Marqueurs linguistiques et compétences mathématiques : une étude exploratoire

Sylvie Normand-Assadi[1], Lalina Coulange[1, 2], Élisabeth Delozanne[3], Brigitte Grugeon[1, 4]

[1]IUFM de Créteil, Rue Jean Macé, 94861 BONNEUIL Cedex, France
{sylvie.normand, lalina.coulange,elisabeth.delozanne}@creteil.iufm.fr
[2]CRIP5 - Paris V, 45-46 rue des Saints-Pères, 75 006 PARIS, France
[3]DIDIREM - Paris VII, 2, Place Jussieu, 75 251 PARIS Cedex 05, France
[4]IUFM d'Amiens 49, boulevard de Châteaudun 80044 AMIENS CEDEX,
brigitte.grugeon@amiens.iufm.fr

Résumé

Nous présentons une étude pluridisciplinaire (linguistique, didactique et informatique) visant à améliorer la modélisation cognitive des élèves en algèbre élémentaire. Nous mettons en parallèle une analyse didactique et des marqueurs linguistiques à partir d'un corpus recueilli lors de l'utilisation d'un logiciel- le logiciel Pépite. L'objectif est d'améliorer l'évaluation des réponses des élèves aux questions ouvertes quand les élèves répondent dans leurs propres mots. Dans un premier temps, nous situons notre étude. Avant de présenter notre méthodologie et l'analyse linguistique des données recueillies. Nous montrons ensuite la pertinence des résultats de cette analyse avec un point de vue de didactique des mathématiques. Nous présentons enfin les perspectives ouvertes par ce travail.

Mots-clé

Gestion des compétences et des connaissances, modélisation cognitive, marqueurs linguistiques, diagnostic de compétences

Abstract

We describe an exploratory empirical study to investigate whether some linguistic markers can improve the assessment of students when they answer to questions in their own words. This work is part of a multidisciplinary project, the Pépite project, that aims to give math teachers software support to assess their students in elementary algebra. We first set this study within the context of the project and in comparison with related work. Then we present our methodology and the data analysis. Next we discuss how the results of this linguistic analysis make sense from a mathematical educational research point of view. The paper ends with promising perspectives.

Key words

Knowledge and skills management, cognitive models, linguistic marks, skills diagnosis.

Introduction

L'analyse des réponses d'élèves exprimées dans leurs propres mots est un verrou sur lequel butent de nombreux projets de e-learning (Rosé et al. 2003b). Beaucoup de travaux s'intéressent à l'analyse automatique des explications produites en langage naturel et des interactions verbales d'étudiants (avec des pairs, avec l'enseignant) ou se centrent sur l'utilisation de modèles linguistiques dans la conception de logiciels d'apprentissage (Aleven, Popescu, Koedinger 2002, Arroyo et al. 2001, Rosé et al. 2003a, Rosé et al. 2003b). Comment prendre en compte la diversité des productions spontanées d'élèves « dans leurs propres mots » dans une analyse informatisée ? En quoi la nature de ces productions peut-elle nous informer sur les apprentissages ou leurs dysfonctionnements ?

Notre étude vise à apporter des éléments de réponses à ces questions dans le contexte très spécifique d'un logiciel d'évaluation diagnostique en algèbre élémentaire.

Cette étude se situe dans le cadre du projet Pépite (Pépite 2004) qui vise à concevoir un logiciel pour évaluer des compétences en algèbre élémentaire. Il s'agit d'analyser les erreurs et les cohérences d'élèves en algèbre en analysant leurs réponses à un test spécialement conçu pour couvrir les principales dimensions de la compétence algébrique de base. L'objectif est, à partir de ce diagnostic de proposer des situations d'apprentissage adaptées aux profils d'élèves ainsi repérés. Ce projet a donné lieu à la réalisation d'un logiciel nommé Pépite qui ne se contente pas de répertorier des erreurs, mais met en évidence aussi les conceptions ou les fonctionnements sous-jacents à la production de ces erreurs (Grugeon 1995, Jean et al. 1999). Pour les didacticiens des mathématiques de notre équipe, il s'avère indispensable de recueillir des réponses où les élèves s'expriment dans leurs propres termes, même si le logiciel n'est pas capable à ce jour d'en faire une analyse automatique complète et fiable.

