Molina, M., Rodríguez-Domingo, S., Cañadas, M. C. y Castro, E. (2017). Secondary School Students' Errors in the Translation of Algebraic Statements. *International Journal of Science and Mathematics Education, 15(*6), 1137-1156. DOI: 10.1007/s10763-016-9739-5

Secondary School Students' Errors in the Translation of Algebraic Statements

Molina, M., Rodríguez-Domingo, S., Cañadas, M. C. y Castro, E.

Abstract.

In this article, we present the results of a research study that explores secondary students' capacity to perform translations of algebraic statements between the verbal and symbolic representation systems through the lens of errors. We classify and compare the errors made by two groups of students: one at the beginning of their studies in school algebra and another one completing their studies on algebra in compulsory education. This comparison allows us to detect errors which require specific attention in instruction due to its persistence and to identify errors that disappear as students advance in their study of algebra. The results and conclusions have pedagogic value to inform instruction and also lead to backed conjectures and research questions to push forwards research on student's translation capacity and students' knowledge of algebraic symbolism.

Keywords: algebraic symbolism, algebra learning, errors, translations, verbal representation system

Introduction and Previous Studies

Algebraic symbolism is a component of school algebra that has a large presence in the secondary education curriculum. Emphasis is placed on its utility, together with the verbal[1] representation system, for the communication and representation of algebraic concepts. The use of both representation systems, as part of the mathematical language, should enable students to express mathematical ideas precisely, communicate their mathematical thinking, solve problems, and model and interpret phenomena from mathematics and other sciences. All these are components of the mathematical competence expected to be developed by students in secondary school education (Ministerio de Educación y Ciencia, 2006; National Council of Teachers of Mathematics, 2000). Students are also expected to be able to move between different representation systems. This capacity is linked to a good understanding of the represented concepts (Gómez, 2007; Janvier, 1987), being better problem solver and having access to a wider set of strategies (Cañadas, Castro & Castro, 2008; Friedlander & Tabach, 2001).

In spite of the strong presence of algebraic symbolism in the secondary education curriculum, which usually prioritizes it over other representations (Bossé, Adu-Gyamfi & Cheetham, 2011a, 2011b), educators and researchers stress the limited mastery students show of this representation system and question the comprehension of algebraic symbolism that students develop (Kieran, 2007; Vega-Castro, Molina & Castro, 2012). Translations between algebraic symbolism and the verbal representation system also present numerous difficulties for secondary students (Cerdán, 2010; MacGregor & Stacey, 1993; Wagner & Parker, 1993; Weinberg, 2007).

[1] We use the term verbal to mean “expressed with words”.

Most studies that attend to the processes of translation in the area of school algebra focus on the tabular, graphical, and symbolic representation systems (Kieran, 2007). These studies show students' difficulties in maintaining the semantic congruence that characterizes these processes, even when students display understanding of the initial and final representations. Several studies focus attention on translation from the verbal representation system to algebraic symbolism, fewer on translation from the latter to the former. Authors like Kaput, Sims-Knight and Clement (1985) and Kaput (1989) have stressed that, in order to perform these translations successfully, students must understand the variables and relationships of mutual dependence between them described in the verbal statement as well as the syntactical characteristics of the symbolic representation system. They must thus alternate syntactical and semantic ways of analyzing both representations during the translation process. Nevertheless, even expert don't use conceptual approaches in some translations from the verbal to the symbolic system according to Kirshner and MacDonald (1992); in some type of sentences just syntactic approaches are sufficient to be successful either directly or after having modified the sentence, without accessing the underlying conceptual structure.

In the context of problem solving, where most of the studies about translations from the verbal representation system to algebraic symbolism have been developed (Cerdán, 2010; González-Calero, Arnau, & Puig, 2014; MacGregor & Stacey, 1993; Wagner & Parker, 1993; Weinberg, 2007), students resist using algebraic symbolism and prefer to use arithmetic strategies and representations (Kieran, 2007). In these cases, the problem is presented through a verbal statement describing a context and some mathematical relationships which must be translated to algebraic symbolism in order to solve the problem. These research studies on secondary and college-preparatory education students, report incorrect translations amounting up to 30-60% of the total number of translations made by the students (the percentage varies depending on the study). One of the most common errors is the *inversion error*. This error consists of representing the opposite relation to the one indicated.

According to Cerdán (2008a, 2008b, 2010), when translating from verbal statements to algebraic symbolism students in college-preparatory education (16-18 years of age): (a) propose various translations, (b) tend to use more letters than the minimum needed, and (d) show common preferences in choosing the quantities to be represented by a letter. This author also detects a polysemic used of letters when the same word is used in the text of the problem to refer to different quantities (e.g., number, age).

Translation from the symbolic to the verbal representation system is a process that has received less attention in research. Posing problems that can be solved through a given equation or system of equations is the methodology used in various studies whose focus is this type of translation (Fernández-Millán & Molina, 2016; Isik & Kar, 2012; Resnick, Cauxinille-Marmeche & Mathieu, 1987). According to Fernández-Millán & Molina (2016), students encounter more difficulties in posing a problem when the equation given includes multiplication of unknowns or coefficients other than one or two. Further, students tend to assign different values to the same unknown when it appears in different members of the equation. As to the invention of problems to be solved using a given symbolic expression, Isik and Kar (2012) identify errors such as assigning unrealistic values to the unknowns in the invented problems, using algebraic symbolism in the statement of the problem, failing to establish a part-whole relationship and lack of a relationship between the equations in a system.

These previous studies identify some of the most frequent errors and difficulties (mostly in a problem solving context), elements of algebraic expressions that seem to increase the difficulty of translations and general skills required for successfully making translations between the symbolic and the verbal representation systems. Moved by these research evidences, we designed the study here reported to advance towards a better understanding of the development of secondary students' capacity to make translations between the verbal representation system and algebraic symbolism and, by exploring and describing this capacity, obtain information about students' development of knowledge of algebraic symbolism. Translations are useful to identify students' learning difficulties and opportunities (Lesh, Post & Behr, 1987) and to measure conceptual knowledge in an implicit way (Rittle-Johnson & Schneider, 2015).

We choose to explore these translations in a non-problem solving context to direct students' attention away from finding an answer and towards the translation process. It also allows to reduce the ambiguity of the verbal representations involved and the complexity of the context. In general, when students translate from the verbal to the symbolic representation system, the presence of unstated and/or irrelevant or confusing information in the statements is a conditioning factor in the difficulty of the translations (Bossé et al, 2011a, 2011b). The difficulty of this kind of translation may also be influenced by the presence and the kind of context implied in the verbal representation given. To date, there is no clear evidence on the nature of this influence: some papers dismiss it (Wollman, 1983) while familiarity of context is a factor recognized in problem-solving processes (Ambrose & Molina, 2014) and even recommended to give concrete significance to the mathematical language (Gómez-Granell, 1989).

