

Natural language as a tool for analyzing the proving process: the case of plane geometry proof

Elisabetta Robotti

Educational Studies in Mathematics
An International Journal

ISSN 0013-1954
Volume 80
Number 3

Educ Stud Math (2012) 80:433-450
DOI 10.1007/s10649-012-9383-0

EDUCATIONAL STUDIES IN MATHEMATICS

AN INTERNATIONAL JOURNAL

EDITOR-IN-CHIEF: ANNA SIERPINSKA

 Springer

ISSN 0013-1954
©2012 EDMAN

 Springer

Your article is protected by copyright and all rights are held exclusively by Springer Science+Business Media B.V.. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your work, please use the accepted author's version for posting to your own website or your institution's repository. You may further deposit the accepted author's version on a funder's repository at a funder's request, provided it is not made publicly available until 12 months after publication.

Natural language as a tool for analyzing the proving process: the case of plane geometry proof

Elisabetta Robotti

Published online: 18 February 2012
© Springer Science+Business Media B.V. 2012

Abstract In the field of human cognition, language plays a special role that is connected directly to thinking and mental development (e.g., Vygotsky, 1938). Thanks to “verbal thought”, language allows humans to go beyond the limits of immediately perceived information, to form concepts and solve complex problems (Luria, 1975). So, it appears language can be studied as a cognitive process (Chomsky, 1975). In this investigation, I study language as a means for making the cognitive process explicit. In particular, I analyze the role of the verbalization produced by pairs of students solving a plane geometry problem. The basic idea of my research is that, during the resolution process of a plane geometry problem, natural language can play roles beyond that of communication: Natural language can be seen as a tool for supporting students’ cognitive processes (Robotti, 2008), and, at the same time, it can also be seen as a researchers’ tool which allows us to shed light on the evolution of students’ cognitive processes. With regard to language as researchers’ tool, I show how natural language (in our case, students’ verbalization during resolution of a plane geometry problem) can be used by the researcher to make explicit, to study, and to describe the development of the students’ cognitive processes during the resolution process. To this end, I present a model I have developed that allows us to identify, in students’ verbalization, different phases of their cognitive processes.

Keywords Natural language · Plane geometry problem · Verbalization

1 Introduction

The reported research study involved field activities in Italy and France as part of a doctoral thesis carried out within the context of Italy–France academic cooperation. In these two countries, the proof is introduced in the geometry domain. Accordingly, I chose to set students the “prove that...” type of geometry problem.

E. Robotti (✉)
DiDiMa srl, spin off ITD-CNR, Genova, Italy
e-mail: elisabetta.robotti@gmail.com

The resolution of a proof in plane geometry is a very complex phenomenon where natural language, geometrical figures, and students' knowledge are all engaged. I focus on natural language in order to analyze the role of the verbalization produced by pairs of students solving a plane geometry problem. Moreover, my analysis focuses on the interaction between figural and conceptual aspects involved in the resolution of this kind of problem. I believe that natural language can support a students' resolution process (Chomsky, 1975; Luria, 1975), moving it from simple observation about the drawing to organization of deductive reasoning. In this sense, natural language can be seen as a tool for supporting students' cognitive processes.

With regard to this issue, I refer to my paper (Robotti, 2008), in which it is shown that natural language provides students with different kinds of help for problem solving. For example, it has been demonstrated that a particular use of natural language guides the students' resolution process, allows them to control that process, and makes it possible to associate "labels" (the combination of words and configurations) to concepts belonging to the student's system of knowledge useful for solving the problem at hand.

I defined these kinds of help as *functions of natural language* (guide function, association function, etc.), emphasizing the conditions under which these functions play an important role in the resolution process (Robotti, 2008). This article focuses on the idea that natural language can also be considered as a researcher's tool that sheds light on the evolution of students' cognitive processes.

2 Theoretical frameworks of reference

2.1 Brief introduction

My hypothesis is that I can analyze students' verbalizations as data and infer considerations and results. To legitimize this hypothesis, I refer to Ericsson and Simon's literature (1980, 1993). They argue that verbal reports can be considered as data under certain conditions. Ericsson and Simon (1980) show that "verbal reports, elicited with care and interpreted with full understanding of the circumstances under which they were obtained, are a valuable and thoroughly reliable source of information about cognitive processes" (p. 247). Ericsson and Simon (1993) show that, if the act of verbalizing subjects' thought processes does not change the sequence of thoughts, then the subjects' task performance should not change as a result of thinking aloud. They found no evidence that the sequences of thoughts (accuracy of performance) were changed when subjects thought aloud as they completed the tasks, compared with subjects who completed the same tasks silently. However, some studies show that think-aloud subjects take somewhat longer to complete tasks, presumably due to the additional time required to produce the overt verbalization of thoughts.

Based on Ericsson and Simon's (1993) theoretical analysis and on Sfard (2001a, 2001b, 2008) and Sfard and Kieran's (2001) more recent studies, I argue that the closest connection between thinking and verbal reports is found when subjects verbalize thoughts generated during actual task completion.

I admit that, in geometry, three registers are always involved: the figural register, the natural language register, and symbolic language. Laborde (1982) has shown the interactions among registers of language. Our point of view is centered, above all, on the relationship between the figural and natural language registers. This led us to recognize the multiple shifts between the graphical domain (drawing) and the theoretical domain (theoretical object) in the discourse produced by students.

In order to characterize the development of a proof in plane geometry, I refer to Laborde's idea, describing that process in terms of the relationship between the graphical space domain (SG) and the theoretical domain (T). To analyze the dialectic between figural (graphical space domain) and conceptual (theoretical domain) in the resolution process, I refer to Duval's (1995a, 1995b) theory concerning different ways to transform a figure, in other words, different ways to operate on graphical entities.

Moreover, in order to show how researchers can use students' verbalization during problem resolution to make explicit, study, and describe the development of students' cognitive processes, I call on the extra-language theory¹ of the psycholinguist Bronckart (1985) and some important aspects of Duval's theory concerning the different ways of *expansion discursives* (development of discourse) by means of which discourse is composed.