.

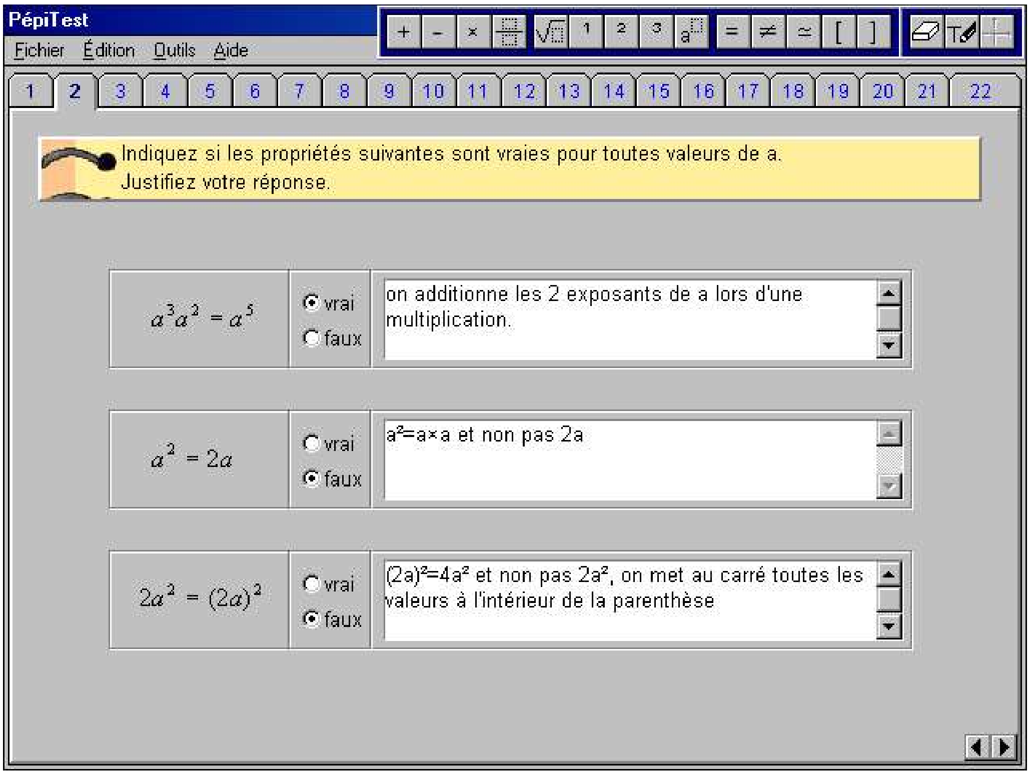

Figure 1 : Réponses d'Arthur à l'exercice 2 de PépitTest

Dans cet article, nous présentons une étude linguistique menée afin d'améliorer, dans le logiciel Pépite, le diagnostic de ce type de réponses, exprimées par les élèves eux-mêmes, dans leur propre langage. Tout d'abord, nous montrons comment cette étude s'inscrit dans l'ensemble du projet. Puis nous présentons notre méthodologie et l'analyse des données recueillies. Nous articulons ensuite les résultats de cette étude linguistique avec un point de vue issu de la didactique des mathématiques. En guise de conclusion, nous dégageons des perspectives de cette étude.

Les justifications d'élèves dans Pépite

Un logiciel (Jean et al. 1999, Delozanne et al. 2003) nommé Pépite a été conçu sur la base des travaux sur l'enseignement de l'algèbre de Grugeon (Grugeon 1995, Artigue et al. 2001, Kieran 1992). Le logiciel est composé de trois modules PépiTest, PépiDiag et PépiProf. PépiTest propose aux élèves un test et recueille l'ensemble de leurs réponses. Le système de diagnostic PépiDiag analyse automatiquement la plupart de ces réponses. PépiProf permet à l'enseignant d'accéder au diagnostic ainsi établi, de le compléter, de le modifier. Le logiciel dresse enfin un profil cognitif qui permet d'obtenir un regard d'ensemble sur la compétence algébrique construite par l'élève concerné. Nous avons testé Pépite avec des enseignants pendant les trois dernières années (Delozanne et al 2003). Ces expérimentations ont montré que les enseignants, sauf exception, n'avaient pas le temps de compléter ou d'étudier le diagnostic détaillé rendu par le logiciel. Sauf en des cas très critiques, ils demandent un diagnostic entièrement automatisé qui permettrait d'obtenir rapidement une vue synthétique sur les compétences de leurs élèves. De plus, dans une optique de e-learning, disposer d'un diagnostic automatique fiable est indispensable. C'est donc un des objectifs que nous nous sommes fixés.