We approach our study of students' capacity to translate algebraic statements between the algebraic and verbal representation system, through the lens of errors. We consider errors as inadequate cognitive schemes and not only as result of lack of knowledge or a slip (Socas, 1997). Previous studies have proven that the study of errors in the process of teaching and learning mathematics permits us to understand the nature of fundamental mathematics notions and the processes for constructing mathematical knowledge (Rico, 1995). Students' errors give information about the difficulties that specific mathematics contents present and suggest pedagogical recommendations that start from the error and move toward the construction of mathematical knowledge (Rach, Ufer, & Heinze, 2013; Rico, 1995; Socas, 1997).

We work with two groups of students: one at the beginning of their studies in algebra (13-14 years old, year 2 of secondary education) and another one completing their studies about algebra in compulsory education (15-16 years old, year 4 of secondary education). We know that experience naturally contributes to increasing students' capability of using symbols with understanding (Pope & Sharma, 2001). Therefore, as students advance in their study of mathematics they will no longer incur in some previous errors but new errors might emerge, both facts are result of reorganizing and developing their knowledge schemes and/or changes in the students' attitudes. Comparing the errors incurred by both groups of students allow us to detect errors which require specific attention in instruction due to its persistence and to identify errors that disappear as students advance in their study of algebra.

Before describing the empirical study developed, we precise some theoretical terms related to the aim of this study.

Representations Systems and Translations

Knowledge in general, and mathematics in particular, requires representations. To think about mathematical ideas, reason about them, and organize the knowledge they provide, it is necessary to have an internal representation of these ideas (Goldin, 2002). External representations are also necessary to express and communicate mathematical ideas, as mathematical concepts take concrete form through these representations (Hiebert & Carpenter, 1992). Research argues a close connection between external and internal representations, and internal representations may be an assimilation of external ones (Castro & Castro, 1997). This paper focuses on external representations (referred just as representations). There are diverse forms of representation for the same concept, and students' mastery of these modes permits greater comprehension of the concept (Goldin, 1998; Kaput, 1992). Representation systems are considered to be a structured set of notations, symbols, and graphs, with rules and conventions, that enable the expression of concepts, properties of the concept, and connections with other concepts (Rico, 2009). The different external representation systems valid for a concept have their own idiosyncrasies; they simultaneously highlight and obscure different properties of the concept (Gómez, 2007; Janvier, 1987).

We attend here to the verbal and symbolic representation systems in the context of school algebra. The verbal representation system is determined by the use of everyday language, sometimes including specific terminology from academic mathematical language. The symbolic representation system used in algebra, also known as algebraic symbolism, is characterized by the written expression of numerals, letters, and signs characteristic of arithmetic and algebra. We use the term algebraic statement to call propositions that can be expressed using algebraic symbolism. An example of an algebraic statement represented verbally is "a number plus its consecutive number is equal to another number minus two", where $x+(x+1)=y-2$ is an algebraic symbolic representation of this statement. The two expressions (verbal and symbolic) are equivalent in meaning.

In this framework, the procedure through which a mathematical object represented by one system comes to be represented in another system is known as a translation between two representation systems (Gómez, 2007). Translation between representation systems consists of transforming the concepts and attributes represented in one system into the corresponding concepts and attributes in another system, to obtain a representation different than the initial one but congruent in meaning. This is a complex process from a cognitive view point. In addition to understanding the representation systems involved, it requires distinguishing the essential information that defines the represented concept to translate it to another representation system and to ignore unnecessary aspects imposed by the system in which the concept is represented (Molina, 2014). A possible referent needs to be identified in the given representation, going beyond the representational mode, and be represented in a different representation system.

Research Objectives and Method

As explained in the introduction of this paper we wonder about which errors students incur when doing translations of algebraic statements between the symbolic and the verbal representation systems in a non-problem solving context. In addition we want to explore which types of errors disappear as students advance in the study of compulsory school algebra and which don´t, as well as if new errors emerge.

These research questions lead to the design of the study here reported and the selection of the participating students. We worked with two groups of secondary students from a Spanish public school: one at the beginning of their studies in algebra (16 students from 13 to 14 years old) and another one finishing their compulsory studies on algebra (26 students from 15 to 16 years old); that is year 2 and 4 of compulsory high school. The school serves a low socio-cultural and economic urban region in Spain and both groups presented a low performance level in mathematics and little motivation and interest in learning and studying mathematics. Both groups can be considered representative in this type of regions.

The specific objectives set to guide the research study are the following ones.

- To classify and describe the errors that both group of secondary students incur when translating algebraic statements from the verbal to symbolic representation systems and vice versa, out of a problem solving context.
- To identify errors that persist and errors that disappear or appear as students complete their compulsory education in algebra.

This is an exploratory and descriptive study (Hernández, Fernández & Baptista, 1991). It is considered exploratory because of the scarcity of studies that explore the translation of algebraic statements from algebraic symbolism to verbal representations, as well as the translation from verbal representations to algebraic symbolism out of a problem solving context. As we have previously explained in a non-problem solving context students' attention is not focused on finding an answer, the ambiguity of the verbal representations is reduced and the possible influence of the familiarity of context is avoided. This justifies the different nature of the translation process considered in our study in comparison to most previous studies and gives this study its exploratory character. It is a descriptive study because it describes the students' capacity to do translations by means of the errors that they incur as well as the differences in the occurrence of these errors in both groups of students. Due to the way we designed our data collection, the results presented are based on simple statistics related to a classification of errors obtained through an inductive process following the grounded theory approach (Corbin & Strauss, 1990). We use the analysis of this data collection, together with claims from previous studies, to make descriptive conjectures about cognitive aspects of secondary students' translation skills which can help to expand our understanding of the students' capacity to address the considered translation processes and which will be of use to inform the design of later studies that test this conjectures.

Data Collection

We planned a data collection process in which the translations of the algebraic statements were presented in a motivating task. We designed algebraic dominoes that enabled us to obtain the data in a game context. Unlike traditional dominoes, ours had algebraic statements expressed in verbal or symbolic form and they did not include double pieces.

The tasks posed to the students simulated a game board with a finished game of dominoes on it (see Figure 1[i]). Parts of some dominoes were blank. Each student was given a copy of figure 1 on an A3-sized sheet of paper. They were asked to fill in the blank parts so that the dominoes would be correctly paired by equivalent algebraic statements expressed in different representation systems. For example, if a domino has the expression "$x+2$" at one end, the end of the domino linked to it should have the expression "a number plus two" or another equivalent to it.

