Newly- and early-immigrated second-language students' knowledge of arithmetic syntax

JÖRAN PETERSSON

The present study investigated how 259 Swedish, grade 9 students, of whom 90 had an immigrant background, achieved on twelve written test items in the content area of number. Four of the twelve test items required good knowledge of arithmetic syntax, such as when it was appropriate to apply order-of-operation rules and the associative and distributive laws of arithmetic operations. On these four test items, the most-recently arrived students showed on average significantly more knowledge than the students who had immigrated when they were younger and had participated in Swedish schools for longer periods of time. The outcome suggests that these two groups of immigrant students in later school years should be considered as separate sub-categories of second-language students when it comes to teaching, assessment and research.

It is known that being a second-language student in the mathematics classroom may constitute an additional challenge when it comes to participating in the classroom activities (Gorgorió & Planas, 2001; Parszyk, 1999). This additional challenge has also been observed in test situations, in which second-language students sometimes cannot demonstrate their mathematical knowledge, since they do not fully understand the question in the test item formulation (Lager, 2006; Norén & Andersson, 2016; Petersson, 2012). Therefore the Swedish national agency for education (Skolverket, 2016) has developed material for teachers' further education in order to map newly-arrived, second-language students' knowledge in school subjects. In Sweden, the term *newly-arrived* is used in both academic and educational discourse though its definition varies with the context (Nilsson & Axelsson, 2013). The two separate student categories newly-immigrated and early-immigrated immigrated second-language

Jöran Petersson
Stockholm University

Petersson, J. (2018). Newly- and early-immigrated second-language students' knowledge of arithmetic syntax. *Nordic Studies in Mathematics Education*, 23 (3-4), 105–122.

students in the present study are defined in the method section and this definition is the same as in Petersson and Norén (2017). Yet, immigrant students' previous mathematical education, either in their original country or in the country to which they immigrated, seems to have received little attention in this research. Consequently, I will focus in this study on test achievement in arithmetic by newly- and early-arrived, second-language students.

Immigrant, second-language students in mathematics education

Earlier research on second-language students typically focused on the additional challenges of participating in the mathematics classroom while also learning the new language of instruction (e.g. Cummins, 2008; Gerofsky, 1996; Parszyk, 1999; Schleppegrell, 2007). This led to concerns about the other resources that students used to make sense of the mathematics they were learning. For example, Moschkovich (2002) emphasized that other forms of communication, such as gestures and use of materials, must be recognised in the mathematics classroom. Campbell, Davis and Adams (2007) broadened the perspective by developing the idea of a *problem space*, a four dimensional model which highlighted different aspect that they considered would affect the cognitive load of second-language students, when working with a mathematical task. The first dimension is the academic mathematical content and how familiar students are with it. The second dimension is the problem solving process in which the students are expected to represent and communicate their reasoning. The third dimension is that natural language plays a crucial role in both in learning new mathematical content and in understanding formulated problems and communicating their solution. The fourth and final dimension is that in applied mathematical problems, the problem context is often rooted in a popular context and life experience that might be local and thus maybe not so familiar to newly-immigrated students. These four dimensions should be seen as an interaction between the formulation of the task itself and the knowledge that the student already has.

In particular, students that immigrate when they are in their later school years may be at the same time experienced mathematics students and beginner second-language students. They may bring many experiences from their previous schooling as first language students into their mathematics classrooms in their new country. These newly-arrived students may bring mathematical content knowledge from the curriculum in their original country, into the new mathematics classroom based on another curriculum, with possibly a different emphasis in the mathematical content and the contexts used for this learning (Campbell et al., 2007). With respect to the mathematics classrooms, the situation of newly-arrived students is a parallel to Cummins' (1979) interdependence hypothesis of developing a school register in their second language. At least at the country-level, Giannelli and Rapallini (2016) found that in the new country immigrants from high-achieving countries (as measured through PISA-results) on average achieved higher than immigrants from low-achieving countries. This effect should also be seen in the light of socio-economic background; that immigrant children from high achieving countries were often children of skilled immigrants from well-developed countries whereas children from low-achieving countries were more likely to be refugees, seeking asylum.

Another aspect, which seems to have an impact of students' test results, is the achievement profiles of the countries that the students move between (Giannelli & Rapallini, 2016; Petersson, 2017). Of newly-arrived students in Sweden, a majority comes from countries in the Middle East and Eastern Europe and in TIMSS 8th grade 2007 and 2011 most of these countries achieve higher in the mathematical content area of algebra than in the content area of number, while the opposite holds for the Nordic countries (Mullis, Martin & Foy, 2008; Mullis, Martin, Foy & Arora, 2012). Olsen (2006) in PISA 2003 found, what he called, a Nordic achievement profile. An achievement profile displays how and if the achievement in one content area is different from the overall average achievement and it can be defined for example as ”achievement in one content area minus the total achievement”. Drawing on Petersson (2013, 2017), the present study focuses on achievement in arithmetic of newly- and early-arrived, second-language students. In doing this investigation, it was decided to reduce the impact of natural language and cultural life experiences on second-language students' achievement on test item formulations, as discussed by Campbell et al. (2007), and focus on the mathematics content dimension.