La première version de PépiDiag analyse bien sûr toutes les réponses aux questions fermées et une partie des réponses aux questions ouvertes du test (lorsque celles-ci correspondent à la production d'une seule expression algébrique). Une première amélioration du système permet l'analyse automatique des réponses constituées de plusieurs expressions algébriques (Prévit, Delozanne et Grugeon 2004). Mais le logiciel n'est pas encore capable à ce jour d'analyser les réponses exprimées en langage « mathurel » (Cf. Figure 1). Par langage mathurel, nous entendons un langage utilisé par les élèves, mélangeant le langage naturel et l'emploi de termes spécifiques aux mathématiques (Rasseneur et al 2002).

Le tableau 1 montre une partie de l'analyse didactique des justifications produites par les élèves sur l'exercice présenté à la Figure 1, en référence à une modélisation cognitive multidimensionnelle de la compétence algébrique élémentaire (Grugeon 1995). Dans cette analyse didactique, pour détecter que les étudiants s'appuient sur des arguments qualifiés de formel scolaire, on relève des expressions : « il faut », « on doit », « on ne peut pas » ou aussi des règles fausses.

En d'autres termes, il s'agit pour ces élèves de respecter des règles formelles, apparemment dénuées de sens pour eux. Cette première classification semblait suffisante pour le diagnostic humain, mais manquait de précision pour être utilisée de façon systématique par un système informatique.
Pour améliorer le diagnostic automatique les concernant, nous avons décidé de mener, d'un point de vue linguistique, une étude exploratoire des données recueillies et d'articuler ce point de vue avec un point de vue didactique.

Niveau de rationalité, preuve par…	**Type de justifications**	**Registre d'expression**	**Exemples de justification d'élèves**
Algèbre	Donner une règle correcte ou une définition	algébrique	$a^2 a^3 = a^{2+3}$ $a^n a^p = a^{n\times p}$
Exemple numérique	Essayer avec un ou plusieurs nombres	numérique	$2 \times 2^2 = 32$
Explications	Donner des explications	langage mathurel	"c'est vrai car les exposants sont additionnés"
			Le produit de deux nombres identiques à exposants différents est ce même nombre mais avec leurs exposants ajoutés tous deux, donc a puissance 2+3
Formel scolaire	Donner une règle formelle scolaire	langage mathurel	"On ne doit pas multiplier les exposants" "C'est une loi fondamentale"

Tableau 1 : Analyse (humaine) de justifications d'élèves associées à l'égalité "$a^2 * a^3 = a^5$" (Grugeon 1995)

Méthodologie

Le type de tâche proposée aux élèves est de « *valider ou invalider l'égalité de deux expressions algébriques par l'assertion vrai/faux puis justifier cette assertion* », correspondant à l'exercice 2 de PépiTest (Figure. 1). Cet exercice est constitué de trois questions, et pour chaque question, la réponse de l'élève est composée d'un choix (vrai/faux) et d'une justification (correspondant aux arguments donnés pour justifier le choix vrai/faux).