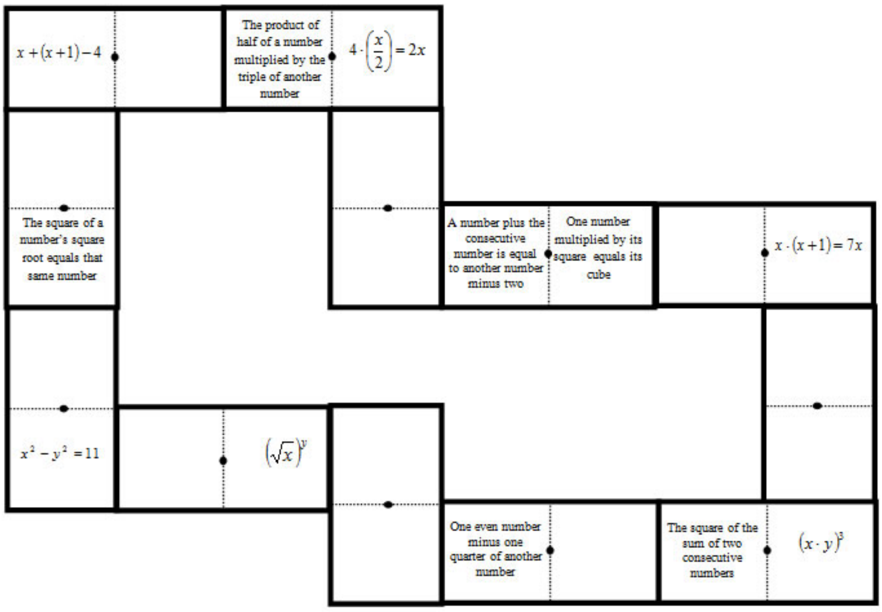

Figure 1. Instrument for the data collection process

The students performed the work as an individual activity in their usual classroom for 55 minutes. The mathematics teacher of the students, member of the research team, gathered the data, that is, the students' written productions in the A3-sized papers.

Design of the Instrument

In choosing which statements to include on the dominoes, we considered algebraic expressions that the students had worked on previously; many of which came from the textbooks used regularly in the classroom. We set different task variables to help us to include statements with diversity of characteristics (see Table 1). We considered the following operations and numerical relationships: sum, difference, product, division, power, square root, and consecutive or even and odd numbers. We proposed twelve statements: six represented in verbal form and six in symbolic form. In each case there was one additive statement, one multiplicative, one involving powers, one additive and multiplicative, one additive and involving a power, and another multiplicative and involving a power. Half of the statements were equations and other were not; half had only one letter and the other half two letters. Similarly, half of the verbal statements were sequential and the other half non-sequential[ii]. All the statements were presented in Spanish to the students (the official language for mathematics instruction at their school). Table 1 presents the 12 statements proposed, as well as their characteristics.

Table 1

Statements and their Characteristics

Representation of algebraic statement	Relations involved	Other features of Statements	Code
Statements in verbal representation			
A number plus the consecutive number is equal to another number minus two	Additive	Sequential, Equation, 2 letters	E3
The product of half of a number multiplied by the triple of another number	Multiplicative	Non-sequential Non-equation, 2 letters	E1
The square of a number's square root equals that same number	Power	Non-sequential Equation, 1 letter	E11
One even number minus one quarter of another number	Additive & multiplicative	Sequential Non-equation, 2 letters	E8
The square of the sum of two consecutive numbers	Additive & power	Non-sequential Non-equation, 1 letter	E7
One number multiplied by its square equals its cube	Multiplicative & power	Sequential Equation, 1 letter	E4
Statements in symbolic representation			
$x+(x+1)-4$	Additive	Non-equation, 1 letter	E12

$4 \cdot \left(\frac{x}{2}\right) = 2x$	Multiplicative	Equation, 1 letter	E2
$\left(\sqrt{x}\right)^y$	Power	Non-equation, 2 letters	E9
$x \cdot (x+1) = 7x$	Additive & multiplicative	Equation, 1 letter	E5
$x^2 - y^2 = 11$	Additive & power	Equation, 2 letters	E10
$(x \cdot y)^3$	Multiplicative & power	Non-equation, 2 letters	E6

Students' Previous Knowledge

Students in year 2 have been introduced to algebra as the part of mathematics which uses letters to express unknown numbers or indefinite values. They have encountered algebraic expressions where algebraic symbolism was used to express: (a) algebraic identities, (b) useful relations to solve problems (equations), (c) general terms of numeric sequences, (d) relations between variables related to different magnitudes (formulas) and (e) general statements about quantities. The different parts of a polynomial algebraic expression (e.g., coefficient, literal part, monomials) had been studied but they did not have experience operating algebraic expressions.

Students in year 4 have studied all algebra included in the Spanish curriculum of compulsory secondary education (Ministerio de Educación y Ciencia, 2006). This comprises doing operations with polynomial expressions, including polynomial fractions, as well as factoring and simplifying them; solving linear and second order equations and inequalities; and doing translations between the verbal and the symbolic representation systems mostly from the verbal to the symbolic and in the context of problem solving.

Both groups of students were expected to master the translation of statements as those included in the algebraic domino, even though at school more attention was given to translations from the verbal to the symbolic system.

Data Analysis and Results

After an inductive process following the grounded theory approach (Corbin & Strauss, 1990), we obtained the categorization presented in Table 2. It allows us classifying the errors identified in the students' productions. In this process, the four members of the research team separately coded the students' productions in order to agree on a common definition of the categories and to increase the reliability of the results. The interpretation of the results was achieved through a joined critical process of analysis backed on the researchers' knowledge about previous research on algebra learning.

The particular names for the categories where inspired by Socas (1997)'s classification of sources for errors, that is: (a) an obstacle (in the sense of Bachelard, 1938, or Brousseau, 1983); (b) absence of meaning: errors with origin in arithmetic (which could be addressed before the study of algebra), errors in applying procedures, and errors due to the particular characteristics of algebraic symbolism; and (c) affective or emotional attitudes towards mathematics (including slips).

Classification of Errors

We distinguish three kinds of errors: (a) relative to the completeness of the statement, (b) derived from arithmetic, and (c) derived from the characteristics of algebraic symbolism (see Table 2).

Table 2

Errors Classification

Category	Subcategory	Code
I. Completeness of statement	Incomplete	I.1
	Extra information	I.2
II. Derived from arithmetic	Division – Product	II.1
	Power – Product	II.2
	Addition – Product	II.3

	Division – Power	II.4
III. Derived from characteristics of algebraic symbolism	Generalization	III.1
	Particularization	III.2
	Letters	III.3
	Structural complication	III.4

The errors according to the *completeness of the statement* refer to whether any symbol or word is lacking or extra in the expression to be correct. In the first case we name the error as "incomplete" (I.1), otherwise the error is named as "extra information" (I.2). For example, to translate the statement $x \cdot (x+1) = 7x$ as "a number times the consecutive number equals seven" is an error of the incomplete type, and to express the verbal statement "the product of half of one number multiplied by the triple of another" as $x \cdot \left(\frac{x}{2}\right) \cdot 3y$ in algebraic symbolism is an error of extra information.

