Reading data from graphs: A study on the role of language

Giovannina Albano, Cristina Coppola, Tiziana Pacelli

Reading data from graphs: A study on the role of language

Giovannina Albano[1], Cristina Coppola[2] and Tiziana Pacelli[2]

1 Università degli Studi di Salerno, DIEM, Fisciano, Italy

2 Università degli Studi di Salerno, DipMat, Fisciano, Italy, ccoppola@unisa.it

The work described in this paper is framed in a larger study focused on written argumentations produced by university students in tasks regarding graphs, conversions between graphs and their analytic properties and relations among different graphs. Using the functional linguistics framework, we analyse difficulties related to an examination task on functions and their derivatives. The use of colloquial and literate registers in mathematics problem solving is the lens through which we analyse and discuss the errors. We draw attention to the difficulties concerning the use of literate registers and how these difficulties influence the errors in the interpretations of texts and figures.

Keywords: Language, colloquial and literate registers, written argumentations.

INTRODUCTION

In this paper we focus on the analysis of written argumentations, produced by undergraduates to justify their answers in problems concerning graphs, relations among various graphs and their coordination with analytic properties.

The study was born in a first year University context, and, in particular, concerns Biology freshman students attending a basic mathematics course. According to the goals of mathematics instruction in applicative domains, such as Biology, the course primarily is aimed at allowing students to interpret and compare graphs of elementary functions. The need for argumentation comes from the need for preventing them from answering at random and for promoting a more-in-depth understanding.

Our research is framed in the context of the language as the key point in learning processes (Sfard, 2001) and of the importance of writing to learn (Morgan, 1998). In this framework, languages are seen as constructors of the meanings themselves. From this viewpoint, the quality of language influences the quality of thinking and this requires educational attention to the correspondence between semiotic activities and linguistic competency of the participants (Ferrari, 2004). In fact, evidence shows how some learning difficulties in mathematics can be ascribed to poor linguistic competence (Ferrari, 2004).

In the following we are going to investigate the protocols produced by a sample of students that were required to compare graphs and to produce an argumentation to justify their choices. A lot of difficulties emerged at different levels, in particular, in data reading, even with visual data.

We interpret these difficulties, which are linguistic or semiotic in nature, using the functional linguistic framework. In particular, we focus on the use of colloquial and literate registers in mathematics problem solving.

Our research question is the following: How the level of linguistic competence, particularly the difficulties concerning the use of literate registers, influences the way of interpreting texts and figures. More in general, in a Vygostkian framework (Vygotskij, 1934), we are interested in investigating how the language does (or does not) support thinking in the interpretations of texts and figures.

THEORETICAL BACKGROUND

Language is growing as one of the most relevant issues for research in mathematics education. Several authors have studied the interactions among the different semiotic systems in mathematics learning.

From the cognitive point of view, Duval's (2006) investigations have highlighted how the coordination of various semiotic representations is the key to comprehension in mathematics.

According to O'Halloran (2005) three main groups of semiotic systems can be devised: verbal language, symbolic notations and figural representations, which are strongly interwoven in doing mathematics.

Sfard (2001) interprets thinking as communication and regard languages not just as carriers of pre-existing meanings, but as builders of the meanings themselves. So, under this perspective, language should heavily influence thinking.

As already said, there is evidence that a good share of students' troubles in mathematics can be ascribed to improper uses of verbal language (Ferrari, 2004). More precisely, students often produce or interpret mathematical texts according to linguistic patterns appropriate to everyday-life contexts rather than to mathematical ones. The difference is not just a matter of vocabulary, grammar or symbols, but it heavily involves the organization of verbal texts, their functions and relationships with the context they are produced within. This is why Ferrari has assumed a pragmatic approach to interpret students' behaviours. This means focusing on the language use and on the different functions it plays rather than just on grammar. In this respect, a further characteristic of verbal language has to be taken into account, the multivariety, that is the use of language in various registers (intending register as a linguistic variety based on use) (Leckie-Tarry, 1995).

According to Ferrari (2013), in mathematics we need both the colloquial registers, to construct the concepts without paying much attention to the form of their representation, and the literate registers, including symbolic representations, which can be considered as extreme forms of literate registers. The latter ones are essential to express relations among concepts, to elaborate generalisations and to explain solving procedures. Ferrari (2004, 2013) has shown how many of the students' errors can be traced back to the use of typical styles of colloquial registers whereas more advanced styles would be necessary.

