THE INTERPLAY BETWEEN SYNTACTIC AND SEMANTIC KNOWLEDGE IN PROOF PRODUCTION: MATHEMATICIANS' PERSPECTIVES

Paola Iannone and Elena Nardi
University of East Anglia
Norwich, UK

We draw on a series of themed Focus Group interviews with mathematicians from six universities in the UK (and in which pre-distributed samples of mathematical problems, typical written student responses, observation protocols, interview transcripts and outlines of bibliography were used to trigger an exploration of pedagogical issues) in order to discuss the interplay between syntactic and semantic knowledge in proof production (Weber & Alcock, 2004). In particular we focus on participants' views of how fluency in syntactic knowledge can be seen as a facilitator of mathematical communication and a sine-qua-non of students' enculturation into the sociocultural practices of university mathematics.

Key words: undergraduate mathematics education, mathematicians, syntactic knowledge, semantic knowledge, proof, enculturation, communication, sociocultural practices

INTRODUCTION

In 2004 Weber and Alcock proposed a theoretical framework for understanding the process through which undergraduate students (and mathematicians) engage with proof. Refining and clarifying what is meant by 'formal' and 'intuitive' reasoning (Weber and Alcock, 2004, p210) the authors suggested that proof production can be of two different kinds: *syntactic proof production* and *semantic proof production.* They define *syntactic proof production* as

> one which is written solely by manipulating correctly stated definitions and other relevant facts in a logically permissible way. [...] A syntactic proof production can be colloquially defined as a proof in which all one does is 'unwrap the definitions' and 'push symbols'. (p210)

and as *semantic proof production*

> to be a proof of a statement in which the prover uses instantiation(s) of the mathematical object(s) to which the statement applies to suggest and guide the formal inferences that he or she draws. (p210)

In this context *syntactic knowledge* and *semantic knowledge* are the abilities and knowledge required to produce syntactic or semantic proofs (p229). The studies from which this theoretical framework emerged are empirical, data-grounded studies and involved observation of undergraduate students, doctoral students and mathematicians as they worked on proving various mathematical statements (typically in Group Theory or Analysis). The participants were asked to 'talk aloud' while writing their proofs and were in some cases interviewed during this process. Amongst the conclusions the authors draw from their studies, is that

> The abilities and knowledge required to produce syntactic proofs about a concept appear to be relatively modest. The prover would need to be able to recite the definition of a mathematical concept as well as recall important facts and theorems concerning that concept. The prover would also need to be able to derive valid inferences from the concept's definition and associated facts. (p229)

while the knowledge required to produce semantic proofs appears to be more complex (p229). The authors conclude that

> Hence, writing a proof by syntactic means alone can be a formidable task. However, when writing a proof semantically, one can use instantiations of relevant objects to guide the formal inferences that one draws, just as one could use a map to suggest the directions that they should prescribe. Semantic proof production is therefore likely to lead to correct proofs much more efficiently. (p232)

In this paper we wish to investigate how syntactic and semantic knowledge concur in proof production. The data we draw from illustrate the perspectives of mathematicians as they reflected on proofs produced by their students (as part of written coursework). In what follows we briefly introduce the study they originated in.

THE STUDY

The data we present originate from a study[1] which engaged mathematicians from across the UK as educational co-researchers; in particular, the study engaged university lecturers[2] of mathematics (more details on the participants to the study can be found in Iannone & Nardi, 2005) in a series of Focused Group Interviews (Wilson, 1997), each focusing on a theme regarding the teaching and learning of mathematics at university level that the literature and our previous work acknowledge as seminal. These themes were:

- Formal Mathematical Reasoning I: students' perceptions of *proof* and its necessity;

---

[1] Supported by the Learning and Teaching Support Network in the UK.

[2] In the text we refer to the participants of the study as Lecturers. Meanings of this term differ across different countries. We use it here to denote somebody who is a member of staff in a mathematics department involved in both teaching and research.

- Mathematical objects I: the concept of *limits* across mathematical contexts;
- Mediating mathematical meaning: *symbols and graphs*;
- Mathematical objects II: the concept of *function* across mathematical topics;
- Formal mathematical reasoning II: students' enactment of *proving techniques*;
- A *Meta-cycle*: collaborative generation of research findings in mathematics education.

