Hewitt, D. (2006) Proceedings of the British Society for Research into Learning Mathematics 26(2) June 2006

THE COMPANY OF WORDS: USING CONCORDANCES TO DEVELOP LANGUAGE IN THE MATHS CLASSROOM

Frank Monaghan
Centre for Language and Communication, The Open University in London

THE PROBLEM WITH LANGUAGE IN MATHS CLASSES

Some thirty years ago, the linguist Michael Haliday argued that:

> The core of the difficulty in the mathematics classroom is that the teacher often understands and takes for granted the whole register of mathematics, and thinks only of the mathematical aspects of these items, whereas for the learner they may also be unfamiliar language - they are 'peculiar' English. It is therefore desirable that the mathematics teacher should be aware of the register of mathematics as a sub-set of English … To this end, mathematics educators and … English Language teachers should collaborate in the production of guidelines, illustrative descriptions and teaching materials concerned with this problem. (Halliday, 1975: 121)

Whilst numerous researchers have tackled the implications of this and sought to explain the 'peculiarities' of maths texts (Pimm, Morgan, Sfard, Barwell, Leung, Street, to name but a few) it remains the case that responses have largely been at a very local level with individual teachers or teams of teachers seeking to address the issue in their particular setting rather than there having been a more systematic approach to curriculum design informed by linguistic analysis. In this paper I will try to outline one possible way forward using a *corpus* of mathematical materials (taken from the SMILE scheme), interrogated using *concordancing* software, more on both of which later, but first I would like to give an illustration of how we can take things for granted about language when we rely on our intuitions. This based on a study by Tversky and Kahneman (1973).

Testing your intuition about language

In a typical sample of text in English, how many more times likely do you think it is that words begin with the letter 'k' than have 'k' as the third letter in them?

When I put this question to a group of mathematics researchers about two thirds of those of them intuited that there would be roughly twice as many words beginning with the letter 'k'. Interestingly, this was roughly the same as found in the Tversky and Kahneman (1973) study where 105 out of 152 people thought there would be more words beginning with 'k'. There are, in fact, likely to be about half as many. This is because we are far more likely to think of the first letter of a word than the third, but once you begin to think about, (bake, cake, fake, hake, lake, make, naked, rake, sake, take, wake…) it becomes clear that the initial assumption is wrong. Once you begin to 'problematize' mathematical English then the peculiarity Halliday identified becomes more obvious.

Some examples

The most obvious level is that of vocabulary. Few maths teachers would be surprised if when a child was asked for the first time what the difference is between 7 and 12 they were given a reply along the lines that one is odd and the other is even rather than the mathematically more expected answer ‘5’. Similarly with ‘similar’. These oddities continue at other levels.

Syntax: we say ‘three quarters ***is*** bigger than five eighths’ rather than ‘are bigger’ because fractions are treated as single entities so in spite of quarters being plural we use the singular form of the verb. Teachers of English as an additional language will have to explain to students whom they have diligently reminded of the need for subject-verb agreement in English why in maths it’s different.

People tend to assume that maths represents a pared down, objective form of language but it is also riddled with metaphor that can lead to confusion. A classic example is ‘multiplication makes numbers bigger’ and its concomitant, ‘division makes things smaller’.

WHAT NEEDS TO BE DONE?

It is clearly desirable, as Halliday said, that maths teacher should know about the register of mathematics if they are going to be able to guide their students through it (and, of course, there are particular challenges facing students learning English as an additional language through maths and *vice versa*. In the remainder of this paper I will try to describe how the use the use of a corpus and concordancing software might help with ‘the production of guidelines, illustrative descriptions and teaching materials’ that Halliday called for.

What is a corpus?

A corpus is simply a body of *authentic language* data. We now tend to think of it as being stored electronically and available for interrogation, e.g. through the use of a concordancer.

What is a concordancer?