Errors *derived from arithmetic* come from incorrect interpretation of signs or operations. We distinguish four subcategories: division–product (II.1), power–product (II.2), sum–product (II.3), and division–power (II.4). The first operation in the name of the subcategory is interpreted as the second operation mentioned. For example, if the statement proposed requires representing $\left(\sqrt{x}\right)^y$ verbally and a student states it as "the square root of one number times another different number", we understand that the student has incurred an error in interpreting the power, since he or she has expressed a product instead (power-product error).

Errors *derived from the characteristics of algebraic symbolism* are specific to the use of the symbolic representation system. In this category, we distinguish four kinds of errors:

- *Generalization errors* (III.1): consisting on generalizing an element or part of the statement. For example, representing -4 as "we subtract an even number".
- *Particularization errors* (III.2): due to the particularization of numbers or specific relationships. For example, translating symbolically "an even number" as 2.

- *Letter errors* (III.3): not distinguishing correctly the use of different letters in a statement. In this case we detect two possibilities: one letter is used to represent different quantities or several letters are used to represent the same quantity.
- *Structural complication errors* (III.4): not interpreting correctly the structure or part of the structure of the algebraic statement. For example, a student expresses symbolically "an even number minus one quarter of another number" as $2x-\frac{4}{y}$.

Analysis and Comparison of the Errors Detected

All but one of the statements that the students produced had errors. The data analysis performed using the classification presented above shows, as expected, a greater number of errors in the year 2 group (134 errors among 16 students) than in year 4 (69 errors among 26 students). The greatest number of errors took placed in translations from the verbal to the symbolic representation system: 52 errors as opposed to 17 in the group from year 4, and 86 errors versus 48 in the group from year 2. Table 3 shows the frequencies for each type of error for each group of students and each direction of translation. New errors did not emerge in year 4 translations in comparison to year 2.

Table 3

Comparison of Errors in each Group and each Direction of Translation

	Symbolic → Verbal		Verbal → Symbolic	
Type of error	Year 4	Year 2	Year 4	Year 2
I.1	3 (18%)	9 (19%)	5 (10%)	23 (27%)
I.2	1 (6%)	7 (15%)	4 (8%)	8 (9%)
II.1	0 (0%)	0 (0%)	2 (4%)	1 (1%)
II.2	7 (41%)	3 (6%)	4 (8%)	4 (5%)
II.3	0 (0%)	1 (2%)	1 (2%)	3 (3%)
II.4	0 (0%)	0 (0%)	1 (2%)	1 (1%)
III.1	4 (24%)	9 (19%)	0 (0%)	9 (10%)
III.2	0 (0%)	2 (4%)	7 (13%)	5 (6%)
III.3	0 (0%)	6 (13%)	13 (25%)	16 (19%)

III.4	2 (12%)	11 (23%)	15 (29%)	16 (19%)
Total	17 (100%)	48 (100%)	52 (100%)	86 (100%)

In the case of the year 2 group, in translating from the verbal representation system to the symbolic, and vice versa, over half of the errors correspond to those classified as derived from the characteristics of the algebraic symbolism, and a third of the errors are relative to the completeness of the statements (see Figures 2 and 3). In the case of the year 4 group, the tendency in the type of error presented does not coincide in the two kinds of translations. In translating verbal statements into their symbolic representations, students' most common errors are those derived from the characteristics of the algebraic symbolism, which constitute two thirds of the errors incurred. The two remaining kinds of errors show similar frequencies. In translating from the symbolic to the verbal representation system, however, the few errors detected incur are distributed almost equally among the three types of errors (see Figure 3).

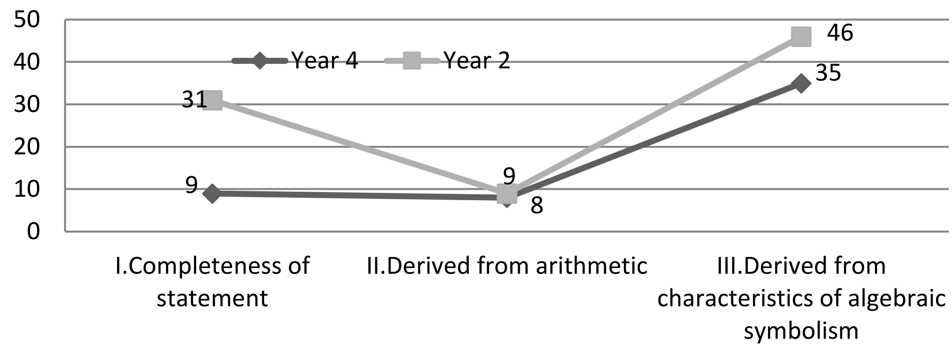

Figure 2. Frequency of error type in translations from verbal to symbolic

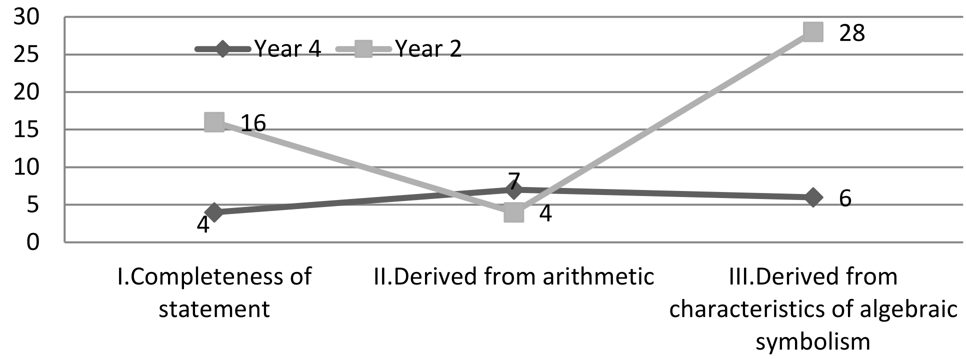

Figure 3. Frequency of error type in translations from symbolic to verbal

To develop this analysis of the errors in greater depth, we now focus separately on both directions of translations.

Errors in translations from the verbal representation system to algebraic symbolism. In considering the different subcategories of errors described above (see Table 3 and Figure 4), we see that the most frequent errors in the year 4 students are types III.4 (structural complication) and III.3 (letters), both of which derive from the characteristics of the algebraic symbolism. If we take these subcategories together, this kind of error accounts for half of the errors in this group of students. We have an example of this kind of error in the case of a student from the year 4 group who translates statement E3 ("a certain number plus the consecutive number, equals another number minus two") as "$x+(x+1)=x-2$". The student uses the same letter to represent different numbers; therefore he incurs a type III.3 error. Another year 4 student makes a type III.4 (structural complication) error when translating statement E7 ("the square of the sum of two consecutive numbers") as "$x+(x+1)=x^2$". The letter errors in this case are all due to a polysemic use of letters.

Although errors of type III.3 (letter) and III.4 (structural complication) also present a high frequently in year 2 (16/86 in both cases), the type of error that occurs most often is I.1 (incomplete), which accounts for approximately one of every four errors in this group of students. For example, a year 2 student expresses statement E7 symbolically ("The square of the sum of two consecutive numbers") as "$(x+x+1)$", omitting the power. The letters errors in this group are mostly due to a polysemic use of letters too.