METHODOLOGY

The sample taken into account for our research consists of 64 students attending the first year of a 3-year BSc degree in Biology and taking part in a 48-hour

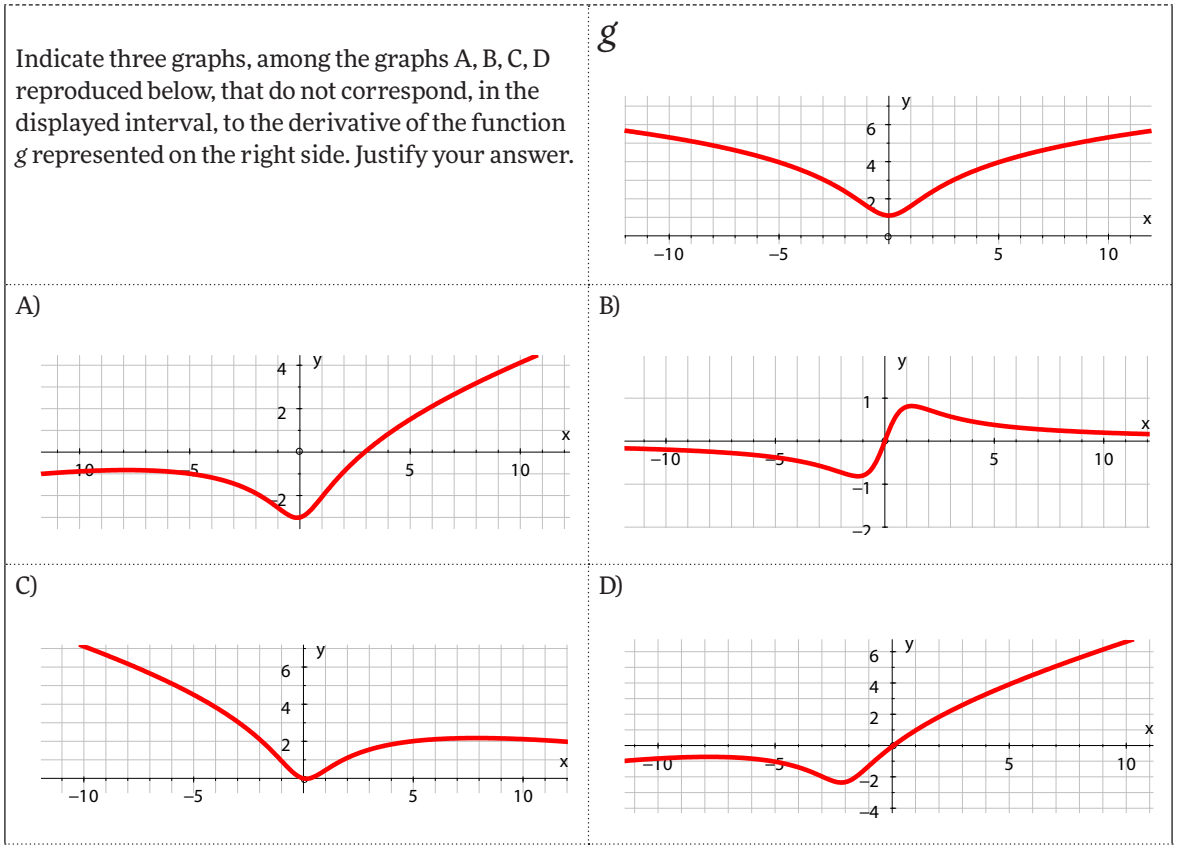

Figure 1: The task in the study

Literate register	Colloquial register
Context of culture: links with other texts and knowledge are activated.	*Context of situation*: the text is linked to the specific situation in which it is interpreted or produced (the exam situation). In this case, students face the problem starting from the text and the images in the task without referring to other knowledge, on the basis of general heuristics (an example is "if f is increasing, then f' is increasing") or looking for analogies (in the graph itself or in other tasks previously carried out).
Lexicalization: words are used according to well-known and shared definitions or descriptions.	*Lack of lexicalization*: everyday meaning of words (that is vaguer) prevails over the one based on definitions (think, for example, of such words as "increasing", "inflexion point", "concavity").
Propositions' explicit structure: subject, complement, conjunctions. Text organization is ruled by grammar and is based on subordination relationships (causative, consecutive, temporal propositions).	*Propositions' implicit structure*: subject or complements are often missing. Text organization is realized by means of spatial nearness of the sentences, or by means of graphical artifices (such as arrows, colours) or textual markers that are vague (for example "and") or randomly (as an example, "therefore" wrongly used).
Stability: text interpretation and production is quite stable, thanks to the explanation and use of verifiable definitions.	*Instability*: text interpretation and production result from factors out of control, e.g. how the students use words, the mental images they activate instead of definitions, previously carried out tasks that they remind of.
Metaphoric-symbolic organization: the link between representations and their content is not natural but conventional or metaphorical.	*Congruent-iconic organization*: organization and form of the representation are not arbitrary but they naturally correspond to organization and form of the meaning.

