INTERPRETING SYMBOLIC STATEMENTS AS TEXTS: AN EXPLORATORY STUDY

Pier Luigi Ferrari Cristina Giraudi
Università del Piemonte Orientale ad Alessandria

This paper is devoted to the problem of the interpretation of mathematical texts[1]*. Some ideas on mathematical language are shortly discussed with the help of some constructs from functional linguistics. Some evidence regarding the interpretation processes of a symbolic text by groups of 10-graders, including both written answers and the transcriptions of spoken interactions is presented and discussed. The outcomes of this study show that students often try to interpret mathematical statements according to everyday-life schemes. This suggests that in school practice mathematical expressions should be dealt with as texts (rather than as abbreviations or local conventions) and that metalinguistic awareness should become one of the the goals of both linguistic and mathematical education.*

Recently various theoretical frameworks have been diffused that enhance the role of languages in the learning of mathematics. This holds specially for the neo-Vygotskian standpoint, which gives great value to communication as a way to promote learning. More recently, in investigations more focused on cognitive aspects, Sfard (2000a, 2000b) interprets thinking as communication and assigns to languages a more complex role than the traditional one: they are not regarded just as carriers of (pre-exixting) meanings, but as builders of the meanings. In a context where communication becomes central, it cannot be regarded but as "an activity in which one is trying to make his or her interlocutor act or feel in a certain way", i.e. an activity pertaining to the realm of pragmatics too. A thorough investigation of the languages of mathematics from the standpoint of pragmatics is far from being developed. Some example of application of pragmatic constructs to the learning of mathematics, such as Grice's[2] Cooperation Principle, may be found in Ferrari (2000). Morgan (1996, 1998) and Burton & Morgan (2000) have carried out investigations on mathematical language from the viewpoint of Halliday's[3] functional linguistics. They focus on some interpersonal aspects of mathematical language (such as the use of impersonal forms in academic mathematics textbooks) but take into account some aspect which are interesting from a cognitive

[1] Through the paper the word 'text' is used according to linguistics to mean any spoken or written instantiation of language, independently from its length or complexity.

[2] Grice (1975)

[3] Halliday (1974, 1985)

viewpoint as well, such as cohesion[4]. In particular Morgan (1998) provides a general description of mathematical language as a set of registers as they are used in mathematical practice (rather than as they are usually accepted by mathematicians) which seems adequate to the needs of research and practice. The use of expressions like 'mathematical language' through this paper assumes Morgan's definitions and discussions. The new functions of mathematical language in education require researchers and practitioners to consider it as a complex system, taking into account all its components (verbal, symbolic, visual, ...) that very often are combined. Ordinary languages and mathematical one are different as regards not only the symbolic or the visual component, but the verbal one too. In the new perspective the verbal component cannot but play a crucial role. Still, the symbolic component has played a major role in the development of mathematical thought and may play an important role in mathematics education. A thorough discussion of the functions of the symbolic component is far beyond the aims of this report. Symbolic notation systems are not important just because they are possibly more precise or less ambiguous than ordinary languages, but because of the computational opportunities they provide. Moreover, they allow people to get rid of some of the meanings embodied in everyday-life words, when it is necessary to build new meanings.

An important feature of educational pratice in mathematics is the need of using the same linguistic forms with different functions: to build and organize mathematical knowledge and to communicate with other people, their experiences and cultures. This plurality of functions may generate conflicts. For example, in contemporary mathematics, it is perfectly acceptable to name 'rectangle' a square shape, whereas this use is inadequate in some contexts as far as it violates pragmatic conventions, such as Grice's Cooperation Principle. In a similar way, in some cases, logical connectives (and, or, if...then, not, ...), no matter whether in verbal or symbolic form, are to be used according to their truth-functional definition (for example, when defining set-theoretical operations), whereas in communication practices within the class they are required to play functions that go far beyond truth-functionality, such as the organization of the texts and of the links between the sentences that occur. If the understanding of texts in mathematics becomes a goal for education, then students have to deal with texts with different format, organization and functions almost at the same time. Different kinds of text may propose different interpretative problems. For example, some of the questions raised by Sfard (2000b) as concerns discursive focus[5] are appropriate. A good share of

[4] According to Halliday (1985, p.309) 'cohesion' refers to non-structural resources designed to establish relations within the text that are not only semantic but also functional.