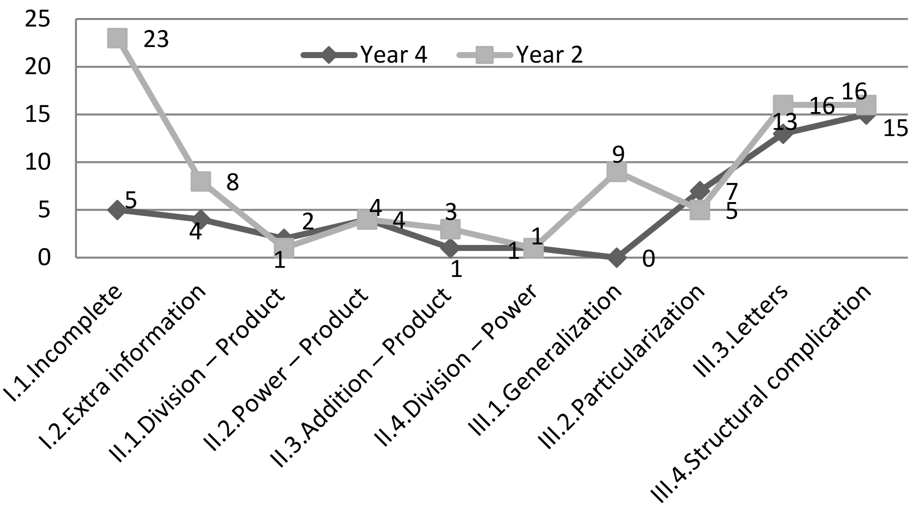

Figure 4. Frequency of error type in translations from verbal to symbolic

In both groups, the errors derived from arithmetic have low frequencies, varying from 1 to 4. The most frequent error is the same one in both cases: power-product. We point to the case of the error III.1 (generalization) which only occurred in the year 2 group, with high frequency in relation to the other errors (see Figure 4).

Even though the design of the data collection does not allow to rigorously identifying individual influences of each tasks variable, we attend to the characteristics of the statements when analyzing the results to suggest possible influences which can be tested in other studies (see Table 4). Even though the influence of various tasks variables might be related, we comment below on separated influences which might be taken place.

Table 4

Errors and Statements' Characteristics in Translations from the Verbal Representation System to Algebraic Symbolism

					Number of errors	
Statements		Statement's characteristics			Year 4	Year 2
E3	Ad	Equation	2 letters	Sequential	7 (13%)	20 (23%)
E1	Mu	Non-equation	2 letters	Non-sequential	15 (29%)	11 (13%)
E11	Po	Equation	1 letter	Non-sequential	0 (0%)	11 (13%)
E8	AdMu	Non-equation	1 letter	Sequential	17 (33%)	23 (27%)
E7	AdPo	Non-equation	2 letters	Non-sequential	12 (23%)	14 (16%)
E4	MuPo	Equation	1 letter	Sequential	1 (2%)	7 (8%)
					52 (100%)	86 (100%)

Note. Ad=additive, Mu=multiplicative, Po=power, AdMu= additive & multiplicative, AdPo= additive & power, MuPo= Multiplicative & power

As can be seen in Table 4, both groups of students incur more errors in the statement E8. In the year 4 group, other statements with high frequency of errors are E1 and E7 (representing 15/52 and 14/52, respectively). In the year 2 group, E3 and E7 are the second ones with higher frequencies (20/86 and 16/86 respectively). Interestingly the statements which are identities (E11 and E4) are the ones with lower frequencies of errors. If we analyze the results according to whether the verbal statements are equations or not, we find a lower presence of errors in the equations, a tendency that is especially marked in the year 4 group (see Table 4). We find an influence of the number of letters that the statement includes only in the year 4 group, where errors are more frequent in the statements that have two letters.

The distinction is hardly noticeable in the year 2 group. The influence of the sequentiality variable is not noticeable in the results.

Errors in translations from algebraic symbolism to the verbal representation system. In translations from algebraic symbolism to their verbal representation, most type II errors have none or a very low frequency (see Figure 5 and Table 3). Error III.2 (particularization) is also very scarce. In addition, in the year 4 group categories I.2 (extra information) and III.3 (letters) have only one error or none at all. In this group, type II.2 error (power-product) is the most frequent (7/17). In the year 2 group, in contrast, the most frequent errors are those of categories III.4 (structural complexity), III.1 (generalization), and I.1 (incomplete) with proportions of 11/48, 9/48, and 9/48, respectively. The letters errors in year 2 are mostly due to assigning the same meaning to different letters.

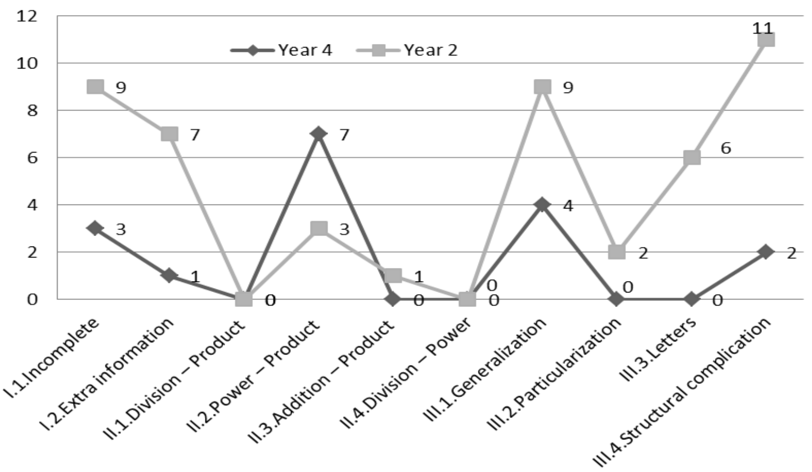

Figure 5. Frequency of error type in translations from symbolic to verbal

Table 5 shows the number of errors in the statements and the characteristics of each statement. The low number of errors that the year 4 students incur in this kind of translation is distributed across all of the statements. The statements E5 and E9 are slightly higher in frequency (5/17 and 4/17, respectively). In the year 2 group, the highest presence of errors occurs in the statement E2 (15/48), followed by E5 (9/48).