Table 1: Categories of errors

module in mathematics. According to the aims of the course, the lectures and the tutoring sessions were not based on a deductive approach (that is, definitions, theorems and applications), but they aimed at involving students in activity of coordination of various semiotic systems (Duval, 2006) concerning functions (formulas, graphs, tables, verbal texts).

In accordance with the lectures, in the exams the students were required to master graphical representations of functions and to be able to coordinate graphical and analytic properties of functions and their derivatives. Moreover, argumentation abilities were required to justify their answer.

In the examination task the question was posed in "negative" mode: given a function graph, the students had to indicate three out of four graphs "not" corresponding to the derivative of the given function, justifying their answer (Figure 1).

Our analysis was focused on the argumentations given by the students, even if the answer was correct (i.e., A, C, D), and we looked at them through the lens of the distinction between colloquial and literate registers. We point out that 23 students answered correctly. The errors that emerged have been classified according to typical linguistic categories (Leckie-Tarry, 1995), as listed in Table 2.

ANALYSIS OF PROTOCOLS

We report some protocols of special representativeness, and explain part of our analysis, using the described categories in Table 1. We present the pictures of the original protocols in Italian and our corresponding English translations.

A09: In y =1 the function is increasing, so its derivative has to be increasing, this rules out the graph B. We rule out C because it has x = 0 and at the point 0 is increasing. We rule out D because for y = 2 the function is increasing, and not passing through 0, for this reason we rule out D. The right graph could be A.

In protocol A09 (Figure 2), there is a quite inaccurate and "careless" language and confusion in the reading of graphs. The student refers to a context of situa-

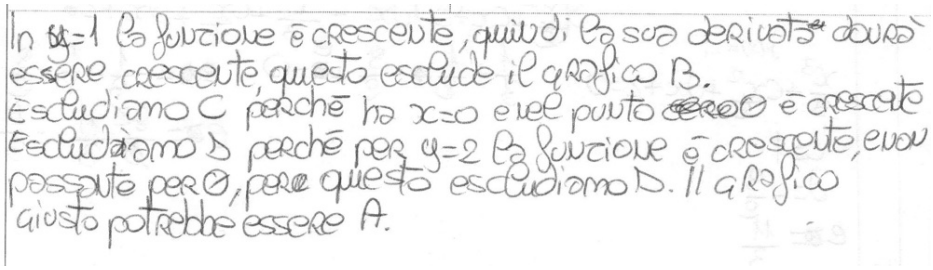

Figure 2: Protocol A09 – Answer: B, C, D

tion using the interpretation that identifies a function's properties with the properties of its derivative. Moreover the structure of propositions is implicit and there are repetitions. Language appears quite vague in "*We rule out C because it has x=0 and at the point 0 is increasing*". Maybe it is meant to say that the function represented in C assumes 0 as the value for $x = 0$. But being increasing in 0, how can this exclude graph C?

Besides, he/she writes "*In y=1 the function is increasing*", where 1 is the *y*-coordinate of the intersection of the graph of the function with the *y*-axis: an iconic organization seems to prevail. The same happens when the student writes "*for y=2 the function is increasing and not intersecting 0*", reading 2 on the *y*-axis and meaning the point in which the function assumes a value equal to 2.

Students, like the one in protocol A09, do not consistently apply some convention to name the points on the graph (e.g. indicating the *x*-coordinate only, or both the *x* and the *y*-coordinates) but they seem to choose the label that is nearer to the point. In other words, the spatial relationship between signs prevails on their defined meanings as well as on the conventions that regulate their use. As shown in Table 1, iconicity is typical of colloquial registers, whereas in literate ones conventions or metaphors are the standard ways of conveying meaning.