Mathematical content dimension

One way of characterising the mathematical content dimension in arithmetic tasks is to determine the extent they require what Linchevski and Livneh (1999) described as ”structure sense”. Structure sense indicates if the tasks require good knowledge of arithmetic syntax, which includes arithmetic laws, such as the commutative law, and arithmetic conventions, such as the order of operations. An order of operations problem could be where the form $a + b \cdot c$ is perceived by some students as $(a + b) \cdot c$. Other potential problems are numerical expressions in the form $a - b + c$, which some students perceive as $a - (b + c)$. This latter example Herscovics and Linchevski (1994) labelled ”detachment from the subtraction symbol”. In Linchevski and Livneh’s (1999) study, one student stated, ”addition and subtraction are at the same level, so we choose what to do first as is more convenient”. In the presence of a detaching subtraction symbol, this student’s perception might lead to an error when parsing the arithmetic syntax if the student finds it more convenient to add b and c first as in $a-(b+c)$ instead of doing the calculations from left to right. In Linchevski and Livneh (1999), another reason to the same syntax parsing error seemed to be that some students perceived memory rules like BOMDAS [Brackets, Of, Multiplication, Division, Addition, Subtraction] as addition having higher priority than subtraction. Linchevski and Livneh (1999) suggested that superfluous brackets emphasizing structure like $a+(b \cdot c)$ and $(a-b)+c$ might support the development of the students’ structure sense for these kinds of numerical (and algebraic) expressions. However, Gunnarsson, Sönnerhed and Hernell (2016) found that using superfluous brackets did not support the students in correctly learning the arithmetic syntax for these cases.

Research question

Arithmetic foundations are laid in early school years (Andrews & Sayers, 2015). This suggests that newly-arrived second-language students may benefit from having a mathematical foundation developed in their first language as compared to their early-arrived classmates, who may have had major parts of their schooling in their second language. The present study tries to explore this by comparing the knowledge in arithmetic syntax of newly-and early-arrived students with the following research question:

> How do the two groups of newly- and early-arrived, second-language students achieve on mathematical tasks focused on arithmetic syntax?

Method

In order to answer the research question, the author constructed a written test and gave it to 259 students in grade 9, being the final grade in Swedish compulsory school. In order to give a clear answer to the research question, the test was designed to have two contrasting sets of test items. One set that focused arithmetic syntax and one that did not. The achievement of the students who were newly- and early-arrived, second-language students was compared using a statistical test. This method section discusses details in designing the student sample, the test instrument and the statistical comparison of the achievement on the test.

Student sample

In research, categorising students, for example, in language groups is sometimes needed to structure the data. However, the categorizing should be seen as contextual and not permanent and Norén and Björklund Boistrup (2013) recommended that categorisations that already exist should be used in order to facilitate comparing the results in different studies. In the present study, the four student categories below were used. These categories were based on whether students were enrolled in either a course for students with Swedish as a second language or a course for students who have Swedish as their first language, enrolment being regulated by the Swedish school act and decided by language experts (SFS 2011:185) and the length of time students were in the Swedish school system. This information was found using a survey, to which the participant students gave informed consent.

Newly2L: Students who were enrolled in the course ”Swedish as second language” and after immigration have entered the Swedish school system during years 8–9 in the school system. These students were categorised as ”newly-immigrated” students.

Early2L: Students who were enrolled in the course ”Swedish as second language” and after immigration have entered the Swedish school system during the first seven years of school. These students were categorised as ”early immigrated” students.

Other2L: Students who were enrolled in the course ”Swedish as second language” and had immigrated before the school starting age or have not immigrated at all, but whose parents may have been immigrants and so the children may not have Swedish as their first language.

Swe1L: Students who were enrolled in the course ”Swedish language” which was designed for students who had Swedish as their first language.

None of the individuals in the Swe1L category had immigrated during school age. Unlike Early2L, the Newly2L students had been in Swedish school system at most for two years. As a consequence they were also more likely to having had a long experience of being mathematics students in their first language in their country of origin. It must be noted that some students in this category may have spent many years in a refugee camp with limited access to education. Notwithstanding, the number of Newly2L was small when compared to the other student subcategories. Therefore, a discussion of the characteristics of the population is needed.