2.2 Relation between verbalization and thinking

With regard to the cognitive psychology viewpoint, I consider Ericsson and Simon's theory. They have developed a processing model whose purpose is to aid interpretation of verbal data obtained from subjects who are engaged in cognitive processes. The authors assume the following hypotheses:

- Information is stored in short-term memory and in long-term memory, which have different capacities and accessing characteristics;
- Any verbalization of cognitive process is based on a subset of the information stored in these memories.

From these hypotheses, Ericsson and Simon draw a major distinction between procedures in which verbalization is a direct articulation of the stored information and procedures in which the stored information is input in intermediate processes, such as abstraction and inference; verbalization is a product of this intermediate processing.

This is the case for cognitive processes involved in the resolution of a plane geometry problem. More specifically, Ericsson and Simon (1993) analyze the verbalization concurrent with task performance. The authors assume "that the verbalization involves either direct articulation of information stored in a language [...]; articulation or verbal recoding of non-propositional information without additional processing [...]; articulation after scanning, filtering, inference, or generative processes have modified the information available [...]" (p. 227). At three different levels, the model predicts that thinking aloud will not change the course and structure of cognitive processes.

Moreover, the authors show that verbalizing information affects cognitive processes only if the instructions require verbalization of information that would not otherwise be attended to.

Bearing these considerations in mind, the design of the field experiments involved students working in pairs, so that communication requirements would motivate them to think aloud. In other words, description or explanation about how they would solve these problems would not be a response to a demand from the researcher but would rather derive

¹ "l'ensemble théoriquement infini de toutes les entités en dehors de la langue.... Par essence, l'activité langagière s'articule à l'extralangage [en définissant] des **espaces** dotés de deux type de pertinence: la pertinence *référentielle*, c'est-à-dire la capacité à devenir un " contenu représenté " de l'activité langagière, et la pertinence *contextuelle*, c'est-à-dire la capacité de contrôler ou gérer le déroulement même de l'activité langagière." (p. 26)

from a request from their schoolmate. The instructions asked them to focus on solving the problem and only to give verbal expression to those thoughts that emerged while generating the solution under normal (silent) conditions.

More generally, in Ericsson and Simon' theory, verbal reports are only one indicator of the thought processes that occur during problem solving. They described other indicators including reaction times (RTs), error rates, patterns of brain activation, and sequences of eye fixations. These kinds of indicators are not the object of my analysis.

Moreover, from the educational point of view, I adopt Sfard's (2001a) idea that "*thinking is a special case of the activity of communicating* [...]" (p. 3, italics in the original). She argued that, "a person who thinks can be seen as communicating with herself. This is true whether the thinking is in words, in images, or in any other symbols." (p. 4).

In accordance with Sfard, I analyze students' thinking via analyses of their recorded verbalizations and their protocols.

As Sfard (2001a) has argued, communication should be viewed not only as an aid to thinking (see Robotti, 2008), but as almost tantamount to thinking itself. The *communicational approach to cognition* is built around this basic theoretical principle. Sfard (2001b) said that "Mathematical discourse is made special by two main factors: first, by its exceptional reliance on symbolic artifacts as its communication-mediating tools, and second, by the particular meta-rules that regulate this type of communication." (p. 46). In accordance with these principles, I have elaborated the idea of different uses of linguistic units. Thus, I refer to the different uses of linguistic units which appear in the discourse developed by students during the resolution of the plane geometry problem in order to identify the development of their cognitive processes.

3 The relationship between the graphical space domain and the theoretical domain in plane geometry

The development of a demonstrative process in plane geometry concerns the relationship between the graphical space domain (SG) and the theoretical domain (T) (Laborde, 1999). Laborde stresses that the passages between these are essential for solving problems in plane geometry. The interplay between theoretical references in the geometrical domain (T) and the graphical entities on which it is possible to operate (SG) is an essential part of the meaning of geometry. Thus, our analysis of the resolution process of a plane geometry problem focuses on the relationships between figural aspects (SG) and conceptual aspects (T). Moreover, in order to analyze the moves between the SG and T domains during the resolution process, I consider Duval's notion of apprehensions of drawing as different ways to operate on the graphical entities.

3.1 The apprehension of a drawing as defined by Duval

It is evident that the resolution of a proving process in plane geometry is based on the information obtained by the visual perceptive system. As Mesquita (1989a, b) states, every drawing can be reconfigured by interpretation of the drawing itself. These reconfigurations may be drawn or mental images, but, either way, they suggest different ways of thinking which may, or may not, be relevant for solving the problem at hand. Duval provides a detailed framework for analyzing the semiotics of geometric

drawings (1995a p. 145–147). In this framework, he identifies different types of what he calls "cognitive apprehension":

The *perceptual apprehension* is the immediate and automatic way to perceive a figure. It allows us to recognize immediately a shape or an object both in the plane and in space.

The *operative apprehension* concerns possible modifications of a figure and its perceptive reorganisation on the basis of those modifications.

The *discursive apprehension* makes clear geometrical properties of the figure other than those written in the task or in the legend. Duval stated that this is a cognitive kind of treatment and, for this reason, it corresponds to deductive thinking.

In the following section, this framework is developed, allowing us to emphasize natural language as a researcher's tool. In this sense, natural language can be used by the researcher to make explicit, study, and describe the development of students' cognitive processes during resolution. To this end, I describe a theoretical linguistic tool that allows us to characterize the development of students' resolution process in terms of the development of their discourse.

The idea is that the practice of a discourse is inseparable from a certain cognitive functioning; it draws both on Vygotsky (1938) in the psycho-pedagogical area and on Duval (1995a) in the didactics of the mathematics domain. Students' verbalization produced during the resolution process (cognitive process) is considered as the discourse that I seek to analyze.