A07: I ruled out the graphs A, B, C. Let's notice that the graph of the function *g* has got horizontal tangent, so $g'(x) = 0$ and I rule out A because it does not pass for the origin; *g* displays an upward concavity so $g'(x)$ is increasing. Therefore I rule out B and C that are decreasing.

Also in protocol A07 (Figure 3) the language is quite inaccurate and there is a strong presence of the characteristics of colloquial register categories (Table 1). The student writes: "*the graph of the function g has got horizontal tangent, so g'(x)=0*", without specifying where there is a horizontal tangent in the graph of function *g*. By writing $g'(x)=0$, does he/she mean that the derivative is identically vanishing? Maybe, in a context of spoken communication, the sentence would be followed by a negotiation with the interlocutor, to explain the exact meaning.

Moreover he/she writes "*g displays an upward concavity*", showing a lack of conceptual control of the image and a lack of lexicalization: maybe the everyday meaning of the word "concavity" influences the answer (it could be that the student is reminded of a "standard" image of parabola). Also in the expression "*I rule out B and C that are decreasing*" a poor conceptual control of the image seems to emerge.

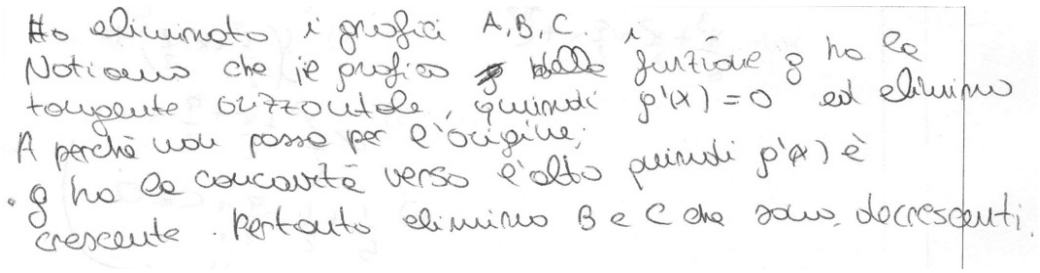

Figure 3: Protocol A07 - Answer: A, B, C

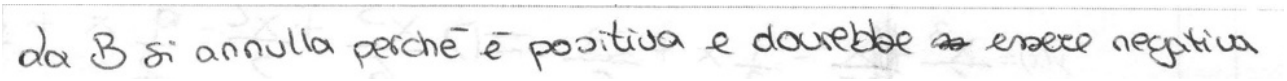

Figure 4: Protocol A21 – Answer: A, B, C

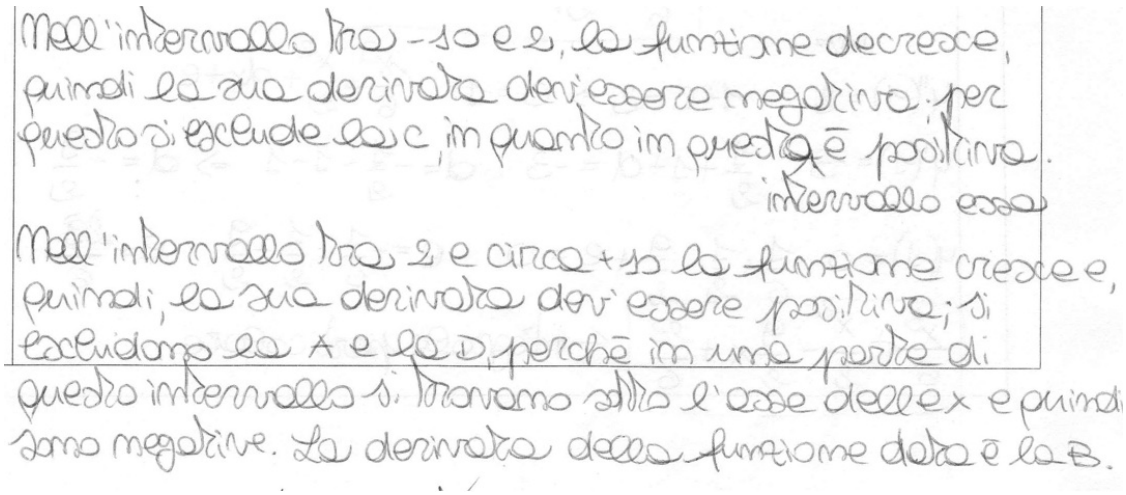

Figure 5: Protocol A55 – Answer: A, C, D

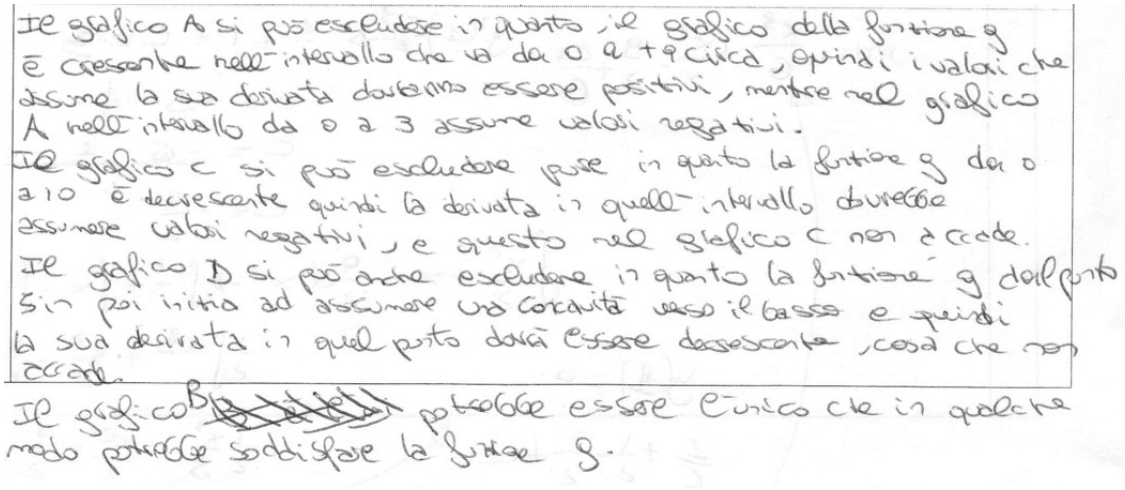

Figure 6: Protocol A06 – Correct answer: A, C, D

A21: B vanishes because it is positive while it should be negative.

Together with the lack in the conceptual control of the image ("*B...is positive*"), we found the protocol A21 (Figure 4) interesting for the use of the verb "vanish" instead of "rule out". The student here uses a word belonging, in some sense, to the mathematical register instead of using a word of everyday language, maybe because the task he/she is carrying out is a "mathematical" task.