Using data from Statistics Sweden (2017), the total population of students in the category Newly2L can be estimated as the number of 16 year old individuals born abroad in a particular year minus the number of 14 year old individuals born abroad two years earlier. With this estimate, Newly2L are about 2.3 % of that age cohort though it also includes about 0.05 % Nordic immigrants that are not likely to have the same language challenges due to the similarity of the Scandinavian languages. To create a random sample of Newly2L with a minimum 20 individuals would require a total sample of 900 students or more. Since a sample of this size is difficult to manage, the choice was made to use a purposeful sample, see table 1, by selecting schools with a high proportion of second-language students. A challenge with a purposeful sample is that it might not have properties similar to the whole population and this need to be discussed before claiming any generality of the results of the present study. The properties of the purposive sample compared to the whole population will be discussed from two points of view, namely achievement profile discussed above and achievement on the Swedish national test in mathematics.

As is the case with the previous research of Gianneli and Rapallini (2016) and Petersson (2017), a purposeful sample needs to provide information on the countries of origin of the students come as this may affect the achievement profiles. The total population of Newly2L consists of about 66 % immigrants from the Middle East and Eastern Europe, 13 % from The Horn of Africa and 21 % from other parts of the world. The total population of Early2L estimated in the same way has a similar geographic distribution though it is not known to what extent the students in this estimate could be described as second-language students. In the present study a purposeful sample of the students in twelve classes in five schools were asked to participate in the study. Of the Newly2L and Early2L students in this study and included in table 1, 70 % are from the Middle East and Eastern Europe, 21 % are from The Horn of Africa and 9 % are from other parts of the world. Thus, the sample can be considered representative of the population of newly- and early-arrived students in Sweden with respect to the region in the world that the students came from.

The sample was also compared with respect to their achievements in part B of the national test (see test instrument below). The achievement measured as a proportion of correct responses was 46 % for immigrants from the Middle East and Eastern Europe, 41 % for those from The Horn of Africa and 38 % for students from other parts of the world. Thus there are no major differences in the mathematics achievements between students from different geographical regions. With respect to mathematics achievement, measured as a proportion of correct points, the sample in the present study is similar to the national random sample. As a single group, the sample of second-language students achieved 46 % on the national test, which is the same result as the random sample collected by the Swedish National Agency for Education (Skolverket, 2012).

Table 1. *Participants' average achievement in Swedish language and mathematics*

Students' background	Newly2L	Early2L	Other2L	Swe1L
Number of students	23	67	56	113
Grade in Swedish ≥ passed	52 %	78 %	86 %	97 %
Mathematics achievement on national test, part B1	49 %	43 %	48 %	56 %

For first language students this figure was 60 %, which is a little higher than the 56 % in the sample (see table 1).

The students in the three second-language categories in table 1 showed that the proportions of passed or higher leaving grade in Swedish language increased with the length of their experience in Swedish schooling.

Test instrument

For the present study, a test was constructed using modified test items from previous Swedish national tests in mathematics for grade 9. Campbell's et al. (2007) problem space with the four dimensions of mathematical content, problem solving process, natural language and life experience, was used as a tool for choosing which parts of the national test to select test items from. The national test in mathematics is mandatory and is used for evaluating the students' knowledge on an individual level as well as at school level and national level. The Swedish national test in mathematics consists of one oral part (A) and three written parts (B, C and D). When evaluated according to the four dimensions of the problem space, part B was deemed more suitable for this study because part B has a lower correlation between mathematics achievement and reading ability than parts C and D (Skolverket, 2009, 2010, 2011). It was also important to reduce the influence on the test outcome in regard to the dimension of culture and life experiences. The test items in parts C and D in the national test are typically formulated in a context with real life applications, while in part B, they typically are put in a "pure" mathematics context (PRIM-gruppen, 2006–2009). Finally, with respect to the dimension of mathematical content in the problem space, part B covers a wider range of test items on number sense, when compared with all the other parts of the national test (PRIM-gruppen, 2006–2009).

In the present study, the test item formulations in table 2 were used. The English translations of the test item are provided in figure 1 in the results. To solve test items S1–S4 (Syntax focused items) in table 2, students need knowledge of arithmetic syntax, such as how and when to apply associative, commutative, distributive laws and the order to operations, while this is not the main focus for test items O1–O8 (Other items) thus contrasting test items S1–S4 with respect to arithmetic syntax. In table 2, the notation ”2009B1#4” referrers to test year 2009, part B1, item number 4, where all the tests are found in PRIM-gruppen (2006–2009). Test items O2 and O4 both asks the solver to determine a fraction of a number, though in O2 the fraction is given in *number format* as 2/3 while in O4 the fraction is given in *word format* as ”hälften” (Eng. ”half”). This makes O2 more complex than O4 with respect to Campbell's et al. (2007) dimension of natural language. Test item O8 was reformulated into a context less embedded in the dimension of life experience.