A concordancer is a piece of software that enables its user to track and display patterns in a body of language by, for example, displaying key words in context (KWiC). Here is an example of a concordance of the word ‘poor’ in ‘A Tale of Two Cities’ book 1:

947	Miss, if the	**poor** lady had suffered so intensely
1884	the love of my	**poor** mother hid his torture from me
1615	stockings, and all his	**poor** tatters of clothes, had, in a long
1577	faded away into a	**poor** weak stain. So sunken and
1001	on your way to the	**poor** wronged gentleman, and, with a
1036	detachment from the	**poor** young lady, by laying a brawny hand

The word under investigation is displayed in bold in the centre with a defined number of words displayed on either side of it to allow the context to be clarified (most concordancers will take the reader to the original text when the word is clicked on so that the full context can be examined). This method of display is useful as it allows the investigator to identify real patterns of use as opposed to assumed ones. For example, if asked what the word 'poor' means, there is a distinct possibility that a first definition would be 'not wealthy', whereas the way it is used here emphasizes a secondary meaning, 'eliciting sympathy'.

Another way of displaying text in a concordancer is by using a frequency table, as in this display of 'feminine' adjectives used in a corpus of 19th century American fiction.

Adjective	**Women**	**Men**
little	112	59
dear	20	15
happy	15	9
pretty	14	9
sweet	13	8
lovely	12	3
pale	11	5
beautiful	9	6
eager	7	1
charming	5	8
delicate	3	7

This sort of display allows the reader to look at distinguishing patterns of use (here gender orientation) or simply to compare frequencies of different words in a text.

WHAT HAS THIS GOT TO DO WITH MATHS?

Concordancers have been used to analyse all manner of texts, literature, journalism, legal proceedings, school history, written and spoken, but I am unaware of much work on school mathematics. As a result, we have gained insights into how such texts are structured and used to create meanings. My ambition is to develop a corpus of school maths materials that will give us some insights into the language of mathematics and how students might be enabled to become more effective users of both. The data are drawn from some 1318 activities (including worksheets, workcards, posters, microprograms, games) of the total 1418 that were available at the time of compilation. This accounted for some 93% of the scheme, the remaining 7% being either unavailable to or held in an unsuitable format (i.e. computer activities without any written support or activities that involved no written text). The corpus itself is approximately 250,000 words. It seems reasonable to argue that, as the corpus included some 93% of the materials in the SMILE scheme and covers all levels of the National Curriculum, it serves as a basis for the analysis of the mathematics register as developed through this particular scheme.

In the following I will give an example of the sorts of things it might be possible to do.

What's a diagonal?

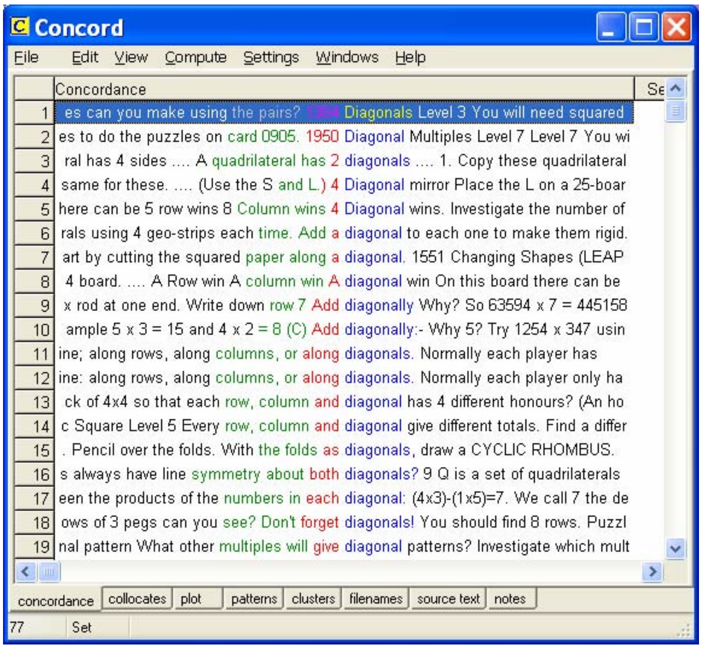
According to one mathematical dictionary:

> A diagonal is a straight line drawn from one vertex of a polygon (or solid figure) to another vertex. (Abdelnoor, 'A Mathematical Dictionary')

According to an ordinary dictionary

> A diagonal line goes in a slanting direction away from another line (COBUILD English dictionary)

I wanted to discover how the word was used in the materials the students were exposed to and the ways in which the two distinct meanings above were reflected in them. Analysis of the word (or in linguistic parlance 'lemma' which includes its multiple forms 'diagonal', 'diagonals', and 'diagonally' revealed that it appeared in as an adjective, 'diagonal moves are not allowed'; as a noun (singular and plural), 'The vector OR is the diagonal of a box…'; and as an adverb, 'You may not hop diagonally'. This demonstrated that the word was used in the scheme both in its mathematical meaning (as an attribute) and in its ordinary meaning (indicating orientation).