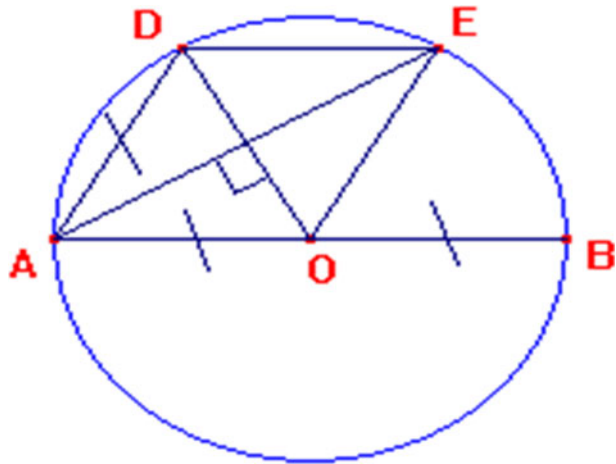
3.2 Duval's construct of "modes of expansion of the discourse"

In order to describe the development of students' cognitive processes in terms of development of their discourse, I characterize, from the linguistic point of view, the moves back and forth between graphical and theoretical aspects that are essential in the resolution of a plane geometry problem. To this end, I consider Duval's theoretical construct concerning different modes of expansion of the discourse (Duval, 1995a). According to Duval, construction of the discourse occurs by different modes of expansion of the discourse. These have a key role in the production of a proof since the proof is a structured whole of propositions that follow predefined rules.

The modes of expansion of the discourse, that is, different modes to link enounced propositions, are *Accumulation* and *Substitution*.

Accumulation mode is characterized by the accumulation of new information. I consider accumulation as the development of discourse made by a juxtaposition of independent propositions concerning information, which, in this case, are of geometrical nature. When the discourse is developed by accumulation, the transition from one proposition to another depends on their content without regard for any rules of order. For example, with respect to Fig. 1, the proposition "AO and AD are equal sides" and the proposition "AE and OD are perpendicular diagonals" are independent propositions until one is the consequence of the other.

By contrast, the *substitution* mode is characterized as a logical progression of the discourse, and it works by inferences. I consider substitution as the discourse progression made by means of a *non-modifiable order of propositions*: The conclusion of a deduction step takes the place of the premise in the following step. For example: "If AO and AD are equal sides, then DAO is an isosceles triangle"; "if DAO is an isosceles triangle, then height

Fig. 1 Drawing of reference

AH is the median of the triangle”; “If the height is the median, then it meets the base at its midpoint”; etc.

When the progression of the discourse comes about by substitution, the passage from one proposition to another does not depend exclusively on the content of the propositions but also on their respective status.

The status of a proposition corresponds to the role that it assumes with respect to another proposition within a global organization of the discourse. The role may be that of hypothesis, of data, of premise, of intermediate conclusion, or final conclusion. The status of some propositions is previously fixed, depending both on the theoretical frame of reference and on the hypotheses that comprise the propositions expressed at the beginning of the discourse. For other propositions, the status is determined by the place where they appear in the discourse. Moreover, Duval (1995a) states that the progression of the discourse also depends on the *sense of the enunciated propositions*. It depends on both the content (which is probable, likely, etc.) and on their epistemic, semantic, and logic value.

According to Duval (1995a), the *epistemic value*² is the degree of certitude or conviction associated with a proposition: The content of a proposition appears immediately evident, certain or only probable, plausible, simply possible or impossible, or absurd, etc. The epistemic value of a proposition is strictly linked to the interlocutors' knowledge system and also to the socio-cultural environment they belong to. The *logic value*³ of the propositions is true or false. Unlike the epistemic value, the logic value of a proposition does not depend only on the understanding of its content but is obtained by specific processes of verification or, also, of proof. Moreover, the logic value of a proposition depends on the status of the proposition inside the sentence: The propositions can be premises, hypotheses, or conclusions.

Summarizing, the epistemic value of a proposition depends on comprehension of its content; in contrast, the logic value of a proposition does not depend on comprehension of the content.

² “le degré de fiabilité que possède ce qui est énoncé dans la proposition : le contenu d’une proposition apparaît évident ou certain ou seulement vraisemblable, ou plausible, ou simplement possible, ou impossible ou encore absurde...” (Duval 1995a, b p. 21)

³ “est le fait que la proposition énoncée est soit vraie soit fausse. A la différence de la valeur épistémique, la valeur logique d’une proposition ne dépend pas de la seule compréhension de son contenu mais elle résulte de procédures spécifiques de vérification ou preuve” (Duval 1995a, b p. 220).

Thus, I consider the progression of propositions as accumulation if they assume mostly an epistemic–semantic value. Instead, if propositions belong to substitution, they assume mostly a logic value. Thus, the progression of the resolution process undertaken when performing a geometrical proof also depends on the change in the values of the propositions: from an epistemic semantic value (probable, certain, etc.) linked to the content, which falls within the accumulation mode, to the logic value (true, false) linked to the status of the propositions, which falls within the substitution mode.

The modalities for identifying this progression are described in the following section, exploiting Bronckart's (1985) idea of linguistic units.

3.3 Use of linguistic units based on Bronckart's theory

In order to recognize the progression of the discursive modes in students' verbalization during the resolution process, I consider the idea of the psycholinguist Bronckart (1985). This uses the occurrences of some linguistic units in the text to define a certain kind of discourse concerning that text. According to Bronckart, linguistic units are pronouns, connectors such as "if...then", verb tenses, auxiliary verbs (such as "aller", in analysis of French texts, or the auxiliary verbs "want", "can", "must"), the adverbs, etc. (Bronckart, 1985, p. 157).

Bronckart provides a text analysis grid devoted to characterizing different kinds of discourse on the basis of presence and frequency of linguistic units (*discourse, scientific discourse, narrative...*).