Discussion of the theme in each interview was initiated by a Dataset that consisted of:

- a short literature review and bibliography;
- samples of student data (e.g.: students' written work, interview transcripts, observation protocols) collected in the course of our previous studies; and,
- a short list of issues to consider. We note that, despite the presence of this list, we gave priority to eliciting participants' own perspectives and kept a minimal role in manipulating the direction the discussions took (Madriz, 2000).

Analysis of the interview transcripts largely followed Data Grounded Theory techniques (Glaser and Strauss, 1967) and resulted in thematically arranged sets of Episodes – see elsewhere (e.g. Iannone & Nardi, 2005) for more details.

The data we present here originate in Episodes from the discussion of the theme *Formal Mathematical Reasoning I: students' perceptions of proof and its necessity.* In these, students' responses to a Year 1 – Semester 1 question that concerned the convergence or divergence of sequences and required the use of the quantified definition of convergence :

> *The sequence* $\{a_n\}_{n\in N}$ *of real numbers converges to a real number* L *as* $n \to\infty$ *if* $\forall\varepsilon>0$, $\exists N$ *in* N *such that* $n\geq N \Rightarrow |a_n - L|<\varepsilon$

triggered a discussion of what type of knowledge students draw on when engaged with proving the convergence of a sequence.

The students had encountered this definition half way through their first semester. In the Episodes we sample the discussion had revolved around two main issues: why it is necessary to teach and use the quantified statement for the convergence of a sequence; and, how formal and informal understanding of the definition of convergence interact for the production of a correct proof. Issues relating to what difficulties students encounter in internalising and manipulating this statement were also touched upon.

Here after discussing a section of the data, we frame our conclusions in Weber and Alcock's terms and attempt to explore how the two types of knowledge they distinguish (semantic and syntactic) coexist in proof production. In particular we focus on some aspects of the role of syntactic knowledge.

WHY DO WE NEED (AS MATHEMATICIANS AND AS TEACHERS) THE FORMAL QUANTIFIED STATEMENT FOR THE CONVERGENCE OF A SEQUENCE?

The participants agreed on the necessity for the students to learn and understand the quantified statement. Moreover, they all recognised the need for the students to learn how to manipulate the quantifiers correctly and how to write meaningful mathematics sentences (i.e. sentences that comply with formal logical reasoning) using them. Various reasons were offered for this. Below we elaborate three: symbolic language as a compensation for shortage of pictorial/geometrical representation (*loco*-visual); symbolic language as the shared medium of communication amongst mathematicians (communicational); and, symbolic language as a tool for manipulating the logic of mathematical arguments (instrumental).

Lecturer E suggested that quantified statements help defining concepts that are not particularly amenable to pictorial/geometrical representation:

> E: You see ... no human can have a … good intuitive geometrical or pictorial view of what the statement "the sequence does not converge" means, for example. […] Or say certainly no one can have a geometrical view of the statement "this function is not uniformly continuous".

Therefore, the symbolic language of quantifiers fulfils the need to express those concepts with as little ambiguity as possible and to compensate for the limited feasibility of pictorial or geometric imagery.

To this reason, Lecturer A adds that symbolic language is the shared language of mathematics:

> A: There is a consensus on what things mean because they are used in context in the lectures, in the seminars. […] I mean, meaning is attached to it […] otherwise it will be almost impossible to ever write an example sheet again because too much has to go into it.

So, in order that the students begin to partake in the discourse of university mathematics as used by the mathematics community and in order for them to be able to communicate their mathematics to other mathematicians, they need to acquire dexterity in using mathematical language. Furthermore the participants were keen to recognise this as a characteristic mainly of pure mathematics: it is in this discipline mainly that symbols acquire progressively more layers of shared meaning:

> E: I think this is a wide problem particularly in pure maths. The process of pure maths as it proceeds, invests more and more meaning in purer and purer symbols. […] And, you know, you can literally write down a one-line statement that would take ten years to explain.

The third main role for symbolic language is more instrumental: symbolic language is used as a tool for writing proofs and for manipulating formal statements. Lecturer E offers the example of writing the negation of the definition of convergence (i.e. the statement 'the sequence does not converge'):

E: Yes, I guess... the power of symbols is something that they [*the students*] learn to manipulate because... to encourage them [*towards*] more algorithmic things. So for example the negation of a quantified statement I think is much easier as a symbolic definition. Because it is an algorithm: you replace 'for all' with 'there exists', you replace 'there exists' with 'for all' and then you get a statement and that is the algorithm.