Table 2. *Test items and sources*

No.	Test item formulation	Source formulation
O1	Hur många minuter är 0,75 h?	2009B1#4; Identical
O2	Vad är 2/3 av 60?	2006B1#7; Vad är hälften av $1\frac{1}{2}$?
O3	Figuren består av rektanglar och trianglar. Hur stor del av figuren är grå?	2007B1#8b; Identical
O4	Vad är hälften av $\frac{1}{5}$?	2006B1#7; Vad är hälften av $1\frac{1}{2}$?
O5	Beräkna 6,32 – 3,44	2006B1#2; Beräkna 15,3 – 8,25
O6	Se tabellen. Hur många grader skiljer det mellan de dagar där temperaturskillnaden är störst?	2009B1#2; Tabell samt frågan: Hur många grader skiljer det mellan de städer där temperaturskillnaden är störst?
O7	Vilket tal ska stå i rutan så att likheten stämmer? 1,365 – ___ = 1,305	2007B1#3; Skriv ett tal i rutan så att likheten stämmer. 1,795 – ____ = 1,705
O8	Skriv talet 1 430 i grundpotensform.	2007B1#11; I Sverige köper vi 120 miljoner tulpaner under vårvintern. Skriv antalet i grundpotensform
S1	Beräkna $3^2 + 2^3$	2013B#5; Beräkna $\frac{10^2}{5^2}$.
S2	Beräkna $\sqrt{9 + 16}$	2007B1#14; Identical
S3	a = 2 och b = 4. Beräkna a(b + 2) + b	2008B1#17; a = 4 och b = -3. Bestäm värdet av a(a + 1) + b
S4	Beräkna 12 – 23 + 9	2010B1#3; Beräkna 15 – 28 + 5

Statistical test

The test results were analysed with respect to achievement (correct/incorrect) and with respect to how they treated the arithmetic syntax as shown in their full responses to the test items. This means a double coding of the students' responses (see e.g. Agnell, Kjaernsli & Lie, 2000). Specifically for second-language students, the double coding made it possible to distinguish between some cases of mathematical errors and errors due to misinterpreting a test item formulation (Petersson & Norén, 2017). For example test item O6 asks for the largest temperature difference between any two days which in theory according to Campbell's et al. (2007) natural language dimension might result in a misinterpretation. For test item O6 in regard to the mathematical content dimension, an error could be from a range of different miscalculations.

The test achievements of early- and newly-arrived, second-language students were compared using a nonparametric test. Since test achievement data can be ranked, they are at least on the statistical ordinal data scale. Yet it might be difficult to give a precise meaning of the difference between the achievements of two individuals. Hence the achievement data was considered as not being on the interval scale. Specifically a Wilcoxon test was used for measuring the statistical significance. The variance in the Wilcoxon test was corrected for the occurrences of equal achievement points of the two student categories. This was done by giving occurrences of equal achievements the average of ranks they would have if there were no equal achievement (Siegel & Castellan, 1988, p. 134). Cliff's d was used for measuring the effect size, which is a sample size independent measure of the substantive significance; which indicates how strong the statistical result is.

Results

The test instrument with the twelve test items in table 2 was given to the 259 students, described in table 1. The test results were assessed as a proportion of correct responses and in regard to the calculation strategies used. In test items S1–S4 in table 2, a linguistic misinterpretation could be traced only to item S3. For test item S3 about 5 % in each of the four student categories, including Swe1L, responded with an expression instead of a value. In test items S1–S4 about 12 % in all four student categories gave *No response* and about 1 % of the errors could not be classified as being mathematical or interpretational. In test items O1–O8, a linguistic misinterpretation could be traced to three test items. In item O3, about 3 % of the second-language students but none of the first language students confused "proportion" with "number" (in Swedish "andel" and "antal" respectively). In item O4, about 9 % of the second-language students and 5 % of the first language students gave a response corresponding to ignoring the formulation "half of". For the word problem in item O6 about 26 % of the second-language students and 12 % of the first language students gave a response corresponding to a linguistically incorrect interpretation of the test item. Examples of this included responding with the largest temperature differences on adjacent days or responding with the days with highest and lowest temperature. In summary, except for test item O6, there were small proportions of linguistic misinterpretations of the test item formulations.

Achievement on the two groups of test items

Figure 1 and table 3 show two things. For the test items O1–O8 taken together, the two categories of Newly2L and Early2L achieved at a similar level on average, with a small advantage for the Newly2L on each test item. For the test items S1–S4 taken together, the Newly2L on average achieved statistical-significance better results than the Early2L. In fact, the Newly2L on average achieved better than all the other student categories, including Swe1L, on test items O8, S2, S3 and S4.

