UN LANGAGE DE PRONONCES DE FORMULES EN MATHEMATIQUES

Dans cet article nous nous intéressons au prononcé (c'est-à-dire à la lecture orale) d'une formule mathématique. Nous pensons que les prononcés des formules forment un langage dont nous analysons les propriétés et la manière dont il est utilisé par les enseignants. Nous proposons des hypothèses sur sa formation et son évolution.

C'est un fait admis par tous aujourd'hui : les problèmes de la communication déterminent le développement des sciences.

La communication dans les Mathématiques est particulière en ce sens qu'elle se fait non seulement en utilisant la langue courante mais aussi des formules écrites *à l'aide de symboles spéciaux.*

Or la communication qui joue un rôle primordial dans l'enseignement a lieu à la fois par écrit et par oral. Les utilisateurs des mathématiques sont amenés à prononcer *des formules. Même le lecteur, seul devant une formule* en fait, dans son esprit, une prononciation muette.

Nous nous sommes spécialement intéressés à ces prononcés de formules. Il semble qu'on puisse penser que leur ensemble, que par la suite nous noterons L_2 [1], constitue à lui seul un langage. Notre but est de dégager des éléments qui tendent à justifier que L_2 est un langage à part entière, de présenter ses propriétés et de proposer des hypothèses sur sa formation et son évolution.

UN LANGAGE RÉSERVÉ A LA LECTURE DES FORMULES MATHÉMATHIQUES?

Une formule est la plupart du temps une file de symboles, c'est-à-dire une suite de symboles écrits les uns à la suite des autres sur une même ligne.

Ex. : a + b.

Il existe des formules dont l'écriture n'est pas une file comme

$$\frac{a}{b} \quad , \quad C_n^p \text{ ou } \binom{n}{p}$$

Prononcer une formule consiste à la traduire de l'écriture mathématique dans L_2 : la traduire au sens habituel du mot, c'est-à-dire traduire chaque symbole dans L_2, puis ranger les traductions des symboles dans un certain ordre (qui n'est pas nécessairement celui dans lequel ils se trouvaient dans la formule).

Ainsi :

$|x|$ peut être lu « valeur absolue de x » les deux barres | | étant d'abord lues, puis la lettre x entre les deux barres.

A ce propos, deux types d'hypothèses par exemple peuvent être formulés :

a) On peut considérer que

»| |» est prononcé « valeur absolue de ».

b) Ou que :

»| |» est prononcé « valeur absolue ».

et que « de » est un mot de liaison rajouté à la lecture.

(A la fin de l'article, nous préférons prendre l'hypothèse a).

En conséquence deux sortes de questions se posent à propos du prononcé :

---

(1) La notation L_2 a été choisie en fonction d'une étude plus large sur les langages utilisés en mathématiques. On peut distinguer :

— la langue courante LC;
— le langage mathématique L_1;
— les écritures de formules E,
— leur prononcé L_2.

— Quel est (ou quels sont) le (ou les) prononcé(s) de chaque symbole? (Vocabulaire de L_2.)

— Quel est (ou quels sont) l' (ou les) arrangement(s) d'une formule? (Syntaxe de L_2.)

Une première constatation : les formules ne se prononcent pas dans la langue courante. Ainsi les formules écrites « $A \cap B$ », « x^2 », sont aussi mystérieuses à un lecteur non familiarisé avec les mathématiques que les lectures orales « a inter b », « icsse deux ».

Les mots prononcés des symboles sont en effet différents des mots de la langue courante ou si ce sont les mêmes, ils ont un sens très particulier dans L_2, qu'ils n'ont pas dans la langue courante.

Ainsi « inter » est un mot qui ne se rencontre que comme préfixe dans les mots français courants.

« union » est employé dans la langue courante mais non avec le sens spécifique et très précis qu'il a dans L_2 (union de deux ensembles).

Le vocabulaire de L_2 n'est donc pas celui de la langue courante, mais il n'est pas non plus le même que celui du langage mathématique écrit (L_1).

Des symboles sont prononcés par des mots non employés à l'écrit, comme :

$\cap$ qui est prononcé **inter** (dans $A \cap B$),

$=$ qui est prononcé **égal** (dans $a = b$).