Pour cet exercice, nous disposons des traces produites par 168 élèves de troisième et seconde ayant passé le test avec le logiciel Pépite. Les justifications sont de différents ordres : justification algébrique, en langage mathurel, mixte, absence de justification. L'étude porte sur les élèves ayant donné au moins une justification exprimée en langage mathurel (52 élèves).
Les justifications écrites par les élèves sont étudiées comme des actes de discours (Searle 1969) accomplis par les élèves dans le contexte de la tâche considérée. Nous relions ce que disent les élèves (actes locutionnaires), à ce qu'ils signifient en disant cela (actes illocutionnaires) et à ce qu'ils accomplissent en disant cela (Austin 1962). Dans le cadre de cette étude, les actes de langage produits par les élèves sont conditionnés par le contexte (ici justifier leurs choix vrai/faux). Nous apprécions la force illocutionnaire des énoncés produits par les élèves en fonction de l'objectif de l'énonciation, correspondant à la tâche prescrite : « valider ou invalider l'égalité entre deux expressions algébriques par le choix vrai/faux et justifier ce choix ». Nous nous inscrivons dans le courant de la pragmatique intégrée (Ducrot 1984) qui associe à la description linguistique les caractéristiques pragmatiques de l'énonciation. Nous relevons ainsi des marqueurs et des structures linguistiques formelles utilisés par les élèves et les interprétons afin d'octroyer à leurs justifications une orientation qui les inscrit dans un registre discursif.

Nous avons dans un premier temps distingué deux groupes d'élèves en fonction de l'exactitude de leurs choix :

- Groupe 1 : élèves ayant fait des choix corrects « vrai/faux » pour les trois questions de l'exercice (24 élèves)
- Groupe 2 : élèves ayant fait au moins un choix incorrect « vrai/faux » en réponse à une des questions posées (28 élèves).

Pour ces deux groupes, nous avons codé le caractère « correct » (C) ou « incorrect » (I) du choix entre vrai et faux de l'élève ainsi que le caractère « correct » (C), « partiel » (P) ou « incorrect » (I) de sa justification.
Nous obtenons ainsi des combinatoires telles que : choix correct / justification correcte (CC), choix correct / justification partielle (CP), choix correct / justification incorrecte (CI), choix incorrect/justification incorrecte (II).Les Tableaux 2 et 3 synthétisent les répartitions au sein des deux groupes d'élèves considérés.

questions	CC	CP	CI	total
1) $a^3a^2=a^5$	5	17	2	24
2) $a^2 = 2a$	14	6	2	23
3) $2a^2 = (2a)^2$	15	5	1	21

Tableau 2 : Combinatoires du groupe 1 (Gr1 : choix corrects pour les trois questions-24 élèves)

questions	CC	CP	CI	II	total
1) $a^3a^2=a^5$	4	9	3	8	25
2) $a^2 = 2a$	8	5	1	10	24
3) $2a^2 = (2a)^2$	4	3	2	5	16

Tableau 3 : Combinatoires du groupe 2 (Gr2: au moins un choix incorrect -28 élèves)

Dans cette étude, nous cherchons à dégager les critères linguistiques qui peuvent nous apporter des informations fiables, relativement à la réussite des élèves à l'exercice (choix corrects/incorrects et justifications correctes/incorrectes/partielles), voire à leurs compétences en algèbre. Pour chaque question, nous relevons des caractéristiques de l'égalité d'un point de vue mathématique qui sont susceptibles d'influencer la nature de ces critères. Puis, pour chaque catégorie de réponses codées, nous mettons en avant les formes linguistiques utilisées par les élèves, qui permettent d'élaborer une typologie des justifications produites.

<table>
<tr><th>Registres de discours</th><th>Niveau d'entrée dans l'algèbre</th><th>Code</th><td rowspan="5">COMPÉTENCE ALGÉBRIQUE</td></tr>
<tr><td>Argumentatif :
L'élève a recours à des relations (conséquence, restriction, opposition)pour expliciter ses arguments.</td><td>Conceptuel :
L'élève manie les concepts</td><td>CC</td></tr>
<tr><td>Descriptif :
L'élève décrit les éléments du contexte posé par l'égalité donnée.</td><td>Contextuel :
L'élève reprend des éléments significatifs du contexte</td><td>CC, CP</td></tr>
<tr><td>Explicatif :
L'élève a recours à la causalité relativement à l'égalité donnée.</td><td rowspan="2">Formel/Scolaire :
L'élève applique ou reprend des règles formelles et non significative</td><td rowspan="2">CP, CI, II</td></tr>
<tr><td>Légal :
L'élève appuie son discours sur une dimension légale en utilisant des verbes modaux.</td></tr>
</table>

Tableau 4 : classification a priori des justifications d'élèves

D'un point de vue didactique, nous distinguons *a priori* différents registres linguistiques employés (argumentatif, descriptif, explicatif, « légal ») par les élèves dans leurs justifications. Nous faisons l'hypothèse que l'emploi de ces registres révèle différents niveaux de discours (conceptuel, contextuel, formel scolaire) qui peuvent être corrélés avec la réussite des élèves à la tâche, voire au développement de leurs compétences en algèbre (Tableau 4).