Table 3 describes the statistical properties of these achievement differences. For the group of test items S1–S4, a Wilcoxon test found the difference in achievement between Newly2L and Early2L was statistically significant with $p = 0.00084$. The effect size was Cliff's $d = 0.52$, which is

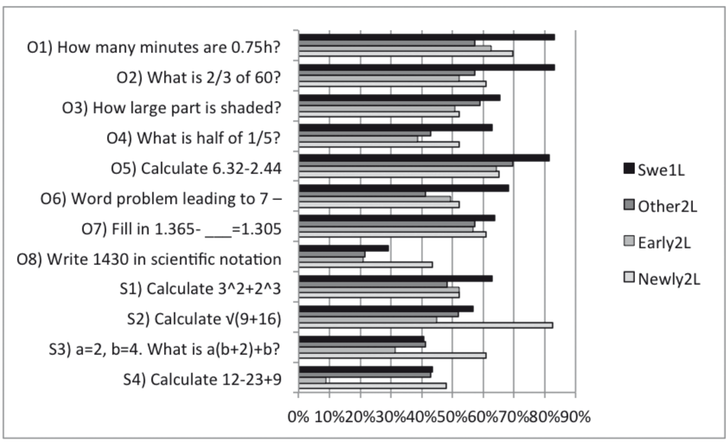

Figure 1. *Achievement on each test item per student category*

Table 3. *Achievement of Early2L and Newly2L for the two groups of test items*

Test item category	Test item category	Early2L	Newly2L	p-value
Test items O1–O8	Average correct responses	49 %	57 %	0.18
	Wilcoxon test rank sum	3146.5	948.5	
Test items S1–S4	Average correct responses	17 %	30 %	0.00084
	Wilcoxon test rank sum	3379.5	715.5	

considered to be a moderate effect size, which means that the strength of the results is between strong and weak. There was no such difference for the group of test items O1–O8. This is interesting since it means that Newly2L and Early2Lachieves similarly on one type of tasks but differently on another type of tasks, indicating that that these two student categories have experienced different mathematics teaching.

Students' arithmetic strategies

The students' correct and incorrect arithmetic strategies for test items S1–S4 are summarised in table 4. The students showed several different solving strategies, some correct and some incorrect. Test item S1 involved the students in calculating $3^2 + 2^3$ and Newly2L and Early2L had similar proportions of correct responses for this test item as shown in table 4. The first type of error in test item S1 was to confuse the meaning of the exponent with multiplication, resulting in $2 \cdot 3$ instead of 2^3. The second type of error in test item S1 concerns the order of operations. Examples of these errors included $(3+2)^{(2+3)} = 5^5$ and $(3 \cdot 2)^{(2 \cdot 3)} = 6^6$. For both the first and second types of errors in test item S1, Newly2L and Early2L had similar proportions of erroneous responses. However, these results differed to those of the other two student categories.

Test item S2 in table 2 was to calculate $\sqrt{9+16}$. Newly2L students generally achieved better than all other groups, including the group of Swedish only students as seen in figure 1. Common errors were to give the square root operation a too short range by calculating $\sqrt{9} + 16 = 19$ or to distribute the square root operation over the numbers to $\sqrt{9} + \sqrt{16} = 7$. These errors were unevenly distributed among the student categories as

Table 4. *Proportions of some typical responses for each of the test items S1–S4*

Test item & response	Newly2L	Early2L	Other2L	Swe1L
S1) Response 5^5 or similar	24 %	26 %	13 %	11 %
S1) Response $2^3 \rightarrow 2 \cdot 3$ or similar	13 %	13 %	27 %	19 %
S2) $\sqrt{25}$ or $5 \cdot 5$	9 %	18 %	2 %	7 %
S2) $\sqrt{9} + 16 = 19$ or similar	0 %	6 %	21 %	11 %
S3) Correct through $2 \cdot 6 + 4$	26 %	3 %	11 %	9 %
S3) Correct through $8 + 4 + 4$	35 %	24 %	29 %	24 %
S3) Correct answer only	0 %	4 %	2 %	8 %
S3) Error in using distributive law	9 %	34 %	30 %	30 %
S4) Erroneous commutation/association	48 %	78 %	54 %	48 %
S4) Unit step error	0 %	4 %	0 %	4 %

shown in table 4. There were also a few cases of misunderstanding the square root operation such as halving $\frac{25}{2} = 12.5$ or squaring at least one of the numbers as in $81 + 16 = 97$.

Test item S3 in table 2 was to evaluate $a(b+2)+b$ given that $a=2$ and $b=4$. To get the correct response, the students calculated either via rules of operation in which *brackets first* as $2 \cdot 6 + 4$ or via the *distributive laws* as $8+4+4$. Table 4 shows that the Newly2L more often used the brackets first and made fewer errors while using the distributive law than did the students in the other categories. Examples of errors in using the distributive law were to only multiply the first number in the brackets as in $2(4)+2+4=14$ or to include all numbers in the multiplication $2(4+2+4)=20$ or to change sign as in $2(4)-2+4=10$. Thus the range – too short or too long – of the brackets was a challenge for some students. For test item S3, table 4 shows large differences in proportions of both correct and erroneous responses for Newly2L and Early2L.