[5] In linguistics various constructs (such as theme/rheme, topic/comment, given/new, ...) have been proposed to deal with these issues. For detailed discussions see for example Halliday (1985) or Leckie-Tarry (1995).

the misinterpretations of texts by college students, for example, are related to failure in the grasping of the focus of the text as a whole rather than in the interpretation of single words or expressions. This may depend from the fact that often in mathematical texts focus is not marked in the same ways as other texts. A closely linked question is cohesion, i.e. the functional links among the various components of a text. Cohesion in mathematical language (in all its components) is usually less explicitly marked than in ordinary language, which may be an obstacle to students' interpretations.

This study is devoted to the ways some groups of 10-graders interpret a text made up of three symbolic expressions. We are interested in the ways students put together the occurring expressions and the interpretation schemes they adopt when the expected interpretation is not adequate from the standpoint of communication.

INTERPRETATION OF A SYMBOLIC TEXT

The following problem has been given, in December, 2000, to four classes of 10 graders.

PROBLEM[6]

The positive integers x, y are given. We know that all of the three following properties hold at the same time

(a) $x^2 < y^2$ (b) $3x > y$ (c) $x^3 > 10^6$

Based on the given data, for each of the following statements find whether it is true or false:

(i) $x \leq y$ (ii) $2 < y$.

Explain your answers.

A PRIORI ANALYSIS

This problem includes various critical points. First of all, there is the interpretation of '≤' in a true statement which violates pragmatic principles. A potential obstacle is the interpretation of inclusive 'or' in a statement which is of the form 'A or B' with A clearly true and B clearly false, which violates elementary pragmatic principles such as Grice's Maxim of Quantity, for 'A' alone would be less expensive and more effective from the viewpoint of communication. Also the coordination of (a) and (c) may result troublesome for a number of students. At this regard it must be remarked that cohesion

---

[6] The original texts of the problem, of students' answers and of the transcriptions of their spoken interactions are in a language other than English. Some of their linguistic features are lost in the traslation process. For example, we are afraid that the English translations of the texts produced in the interactions are not in the same register as the original ones. Nevertheless, in this paper we investigate aspects which are not too much affected by the translation. Anyway, the original versions of the materials are at disposal of anyone interested.

among (a), (b) and (c) is explicitly pointed out (at the meta-textual level) in the preceding verbal text ('all of the following ...', 'at the same time', 'based on the given information'), but this may be not enough, since neither specific algorithms nor standard linguistic markers are available. The problem has been designed in order to prevent students from applying some standard algorithm with little control, forcing them to use methods based on the interpretation of the given statements.

METHODOLOGY

The problem has been given to 4 classes of 10-graders (76 students altogether). Students have worked about 30' individually (producing written answers) and other 30' in small groups (2-3 students of different skill levels). Copies of their written individual answers were available to students during the interaction; the work of all the groups of one class has been recorded. We present some quantitative data on the whole sample and investigate the behavior of one group of 2 students more closely. Of course, the data of the first kind are gathered from texts actually written out by the students whereas the others are transcriptions of spoken interactions.

INDIVIDUAL ANSWERS

Question (i)

57 students claim that (i) is false, 16 that it is true and 3 give no answer. Negative answers mostly refer to the fact that "it cannot happen that $x=y$". Some other students are puzzled by the occurrence of '$\leq$' in (i), whereas in the data occur '<' or '>' only.

Question (ii)

44 students claim that (ii) is true, 11 do not answer, 8 claim that it is false and 13 claim they have not data enough to give a definite answer. Altogether 21 students seemingly do not recognize the links between (a) and (c). Moreover the number of non-answers is larger than in question (i). Most of the answers to (ii) have been given with no explanations. Among the explanations given we mention: "I cannot know if $2<y$, it depends on the values of x, y" or "I have no data on y" or "$2<y$ is false because if $x=1$, then y could be 2".

Let us see some more examples.

Valentina answers to (i)*: "False because x is never equal to y", whereas to (ii) answers:* "$x>100$ (c), $x^2<y^2$ (a) $\Rightarrow y>2$, since x, y are positive"

Ivano: *"(i) is false because* $\sqrt{x^2<y^2} = x < y$ *, (ii) is false too because y could be 2; if (ii) were with '=' it would be true"*

Andrea: *"(i) is false, because in (a) there is < whereas in (i) there is =. I do not know whether (ii) is true or false because I have data on x and not on y"*

Sergio: *"(i) is false.* $y = \sqrt{x^2 + k} = x + \sqrt{k}$ *but this do not imply 2<y"*

Deborah: *"(i) cannot be, because in the hypothesis there is* $x^2 < y^2$*; if it were* $x^2 \leq y^2$ *it would be true. (ii) is true"*

Enrica: *"(i) is false because y cannot be equal to x because if I square both they would be still equal; (ii) is true bacause else x would be less or equal to 1 and the third piece of information would be false"*

It is noteworthy that among the 'improper' answers to (i) a good share are explained quite clearly (showing some command of mathematical notations), whereas it is difficult to find well explained answers to (ii). Moreover, no student refers to (b) in his or her answer to (i) and (ii).