Since my aim is not to distinguish different kinds of discourse, I do not postulate a biunique correspondence between the presence and frequency of certain linguistic units and the discourse concerning the proving process in plane geometry. Rather, I focus on Bronckart's idea that "les caractéristiques des conditions de production devraient permettre de prévoir les caractéristiques morphosyntaxiques du text" [the characteristics of the conditions production should allow to predict the morphosyntactic characteristics of the text] (discourse) (p. 67).

Bearing this in mind, I have identified some of the main linguistic units characterizing discourse produced during the resolution process of a geometry problem. I have subsequently characterized their *different uses* when they appear in accumulation mode or in substitution mode. In this way, both the occurrences and the different uses of the linguistic units allow us to identify the discourse in terms of progression from accumulation mode to substitution mode.

To Bronckart's list of linguistic units for defining a scientific discourse (pp. 74, 75), I add new linguistic units characterizing the discourse developed during the resolution process of a plane geometry problem. A description of the experimentation work and the methodology adopted is provided in the following section.

4 Methodology

This research is part of doctoral research developed both in Italy and France (co-tutored thesis). For this reason, the experimentation has been performed both in Italian and French schools. Students are asked to solve a geometrical problem; this means that they are not involved in a real learning activity. Our aim is to analyze their verbal production in addition to their written protocols.

To foster communication, the students worked in pairs, both to solve the problem and to write the solution. Actually, the students were asked to produce only one written solution. This meant that they had sufficient time for a deep verbalization activity. The resolution processes of each pair of students were recorded so that the verbal reports could be analyzed in addition to the written protocols.

4.1 Problem assigned

The text of the geometry problem that the students were set during the experimentation is given below. The drawing referring to this problem is in Fig. 1. Note that the drawing was not provided to the students.

C is a circle with center O and diameter AB. D is a point on the circle so that $AD=AO$.

The perpendicular to OD passes through point A and intersects the circle at point E.

Prove that OADE is a rhombus.

4.2 Linguistic units and their uses in accumulation and substitution mode

In this section, I report a table of linguistic units associated with the accumulation and substitution mode, specifying particular uses for each.

To compose this table, I added new linguistic units to Bronckart's list of linguistic units that characterize a scientific discourse. These new units appear in the discourse during the proving process of a plane geometry problem. Then, as a result of experimentation, this table was expanded and completed. Here, I present the expanded table.

Linguistic units	Uses linguistic units in accumulation mode	Uses linguistic units in substitution mode
« if...then », «since...then»	<p>“so, thus”, “so necessarily”, “since” etc.</p> <p>Use conversational</p> <p>They introduce relations among the information of a list as juxtaposition of propositions. They give a linear order to this information. They can be obtained by a deductive step (inference) or inside the domain of interpretation of the drawing. For example, by interpreting the drawing of the parallelogram we can obtain the implication: “if the sides are equal, then they are parallel”. This implication is not the result of the application of the theorem, but is simply the result of a visual observation about the drawing. In this case, the linguistic units “since, thus, if... then...” can be coupled with verbs referring to the perception such as “to see, to look at...” The</p>	<p>And the other synonyms used in the same deductive structure. These highlight a deductive step.</p>

	<p>linguistic unit “if...then...” and also “since, thus...” are usually associated with one of two modes of use, conversation kind or deductive kind, which are not substitutions. Interrogative sentences introduced by “why, how...” such as “Why do you say that...?” or “How do you know this...?” are addressed to an interlocutor (and not to the referent, as in the case of the interrogative sentences in substitution mode).</p>	
<p>Conjunctions such as “and, or, ...”</p>	<p>They link propositions of the same nature (coordinating conjunction) or link subordinated and principal propositions (subordinate conjunctions).</p>	<p>And other such as “because, since,...” used to introduce a justification to the deductive step. In particular, these can introduce the theorem used to legitimize the step.</p>
<p>Temporal adverbs such as “before”, “after”, “then”</p>	<p>Conjunctions allow us to add information in a linear way.</p> <p>They assign an order to the information (usually the order inside the discourse). In other words, temporal adverbs make it possible to juxtapose propositions: “we have... and then... before that, it would be better to look at whether...”</p>	<p>Express the order of the operations of substitution</p>
<p>Terms such as “this, this one, that...”</p>	<p>Temporal adverbs (still, as then, after...) in accumulation mode define an information sequence that forms a list.</p> <p>These terms, often expressed with gestures, are used in <i>deictic</i> replacing of nouns or geometrical properties. They appear in the <i>perceptual apprehension</i> and <i>opératoire</i> of the drawing. These terms may also appear in substitution mode but, in that case, they appear for a different use.</p>	<p>These terms are used in <i>anaphora</i> to “condense” many data. They are used in deductive steps where the premises are already explicated and there are, for example, too many numbers to be explicated without losing the thread of the argument. For example, “if this is true, then...”</p>
<p>Mathematical terms, such as the name of some polygons (parallelogram, triangle, etc.) and some elements of the figure (that is gestalts of smaller dimension compared with the figure, for example, the height, median, perpendicular bisector of a side).</p>	<p>These terms usually denote the object of the problem or the drawing, not the theoretical object.</p>	<p>These terms are linked to theoretical objects and not to the objects of the problem.</p>
<p>Verbs related to or involving perception, such as “to look at, to see” or “to move, to shift...” or “draw, trace...” (for example, “shift the drawing down the page).</p>	<p>These verbs involve the perceptual rather than the conceptual plane and can be observed in perceptive and operative transformation of the drawing. They focus attention on drawing.</p>	