Test item S4 in table 2 was to calculate $12-23+9$. The most common erroneous response was ”20”, which 116 students had arrived at by calculating either $(23-12)+9$ or $(23+9)-12$. These responses can be considered to be an erroneous use of the associative $(23+9)$ and the commutative rule $(23-12)$ under the condition of a detaching subtraction symbol. Twenty students gave the erroneous response ”-20”, after calculating either $12-(23+9)$ or $12-23=-11$ followed by $-(11+9)$ where the minus symbol was erroneousely distributed over the addition. Seven students responded ”2” after calculating $23-(12+9)$ or stating that $-11+9=2$. One student calculated $(12-9)+23=26$. All these responses demonstrated gave an erroneous meaning to the arithmetic operation of subtraction. For test item S4, table 4 shows that Early2L students made larger proportions of these errors than any of the other student categories. Another kind of error was to subtract by counting backwards including *both* start position and stop position as illustrated in figure 2. Since this strategy results in a difference becoming one unit too large, this calculation error was denoted *unit step error.* Eight students made this error

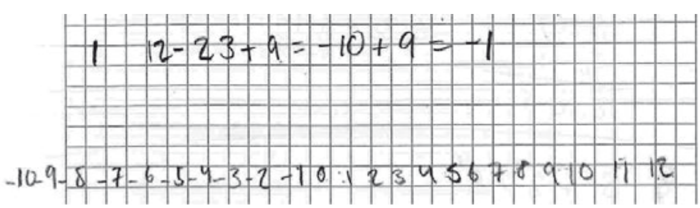

Figure 2. *Unit step error: a student counts backward including both start and stop position*

and of these about half of them also used inappropriate interpretations of commutative and associative properties as described earlier.

Figure 2 illustrates how one student calculated 12 – 23 = -10 by counting 23 steps backwards including both start position 12 and stop position -10 thus getting the stop position as an intermediate result instead of the correct difference -11. Another student wrote the ordered numbers ranging from -23 to 12 and gave the response 19. Four other responses of this kind were to state that -11 + 9 = -3 and to give the answers -19, 1 and 21 respectively. Also for the integer subtraction in test item O6, 5 % of errors were classified as unit step errors in all four student categories.

Discussion

Given the earlier definition of the two student categories Newly2L and Early2L, the present study focused on how these two student categories achieved in a test which had been modified in accordance with Campbell's et al. (2007) problem space, so that there were reduced dimensions of linguistic and life experiences. This was done by focusing on the dimension of mathematical content. As shown in table 1, on average Newly2L achieved lower in Swedish language than the other student categories yet this was not the case in this mathematics test.

The main result, given in table 3 and figure 1, was that the Newly2L achieved statistical significance better results, than Early2L for the group of test items (S1–S4) focusing on arithmetic syntax, while there was no significant difference for the group of other test items (O1–O8). The analysis of the students' responses showed that when compared with Early2L, the Newly2L had lower proportions of errors related to arithmetic syntax as displayed in table 4. Particularly for test item S4, errors due to detachment from the subtraction symbol, as described by Herscovics and Linchevski (1994), was frequent in all student categories, but in particular among Early2L. A small proportion of students made unit step errors in test items O6 and S4, when they subtracted integers, as illustrated in figure 2 (see also Petersson, 2012). The interpretation is that in the present study, when compared with Early2L, Newly2L showed a more solid knowledge of arithmetic syntax.

It may be that Newly2L have had on average longer experience than Early2L as first language mathematics students before immigrating to Sweden. Moreover, Early2L have had on average longer experience of being taught mathematics in their second language. This may have resulted in additional challenges in successfully participating in the mathematics classroom learning activities as a consequence (Campbell et al., 2007; Gerofsky, 2006; Lager, 2006; Mullis et al., 2008; Mullis et al., 2012; Petersson, 2012; Norén & Andersson, 2016; Parszyk, 1999). This could explain why Newly2L achieved better than Early2L on the mathematical content dimension in test items S1–S4 emphasizing arithmetic structure.

But Newly2L also achieved similarly to Early2L and lower than Swe1L on the other test items O1–O7 as well as similarly or better than Swe1L on test items S1–S4. These two different patterns in the achievements of Newly2L suggest that there might be a second component in explaining the results. One suggestion is that Newly2L may have brought culture and life experiences relevant for the mathematics classroom and coloured by their previous schooling in a way similar to the students in the study of Giannelli and Rapallini (2016). In other words this means that Newly2L have experienced more or better teaching in the topic of arithmetic structure. Moreover, these experiences might colour their achievement profile differently to the general "Nordic profile of mathematics achievement" described in Olsen (2006). However, it would need further investigations to confirm this suggestion, for example by comparing classroom teaching or textbook content.