INTERACTIONS

Le us examine the transcriptions of the interactions of the group made up by Valentina (a girl with excellent grades in all subject matters including mathematics) and another girl named Ines (rated at average level).

Valentina: *"The first is false because x cannot be equal to y"*

Ines: *"If the square is less, the number too is less"*

V: *"Hmm, here there is 'less or equal'"*

I: *"It is the same!"*

V: *"It is not equal!"*

I: *"But it works all the same!"*

V: *"Why am I to write 'equal' if it is less?'*

I: [a bit vexed] *"Oh, it is like the elevator: there is 'Maximum weight three hundred kg' but you take it even when you are alone.* [laughing] *You do not weigh three hundred kg!"* [Valentina is small and slim]

V: *"Of course not. Maybe you are right. The second is true."*

I: *"We have no information on y"*

V: *"The cube of x is ten to the sixth. So x is equal to one hundred. If x is at least one hundred, y must be one hundred one, at least."*

I: *"y could be one and the statement would be false*"

V: *"If y were one, x would be zero, the cube of zero is zero"*

I: *"It could be:* [points at the occurrences of *x* and *y* in (a), (b) (c)] *x is zero, y is one, x is one hundred one"*

V: *"x is always the same. Okay, we know that y is more than x, and x is more than one hundred one, then y is more than one hundred one"*

DISCUSSION

A palpable outcome is students' uneasiness in recognizing that $x \leq y$. Most of them adopt the argument that x cannot be equal to y, so pointing at the communicative inadequacy of the formula rather than at its claimed falsity. Valentina's answer is quite clear also because she shows a good command of mathematical notations and steadily applies even 'ab absurdo' arguments. It seem reasonable to conjecture that these students (rightly) feel the inadequacy of the statement $x \leq y$ which violates not only everyday-life pragmatic rules but even implicit rules of school practice: usually to answer a question it is not accepted just a true statement, but the statement which is the most adequate to the question and the related context is required. Other answers point out a further aspect: the difference between the relation '<' occurring in (a) and '≤' occurring in (i). In this case it is questioned the adequacy not just of (i) but rather of the whole text.

As concerns (ii), the answers of students who fail in linking the question to both (a) and (c) can be classified into two groups. Who answers 'false' most likely focuses on (a) only and remarks that y could be 2. (b) is neglected by almost all students, maybe because in it occurs '>' in place of '<' or '≤'. Most likely (c) is not taken into account by some students as it does not involve y explictly. These answers seem depend on the lack of linguistic markers of cohesion between (a), (b), (c). The lack of cohesion induces some students to assign to each statement its own topic. The topic of (c) alone cannot be other than x, which is the only variable occurring. Some students try to apply algorithms to put data together.

The different features of mathematical language (in all its components) from the functional (not only grammatical or lexical) viewpoint may explain a number of students' difficulties. Most of the students refuse $x \leq y$ as inadequate compared to the data available. Even the troubles with question (ii) may be explained in a similar way. Ordinary language provides a number of ways to mark the links between the statements in a text (intonation in spoken texts, vocabulary, pronouns, connectives, ... in all the texts) whereas mathematical one (in its symbolic and often verbal component) cohesion is usually marked in other ways, such as the spatial disposition of the formulas or the availability of specific algorithms or the repetition of some symbol or letter. This happens for example in the solution of linear systems: standard methods automatically take into account all the equations involved, that are identified mainly by their spatial disposition and by the occurrence of brackets or braces.