Table 5

Errors according to the Statement's Characteristics in Translations from Algebraic Symbolism to the Verbal Representation System

Statements	Statement's characteristics			Number of errors	
				Year 4	Year 2
E12	Ad	Non-equation	1 letter	2 (12%)	6 (13%)
E2	Mu	Equation	1 letter	1 (6%)	15 (30%)
E9	Po	Non-equation	2 letters	4 (24%)	6 (13%)
E5	AdMu	Equation	1 letter	5 (29%)	9 (18%)
E10	AdPo	Equation	2 letters	3 (18%)	6 (13%)
E6	MuPo	Non-equation	2 letters	2 (11%)	6 (13%)
				17 (100%)	48 (100%)

Note. Ad=additive, Mu=multiplicative, Po=power, AdMu= additive & multiplicative, AdPo= additive & power, MuPo= Multiplicative & power

If we analyze the errors by distinguishing whether the verbal statements are equations or not, only the year 2 group shows influence of this task variable, with higher presence of errors in equations. In this group, statements with one letter have double frequency of errors than those with two letters. The only identity included (E2) presents a high number of errors in year 2 group but only one in year 4.

Discussion of Results and Conclusions

This article presents a classification of the errors that students in two groups at different levels of secondary education incur when translating algebraic statements between the verbal and the symbolic representation systems out of a problem solving context. The kinds of errors that make up this classification and their breakdown into subtypes are a contribution to existing research on translations. The diversity of errors detected suggest that different causes are at the heart of each error. Here we discuss plausible causes for these errors; they are based on our knowledge of previous studies and conditioned by the classification of errors elaborated. They need to be considered as conjectures to be explored in future studies where theoretical frameworks and/or more in-depth analysis can be developed. In particular, the consideration of semiotic approaches to explore the errors identified and the conjectures stated can help in the effort of capturing and explaining the cognitive complexity of doing translations (Hoffman, 2006). “A semiotic perspective of mathematical activity provides a way of conceptualising the teaching and learning of mathematics driven by a primary focus on signs and sign use” leading to an alternative viewpoint (Ernest, 2006, p. 68). The multiplicity of semiotic frameworks currently in use in mathematics education can provide diverse interpretations to the errors we detected. Our classification of errors attend to the mathematics content and distinguish if errors might be addressed before the study of algebra or they are linked to algebraic contents or symbols.

Plausible Causes of the Detected Errors and their perseverance

Several of the errors related to the characteristics of algebraic symbolism can be interpreted as a consequence of the precision that characterizes algebraic symbolism and mathematics language in general. In translations from the verbal to the symbolic representation system, previous studies have described various phenomena which evidence lack of precision in students’ use of the symbolic and verbal representation systems. In the initial steps of problem solving, Mitchell (2001) have observed that students change the words in the text of the problem in a way that affects its meaning. This behavior, named “wordwalking, leads to interpreting the relations described in a verbal statement differently. González-Calero et al. (2013) claim that students are not precise when they specify what each letter of a symbolic expression represents in problem solving, their definition of variables tend to be fairly ambiguous (e.g., *x*=cars). Cerdan (2010) also detect a lack of precision in students´ analysis of verbal statements as they only attended to some words in the texts when referring to the quantities, considering equal those quantities whose descriptions share a word.

Even in our study where ambiguity was reduced by considering a non-problem solving, students struggle with the precision of algebraic symbolism. The use of the same literal sign to represent different unknown quantities, a persistent error in translations from the verbal to the symbolic representation system, assigning different meanings to a letter when translation to the verbal representation system, and the particularization and generalization errors can be interpreted as lack of precision in students use of algebraic symbolism. In a later study, Rodríguez-Domingo (2015) have observed that year 2 students do not consider wrong to express part of algebraic statement more generally although they acknowledge that other statements may be considered as “better” translations. Starting from this assumption, assigning a value to an unknown quantity (particularization errors) may be for students an accepted change which help them to handle or avoid the uncertainty expressed in the statement.

In relation to these errors, the comparison of year 2 and 4 students’ errors suggest some progression when going from the symbolic to the verbal representation system but advancing in the study of algebra does not seem to help significantly to acquire a precise use of algebraic symbolism. Further studies focus on this characteristic of algebraic symbolism are need to inquire about its acquisition by students.

Structural complication and arithmetic errors suggest a lack of understanding of the quantitative relations represented in the statements (Kaput, 1989; Kaput et al., 1985). In translations from the symbolic to the verbal representation systems they also evidence difficulties in recognizing the structure of an algebraic expression. When parsing an algebraic expression (Kirshner, 1989), students need to combine various skills such as considering part of the expression as a whole, identifying relations between different parts of the expression and recognizing familiar structures. According to studies which focus on these skills under the denomination of structure sense (Vega-Castro, Molina & Castro, 2012; Hoch & Dreyfus, 2005), their successful development requires extended experience and intense attention in instruction. In both direction of translations, deficiencies in students' structure sense together with the absence of alteration of syntactical and semantic ways of analyzing both representations during the translation process (Kaput, 1989; Kaput et al., 1985), limit students skills to detect and correct not only structural complication and arithmetic errors but also errors related to the completeness of the statements.

It is interesting to notice that in translations from the verbal to the symbolic representation system arithmetic errors were related to all the operations considered, but were mostly related to the product and power operations when translating in the other direction. Unfortunately we do not have an explanation for this difference.

The comparison of errors detected in each group show greater competence among students in year 4 in recognizing the structure of symbolic expressions, which makes it easier for them not to incur errors of omission or to include extra information and gives them better capability to translate from the symbolic to the verbal representation system. Further exposure to algebraic statements had a positive influence in students' structure sense, however, the persistence of errors derived from arithmetic suggest a lack of progression in students understanding of some quantitative relations.

We see that the translation processes that present the greatest frequency of errors are those that students work with most in school practice: translation from the verbal to the symbolic representation system. Future studies are needed in order to provide an explanation for this result. The influence of factors such as the greater precision of algebraic symbolism as compared to verbal language (Socas, 1997), the "wordwalking" phenomenon (Mitchell, 2001), and the possible need to reorganize the information before they can be translated into algebraic symbolism (MacGregor & Stacey, 1993, Kirshner & MacDonald, 1982) need to be tested.

Plausible Causes of Influences of Tasks Variables

Considering the task variables involved in the design of the instrument, we find that in translations from the symbolic to the verbal representation system, the equations and those statements with one letter seem to present the greatest translation difficulties to year 2 students. These influence may be due to the less extensive experience that these students have with equations (they have not studied strategies for solving equations yet), the fact that equations represent more complex quantitative relations as result of stablishing an equality and students' tendency to incorporate more letters than needed.

In translations from the verbal to the symbolic representation system, however, the students from both groups appear to incur more errors when the statements were not equations. Year 2 students had similar experience with equation and non-equations however, this was highly unbalanced in year 4. Students' necessity of closure already reported in other studies may be a reason underneath this result (Drijvers y Hendrikus, 2003; Kieran, 1981).

Fewer errors were detected in statements expressing an identity when being translated from the verbal to the symbolic representation system. It would be interesting to explore in future studies if identities are expressions easier for secondary students to understand than other algebraic expressions. This might be the case as they represent relations that are (or can be) known to be true by the students.

Implications for teaching

The specific results of this study and their discussion inform about the students' ability to make translations and their development as they advance in their learning of algebra. The precision of algebraic symbolism, the students' structure sense and the understanding of the quantitative relations represented in the statements are mentioned as key factors influencing the studied translations which can help to explain the detected errors and, therefore, deserve careful attention in teaching.