Dans le langage mathématique écrit, on utilise le mot « intersection », et l'expression « égal à ».

Vocabulaire particulier, mais aussi syntaxe particulière à L_2, qu'on ne retrouve dans la langue courante, ni dans le langage mathématique écrit.

L'écriture des formules mathématiques étant très concise par nature grâce à l'utilisation de symboles, son prononcé essaie de l'être.

Aussi y a-t-il dans L_2 de nombreux phénomènes visant

— ellipses du verbe,

— suppression d'articles,

— suppression de prépositions.

Les mots restants étant pratiquement les mots indispensables à la compréhension.

Toutes ces constatations amènent à penser que L_2 a un vocabulaire et une syntaxe spécifiques, donc qu'il est à lui seul un langage différent des autres.

Exemples :

Phénomène	Écriture mathématique	Prononcé	Suppression de
Ellipses du verbe	$p > q$	« p supérieur à q »	(p) est (supérieur à q)
	$a \in m$	« a élément de M »	(a) est (élément de M)
Suppressions de prépositions, de conjonctions et d'articles (elles ont lieu souvent ensemble)	$R(x, y)$	« R de x y »	(R de x) et de (y)
	$(x + y)^4$	« x plus y puissance quatre »	(x plus y) à la (puissance quatre)
	$(x + y)^2$	« x plus y carré »	(x plus y) au (carré)
	$\vert x \vert$	« valeur absolue de x »	la (valeur absolue de x)

L_2 est-il un langage universel?

L'existence de L_2 n'est pas ressentie par tous les utilisateurs des mathématiques, en particulier par tous les enseignants. Ce n'est que depuis peu de temps qu'il est fait mention du prononcé des formules dans les manuels de l'enseignement secondaire.

En effet, à l'oral il arrive aux lecteurs de paraphraser une formule par une phrase du langage mathématique et non de la prononcer.

Au lieu de lire « $a \in M$ » « a élément de M », ils énoncent par exemple « l'ensemble M a pour élément a » ou « a est un élément de M ».

Ils traduisent directement la formule par une phrase complète du langage mathématique, ils ne la lisent pas.

On peut trouver cette habitude même dans les manuels. Nous avons lu par exemple :

« E est un sous-ensemble de F » s'écrit « $E \subset F$ » ce qui donne à penser qu'inversement la formule « $E \subset F$ » se lit : « E est un sous-ensemble de F », alors que cette phrase est une phrase du langage mathématique et que « $E \subset F$ » se prononce habituellement « E inclus dans F ».

L_2 dans ces cas est donc inutilisé et c'est regrettable parce que le prononcé de la formule est plus près de son écriture que sa paraphrase et encore parce qu'il y a un nombre très grand de paraphrases d'une même formule, ce qui risque de créer des difficultés d'ordre pédagogique, les élèves se perdant dans les multiples paraphrases d'une formule.

Cette absence de prise de conscience de l'existence de L_2 provient certainement de ce que L_2 n'est pas un langage fixé, normalisé comme le langage des écritures des formules. Les constatations suivantes le prouvent.

1. — Certains symboles ont différents prononcés : « $\forall$ » se prononce « pour tout » ou « quel que soit ».

2. — Certains symboles ont différents prononcés suivant la formule où ils sont employés.

Comme l'exposant des puissances :

« a^4 » est lu « a quatre »

mais

« $(x + y)^4$ » est lu « x plus y puissance quatre ».

Ou comme le signe de multiplication (omis à l'écriture) :

« a b » est lu « a b »

mais

« $(a + b)(c + d)$ peut être lu

« a plus b facteur de c plus d ».

Ou comme les parenthèses :

elles ne sont pas prononcées dans

« $\sin(a + b)$ » qui est lu « sinus a plus b »

mais on peut considérer qu'elles le sont sous la forme de la préposition « de » dans

« $f(x)$ » qui est lu « f de x ».

3. — Une même formule n'a pas un prononcé unique dans L_2. Une simple écriture comme « 3×4 » est certainement lue dans une assemblée d'une dizaine de personnes au moins de deux façons différentes parmi « trois fois quatre », « trois multiplié par quatre », « trois que multiplie quatre »...

