

Answers to Dr. Ferenc Kemény's opponent review

Thank you for your precious comments and suggestions on my PhD thesis. I am also very grateful for your words of appreciation on the overview of the related literature and the applied research methods. Your questions, which can contribute to the further refinement of this study, made me think my research over, and here I take the opportunity to answer them.

Question 1: How CLIL learners and general language programme participants could be integrated to the monolingual-bilingual continuum.

Considering your suggestion as for the need of a more thorough integration of the two halves of the introduction, I assume that the integration of the general language programme and the CLIL participants to the monolingual-bilingual continuum is not a clear case for a few reasons. Cook and Singleton (2014:6) posit that all learning (using) of another language changes people's thinking to some extent since learners (users) 'no longer see the world in quite the same way as monolinguals'. However, in the research the importance of mapping participants' language background is not often emphasized; therefore participants' pure monolingualism cannot be taken for granted. In addition, there is a scarcity of studies focusing on the cognitive or linguistic consequences of using an L2 at different levels.

According to Grosjean's (2016) psycholinguistic approach, bilinguals use their languages with different people, in different domains and modalities. Even if bilinguals have different proficiency levels in the four language skills of their languages, they can use their languages to achieve their goals, which he defines as functional bilingualism. Proficiency, fluency, frequency, and the number of domains in which the languages are used are strongly correlated and have an impact on the level of bilinguals' fluency: the more situations a language is used in, the higher level of mastery can be achieved.

Considering these viewpoints and the available information on the participants' language background, it can be concluded that CLIL learners must have preceded their peers by studying according to a general curriculum in the monolingual-bilingual continuum at the time of data collection.

Question 2: Would the author expect that controlling for parental education/SES modifies the results of the current study?

My decision on investigating participants' socio-economic status can be traced back to two facts: first, some authors (Verspoor et al., 2015; Dallinger et al., 2016) suggest the careful consideration of specific individual factors (like socio-economic status) on the learning outcomes and secondly, in the international literature, CLIL programmes are often accused of being elitist and excluding unprivileged individuals. Controlling for this variable contributed to the confirmation of the comparability of the two groups.

Question 3: How can the nature of accents/dialects be characterized in Hungarian?

I absolutely agree with your comment related to the questions on accent in the LEAP-Q questionnaire. As you pointed out the detection of L1 is less effortful in L2 than that of L2 in

L1. However, Lengyel (2021) found that neither bidialectism nor accents in the Hungarian language used within the Hungarian borders differ so much that this would affect the L2 pronunciation.

Since the data of this questionnaire only served as background information, questions related to accent were left unchanged for methodological reasons. In order to make the question more meaningful and transparent to the age group, I simplified the term of 'accent' to 'pronunciation' of the particular language. This explanation reflects the results, that is, all participants are happy with their Hungarian pronunciation, while they are not pleased with their English pronunciation.

Question 4: Would the Author expect any changes in the pattern of results when the criteria are controlled for (in the d2-R test)?

You draw my attention to the signal detection theory as a possibility to control for differences in the instruction (criterion) of the d2-R test. However, I would exclude the assumption that the results might have been influenced by the different wording of instructions, since during the test taking sessions I strictly followed the rules declared in the test manual, more specifically, I read aloud the instructions word by word.

Question 5: Could the Author clarify data treatment in general?

I highly appreciate your remarks on the clarification of data presentation. Regarding data treatment I compared both group-wise and individual data, however, not all of them are presented in the text. With the parallel calculation of group-wise and individual data I aimed at ensuring that the slight difference in the sample sizes would not have caused the distortion of data. However, this difference between the two groups was generated by life. Adjusting the sample sizes to one another would have caused the loss of precious data that I intended to avoid.