Limitations of the present study and suggestions for further research

The two sub-groups of Newly2L and Early2 are small in number in the present study and a larger study would be useful in confirming the results more generally. However, I believe that §the results have some degree of generality since the second-language students in the present study achieved similarly results to a nation-wide random sample of second-language students, see table 1. It is also possible that the achievement differences might have been smaller if the test items had been more deeply connected to the dimension of culture and life experiences (Campbell et al., 2007) or in the dimension of natural language had been given a more text intensive formulation as in test item O6 (e.g. Gerofsky, 2006; Lager, 2006).

It may be useful for future research to focus on Early2L and Newly2L in late school years as two separate categories with slightly different know-ledge profiles and educational needs. Newly2L are likely to have been taught mathematics in their first language longer than Early2L. Drawing on for example Cummins' (1979) interdependence hypothesis, Newly2L may have had better opportunities to build a mathematical foundation. This may not be the case for Early2L, who in school may need increased support in laying a mathematical foundation. In the present study, Early2L students had large proportions of conceptual errors. Therefore a specific suggestion for further research is to study their opportunities for learning mathematics as being newly-arrived, second-language students in early school years, with the aim to develop teaching methods that better include Early2L in learning central mathematical concepts.

Acknowledgements

Thank you to the anonymous reviewers and my colleagues, who gave beneficial advice on earlier versions of this manuscript and for the editors' support. I thank Professor Astrid Pettersson, Department of mathematics and science education, Stockholm university, for giving access to the random sample from the national test results.

References

Agnell, C., Kjaernsli, M. & Lie, S. (2000). Exploring students responses on free-response science items in TIMSS. In D. Shorrocks-Taylor & E. W. Jenkins (Eds.), *Learning from others* (pp. 159–187). Dordrecht: Kluwer Academic.

Andrews, P. & Sayers, J. (2015). Identifying opportunities for grade one children to acquire foundational number sense: developing a framework for cross cultural classroom analyses. *Early Childhood Education Journal*, 43 (4), 257–267. doi: 10.1007/s10643-014-0653-6

Campbell, A. E., Davis, G. E. & Adams, V. M. (2007). Cognitive demands and second-language learners: a framework for analyzing mathematics instructional contexts. *Mathematical Thinking and Learning*, 9 (1), 3–30. doi:10.1080/10986060709336603

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 49 (2), 222–251. doi: 10.3102/00346543049002222

Cummins, J. (2008). BICS and CALP: empirical and theoretical status of the distinction. In H. N. Hornberger (Ed.), *Encyclopedia of language and education* (pp. 487–499). New York: Springer.

Gerofsky, S. (1996). A linguistic and narrative view of word problems in mathematics education. *For the Learning of Mathematics*, 16 (2), 36–45.

Giannelli, G. C. & Rapallini, C. (2016). Immigrant student performance in math: Does it matter where you come from? *Economics of Education Review*, 52, 291–304. doi:10.1016/j.econedurev.2016.03.006

Gorgorió, N. & Planas, N. (2001). Teaching mathematics in multilingual classrooms. *Educational Studies in Mathematics*, 47 (1), 7–33. doi:10.1023/A:1017980828943

Gunnarsson, R., Sönnerhed, W. W. & Hernell, B. (2016). Does it help to use mathematically superfluous brackets when teaching the rules for the order of operations? *Educational Studies in Mathematics*, 92 (1), 91–105. doi: 10.1007/s10649-015-9667-2

Herscovics, N. & Linchevski, L. (1994). A cognitive gap between arithmetic and algebra. *Educational Studies in Mathematics*, 27 (1), 59–78.

Lager, C. A. (2006). Types of mathematics-language reading interactions that unnecessarily hinder algebra learning and assessment. *Reading Psychology*, 27 (2-3), 165–204. doi:10.1080/02702710600642475

Linchevski, L. & Livneh, D. (1999). Structure sense: the relationship between algebraic and numerical contexts. *Educational Studies in Mathematics*, 40 (2), 173–196. doi:10.1023/A:1003606308064

Moschkovich, J. (2002). A situated and sociocultural perspective on bilingual mathematics learners. *Mathematical Thinking and Learning*, 4(2-3), 189–212.

Mullis, I. V. S., Martin, M. O. & Foy, P. (Eds.). (2008). *TIMSS 2007 international mathematics report: findings from IEA's trends in international mathematics and science study at the fourth and eighth grades*. Chestnut Hill: TIMSS & PIRLS International Study Center.

Mullis, I. V. S., Martin, M. O., Foy, P. & Arora, A. (2012). *TIMSS 2011 international results in mathematics*. Chestnut Hill: TIMSS & PIRLS International Study Center.

Nilsson, J. & Axelsson, M. (2013). Newly arrived students' experiences of pedagogical and social provision in introductory and regular classes. *International Electronic Journal of Elementary Education*, 6 (1), 137–164.