4. — Des formules construites exactement de la même manière ont des prononcés de construction différente.

« $\log x$ » est lu « logarithme de x »

« $\sin x$ » est lu « sinus x »

Il s'agit de deux symboles fonctionnels portant sur le même élément, mais dans le premier cas est introduite la prépostiion « de », et non dans le second.

Il ne fait pas de doute que le prononcé dépend d'usages, de coutumes quelquefois très anciennes, ainsi que des habitudes de l'orateur, du niveau en mathématiques de son auditoire, des conventions qui ont pu être faites.

Ainsi à des élèves qui viennent d'aborder l'associativité de l'addition, le professeur insistera sur la place des parenthèses dans le prononcé de « $a + (b + c)$ » en lisant par exemple « a plus entre parenthèses b plus c ».

Plus tard, pour les mêmes élèves, il ne jugera plus nécessaire de le faire.

La formule « u . v » est prononcée

« u scalaire v » ou « vecteur u scalaire vecteur v » suivant que l'on désire ou non insister sur le fait qu'il s'agit d'un produit d'éléments d'un espace vectoriel.

La lecture d'une matrice pose des problèmes. Doit-on la lire ligne par ligne ou colonne par colonne?

L'orateur pourra passer une convention avec son auditoire.

S'il n'y a pas unicité de prononciation d'une même formule, ou d'un même symbole, un prononcé peut malheureusement prêter à plusieurs interprétations possibles.

Ainsi « e puissance x plus un » est le prononcé de

Ainsi « e puissance x plus un » est le prononcé de

« $e^x + 1$ » (1)

comme de

« e^{x+1} » (2)

la seule différence ayant lieu pour les habitués au niveau d'un soupir (ou d'une demi-pause) qui se place avant plus pour (1), avant x pour (2).

Ces ambiguïtés apparaissent souvent (comme dans cet exemple) pour les expressions parenthésées, les parenthèses étant omises à la lecture, ou plutôt la parenthèse ouvrante étant traduite par un silence, la parenthèse fermante n'étant pas prononcée, ni par un mot, ni par un silence.

L_2 est donc un langage bien subtil et d'un maniement difficile, puisqu'il utilise même des procédés qui pour l'instant paraissaient réservés à l'art musical! Il apparaît comme peu normalisé.

Que font les mathématiciens et surtout les enseignants face à ces subtilités?

Une enquête auprès des enseignants.

Pour avoir une idée de l'implantation et de l'usage de L_2, nous avons envoyé à une centaine d'enseignants du premier et du second cycle du secondaire et du supérieur un questionnaire qui leur demandait d'indiquer comment ils prononçaient de façon spontanée [2] les formules suivantes :

$a + (b + c)$
$3 \leqslant 4$
3×4
$\{1, 2, 3\}$
$P \cup Q \supset P$
$P \cap Q \subset P$
$R \circ S = T$
$C_E A$

(2) En fait, nous ne pouvons juger le degré de spontanéité des réponses. Il est difficile de donner, hors du contexte, le prononcé d'une formule, comme on le ferait au cours d'une leçon de mathématiques. Il est à craindre que les réponses ne soient pas spontanées, comme l'indiquent les ratures que nous avons pu y trouver.

$\vec{u} . \vec{v}$

$\vec{u} \wedge \vec{v}$

$x \equiv \frac{\pi}{3} [2\pi]$

$\vec{X} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$

$\begin{pmatrix} ab \\ cd \end{pmatrix} \begin{pmatrix} ef \\ gh \end{pmatrix}$

$\forall n \in N \exists p \in N \quad p > n$

Nous avons distingué dans les 51 réponses reçues plusieurs types de formules.

Type 1 : Les formules qui recueillent un prononcé commun presque à l'unanimité.

Type 2 : Les formules qui recueillent un prononcé commun à la majorité absolue.

Type 3 : Les formules qui recueillent plusieurs prononcés chacun avec une proportion voisine de $\frac{1}{3}$.

Type 4 : Les formules qui recueillent presque autant de prononcés que de réponses. Il s'agit d'ailleurs plutôt de paraphrases.

Que conclure des réponses?