This is more easily done by using the algorithmic negation, (or *negation by rules*, Dubinsky et. al 1988) using directly De Morgan type laws.

This final suggestion points to the importance, in the participating mathematicians' perception, of possessing syntactic knowledge. As in Weber and Alcock (2004), this is not meant to be merely procedural dexterity with manipulating strings of symbols, but also ability to construct sentences that follow the laws of formal logic and to unpack and use a definition. In the words of Lecturer E:

E: I mean I try in Analysis to convey the idea, if you like, that by definition the last line of the proof is the definition of something. And that view is not... that approach is not viewed across their homework. [...] And that itself is a difficult idea and is one of the things that first year Analysis should be teaching them. And on the other hand, yes, I mean it is not... of course that is not how we do mathematics and that is not how they have done mathematics so... it is difficult.

The emphasis that Lecturer E places on the need to acquire syntactic knowledge – and to learn how to present proof in ways that are acceptable within the mathematics community – is revealing. While advocating this need he also acknowledges that this very ordered and established way of presenting mathematics does not represent the process by which mathematicians do mathematics. He thus emphasises the role of syntactic knowledge as a tool of mathematical communication (see also Hanna, 2000, p8) and how syntactic knowledge in proof production serves this role.

IS POSSESSING SYNTACTIC KNOWLEDGE ENOUGH? THE ROLE OF SEMANTIC KNOWLEDGE

The mathematicians in our study constantly teamed up syntactic fluency (e.g. with using quantifiers) with what they often referred to as 'construction of meaning':

A: This is the definition and that is the meaning, and the meaning I construct is equivalent to the definition.

But how does this 'construction of meaning' interact with the (formal) definition and how does this meaning come into being? There is agreement amongst them that this meaning should be constructed with the help of mental images and verbal explanations of the definition, and that those are fundamental parts of being able to work with such statements. Lecturer A refers to his experience as a learner and as a teacher of mathematics:

A: I find it very difficult to work with statements which have quantifiers [...] So the only way for myself in which I can unravel such things is that I have to build up

> a mental picture by which I know, ok, … this is what is going on. […] So when it comes to convergence I think that the primary notion for the students is asking that no matter what I specify the ε region about the L, from a certain point onward everything fits inside this box. […] Unless that is the direct connection between images that you have and formalisation I think you are lost. If you just are juggling around ε and δ then it is a completely unworthy process. […] When it comes to convergence [*this*] is something that is very private: some people work like this and some don't. And I can well imagine that there are students that can work along a string of quantifiers they just do what they are told. You can view this as the recipe, you can do this, you do this and you do this…

So, at least for Lecturer A, fluency in syntactic knowledge must come hand in hand with engagement with the meaning behind the symbols in question, namely an analogous fluency in semantic knowledge. Furthermore he acknowledges the highly personal nature of this enterprise and offers the example of how he himself deals with strings of symbols and with formal logic reasoning: others, he says, may engage in this process quite differently but for him a simultaneous syntactic and semantic engagement is absolutely central.

In other words, for example those of Tall & Vinner's (1981) Concept Image and Concept Definition, from these mathematicians' perspective (and particularly from Lecturer A's quote at the beginning of the paragraph) it appears that the conflict between "meaning" and "definition" (lack of "equivalence") is a crucial source of difficulty in proof production (and mathematical understanding more broadly) for students – as well as for professional mathematicians. Moreover, again framing the above quotes in Tall & Vinner's terms, the interplay between Concept Image and Concept Definition is a highly personal affair, depending on previous mathematical experiences but also, in the case of professional mathematicians, on their specific field of expertise.