Analyse des données

Question 1: $a^3\,a^2 = a^5$

Caractéristiques mathématiques de l'égalité

Cette égalité est vraie. De plus, elle est très proche d'une règle algébrique ($a^m\ a^n = a^{m+n}$) qui fait partie du curriculum officiel : on la trouve dans chaque manuel scolaire et elle apparaît fréquemment dans les cours d'algèbre.

Analyse linguistique

Voici pour cette première question, les différentes catégories de productions représentées selon les codes adoptés.

CC – Choix correct / Justification correcte : argumentatif et conceptuel

L'élève utilise des conjonctions de coordination qui lui permettent de marquer les relations auxquelles il a recours du type *conséquence*, *restriction*, avec des conjonctions de coordination du type « mais », « donc ». Il inscrit son discours dans le *registre argumentatif* et ancre son raisonnement dans le *conceptuel*.

Exemple: « Le produit de deux nombres identiques à exposants différents est ce même nombre mais avec leurs exposants ajoutés tous deux, donc a puissance 2+3 »

CC – Choix correct / Justification correcte : descriptif et contextuel

L'élève utilise une phrase complexe construite à partir d'une proposition principale et d'une proposition subordonnée situationnelle soit juxtaposée soit enchâssée, en utilisant des conjonctions temporelles ou spatiales : lorsque, quand, lors, dans. La première proposition porte « l'acte » (on ajoute, on additionne) la seconde pose le « contexte » (lors d'une multiplication, dans une multiplication) qui est défini et nécessaire à la réalisation de l'acte. Il inscrit ainsi son discours dans le registre *descriptif* situationnel et ancre sa justification au niveau *contextuel*.

Exemple: « quand on multiplie des mêmes nombres avec des puissances, on adition les puissances et le nombre reste inchangé »

CP – Choix correct / Justification partielle : descriptif et contextuel

L'élève utilise une phrase complexe identique à celle décrite précédemment. Cependant, sa justification est considérée comme partielle dans la mesure où l'élève ne fait pas mention de toutes les conditions nécessaires à l'application de la règle donnée dans le cours (produit de deux puissances d'un même nombre). Ceci peut s'expliquer par le fait que l'élève s'appuyant sur le contexte donné par l'item, ne s'attache qu'aux éléments variants d'un membre à l'autre de l'égalité (les exposants 3, 2 et 5), pour délaisser l'élément invariant (le nombre indéterminé *a*). Pour ce type de justifications, les marqueurs linguistiques employés sont identiques à ceux de la catégorie précédente : lorsque, quand, dans.

Exemple : « quand on multiplie des nombres avec des puissances il faut additionner les puissances »

Code, Registres de discours, Niveau d'entrée dans l'algèbre, Marqueurs	**Exemple**
CC, *Argumentatif,* Conceptuel *Mais, donc* (3 Gr 1)	Le produit de deux nombres identiques à exposants différents est ce même nombre mais avec leurs exposants ajoutés tous deux, donc a puissance 2+3
CC, *Descriptif*, Contextuel *Quand, lorsque, dans* (5, 1 Gr 1, 4 Gr 2)	quand on multiplie des mêmes nombres avec des puissances, on adition les puissances et le nombre reste inchangé
CP, *Descriptif*, Contextuel *Quand, lorsque, dans* (15, 12 Gr 1, 3 Gr. 2)	Dans les multiplication à puissances, on additionne les exposants
CP, *Explicatif*, Formel scolaire *Car, vrai car* (6 Gr. 2)	Dans les multiplication à puissances, on additionne les exposants
II, *Légal*, Formel, scolaire *Il faut, il ne faut pas, on doit* (4 Gr 2)	il ne faut pas additionné les puissances mais les multiplier

Tableau 5 : réponses associées à l'égalité $a^3 a^2 = a^5$

CP – Choix correct / Justification partielle : explicatif, légal et formel scolaire :

Tout comme dans le cas précédent, les élèves considèrent seulement les éléments variant du membre de droite au membre de gauche de l'égalité. Mais au lieu de replacer ces arguments relativement au contexte posé par l'égalité, ils ont recours à la causalité marquée dans le discours par « car », « c'est vrai car ». Certains d'entre eux utilisent des verbes modaux qui indiquent la faisabilité ou la possibilité, voire l'obligation (« il faut »).