Regarding data treatment, in general, I strictly kept the rules of GDPR. I had the parents' written consent and learners had the opportunity to quit the process at any time. They were given codes both on paper-based and electronic documents, so they could not have been identified by any unauthorized persons. No sensitive data had been collected or stored. Parents and learners were informed about the aim of the data collection and the method of data storage. I ensured that in the course of publications, reports, presentations and statements, information covered by the obligation of professional secrecy could not have been disclosed to any unauthorized persons.

Question 6: Could the Author detail the results of the PCAs?

Thank you for your question related to the Principal Component Analysis. The reason why I did not emphasize the results of the PCA was twofold: first, I considered them as background information and secondly, since in Creswell's Sequential Explanatory Design (CSED) the quantitative method serves as a starting point that is refined (deepened) by the qualitative method, the involvement of a more complex statistical analysis is not required.

Regarding the CLIL group's L1 skills, four variables created the principal (latent) component in the PCA: speaking, listening, reading and writing. The optimal model consisted of one principal component, which contained 68% of the original information of the other four

variables. (If two components are involved, they contain 82% of the original information of the four variables. In case of three components: 92% and in case of four components: 100%). In other words, the results of the PCA revealed that CLIL group participants found speaking, listening, reading and writing the most determinative skills while practising their L1.

In case of the activities done in CLIL learners' L1, twelve variables have been involved in the PCA, of which six (watching films, chatting, learning, chatting with the family, info search and watching TV) created the principal components that I interpreted as the most dominant activities done by the CLIL learners in their L1. The related data:

Component	Cumulative %
1	50,715
2	64,674
3	76,467
4	85,491
5	93,013
6	100,000

Regarding the CLIL group's L2 skills, four variables created the principal component in the PCA: speaking, listening, reading and writing. The optimal model consisted of one latent component, which contained 77% of the original information of the other four variables. (If two components are involved, they contain 87% of the original information of the other four variables; if three components are involved, they contain 95% of the original information of the other four variables.) In other words, the results of the PCA revealed that CLIL group participants found speaking, listening, reading and writing the most determinative skills while practising their L2.

Component	Cumulative %
1	77,373
2	87,211
3	94,557
4	100,000

In case of the activities done in CLIL learners' L2: twelve variables have been involved in the PCA, of which five (reading, listening to music, watching films, using the internet and info search) created the principal components that I interpreted as the most dominant activities done by them in their L2. The related data:

Component	Cumulative %
1	57,265
2	68,368
3	89,424
4	96,749
5	100,000

Regarding the control group's L1 skills, I stated that all the five variables (speaking, listening, reading, writing and accent) formed the principal component. In other words, the results of the PCA revealed that control group participants found speaking, listening, reading and writing as the most determinative skills while practising their L2.

Component	Cumulative %
1	65,625
2	81,600
3	92,079
4	96,494
5	100,000

In case of control group activities done in their L1, I stated that seven variables formed the principal component, namely reading, watching films, chatting, learning, chatting with the family, info search and watching TV, which I interpreted as the most dominant activities done in their L1.

Component	Cumulative %
1	46,438
2	62,714
3	76,947
4	86,212
5	93,269
6	96,972
7	100,000

Regarding the control group's L2 skills, I stated that four variables (speaking, listening, reading, writing) formed the principal components. In other words, the results of the PCA revealed that control group participants found speaking, listening, reading and writing the most determinative skills while practising their L2.

Component	Cumulative %
1	83,977
2	91,870
3	96,233
4	100,000

Regarding the control group's L2 activities reading, watching films, internet use, chatting, online games, writing texts, use of Skype and info search formed the principal components that I interpreted as the most dominant activities done in their L2.

Component	Cumulative %
1	47,024
2	61,031
3	73,010
4	81,572
5	87,579
6	93,092
7	96,954
8	100,000

I hope my answers refine the picture and you can accept them. Once again, I do appreciate your questions, which made me think my research over again. Thank you for the time and energy that you invested in the review of my dissertation.

Székesfehérvár, 27th August, 2021

Sántha-Malomsoki, Ágnes