A55: In the interval between -10 and 2, the function decreases, so its derivative has to be negative; for this reason the C is ruled out, for in this interval it is positive. In the interval between 2 and about +10 the function increases and, so, its derivative has to be positive; A and D are ruled out, because in a part of this interval they are below the *x*-axis and so they are negative. The derivative of the given function is B.

The argumentation used in protocol A55 (Figure 5) is quite accurate, but there is a strong presence of iconicity. The student writes "*In the interval between -10 and 2, the function decreases*" reading -10 on the *x*-axis and 2 on the *y*-axis which are explicitly written in the diagram (while 0 on the origin is not written). Maybe he/she chooses the most evident label, or the label nearest to the point he/she wants to indicate. He/she writes again: "*In the interval between 2 and about +10 the function increases*", showing once more an iconic interpretation of the diagram.

In ruling out graph D ("*A and D are ruled out, because in a part of this interval they are below the x axis*"), the student shows not only a lack of conceptual control on the image, but, maybe, again iconicity: it seems that the student considers the interval between -2 and +10, instead of the interval +2 and + 10; he/she refers to the

label -2 or to the label 2, indifferently, as if both the labels indicate the same point.

A06: Graph A can be ruled out because the graph of function *g* is increasing in the interval going from 0 to about +9, so the values of its derivative have to be positive, whereas in graph A in the interval from 0 to 3 it assumes negative values. Graph C can be ruled out too because function *g* from 0 to 10 is decreasing, so the derivative in that interval should assume negative values, and this does not happen in graph C. Graph D can be ruled out too because function *g* from point 5 on begins displaying a downward concavity so its derivative in that point has to be decreasing, and this does not happen. Graph B could be the only one satisfying in some way function *g*.