<p>Modals such as “have to”, “must...” frequently coupled with verbs such as “to prove, to demonstrate, to say...”</p>	<p>They are present for contract reasons, instead of theoretical needs (for example, the modal does not introduce a deductive step). The specific use of the modal related to the accumulation mode can be identified by “what is around” the modal itself inside the sentence. For example, in the sentence “we can see that AO is symmetric to DE [sides of the parallelogram OADE], we have to prove that...” the modality is not a theoretical need, since it is not introduced by a theorem of reference, but by perceptive evidence. Thus, it is a need of the contract: it is not legitimate to say, “we can see”, because the didactical contract obliges the student to “prove” the statements.</p>	<p>They are used for theoretical reasons and not for contract reasons. They highlight a chain of deductive steps.</p>
<p>Pronouns “I, we, ...”</p>	<p>The use of “I” in the accumulation mode reveals the presence of the enunciator. In some cases, “we” is added to include other actors-enunciators.</p>	<p>They denote a universal enunciator, an anonymous representative of the scientific community engaged in the social interaction, classmates, the pair of students, the mathematical community ...</p>
<p>Adverbs expressing an affirmative judgment such as “certainly, obviously, necessarily...” or a negative judgment “never...”</p>	<p>These assign, within the accumulation mode, a certain/uncertain value to the statement, which is a semantic epistemic value depending on the content of the sentence.</p>	

A particularly interesting aspect for our analysis regards linguistic units that assume a “pivot role”, since they support the passage from accumulation mode to the substitution mode. Following our idea that the passage between modalities of discourse expansion can be induced by the same linguistic unit but by means of different uses, I look for linguistic units present in students’ verbalizations in both discourse expansion modes. Moreover, if the same linguistic unit is present in the two modes of discourse expansion and in each of them its use is characteristic of that mode of discourse expansion, then, I can reasonably assume that the attendance of the linguistic unit in the accumulation mode may have favored the presence of the same linguistic unit in the substitution mode.

5 Results

I have analyzed students’ verbalizations developed in both Italian and French classes using the previously described linguistic tools. This analysis produced interesting results concerning some regularity in the students’ verbal behaviors, named action models. As I show in the following, these models are linked to the different modes of discourse progression (substitution

and accumulation) and are characterized by the linguistic tools defined in the previous sections. The main action models are named: list model and aim model.

5.1 “List” action model

The “list” action model is characterized by the creation of an “environment of work”. This is done by collection of information. As soon as the list becomes too long, it is processed by organization of collected information.

The following extract gives an example of a list:

T: the crossed diagonals AE and OD meet at their midpoints and make a right angle, and AO is equal to OE, OE is equal to AD.

S: We have DO, it is a radius of the circle too.

The collected information is as follows:

- The crossed diagonals AE and OD meet at their midpoints
- The crossed diagonals AE and OD make a right angle
- AO is equal to OE
- OE is equal to AD
- DO is a radius of the circle

The list is composed of a juxtaposition of information and is identifiable by the juxtaposition of independent propositions.

In general, the criteria that allow us to recognize a list of information include the following:

- The juxtaposition of independent propositions linked by conjunctions such as: “and, moreover, after this, then, ...” or by sentences such as: “we know that... and...”;
- The presence of verbs linked to perception such as “to look, to see, ...”;
- Dominant presence of the simple present verbal ending;
- The presence of temporal adverbs “then, after this...” which assign a temporal order to the information in the list;
- The presence of deictic words or deictic adverbs such as: “here, over there, ...” or “this, this one, that, that one,...” usually coupled with gestures;
- The presence of modals such as: “must ... have to...” in the conditional;
- The presence of non-justified inferences. Usually, these are obtained by the perceptual apprehension of the drawing. These inferences are highlighted by linguistic structures which seem to be of the kind “if...then...”, but are actually engaged connectors such as: “since we know that... then necessarily,...” The following intervention is an example:

T: Since we know, ... since we have the diagonal OD, I draw the continuation of the line OD, we have the parallel, no, the perpendicular line AE, since it's a circle, and since we know that AO, OD and OE are radii of the circle and that OA is equal to AD and that OE is equal to AD too, then necessarily DE is equal too.

Connectors of this kind have the sole role of introducing a new information item in the list; they are not employed for connecting two or more information items in the list. Moreover, the word *necessarily* belongs to the modalities of enunciation of the statement, which assigns an epistemic semantic value linked to the content of the sentence;

- The presence of inferences obtained by a deductive step which have a local action.

We can observe that the described criteria are a subset of linguistic units characterizing the accumulation mode. For this reason, the “list” model of action can be considered as part of the accumulation mode.

As far as the substitution mode is concerned, it is necessary to discriminate between the elements of the list to obtain the necessary and sufficient premises to build a sequence of inferences.

5.2 How can a list be composed and manipulated?

The phases of creation of a list are:

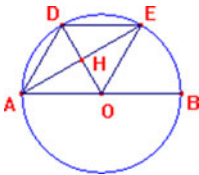
- *Collection of information.* The information can be taken from the data of the statement; they can be obtained by manipulation or interpretation of the drawing (*perceptual apprehension, operative apprehension*), and by inferences. The information can also be obtained by a deductive step not belonging to a sequence of deductive steps.
- *Grouping of information.* This corresponds to repetition of the information in sequence.
- *Manipulation of the List*, which concerns:
 - adding other information;
 - removing information. This makes it possible to remove one or more information items from the list by inference. As a matter of fact, control of the list becomes necessary when it becomes too long to manage. To shorten the list without losing information, a set of information can be replaced with an equivalent one by means of a deductive step;
 - ordering the information. This involves repeating the information items in the list in an order that differs from the one in which they were collected.

In the following, I describe these phases in greater detail.

5.3 Collection of information

In geometry, there is a natural source for obtaining information: the drawing. Students can obtain information by perception of the drawing (*perceptual apprehension*) or by possible modifications of the drawing (*operative apprehension*) and by inferences.

The following extracts show how information is obtained by modification of the drawing (*operative apprehension*) and by its perception (*perceptual apprehension*).



K: Right, AD is equal to AO, AO is radius and EO is a radius of the circle, so necessarily, AD is parallel to this

The information *AD parallel to OE* is obtained by the combined actions of perception (*perceptual apprehension*) and modification of the drawing (*operative apprehension*). The information *AO is radius* and *EO is radius* are obtained by *discursive apprehension* of the drawing on the basis of the data present in the task. The presence of the verb “to be” in the present tense highlights the epistemic value of the propositions whose content appears certain. We observe that the present tense of the verb “to be” in the proposition “so necessarily, *AD is parallel to this*” assigns the same “certain” epistemic value as the previous propositions because it is obtained from the drawing by perceptive evidence.