Although arithmetic errors were not very frequent, they are an obstacle to students' progress in understanding. Therefore, we point the need to tackle the confusion that occurs in interpreting the operations of powers, multiplication, addition, product, and division to help the students correct these errors. This can be addressed not only in algebraic but also in arithmetic contexts.

Instruction can use the greater facility that students show in translating algebraic statements from algebraic symbolism to a verbal representation as a means of support for developing students' understanding of algebraic symbolism and improving the processes of the inverse translation. From the perspective of posing problems, we can take advantage of this greater facility by asking the students to pose problems from algebraic statements expressed symbolically and, then, to tackle translation from the verbal to the symbolic representation system in order to solve the problem. Integrated study of posing and solving of problems can potentially help students to become aware of the greater precision and synthetic capability of algebraic symbolism in comparison to verbal language.

The identification of the influence of the task variables, although it should be confirmed in future studies, is also useful for professors and textbook editors so that attention can be directed to characteristics that increase the difficulty of translation processes and more practice can be provided.

The students easily understood the data collection instrument because all of them have experience playing domino. We acknowledge that the design of this instrument changed the students' attitude and, therefore, it might have positive impact on the results. Nevertheless, the process of translating between the verbal and the symbolic representations proved difficult for both groups. Clinical interviews are needed to explore the thinking of individuals beyond the particular students participating in our study, before sound argument can be made about the conceptual basis underlying their errors and about cognitive process development.

Even though the data come just from two specific groups of students, the results are considered relevant to inform further studies on secondary students' capacity to translate algebraic statements thanks to its descriptive character and the scarcity of previous studies on this type of translations in a non-problem solving context. This study provide rich information to inform the design of further studies as well some interesting conjectures and questions that will help to push forward research on translations between representation systems.

References

Ambrose, R., & Molina, M. (2014). Spanish/English bilingual students' comprehension of story problem texts. *International Journal of Science and Mathematics Education. 12*(6), 1469-1496. DOI 10.1007/s10763-013-9472-2.

Bachelard, G. (1938). *La formation de l'esprit Scientifique [The formation of the scientific spirit]*. Paris: Vrin.

Bossé, M. J., Adu-Gyamfi, K., & Cheetham, M. R. (2011a). Assessing the difficulty of mathematical translations: synthesizing the literature and novel findings. *International Electronic Journal of Mathematics Education, 6*(3), 113-133.

Bossé, M. J., Adu-Gyamfi, K., & Cheetham, M. R. (2011b). Translations among mathematical representations: Teacher beliefs and practices. *International Journal of Mathematics Teaching and Learning, June*.

Brousseau, G. (1983). Les obstacles épistémologiques et les problèmes en mathématiques [The epistemologic obstacles in mathematics problems]. *Reserches en Didactique des Mathematiques, 4*(2), 165-198.

Cañadas, M. C., Castro, E., & Castro, E. (2008). Patrones, generalización y estrategias inductivas de estudiantes de 3º y 4º de Educación Secundaria Obligatoria en el problema de las baldosas. *PNA, 2*(3), 137-151.

Castro, E., & Castro, E. (1997). Representaciones y modelización. In L. Rico (Ed.), *La Educación Matemática en la enseñanza secundaria* (pp. 95-124). Barcelona, Spain: Horsori.

Cerdán, F. (2008a). *Estudio sobre la familia de problemas aritméticos-algebraicos [Study about the family of arithmetic-algebraic problems]*. Valencia, Spain: Server de Publicacions de la Universitat de València.

Cerdán, F. (2008b). Las igualdades producidas en el proceso de traducción algebraico: estudio de las igualdades correctas [The equalities produced in the algebraic translation process: study of correct equalities]. In R. Luengo, B. Gómez, M. Camacho, & L. Blanco (Eds.), *Investigación en educación matemática XII* (pp. 257-272). Badajoz, Spain: Sociedad Española de Investigación en Educación Matemática.

Cerdán, F. (2010). Las igualdades incorrectas producidas en el proceso de traducción algebraico: un catálogo de errores [Incorrect equalities produced in the process of algebraic translation: a catalogue of errors]. *PNA*, *4*(3), 99-110.

Corbin, J., & Strauss, A. (1990). *Basics of qualitative research: Grounded theory procedures and techniques.* Londres: Reino Unido: Sage.

Drijvers, P. y Hendrikus, M. (2003). *Learning algebra in a computer algebra environment: design research on the understanding of the concept of parameter.* Tesis doctoral no publicada, Universidad de Utrecht, Utrecht.

Ernest, P. (2006). A semiotic perspective of mathematical activity: The case of number. *Educational Studies in Mathematics, 61*(1), 67–101

Fernández-Millán E. & Molina, M. (2016). Indagación en el conocimiento conceptual del simbolismo algebraico de estudiantes de secundaria mediante la invención de problemas [Inquiry into secondary students' conceptual knowledge of algebraic symbolism by means of problem posing]. *Enseñanza de las Ciencias, 34*(1), 53–71. doi: 10.5565/rev/ensciencias.1455

Friedlander, A., & Tabach, M. (2001). Promoting multiple representations in algebra. In A. A. Cuoco (Ed.), *Yearbook of the National Council of Teachers of Mathematics: The roles of representation in school mathematics* (pp. 173-185). Reston, Virginia: National council of teachers of mathematics.

Goldin, G. (1998). Representational Systems, Learning, and Problem Solving in Mathematics. *Journal of Mathematical Behavior*, *17*(2), 137-165.

Goldin, G. (2002). Representation in mathematical learning and problem solving. In L. English (Ed.), *Handbook of international research in mathematics education* (pp. 197-218). Mahwah, NJ: Lawrence Erlbaum Associates.

Gómez, P. (2007). *Desarrollo del conocimiento didáctico en un plan de formación inicial de profesores de matemáticas de secundaria [Development of didactical knowledge in a initial training plan for secondary mathematics teachers]*. Doctoral thesis. Granada, Spain: Universidad de Granada.

Gómez-Granell, C. (1989). La adquisición del lenguaje matemático: un difícil equilibrio entre el rigor y el significado [The adquisition of mathematical language: a difficult equilibrium between rigor and meaning]. *CL & E: Comunicación, Lenguaje y Educación, 3-4*, 5-16.

González-Calero, J. A., Arnau, D., & Puig, L. (2013). Dificultades en la construcción de nombres de cantidades durante la resolución algebraica de problemas verbales por estudiantes de primaria [Difficulties in the construction of names of quantities during algebraic solving of verbal problems by elementary students]. In A. Berciano, G. Gutiérrez, A. Estepa, & N. Climent (Eds.), *Investigación en Educación Matemática XVII* (pp. 301-310). Bilbao, Spain: Sociedad Española de Investigación en Educación Matemática.