The spoken interaction between Valentina and Ines points out some interesting processes. Valentina, who usually takes good grades in mathematics, gets stuck because of the occurrence of '$\leq$'. Maybe Ines, who generally takes lower grades in mathematics than Valentina, is not completely aware of the question raised by her friend, but her indifference for the distinction between '<' and '$\leq$' and her efforts to represent the data verbally play a positive role. Ines' attitude is clearly agonistic, as she seems to be moved mainly by the wish of prevailing against Valentina. As regards question (ii) Valentina, in order to explain her answer, provides a sequence of examples that use statements that are not consequences of the data ("the cube of x is ten to the sixth", "so $x$is equal to one hundred") afterwards rectified by others ("if x is at least one hundred"). Ines clearly does not grasp cohesion among the data and interprets the two occurrences of x as different numbers. Both the parts of the interaction enhance some features of verbal language. Ines and Valentina are both inaccurate in their interpretations. Ines tries to use 'less' to interpret both '<' and '$\leq$', which is inaccurate, but succeeds in drawing Valentina's attention on some aspects of the meanings involved that are relevant to the answer. Moreover she uses an example (the elevator) which is not closely related to the problem, but where the pragmatic function of the warning 'Maximum weight three hundred kg' is made straightforward by the situation and one's everyday experience. In other words, the example of the elevator is pragmatically rather than semantically related to the problem situation. Also Valentina, as remarked above, is inaccurate in her examples. Her efforts to give x and y values compatible with the data seem useful steps toward the solution. In both cases, verbal language (in a spoken register) allows her to make inaccurate statements and rectify them afterwards without too much danger. If Valentina had written down her examples in symbolic form ("$x^3=10^6$", "$x=100$") and had applied to them standard algebraic transformations, she could have lost the control of the function of her productions. In other words verbal language (in both spoken and written registers) not only provides much more opportunities to mark some of the functional features of the texts (topic, cohesion, ...) but is also more flexible than symbolic one, as it allows people to produce inaccurate statements and to rectify them afterwards, or to mark them as conjectures, or examples, or other. Very often I find college students who write down formulas that are not consequences of the assumptions but only examples. Unfortunately, they often forget the functions of their writings, and apply to them algebraic transformations, and derive false conclusions. Their behaviors, that are sometimes labelled as 'incorrect applications of rules' could more effectively be regarded as examples of failure in the control of the functions of the texts produced.

TEACHING IMPLICATIONS

A possible interpretation of this data is: experiments of this sort have no relevant teaching implications as the problems assigned are tricky and unfair. This opinion is compatible with traditional teaching practices that mainly enhance the learning of standard procedures in standard formats. If we give a central role to communication the role of languages becomes more relevant. In particular it seem reasonable that students should interpret simple texts in mathematical language including those containing symbols, even if they are not in standard format (as happens in everyday-life communication). The outcomes of this experiment point out the need that students command the transitions between different languages or registers, with the related functional properties and conventions. This suggest that mathematical expressions should be studied as texts rather than just as local conventions or abbreviations. This implies a better coordination between the teaching of languages and the teaching of mathematics and a stronger focus on aspects like metalinguistic awareness, i.e. awareness of form and functions of a text, in addition to its meaning, as suggested by MacGregor & Price (1999). Of course, further research is needed to refine these ideas and to design the teaching methods more suitable to attain the goals suggested above, but we believe that anyway mathematical language should be considered in the context of actual interactions (rather than as a separate code) with all its components.

REFERENCES

Burton, L. & Morgan, C.: 2000, 'Mathematicians Writing', *Journal for Research in Mathematics Education*, 31/4, 429-453.

Ferrari, P.L.: 2000, 'Cooperative principles and linguistic obstacles in advanced mathematics learning' in Inge Schwank (ed.), *Proc. of the I Conference of the European Society for Research in Math. Education,* Vol.2, http://www.fmd.uni-osnabrueck.de/ebooks/erme/cerme1-proceedings/cerme1-proceedings.html

Grice, H.P.: 1975, 'Logic and conversation', in Cole, P.&J.L.Morgan (Eds.), *Syntax and semantics: Vol.3. Speech acts* (pp.41-58), New York: Academic Press.

Halliday, M.A.K.: 1974, 'Some aspects of sociolinguistics', *Interactions between Linguistics and Mathematical Education Symposium,* Paris, UNESCO.

Halliday, M.A.K.: 1985, *An introduction to functional grammar,* London, Arnold.

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

MacGregor, M. & E. Price: 1999, 'An Exploration of Aspects of Language Proficiency and Algebra Learning', *Journal for Research in Mathematics Education,* 30-4, 449-467.

Morgan, C.: 1996, '"The language of mathematics": towards a critical analysis of mathematical texts', *For the Learning of Mathematics,* 16/3.

Morgan, C.: 1998, *Writing Mathematically. The Discourse of Investigation,* London, Falmer Press.

Sfard, A.: 2000a, ' Symbolizing Mathematical Reality Into Being--Or How Mathematical Discourse and Mathematical Objects Create Each Other', in Cobb, P., E.Yackel and K.McClain (eds.), *Symbolizing and Communicating in Mathematics Classrooms,* Lawrence Erlbaum Associates.

Sfard A.: 2000b, 'Steering Dis(Course) Between Metaphors and Rigor: Using Focal Analysis to Investigate an Emergence of Mathematical Objects', *Journal for Research in Mathematics Education*, 31/3, 296-327.