The combined actions of *perceptual apprehension* *perceptive* and *operative apprehension* of the

<p>S: Since DA, AO and OE are equal, then necessarily DE is equal... it is equal because it shapes (forms) a quadrilateral.</p>	<p>drawing allows S to obtain the information: “DE is equal to the segments DA, OA and OE”.</p>
<p>G: Look!, This is symmetrical to this one (AO and DE) so it is the same.</p>	<p>The presence of “<i>necessarily</i>” gives to the proposition a “certain” semantic epistemic value linked to the content of the phrase.</p>
<p>D: ... since this (DA) is equal to this one (AO), and DE is equal to AO, then DE is equal to OE</p>	<p>The information “AO is symmetrical with regard to the segment DE” is obtained by the perceptual <i>apprehension</i> <i>perceptive</i> and operative <i>apprehension</i> of the drawing and they are highlighted by the use of the perceptive verb “to look”.</p> <p>The perceptual <i>apprehension</i> and operative <i>apprehension</i> of the drawing are highlighted by the use of deictic terms such as <i>this</i> and <i>this one</i>. They refer to a pure designation of the referential function. The linguistic units “since... then” introduce a relation among information in the list.</p>

We note that connectors such as “Since, ... then...”, “so necessarily...”, “it’s necessary that...so necessarily...” refer to the connector “if...then...” without inserting an inference justified by a third statement (as a matter of fact, the inference is obtained by *perceptual apprehension* and *operative apprehension* of the drawing).

Other linguistic units such as “it is sure that...”, “it seems that...”, “I think that...”, “It’s obvious that...”, “It’s probable that...”, “It is possible that...” are present in the list model since the list is associated with the *accumulation* mode of expansion of the discourse. They assign to the sentence an epistemic value. Moreover, since the list is frequently associated with manipulation of the drawing, these linguistic units assign to the sentence a semantic epistemic value but not a status: the students assign to the propositions a “true” (certain) epistemic value but not a “true” logic value since the list cannot be assigned a status of premises or of conclusion to the propositions (except for local inferences). In other words, the subject does not distinguish true logic from the epistemic.

We observe that the modals such as: want, must, have to are used for contract reasons and not for theoretical reasons, which can be identified during verbalization of a theorem (as I describe in the aim model). The information can be obtained both by interpretation of the drawing and inferences. An example of this can be seen in the following interventions:

E: AO is radius, OE is radius, **so** AO is equal to OE since they are radii. [...]

E: In this way, the triangle AHO is a right triangle [...]

E: AO is equal to AD, so the triangle AOD is isosceles. Better, it is equilateral. So the triangle AHO is a particular triangle, a triangle 30, 60 and 90 [degrees]

The information “AO radius” and “OE radius” are obtained by manipulation of the drawing (operative *apprehension* and *discursive*), since they are not in the data of the sentence.

The information “AO=OE” is obtained by inference through the previously collected information. I identify an inference by the presence of “so”.

5.4 Grouping information into a list

This is performed by repetition of the information collected during the process, following the order in which the information was made explicit.

5.5 Manipulation of the list

Sometimes students derive an implication from information in the list but nonetheless maintain the two information items in the list as well as the obtained implication. In this case, list manipulation does not result in reduction of information; on the contrary, the list grows longer and longer. For example, from the information P: “ $AO=AD$ ” belonging to the data of the statement, it is possible to obtain the information Q: “OAD is an isosceles triangle”. Frequently, students maintain the information P and the inferred information Q inside the list. Nevertheless, when a list becomes too long to manage, students can manipulate it, deleting some information. They can delete one or more items by inference, which can lead to a loss of information. For example, the information «OA radius» and «OE radius» can be replaced by the information « $AO=OE$ ».

Reorganization of list information becomes necessary when the information assumes a different role from that assumed at the moment of its specification. The role of the information and its position inside the list depend on the aim of the list itself: simple collection of information or guided search of information. I describe this aspect in greater detail in the following section.

5.6 Aim of the list

The list has two main aims: The first is to start a deductive process (for example in order to recall a theorem useful for guiding the process); the second, which follows verbalization of the statement of the theorem, is simply to verify the premises of the theorem. I adopt the term “research aim” and “verification aim” to distinguish these aims. Clearly, manipulation of the list depends on the aim involved. In order to provide an example of both aims, I present an extract where I highlight the change of aim as regards the objective defined by the student. At the beginning, the student collects information to evocate (recall) a useful theorem. Then the student collects information to verify the hypothesis of the recalled theorem (see “Function de guide”, Robotti 2008).

34.J: There are crossed <u>diagonals that intersect at their midpoints and that are perpendicular</u>	List with research aim
41. K: [...] <u>it's a parallelogram</u> , I think it's a parallelogram	Perpendicular crossed diagonals
42. J: <u>The opposite sides are parallel, and this is a radius (OE)</u> . We know that is a parallelogram, we have to prove that is a rhombus, it's a particular parallelogram. So, to prove that these sides are parallel...	Crossed diagonals that intersect at their midpoints
Well, <u>AD equal to AO</u> , <u>AO is radius and EO is radius of the circle, thus necessarily, AD is parallel to OE</u>	-the quadrilateral is a parallelogram
45. K: If crossed diagonals are perpendicular and they intersect in their midpoint, it's a rhombus.	Opposite and parallel sides
54. J: Or, it's sure that <u>OD is a radius</u> .	OE=radius
55. K: We know that <u>OD is perpendicular to AE</u> because...	AD=AO
60. J: First of all, we know that they intersect in their midpoint ... wait, we have to think a moment before... <u>AO=AD</u>	AO=radius