The student of protocol A06 (Figure 6) gives the correct answer (he/she rules out graphs A, C and D), produces a detailed conversational argumentation and shows a good capability of reading of graphs and linking data to conclusions. Nevertheless also in his/her protocol, we can find an expression such as: "*function g from 0 to 10 is decreasing*". He/she expresses like one looking at the graph from the origin to the left and from the origin to the right.

Indeed this behaviour emerges from several protocols: a lot of students privilege their own point of view with respect to definitions and to the shared and well-known rules of using representations.

DISCUSSION AND CONCLUSIONS

From the protocols' analysis three macro-groups of students' behaviours seem to emerge. These macro-groups are based on the categories of errors (in Table 1), categories that in students' behaviours are continuously interwoven.

The first macro-group is represented by the protocols of those students that do not apply the usual conventions (the "grammar") of the Cartesian plane but use the image they see. Students belonging to this group do not refer to the *context of culture* evoked by the representations but to the *context of situation*.

In some protocols, as an example, we can observe that, in order to designate the points of the Cartesian plane, the students tend to label the points only by their ordinate value. Probably, they choose this value just because it is the label nearest to the point. Conversely some students use a point of the plane in order to indicate the starting point of an interval on the *x*-axis (see Protocol A55 in Figure 5).

In this macro-group we can find also other students' behaviours: for example students exploring the graph starting from the origin of the Cartesian axis and considering as "positive" the direction from 0 to $-\infty$; other students take into consideration in different graphs points that look the same (as an example, points that are minimum) instead of comparing points that have the same *x*-coordinate. These students remain in an *iconic interpretation of the diagram* and privilege their own point of view (related to the context of situation, i.e. the diagram drawn on the paper as a material object), instead of the recognized rules or conventions for using representations (which are obviously related to the context of culture).

A second macro-group is represented by the protocols of students having some difficulty in the *use of the vocabulary*. For these students some words, such as "positive", "increasing", etc., have a vague and unstable meaning. As a consequence they make errors also in reading the graphs. In such a situation students write sentences such as "*I rule out B because it should be positive for* $x > 0$", probably having in mind the concept "increasing".

Another behaviour related to this macro-group is mixing up the subjects in sentences like "*the derivative vanishes* $\leftrightarrow$ *the function vanishes*".

All these behaviours are typical of the use of colloquial registers.

The last macro-group is represented by the protocols of students having some problems in the *organization of the text*, i.e. the meaning of the words related to their position in the text.

As an example we can observe some protocols in which the students use the expressions "increasing", "positive", "upward concavity" for describing properties related to an interval even if these expressions are true only in a part of the interval.

This behaviour is linked both to that one, above described, characterized by taking into consideration the first image or the first concept that seems to work well without making further inferences, and to how the students tend to organize the text in the colloquial registers. In the last case they are not used to specify the complements because in a context of spoken communication the meanings could be negotiated later. In this way the meanings become unstable and vague and as a consequence it is hard reading the images and making further argumentations.

The outcomes of the analysis seem to suggest the need for a stronger educational attention to linguistic issues in mathematics. It is fundamental the capability of using both literate registers, which support scientific thinking, and colloquial ones, which are indispensable too in the construction of mathematical concepts. Thus switching back and forth between colloquial and literate registers is a crucial process in mathematics learning. These skills are not natural but have to be developed and fostered, by specifically planned teaching activities, from primary school onwards.

REFERENCES

Duval, R. (2006). The cognitive analysis of problems of comprehension in the learning of mathematics. *Educational Studies in Mathematics, 61*, 103–131.

Ferrari, P. L. (2004). Mathematical language and advanced mathematics learning. In M. Johnsen-Høines & A. Berit-Fuglestad (Eds.), *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 383–390). Bergen, Norway: PME.

Ferrari, P. L. (2013). Rappresentazione e comunicazione in matematica: le potenzialità dell'e-learning. *L'Educazione Matematica, 3*(1), 63–70.

Leckie-Tarry, H. (1995). *Language and context. A functional linguistic theory of register.* London: Pinter.

Morgan, C. (1998). *Writing mathematically. The discourse of investigation.* London: Falmer Press.

O'Halloran, K. L. (2005). *Mathematical discourse. Language, symbolism and visual images*. London: Continuum.

Sfard, A. (2001). There is more to discourse than meets the ears: looking at thinking as communicating to learn more about mathematical learning. *Educational Studies in Mathematics*, *46*, 13–57.

Vygotskij, L.S. (1934). *Thought and language.* Moscow-Leningrad: Sozekgiz.