Par ces recours, l'élève inscrit son discours dans le registre *explicatif*, voire *légal* et ne formule que partiellement la règle sans mentionner le contexte d'application. Ces justifications sont considérées comme *formelles scolaires:* en d'autres termes, les élèves concernés interprètent cette égalité comme respectant les lois formelles du calcul algébrique, sans pour autant lui donner du sens.

Exemples : « c'est vrai car on additionne les 2 exposent », « car il faut additionner les puissants »

II – Choix incorrect / Justification incorrecte : légal et formel scolaire :

L'élève utilise des verbes modaux tels que « falloir » et « devoir » pour justifier son choix. Ces verbes inscrivent le discours de l'élève dans une dimension *légale* où le choix n'est pas posé quant à la réalisation du possible mais dicté par une règle incorrecte parfois formulée (comme $a^m a^n = a^{mn}$).

Exemple : « on doit faire une soustraction entre les deux chiffres du haut »

Question 2: $a^2 = 2a$

Caractéristiques mathématiques de l'égalité

L'égalité donnée est fausse. Par ailleurs, elle ne ressemble à aucune règle classique donnée dans les cours d'algèbre.[1].

Analyse linguistique

CC – Choix correct / Justification correcte : argumentatif (opposition) et conceptuel :

L'élève utilise la phrase complexe construite à partir d'une proposition principale et d'une proposition subordonnée circonstancielle. Les deux propositions sont articulées par une locution conjonctive d'opposition : « tandis que », « alors que », « et non pas ». L'élève marque explicitement la relation d'opposition existante entre les deux propositions qu'il formule. Il inscrit son discours dans le registre argumentatif qui révèle un niveau conceptuel.

Exemple : « a^2 signifie a×a alors que 2a signifie a×2 »

CC – Choix correct / Justification correcte : argumentatif (coordination) et conceptuel :

Comme pour la catégorie précédente, les élèves utilisent une phrase complexe mais cette fois, les propositions principale et subordonnée sont liées par une conjonction de coordination « et ». Le lien entre les deux propositions est établi mais non spécifié, contrairement au cas précédent où l'opposition est explicitement marquée. Exemple : « car le premier ça fait a fois a et le deuxième ça fait 2 fois a »

CP – Choix correct / Justification partielle : descriptif et contextuel

Dans cette catégorie, le lien avec un membre de l'égalité devient implicite : seul un des deux membres est considéré. L'élève ne porte son intérêt que sur l'expression algébrique de ce membre et décrit des expressions algébriques équivalentes en les introduisant par « c'est », « ça fait ». Il inscrit son discours dans un registre descriptif et le niveau reste contextuel.

Exemple : « c'est « a+a » qui est égal à 2a. »

II – Choix incorrect / Justification incorrecte : explicatif et formel

L'élève a recours à la causalité marquée dans son discours par l'utilisation de « car » ou « c'est vrai car ». Son discours s'ancre dans le registre explicatif en utilisant des arguments formels erronés.

Exemple : « car le a au carré vaut bien deux fois a »

1 En ceci elle est différente d'une égalité fausse comme $a^3 a^2 = a^6$ qui ressemble à la règle $a^m a^n = a^{m+n}$.

Code, Registres de discours, Niveau d'entrée dans l'algèbre, Marqueurs	Exemple
CC, *Argumentatif opposition,* Conceptuel *tandis que, et non pas, alors que* (11, 9 Gr 1, 2 Gr 2)	a^2 signifie a×a alors que 2a signifie a×2
CC, *Argumentatif coordination,* Conceptuel *et* (9, 5 Gr 1, 4 Gr 2)	car le premier ça fait a fois a et le deuxième ça fait 2 fois a
CP, *Descriptif,* Conceptuel *C'est, ça fait* (5, 3 Gr 1, 2 Gr 2)	c'est « a+a » qui est égal à 2a.
II, *explicatif,* formel *car, c'est vrai car* (6 Gr 2)	c'est vrai car la lettre a qui est élevé au carré donne 2a (a×a = 2a).