64. J: So, <u>AE is perpendicular to OD</u> and we try to prove that they intersect in their midpoint...	OE=radius
68. J: Perhaps, there is something with vectors	AD//(OE)
69. K: Yes, but I don't remember, I don't remember properties. I think of a point equal to EA [...] something, ... I don't remember...	<p>The collection of the information is not guided by the statement of the theorem. For this reason, it is a collection of juxtaposed information.</p> <p>List with verification aim</p> <p>Verbalization of the theorem</p> <p>OD=radius</p> <p>OD⊥AE</p> <p>AO=AD</p> <p>AE ⊥OD</p> <p>The search for information is guided by the need to prove the premises of the theorem: "crossed diagonals that intersect in their midpoint". Via this search, the student obtains a "label" that is a junction between a particular configuration (the crossed diagonals) and a particular word (midpoint), which makes it possible to evoke the vector addition theorem (see Robotti, 2008).</p>

I join the research aim list with the accumulation mode and the verification aim list with the substitution mode. Analysis of students' productions allows us to observe that the verification aim list is sometimes obtained by starting from the research aim list and modifying it (as a matter of fact, in the resolution process, a substitution mode follows an accumulation mode). Modifications to the list can involve reordering or deletion of the information inside the list. They are guided by the verbalization of the theorem, as in the following example.

-
2. E: [...] AE is perpendicular to OD, AO=AD. Proving that OADE is a rhombus ... AO is a radius of the circle, OD is radius of the circle, OE is radius of the circle, AD=AO data, thus the rhombus ... is a quadrilateral with four equal sides
7. A: I don't understand what you want to do!
8. E: AO, OE and OD are radii of the circle, and they are equal... yes, OD isn't useful!
9. A: Why isn't it useful?
10. E: Because it is a diagonal... and then AD is equal to AO...
-

The information "OD is a radius" and the information which links OD to the segments AO and OE by an equality relation are deleted from the list by means of the verbalization of the theorem (intervention [2]). The student does not need this information to verify the following theorem hypothesis: The quadrilateral has four equal sides.

It is evident that only information useful for verifying the chosen theorem is kept in the list: This results in students selecting information from the list. I can assume that verbal language has the function of selective memory on (about) the information of the list. As a matter of fact, in geometry tasks, interpretation of the drawing plays a very important role because it makes it possible to obtain information (some spatial properties of the drawing can be interpreted as properties of the theoretical object). For this reason, I consider the drawing as a memory support.

Nevertheless, I observe that the drawing is characterized as a whole memory, a sort of overview of memory (Laborde & Capponi, 1994). By contrast, verbal language is characterized as a selective memory. For this reason, verbal language allows students to select what to add to the list from the information they have obtained from the drawing.

5.7 “Aim” model of action

The “aim” model of action develops from the question of the task in order to trigger the resolution process. The following intervention is an example:

C.: well, for a rhombus we have to say that crossed diagonals intersect in their midpoint and they are perpendicular.

This model is characterized by the search for the “cause” starting from the “effect”: When the cause is defined, the transformation of an “aim” structure of the discourse into a “conditions–consequence” structure of the discourse can be performed. In the “aim” model, the action is based on the search for the necessary and sufficient conditions in order to obtain the effect, in this particular case, the rhombus. For this reason, students consider the definition of the rhombus or the theorems concerning the rhombus that allow them to define the necessary and sufficient theoretical elements (cause) for characterizing the rhombus.

In general, the operative apprehension and the interpretation of the drawing allow students to collect these conditions. Obviously, the operative apprehension and the interpretation of the drawing are guided by verbalization of the theorem statement represented by the “aim” structure.

For example, the question of the task “prove that the quadrilateral is a rhombus” can be transformed into an “aim” structure: “because the quadrilateral is a rhombus (effect), the crossed diagonals must be perpendicular and they must intersect in their midpoint (cause)”. The geometrical relations needed for the quadrilateral to be a rhombus are perpendicular crossed diagonals that intersect in their midpoint. Starting from this, the operative apprehension and the interpretation of the drawing are used to identify (i.e., to recognize in the drawing) two segments that are crossed diagonals of a quadrilateral that are linked by a perpendicular relation and that intersect in their midpoint.

In this way, the resolution process starts from the theoretical reference and then focuses on manipulation of the drawing (operative apprehension). Among the criteria that make it possible to recognize an “aim” model of action is the presence of the modals “have to”, “must” linked to verbs such as “to say, to prove, to demonstrate...” I observe that these linguistic units characterize the “substitution” mode if they are used for theoretical needs, as I have previously described.

6 Conclusions

In this research study, I have analyzed the language produced by pairs of students while solving plain geometry problems, with particular attention to situations when the interaction between figural and conceptual aspects is brought into play. The aim of the analysis has been to show how students’ natural language can be used as a means to reveal evolution in students’ cognitive processes, especially in proving processes, as explored here.

To analyze students’ verbalization during the resolution process of a plane geometry problem, I have identified some significant relationships between figural and conceptual

aspects, characterizing each of these in relation to students' apprehension of the drawing. Specifically, I refer to Duval's notion of apprehension of the drawing (Duval, 1995a). Duval defines three kinds of drawing apprehension: perceptual apprehension centered on perception, which allows students to immediately perceive the shape; operative apprehension centered on the possible modifications to the drawing that the student can make; and discursive apprehension corresponding to explicit geometrical properties other than that provided by the figure or data. These different kinds of apprehensions of the drawing, together with "moves" to and from figural and theoretical aspects, can be put into relation with the two different means of discursive progression that Duval defined as "accumulation" and "substitution". Accumulation is defined as a juxtaposition of independent propositions linked to each other only by their content. Substitution is characterized as logic progression in discourse that allows the passage between propositions to be made by inferences; it does not depend exclusively on their content but also on their status in the phrase (premises, conclusion, etc.). The modes of discourse progression are also identified by the value of propositions, epistemic semantic value being linked to accumulation mode and logic value being linked to the substitution mode. This allows us to characterize different modalities for expressing the structure of reasoning. Moreover, drawing on the idea of the psycholinguist Bronckart (1985), I have characterized the passages to and from theoretical and graphical referents, contextualizing them in the evolution of the discursive structure (passages from accumulation to substitution) by means of the recurrences of certain linguistic units and, in particular, of their different use.