Type 1 : Les formules de ce type sont des formules courtes (au maximum 4 symboles), avec un symbole opératoire ou relationnel portant sur deux lettres dans tous les cas, sauf pour « $R_oS = T$ ».

Ces formules sont utilisées dans toutes les classes du secondaire.

Type 2 : Ce sont des formules courtes mais moins utilisées dans le secondaire, ou plus longues.

Type 3 : Ces formules ont au moins 5 symboles.

Elles sont moins utilisées que les précédentes ou comportent des symboles difficiles à prononcer comme les parenthèses et les accolades. Remarque : la formule $\{1, 2, 3\}$ n'a jamais recueilli le prononcé qui semble le plus court, « l'ensemble d'éléments un deux trois ».

Type 4 : Ces formules sont plus longues; deux d'entre elles ne sont pas des files et posent une difficulté supplémentaire : l'ordre de lecture des symboles.

Les prononcés obtenus dans les réponses sont en fait des paraphrases construites comme des phrases complètes avec verbes, articles, prépositions.

Ci-dessous le classement des formules.

Type	Écriture de la formule	Prononcé(s) obtenu(s)	Obtenu à
1	$3 \leqslant 4$	Trois inférieur ou égal à quatre	80 %
	3×4	Trois multiplié par quatre	80 %
	$R_oS = T$	r ronds s égal t	80 %
	$C_E A$	Complémentaire de A dans E	80 %
2	$\vec{u} . \vec{v}$	u scalaire v	60 %
	$\vec{u} \wedge \vec{v}$	u vectoriel v	60 %
	$P \cup Q \supset P$	p union q contient p	60 %
3	$a + (b + c)$	a plus b plus c	40 %
		a plus b plus c entre parenthèses	35 %
	$P \cap Q C P$	p inter q est inclus dans p	35 %
		p inter q inclus dans p	27 %
	$x \equiv \frac{\pi}{3} [2\pi]$	x congru à π sur 3 modulo deux π	27 %
		x est congru à π sur 3 modulo deux π	23 %
	$\{1, 2, 3\}$	l'ensemble un deux trois	27 %
		l'ensemble formé des éléments un deux trois l'ensemble constitué des éléments un deux trois	20 %
4	$x = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$		
	$\begin{pmatrix} ab \\ cd \end{pmatrix} \begin{pmatrix} ef \\ gh \end{pmatrix}$		
	$\forall n \in N \exists p \in N p > n$		

Plus la formule est longue, plus la divergence est grande sur son prononcé et plus le prononcé prend l'allure d'une phrase complète. Ainsi dès le type 3, le prononcé avec ellipse du verbe recueille autant de voix que celui sans ellipse.

A ce propos, remarquons que :

« $P \cup Q \supset P$ » appartient au type 2

et que

« $P \cap Q \subset P$ » appartient au type 3

alors que ce sont deux formules d'écritures très voisines.

En effet, le symbole $\supset$ se lit par un verbe « contient » tandis que le symbole $\subset$ se lit par un verbe suivi d'un adjectif « est inclus », ce qui offre deux possibilités : ellipse du verbe être ou conservation du verbe être. C'est pourquoi il y a moins de divergence sur le symbole $\supset$.

Cependant un accord très important a lieu sur les symboles d'usage fréquent et les formules les utilisant. Il n'est personne qui ne prononce aujourd'hui « $+$ », « $-$ », « $=$ » autrement que « plus », « moins », « égal », « $a + b$ », « $a - b$ », « $a = b$ » autrement que « a plus b », « a moins b », « a égal b »; ce qui n'était pas vrai au début du XX° siècle. En effet dans les manuels des classes du secondaire des environs de 1900, L_2 n'était pas mentionné, et dans le langage mathématique écrit étaient employées des expressions comme : « la somme de a et de b » ou « la somme de a et de la somme de b et de c effectuée ».

Tout ceci nous incite à penser que L_2 est un langage qui existe, mais qui s'installe et prend sa forme définitive petit à petit.

En témoignent les manuels d'enseignement qui indiquent après chaque formule écrite à l'aide de nouveaux symboles son prononcé entre guillemets, ce qui insiste sur le fait que le prononcé relève d'un langage différent des autres employés dans le manuel.