INTERPLAY BETWEEN SYNTACTIC AND SEMANTIC KNOWLEDGE IN PROOF PRODUCTION

The participants reflect on the existence of formal definition of a concept and its informal understanding as follows: in the process of 'creating meaning for a concept', as one of them calls it, the need for drawing upon both semantic and syntactic knowledge, often simultaneously, emerges. When only one type of knowledge is used, results are often unsatisfactory. After having discussed students' homework from the first weeks of the Analysis course, where the students were asked to apply the formal definition of convergence of a sequence to find out if the given sequence converged or not, Lecturer D remarks:

D: Again… for example that student of mine who said, you know, why… why does it [*applying the formal definition*] prove convergence? The impression I get is that she would end doing it all, all the side calculations and everything, but she was approaching it because she knew this is what you

> are supposed to do to prove convergence, but she didn't really understand why she was doing it, I think.

So, being able to handle and apply, even correctly, the 'formal machinery' (as Lecturer A calls syntactic knowledge) of the convergence of a sequence is not enough to claim understanding of it. From the above it appears that this student has acquired a 'formally operable' definition of convergence of a sequence (Bills and Tall, 1998) in that she is able to apply it correctly to a given situation. But she has not yet created the meaning which grants deeper understanding of the concept of convergence – or more precisely in what ways this string of symbols that her lecturer refers to as the definition of convergence relates to her perception of what convergence is – and which will enable her to apply the same definition in other contexts. With regard to this point Lecturer A responds to Lecturer D as follows:

> A: It is the situation between the formal and the informal, I think. I mean… unless the student reaches ever the informal concept I think … to my mind it should be first very deeply ingrained in the student. And then it should be a justification in order to make sure that this is really doing what it ought to do this formal machinery. And they need to be able to jump from one to the other concept … and I think this is also how they find the N [*in the definition of convergence see above*]. Because, how do you find the N, how do you pluck it out of the air? You have to have some informal reasoning, some intuition, draw some pictures, do some side calculations and then you say oh, maybe given this ε the N maybe this.

Just learning the formal machinery is of course devoid of meaning. However also relying exclusively on this 'intuition' and 'pictures' is not enough: in fact it can be misleading What fluency with syntactic knowledge offers here is a shield against such misguidance, a tool of closer scrutiny through which one can establish that conclusions can actually be inferred formally and logically:

> *Interviewer*: Can I just ask a question, I mean, it is very close to […] what you were wondering about. They [*the students*] say, to me is pretty obvious that one over n, you know, the larger the n is the smaller the one over n is… so it goes to zero. And why do I need to…

> A: Why to bother. In fact at that moment you should say ok, this is your informal understanding, you are expecting something that is correct. But then maybe you want to say, well, there is an exotic example … informally you also draw this conclusion however you are wrong. And so why is that? So that justifies the formal apparatus to sort out what is right and what is wrong.

Therefore, syntactic and semantic knowledge have both to interact while the learner is engaging with proof production (whether the learner is intended here as a student or as a mathematician producing new results in mathematics). Drawing on only one of the two jeopardises both construction of meaning and successful proof production.

DISCUSSION

In sum participants expressed the view that syntactic knowledge

- helps defining and clarifying concepts that escape pictorial representation without ambiguity
- is the "shared language" of mathematics and as such it acquires a socio-cultural dimension: in order for students to enter the mathematical community and communicate their mathematical findings to others they need to become fluent in its language
- can be an effective tool for proof production
- acts as a checking device for intuition and for semantic knowledge.

While semantic knowledge

- guides syntactic knowledge in proof production and it is of great importance when there are parts of proofs that require an act of choice on the part of the prover (in the example of convergence of a sequence, semantic knowledge guides the choice of N in the definition)
- grants deeper understanding of the mathematical concepts considered
- grants flexibility in applying known concepts to new situations.

In addition to the above what also emerges from the data we presented is that resorting to one type of knowledge alone in proof production is limiting, even potentially misguiding and ineffective. In fact, our data seem to point to a cyclic process based on drawing on syntactic and semantic knowledge in turn and often simultaneously. Syntactic knowledge is needed both to guarantee unambiguous use of the definition and as a tool that helps manipulate and produce a formal argument. In turn, semantic knowledge is needed to guide the syntactic proof production by drawing on insight into the main properties of the mathematical objects involved.

Semantic knowledge is of great importance, for example, when an act of creativity or choice is involved in proof production. The example the mathematicians referred to is how to find an N for a given epsilon when trying to prove the convergence of a sequence by referring to the formal definition. It may be the case that finding such N can be done in some cases through applying algorithmic procedures (e.g. solving backwards $|a_n - L|<\varepsilon$ given a particular epsilon). However as the mathematicians above reported students were often puzzled about where N came from; we believe that what they were actually reporting there, what is underlying the students' puzzlement is a certain degree of breakdown between syntactic and semantic knowledge.