Norén, E. & Andersson, A. (2016). Multilingual students' agency in mathematics classrooms. In A. Halai & P. Clarkson (Eds.), *Teaching and learning mathematics in multilingual classrooms* (pp. 109–124). Rotterdam: Sense Publishers.

Norén, E. & Björklund Boistrup, L. (2013). What is a group? Theoretical considerations when researching affordances for multilingual students' mathematical learning. In M. Berger, K. Brodie, V. Frith & K. le Roux (Eds.), Proceedings of the 7th international mathematics education and society conference (pp. 431–440). Capetown: MES 7.

Olsen, R. V. (2006). A Nordic profile of mathematics achievement: myth or reality? In J. Mejding & A. Roe (Eds.), *Northern lights on PISA 2003: a reflection from the Nordic countries* (pp. 49–61). Copenhagen: Nordic Counsil of Ministers.

Parszyk, I.-M. (1999). *En skola för andra: minoritetselevers upplevelser av arbets- och livsvillkor i grundskolan* (Dissertation). Stockholm: HLS.

Petersson, J. (2012). Rare mathematics – a needle eye for teachers of second language learners. In G. H. Gunnarsdóttir, F. Greinsdóttir, G. Pálsdóttir, M. Hannula, M. Hannula-Sormunen, E. Jablonka, et al. (Eds.), *Proceedings of Norma 11, the 6th Nordic Conference on Mathematics Education* (pp. 483–492). Reykjavik: University of Iceland Press.

Petersson, J. (2013). Late-arrived immigrants in school and performance in arithmetic for negative numbers. In *A. M. Lindmeier & A. Heinze (Eds.), Proceedings of the 37th conference of the International Group for the Psychology of Mathematics Education* (vol. 5, pp. 252–252). Kiel: PME.
Petersson, J. (2017). First and second language students' achievement in mathematical content areas. *Nordic Studies in Mathematics Education*, 22 (2), 33–50.
Petersson, J. & Norén, E. (2017). To halve a fraction: an issue for second language learners. *Education Inquiry*, 8 (3), 173–191. doi: 10.1080/20004508.2016.1275187
PRIM-gruppen (2006–2009). *Subject test in mathematics, grade 9, part B1*. Stockholm University. Retrieved from https://www.su.se/primgruppen/matematik/årskurs-9/tidigare-prov
Schleppegrell, M. J. (2007). The linguistic challenges of mathematics teaching and learning: a research review. *Reading & Writing Quarterly*, 23 (2), 139–159. doi:10.1080/10573560601158461
SFS 2011:185. *Skolförordning* (5 kap. Utbildningen, Svenska som andraspråk, 14§) [Swedish code of statutes, school regulation]. Retrieved from www.riksdagen.se/sv/Dokument-Lagar/Lagar/Svenskforfattningssamling/Skolforordning-2011185_sfs-1994-1194/
Siegel, S. & Castellan, N. J. (1988). *Nonparametric statistics for the behavioral sciences* (2nd ed.). New York: McGraw-Hill.
Skolverket. (2009). *Subject test in mathematics, grade 9 spring 2008* [Ämnesproven 2008 i grundskolans åk 9 och specialskolans åk 10]. Stockholm: Skolverket. Retrieved from http://www.skolverket.se/publikationer?id=2169
Skolverket. (2010). *Subject test in mathematics, grade 9 spring 2009* [Ämnesproven 2009 i grundskolans årskurs 9 och specialskolans årskurs 10]. Stockholm: Skolverket. Retrieved from http://www.skolverket.se/publikationer?id=2372
Skolverket. (2011). *Subject test in mathematics, grade 9 spring 2010* [Ämnesproven 2010 i grundskolans årskurs 9 och specialskolans årskurs 10]. Stockholm: Skolverket. Retrieved from http://www.skolverket.se/publikationer?id=2519
Skolverket. (2012). *Subject tests 2011 schoolyear 9* [Ämnesproven 2011 i grundskolans årskurs 9 och specialskolan årskurs 10]. Stockholm: Skolverket. Retrieved from http://www.skolverket.se/publikationer?id=2733
Skolverket. (2016). *Webbkurs – kartläggning av nyanlända elevers kunskaper* [web-course, mapping of newly arrived students' knowledge]. Stockholm: Skolverket. Retrieved from http://kartlaggning.skolverket.se/
Statistics Sweden. (2017). *Statistikdatabasen*. Retrieved from www.statistikdatabasen.scb.se

Jöran Petersson

Jöran Petersson is a senior lecturer at Stockholm University, Sweden. He wrote his PhD-dissertation on test achievement of second language students in the last year of compulsory school. Presently he is doing post-doctoral research on how foundational number sense appears in text-books and homework in the first year of compulsory school. Jöran also has an interest in mathematical modelling.

joran.petersson@mnd.su.se