Quelle forme prendra L_2 dans les années à venir. Des hypothèses.

	a	+	b
est lu	↓	↓	↓
	a	plus	b

	a	—	b
est lu	↓	↓	↓
	a	moins	b

	a	=	b
est lu	↓	↓	↓
	a	égal	b

Ces formules dont le prononcé est unique sont des files. Les prononcer consiste à prononcer chaque symbole en parcourant la file de gauche à droite.

Mais les prononcés des formules suivantes peuvent être considérées comme obtenus de la même façon, le prononcé d'un symbole consistant en plusieurs mots.

	A	$\subset$	B
est lu	↓	↓	↓
	A	inclus dans	B

	a	$\leqslant$	b
est lu	↓	↓	↓
	a	inférieur ou égal à	b

Nous pensons qu'avant d'aboutir à cette forme proche de l'écriture mathématique, le prononcé d'une formule passe par différentes phases. Nous distinguons.

— **La première phase :** phase de la paraphrase (type 4)

« $P \cap Q \subset P$ »

est lu

« l'ensemble intersection de P et de Q est inclus dans P »

ou

« l'intersection des ensembles P et Q est incluse dans P ».

— **La deuxième phase :** la formule est prononcée non plus paraphrasée; les symboles sont prononcés dans L_2, dans l'ordre de gauche à droite où ils apparaissent dans la formule si c'est une file, mais le prononcé garde l'allure d'une phrase.

« P	$\cap$	Q	$\subset$	P »
↓	↓	↓	↓	↓
l'ensemble P	inter	l'ensemble Q	est inclus dans	l'ensemble P

Chaque symbole est prononcé par un groupe de mots.

— **La troisième phase :** à l'usage les mots qui n'ajoutent rien à la compréhension disparaissent.

« P	$\cap$	Q	$\subset$	P »
↓	↓	↓	↓	↓
P	inter	Q	inclus dans	P

Le prononcé est d'autant plus stable que toute suppression de mots du prononcé nuirait à sa compréhension.

Dans l'exemple cité le « dans » ne peut être supprimé à cause du risque de confusion à l'oral entre « inclus » et « inclut ».

La formation du prononcé stable à partir de la formule serait donc la suivante :

— les symboles sont lus de gauche à droite si la formule est une file;

— les lettres sont lues dans la langue courante;

— les symboles relationnels ou opératoires sont lus par un ou plusieurs mots qui ne sont pas nécessairement consécutifs, cet ensemble de mots étant l'ensemble minimal pour la compréhension. Si ces mots ne sont pas consécutifs, leur place est déterminée.

Ainsi :

« $C_E A$ » est lu « complémentaire de A dans E » le symbole « C » étant lu « complémentaire de... (1) dans... (2) ».

En (1) est la lettre à droite de C.

En (2) est la lettre en bas à droite de C.

Beaucoup de formules sont prononcées de cette façon mais les symboles ont des prononcés non minimaux [3], souvent plusieurs prononcés ou des prononcés différents suivant la formule où ils se trouvent.

On peut espérer, puisqu'il en est déjà ainsi pour les formules usuelles, que les prononcés évolueront vers une forme stable, afin de permettre une relative normalisation dans ce domaine, ce qui serait souhaitable pour élèves et enseignants.

Cette étude n'a été faite que relativement à la langue française. Il serait intéressant de savoir ce que sont les L_2 d'autres langues; en particulier si ce sont les mêmes formules qui ont un prononcé fixé, si les mêmes hypothèses sur l'évolution des prononcés peuvent être faites.

Il serait également intéressant d'évaluer la part de l'influence de la langue mère sur L_2. Peut-être serait-il souhaitable que L_2 obéisse en partie à des règles de construction communes à tous les pays, comme l'écriture des formules mathématiques?

Cette étude a été faite en liaison avec une équipe de Grenoble d'enseignants du secondaire et du supérieur, en mathématiques, lettres et linguistique, qui travaille sur les langages des mathématiques.

Colette LABORDE.

Assistante de mathématiques.
U.S.M.G. Grenoble.

(3) Comme $\leqslant$ qui est lu « inférieur ou égal ». On pourrait proposer un prononcé consistant en un seul mot, comme « infégal » ou « infeg ».