Furthermore when informal understandings seem to lead to inaccurate deductions syntactic knowledge can re-direct these understandings and shed light on aspects of the argument not necessarily accessible through intuition. Finally, if we consider the need to produce proofs in the format that can be shared amongst mathematicians, syntactic knowledge functions as a communication tool that serves exactly this purpose.

CONCLUSIONS

In this paper we have discussed how the mathematicians in our study articulated the roles of syntactic and semantic knowledge in proof production, and how they consider their students' acquisition and use both types of knowledge a priority. While discussing the interplay between semantic and syntactic knowledge it appears that, based on their experience as both teachers and learners of mathematics, the participants believe that both types of knowledge need to concur to produce successful proofs and that resorting to only one type of knowledge is not enough. From the data presented above it also emerges that syntactic knowledge has a role within the mathematical community as a tool of communication. In other words, it represents the *genre speech* of this community. Here we use the term *genre speech* in the sense of Bakhtin (1986) and as explored further by Van Oers (2002):

> The genre is primarily a social tool of a sign community for organising a discourse in advance and often even unwittingly. It is a style of speaking embodied in a community's cultural inheritance, which is passed to members of that community in the same way as grammar is passed on. (p69)

Therefore syntactic knowledge contributes to mathematics as a social activity by becoming its genre speech, the common language that everyone in the community understands and uses for exchanging ideas and results. We concur with Otte (1990) who emphasises this social role of syntactic knowledge when he writes about proof presentation (referring to how proofs appear in mathematical publications, largely as chains of symbols that convey the logical deductions underlying formal mathematical reasoning):

> It is in this way that proofs are both mechanical procedures and social processes. ...although intuition is commonly worshipped in contrast to proof as the highest form of knowing, this attitude is in danger of depriving man of his social nature and thereby of his character as a human subject. (p62)

REFERENCES

Bakhtin, M. M.: 1986, *Genre Speeches and Other Late Essays*, University of Texas Press.

Bills, L. and Tall, D.: 1998, Operable Definitions in Advanced Mathematics: the Case of the Least Upper Bound, *Proceedings of the 22nd International Conference, Psychology of Mathematics Education,* Vol. 2, Stellenbosch, South Africa, 104-111.

Dubinsky E., Elterman F. and Gong C.: 1988, The Students' Construction of Quantification, *For the Learning of Mathematics,* 8(2), 44-51.

Glaser, B. G. and Strauss, A. L.: 1967, *The discovery of grounded theory: strategies for qualitative research*, New York: Aldine de Gruyter.

Hanna, G.: 2000, Proof, explanation and exploration: an overview, *Educational Studies in Mathematics*, 44, 5-23.

Iannone, P. and Nardi. E.: 2005, On the pedagogical insight of mathematicians: 'Interaction' and 'Transition from the concrete to the abstract', *Journal of Mathematical Behavior*, 24, 191-215.

Madriz, E.: 2000, Focus groups in feminist research, in N. K. Denzin and Y. S. Lincoln (ed.) *Handbook of Qualitative Research,* second edition, Sage Publications, 835-850.

Otte, M.: 1990, Intuition and formalism in mathematical proof, *Interchange,* 21(1), 59-64.

Tall, D. and Vinner S.: 1981, Concept image and concept definition in mathematics, with special reference to limits and continuity, *Educational Studies in Mathematics,* 12, 151-169.

Tall, D.: 1991, *Advanced mathematical thinking*, Dordrecht / Boston / London: Kluwer Academic Publishers.

Van Oers, B.: 2002, Educational forms of initiation in mathematical culture, in Kieran C. and Sfard A. (ed) *Learning Discourse, Discoursive Approaches to Research in Mathematics Education,* Kluwer Academic Publisher, 59-85.

Weber, K. and Alcock, L.: 2004, Semantic and syntactic proof productions, *Educational Studies in Mathematics,* 56, 209–234.

Wilson, V.: 1997, Focus Group: a useful qualitative method for educational research? *British Educational Research Journal,* 23(2), 209-225.