By looking at the frequency of occurrence I was able to discover some other interesting features. The corpus was structured in such a way that I was able to track what NC levels the word appeared at and so track its development through the materials. This threw up some interesting issues. For example, the word did not appear at NC level 2 in the scheme, suggesting that there might be a need or opportunity to plan in another activity. Further, there were particular clusters where it seemed to feature more heavily than at others (i.e. between levels 4 to 7), which may reflect its proper place within the hierarchy of mathematical concepts. What was most interesting, however, was that the lemma occurred 66 times throughout the scheme and of these 32 were in its technical, mathematical sense and 34 were in its non-mathematical sense. This ran counter to my intuition that a mathematics scheme would privilege mathematical meanings over ordinary meanings. That it doesn't has serious potentially implications for how students experience and understand mathematical terminology since the most common form they encounter is likely to have the most salience.

Students' understanding of mathematical terms

Here is an extract from an exchange between a group of Year 8 students. They were working on an activity that involved them identifying a shape from a set of attribute cards.

> **Jack:** [reading attribute card] 'At least one reflex angle'. No one knows what that is so you can't stop me.
>
> **Lucy:** No, wait. What is a reflex angle?
>
> **Jack:** I don't know. If you bop it then it'll kick you up.
>
> **Lucy:** No, wait. What's a reflex angle?
>
> **Jack:** It's when someone taps your knee and your leg jumps up.
>
> **Nadia:** Shall I go and ask the librarian?
>
> **Lucy:** One obtuse angle. What's obtuse? Nadia, while you're there, ask him what obtuse means.

However groan-worthy Jack's comment is (and clearly he loved it enough to use it twice and clearly Lucy was so used to his sense of humour that she completely ignored him) it does illustrate that if in doubt students are likely to fall back on a meaning that they know and where words in the mathematical register have a specialist meaning that is different from their meaning in ordinary English then problems can arise. When I interrogated the corpus to find out how 'reflex' was used I discovered that it made its first appearance at level 6, and none of the students were currently working at that level. So again, the concordance was useful in identifying a gap in the resources.

When Nadia returned from the librarian, the following exchange took place:

Nadia: [reading a written note she'd made] 'A reflex angle is an angle that is more than 180º. An obtuse angle is greater than 90º and less than180º.'

Jack: So it's the opposite of a right angle. It's an isosceles trapezium.

Mark: They're not parallel because that one's longer than that.

Jack: So?

Mark's final comment was also very revealing about his understanding of attributes; it would seem from this that for him parallel lines have to be equal in length. The corpus proved helpful here too. I had annotated all activities that had illustrations on them and so was able to go back and look at the salient features of these illustrations. A good example of this was the representation of rectangles There were 32 activities containing 48 images of rectangles. The modal ratio of the sides was 1:2, the mean ratio was 1:2.19. Roughly two thirds of the images depict the rectangle with the width greater than the height and all but a handful are oriented perpendicular to the page. This might account for comments such as these, taken from students who were asked to define the difference between a rectangle and a square:

Marcus: A rectangle is a four sided shape but it is longer

Celina: The lenth of square is more short then the rectangle

Andrew: A rectangle has two sides (the bottom and top) wich are wider than the two sides

Again, it would seem reasonable to argue that students' understandings of mathematical attributes are at least partly constructed by their exposure to images or the language used in the materials they are exposed to.

The challenge is whether we can develop and use corpora such as this one to identify gaps and opportunities for fresh interventions that teach our students what we want them to learn in a more linguistically (and therefore mathematically?) more principled way. A further ambition, with the needs of students learning mathematics through English as an additional language in mind, is to develop a language curriculum through maths. This would be to seriously address Halliday's challenge though whilst it is necessary it may not be sufficient. To address broader aspects of the child's learning experience, for example, the analysis would need to be extended to include affective aspects of the curriculum such as how it reflects cultural diversity – how many non-English names there are, what jobs people have compared to the mathematics they use, how disability is reflected in the materials and so on. It is through close examination of the particular that we may finally illumine the peculiar.

REFERENCES

Tversky, A. and Kahneman, D, (1973) 'Availability: A heuristic for judging frequency and probability'. Cognitive Psychology, 5, 207-32