This model for analysing the verbalization of students working in pairs has allowed us to identify some regularities in their verbal behavior. These can be associated with the various modes of progression in discourse connected to demonstration. The regularities have been called "action models", and they have been identified: the list model and the aim model. I have shown that each of these is characterized by a mode of discourse progression: The list model is related to accumulation mode, while the aim and hypothetical–deductive models are related to substitution mode. So, it has been possible to verify that progression in problem resolution comes about via progression in the modes of discourse expansion; students pass from a simpler mode of discourse progression like accumulation to a more complex one like substitution. This evolution comes about through changes in the value of the proposition: from a content-related value to one related to the proposition status. Importantly, in this evolution, a referential change occurs to the enounced propositions: The list action model is primarily linked to manipulation of the drawing, while the aim action model is mainly linked to the theoretical referent.

On the basis of the arguments made here, I believe that the reported research has yielded significant results in terms of research methodology. As a matter of fact, the theoretical tools developed for the analysis have proved useful for analysing students' verbalization processes. They help to reveal evolution in the cognitive processes involved when students seek to solve plain geometry problems, expressed through evolution in discourse.

This continuing research work will move on to consider whether the results obtained in the geometry field can be extended to algebra. The specifics of algebra will be examined, and light will be shed on the affinities and contrasts with the work performed in geometry. Furthermore, our work will move on to consider whether the results obtained may have didactic implications: I can assume that the action models can also be used to guide and shape the verbalization processes underway. By supporting certain verbal behaviors in class activity (e.g., during mathematical discussion, Bartolini-Bussi, 1998a, b), the teacher could support the passage from accumulation to substitution in discourse progression. In this sense, the models would be tools that the teacher could use in the classroom and could

therefore be considered teaching tools. The possibility of orienting discussion towards specific action models as need arises means that teachers have an effective tool for operating in the zone of proximal development (Vygotsky, 1938, *Pensée et langage*, chapter 6).

References

- Bartolini-Bussi, M. (1998a). Verbal interaction in mathematics classrooms: A Vygotskian analysis. In H. Steinbring, M. G. Bartolini-Bussi, & A. Sierpinska (Eds.), *Language and communication in the mathematics classroom* (pp. 65–84). Reston, VA: NCTM.
- Bartolini-Bussi, M. (1998b). Joint activity in the mathematics classroom: A Vygotskian analysis. In F. Seeger, J. Voigt, & U. Waschesho (Eds.), *The culture of the mathematics classroom* (pp. 13–49). U. P: Cambridge.
- Bronckart J.P. (1985). *Le fonctionnement des discours*. [The operation of the speeches] Paris: Delachaux et Niestlé.
- Chomsky, N. (1975). *Reflections on language*. New York: Pantheon.
- Duval R. (1995a). *Sémiosis et pensée humaine*. [Semiosis and human thought] Bern: Peter Lang.
- Duval, R. (1995b). Geometrical pictures: Kinds of representation and specific processings. In R. Sutherland & J. Mason (Eds.), *Exploiting mental imagery with computers in mathematics education*. Berlin: Springer.
- Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, 87(5), 215–252.
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.
- Laborde, C. (1982). *Langue naturelle et écriture symbolique*. [Natural language and symbolic writing] Thèse de Doctorat. Grenoble: Université Scientifique et Médicale Institut National Polytechnique de Grenoble
- Laborde, C. (1999). Dynamic geometry software as a window on mathematical learning: Empirical research on the use of Cabri-geometry. In A. Gagatsis & G. Makrides (Eds.), *Proceedings of the Second Mediterranean Conference on Mathematics Education*. Nicosia, Cyprus: Cyprus Mathematical Society, 1 (pp.142-160).
- Laborde, C., & Capponi, B. (1994). Cabri-géomètre constituant d'un milieu pour l'apprentissage de la notion de figure géométrique. [Cabri-geometry is an environment for learning the concept of a geometric figure]. *Recherche en Didactiques des Mathématiques*, 14(12), 165–210.
- Luria, A. R. (1975). *Linguaggio e sviluppo dei processi mentali nel bambino* [Speech and the development of mental processes in the child]. Firenze: Giunti-Barbera.
- Mesquita A. (1989a). Sur une situation d'éveil à la déduction en géométrie. [A situation of introduction to the deduction in geometry] *Educational Studies in Mathematics*, 20, 95–109.
- Mesquita A. (1989b). *L'influence des aspects figuratifs dans l'argumentation des élèves en géométrie : Eléments pour une typologie*. [The influence of figurative aspects in the pupils' argumentation in geometry: Elements for a typology] Thèse de l'Université Luis Pasteur, Strasbourg, Publication de l'Institut de Recherche Mathématique Avancée.
- Robotti, E. (2008). Les rôles du langage dans la recherche d'une démonstration en géométrie plane [Role of language in the research of plane geometry proof]. *Recherche en Didactique des Mathématiques*, 28(2), 183–218.
- Sfard, A. (2001a). Learning mathematics as developing a discourse. In: R. Speiser, C. Maher, C. Walter (Eds.), *Proceedings of the 21st Conference of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 23–44). Columbus, OH: Clearing House for Science, Mathematics, and Environmental Education.
- Sfard, A. (2001b). 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.
- Sfard, A. (2008). *Thinking as communicating: Human development, the growth of discourses, and mathematizing*. Cambridge, UK: Cambridge University Press.
- Sfard, A., & Kieran, C. (2001). Cognition as communication: Rethinking learning-by-talking through multifaceted analysis of students' mathematical interactions. *Mind, Culture, and Activity*, 8(1), 42–76.
- Vygotsky, L. (1938) *Pensée et langage*. [Thought and language] (1997) Paris: La Dispute.