668.

Lazar, J. N., Stern, A., & Cohen, R. (2014). Decision-Making in a Foreign Language Reduces Emotional Arousal. *Psychology, 5*, 2180–2188.

Preacher, K. J. (2001). Calculation for the chi-square test: An interactive calculation tool for chi-square tests of goodness of fit and independence [Computer software]. Available from http://quantpsy.org.

Robinson, P. (2001). Individual differences, cognitive abilities, aptitude complexes and learning conditions in second language acquisition. *Second Language Research 17, 4*, 368–392.

Robinson, P. (2005). Aptitude and second language acquisition. *Annual Review of Applied Linguistics, 25*, 46–73.

Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences, 20*, 110–122.

Tversky, A., & Kahneman, D. (1981). The Framing of Decisions and the Psychology of Choice. *Science, New Series*, *211*(4481), 453–458.

Weissheimer, J., & Mota, M. B. (2009). Individual differences in working memory capacity and the development of L2 speech production. *Issues in Applied Linguistics, 17*(2), 93–112.

Appendix 1

The CRT test, modified

NT

Odpowiedz na pytania niżej. Masz na to 4 minuty (1 min / zadanie).

1. Rakietka do badmintona i lotka kosztują razem 110 zł. Rakietka jest o 100 zł droższa od lotki. Ile kosztuje lotka?

   5 zł 10 zł

2. Staw zarastał rzęsą w taki sposób, że codziennie obszar zarośnięty podwajał się. Po 48 dniach staw zarósł całkowicie. Którego dnia staw był zarośnięty w połowie?

   24-tego 47-tego

3. Pięć maszyn piekarskich potrzebuje pięciu minut, żeby zagnieść z ciasta pięć bułek. Ile czasu potrzebuje 100 maszyn, żeby zagnieść 100 bułek?

   5 min 100 min

4. Niektóre kwiaty szybko więdną. Róże to kwiaty. Niektóre róże szybko więdną – tak czy nie?

   tak nie

FL

Answer the following questions. You have 4 mins (1 min / task).

1. The baseball set – a bat and a ball – cost $11. The bat is $10 more expensive than the ball. How much is the ball?

   50 cents $1

2. In a lake, there is a patch of lily pads. Every day, the patch doubles in size. It takes 48 days for the patch to cover the entire lake. How long would it take for the patch to cover half the lake?

   24 days 47 days

3. It takes 5 machines 5 minutes to make 5 widgets. How long would it take 100 machines to make 100 widgets?

   5 mins 100 mins

4. Some flowers fade quickly. All roses are flowers. Some roses fade quickly – true or false?

   true false

ACKNOWLEDGEMENTS

I would like to thank the 84 students from the English Studies programme at the Pedagogical University in Cracow, Poland, for taking part in my research; Ann Cardwell for proofreading my text; and my colleague and friend Joanna Bielska for her help considering statistics.