Hernández, C. R., Fernández, C., & Baptista, P. (1991). *Metodología de la investigación [Research methodology]*. México DC: McGraw Hill.

Hiebert, J., & Carpenter, T. (1992). Learning and teaching with understanding. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning.* (pp. 65-97). Nueva York, NY: MacMillan.

Hoch, M., & Dreyfus, T. (2005). Students' difficulties with applying a familiar formula in an unfamiliar context. In H. L. Chick & J. L. Vincent (Eds.), *Proceedings of the 29th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 145-152). Melbourne, Australia: University of Melbourne.

Hoffman, M. H. G. (2006). What is a “semiotic perspective”, and what could it be? Some comments on the contributions to this special issue. *Educational Studies in Mathematics, 61*(1), 279–291
Isik, C., & Kar, T. (2012). The analysis of the problems the pre-service teachers experience in posing problems about equations. *Australian Journal of Teacher Education, 37*(9), 93-113.
Janvier, C. (1987). Translation processes in mathematics Education. In C. Janvier (Ed.), *Problems of representation in the teaching and learning of mathematics*. Hisllsdale. NJ: Lawrence Erlbaum Associates.
Kaput, J. (1989). Linking representations in the symbolic systems of algebra. In S. Wagner & C. Kieran (Eds.), *Research agenda for mathematics education: Research issues in the learning and teaching of algebra* (pp. 167-194). Reston, VA: National Council of Teachers of Mathematics.
Kaput, J. (1992). Technology and mathematics education. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 515-556). New York, NY: MacMillan.
Kaput, J., Sims-Knight, J., & Clement, J. (1985). Behavioral objections: A response to Wollman. *Journal for Research in Mathematics Education, 16*(1), 56-63.
Kieran, C. (1981). Concepts associated with the equality symbol. *Educational Studies in Mathematics, 12*(3), 317-326.
Kieran, C. (2007). Learning and teaching algebra at the middle school through college levels: Building meaning for symbols and their manipulation. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 707-62). Charlotte, NC: Information Age Publishing.
Kirshner, D. (1989). The visual sintax of algebra. *Journal for Research in Mathematics Education, 20*(3), 274-287
Kirshner, D. & McDonald J. (1992). *Translating English Sentences to Algebraic Notation: methodological and epistemological considerations.* Unpublished document. Lousiana State University.
Lesh, R., Post, T., & Behr, M. (1987). Representations and translations among representations in mathematics learning and problem solving. In C. Janvier (Ed.), *Problems of representation in the teaching and learning of mathematics* (pp. 33-40). Hillsdale, NJ: Lawrence Earlbaum Associates.
MacGregor, M., & Stacey, K. (1993). Cognitive models underlying students' formulation of simple linear equations. *Journal for Research in Mathematics Education, 24*(3), 217-232.
Ministerio de Educación y Ciencia (2006). Real Decreto 1631/2006, de 29 de diciembre, por el que se establecen las enseñanzas mínimas correspondientes a la Educación Secundaria Obligatoria [Royal Decree 1631/2006, of 29 December, by which the corresponding core curriculum for compulsory secondary education is established]. *Boletín Oficial del Estado, 5*, 677-773.
Mitchell, J. M. (2001). Interactions between natural language and mathematical structures: the case of “wordwalking”. *Mathematical Thinking and Learning, 3*(1), 29-52.
Molina, M. (2014). Traducción del simbolismo algebraico al lenguaje verbal: indagando en la comprensión de estudiantes de diferentes niveles educativos. *Gaceta de la Real Sociedad Matemática Española, 17*(3), 559-579.National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics.* Reston, VA: Author.
National Council of Teachers of Mathematics (2000). P*rinciples and standards for school mathematics.* Reston, VA: Author.

Pope S., & Sharma, R. (2001). Symbol sense: Teacher's and student's understanding. *Proceedings of the British Society for Research into Learning Mathematics*, *21*(3).
Rach, S., Ufer, S., & Heinze, A. (2013). Learning from errors: effects of teachers' training on students' attitudes towards and their individual use of errors. *PNA, 8(*1), 21-30.
Resnick, L., Cauzinille-Marmèche, E., & Mathieu, J. (1987). Understanding algebra. In A. Sloboda Thon & D. Rogers (Eds.), *Cognitive processes in mathematics* (pp. 169–225). Oxford, United Kingdom: Clarendon Press.
Rico, L. (1995). Errores y dificultades en el aprendizaje de las matemáticas [Errors and difficulties in mathematics learning]. In J. Kilpatrick, P. Gómez, & L. Rico (Eds.), *Educación Matemática. Errores y dificultades de los estudiantes. Resolución de problemas. Evaluación. Historia* (pp. 69-108). Bogotá, Colombia: una empresa docente.
Rico, L. (2009). Sobre las nociones de representación y comprensión en la investigación en Educación Matemática [About the notions of representation and understanding in mathematics education research]. *PNA, 4*(1), 1-14.
Rittle-Johnson, B. & Schneider, M. (2015). Developing Conceptual and Procedural Knowledge of Mathematics. In R. C. Kadosh y A. Dowker (Eds.), Oxford Handbook of Numerical Cognition (pp. 1118-1134). Oxford, UK: Oxford University Press.
Rodríguez-Domingo, S. (2015). Traducción entre los sistemas de representación simbólico y verbal: un estudio con alumnado que inicia su formación algebraica en secundaria [Translations between the symbolic and verbal representation systems: an study with students at the beggining of their algebraic education in secondary]. Tesis doctoral, Universidad de Granada, Granada.
Socas, M. (1997). Dificultades, obstáculos y errores en el aprendizaje de las matemáticas en la educación secundaria [Difficulties, obstacles and erros in mathematics learning in secondary education]. In L. Rico (Eds.), *La Educación Matemática en la enseñanza secundaria* (pp. 125-154). Barcelona, Spain: Horsori.
Vega-Castro, D., Molina, M., & Castro, E. (2012). Sentido estructural de estudiantes de bachillerato en tareas de simplificación de fracciones algebraicas que involucran igualdades notables. *Relime, 15*(2), 233-258.
Wagner, S., & Parker, S. (1993). Advancing algebra. In P. S. Wilson (Ed), *Research ideas for the classroom. High school mathematics* (pp. 119-139). New York, NY: Macmillan.
Weinberg, A. (2007). New perspectives on the student-professor problem. In T. Lamberg & L. R. Wiest (Eds.), *Proceedings of the 29th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 164-170). Stateline, NV: University of Nevada.
Wollman, W. (1983). Determining the sources of error in a translation from sentence to equation. *Journal for Research in Mathematics Education, 14*, 169-181.

[ii] The algebraic statements which appear represented verbally have been translated to English by the authors. The translations are faithful to the original version which resembles the algebraic statements appearing in mathematics books used in Spain, in particular those used by the students in this study.

[ii] A sequential statement is one that corresponds to the sequential reading of an algebraic expression.