Tableau 6 : réponses associées à l'égalité $a^2 = 2a$

Question 3: $2a^2 = (2a)^2$

Caractéristiques mathématiques de l'égalité

L'égalité est fausse. Comme la précédente elle ne ressemble à aucune règle classique donnée dans les cours d'algèbre. Notons cependant que le membre de droite de l'égalité contient des parenthèses : les professeurs de mathématiques explicitent souvent le rôle des parenthèses dans le calcul numérique ou algébrique.

Analyse linguistique

CC – Choix correct / Justification correcte : argumentatif (opposition) et conceptuel :
Comme pour la question précédente, l'élève utilise une phrase complexe comprenant une proposition principale et une proposition subordonnée. Les propositions sont articulées par une locution conjonctive qui marque l'opposition entre les deux membres de l'égalité donnée : « tandis que », « alors que », « et non pas ». L'élève met parfois en avant le rôle joué par les parenthèses dans cette opposition. Son discours se situe dans le registre argumentatif et est de niveau conceptuel.
Exemple : « Dans la première partie de l'équation, seul a est au carré alors que dans la deuxième, le produit de 2a est au carré »

CC – Choix correct / Justification correcte : argumentatif (coordination) et conceptuel :
L'élève utilise à nouveau une phrase complexe de structure similaire à la précédente mais les propositions sont liées par une conjonction de coordination « et ». Il juxtapose les propositions, en considérant séparément chaque membre de l'égalité, sans expliciter d'opposition.
Exemple : « car $2a^2$, c'est a qui est au carré. Et $(2a)^2$, c'est 2a qui est au carré. »

Code, Registres de discours, Niveau d'entrée dans l'algèbre, Marqueurs	Exemple
CC, *Argumentation opposition*, Conceptuel *tandis que, et non pas, alors que* (14, 12 Gr 1, 2, Gr 2)	Dans la première partie de l'équation, seul a est au carré alors que dans la deuxième, le produit de 2a est au carré
CC, *Argumentation coordination*, Conceptuel *Et, donc* (5, 3 Gr 1, 2 Gr 2)	car $2a^2$, c'est a qui est au carré. Et $(2a)^2$, c'est 2a qui est au carré.
CP, *Descriptif restriction*, Contextuel *C'est juste, seulement* (4, 2 Gr 1, 2 Gr 2)	comme il n'y a pas de parenthèses, c'est seulement la valeur « a » que l'on multiplie par elle-même
II, *Explicatif,* Formel *car* (2 Gr 2)	car les deux résultats sont égaux.
II, *Légal*, Formel scolaire *On peut, on a le droit* (2 Gr 2)	on a le droit de mettre des parenthèses à un chiffre

Tableau 7 : réponses associées à l'égalité $2a^2 = (2a)^2$

CP – Choix correct / Justification partielle : descriptif (restriction) et contextuel :
Le lien avec l'un des membres de l'égalité reste implicite. L'élève considère un seul des deux membres de l'égalité : il s'intéresse seulement au membre de droite, en introduisant sa description par « c'est ». L'élève explicite souvent le fait que le carré ne s'applique qu'à la variable *a* (à cause de l'absence de parenthèses) en utilisant des termes comme « juste », « seulement ». Son discours est descriptif et ancré au niveau contextuel.
Exemple : « comme il n'y a pas de parenthèses, c'est seulement la valeur « a » que l'on multiplie par elle-même.»

II – Choix incorrect / Justification incorrecte : legal et formel scolaire:
L'élève utilise souvent des verbes modaux, comme « pouvoir », ou « avoir le droit » pour justifier son choix incorrect, en se centrant sur le rôle joué par les parenthèses. Par ces recours, il donne une dimension légale à son discours qui le situe au niveau formel scolaire.

Exemple : « on peut metrtre une parenthese, cela ne change rien sauf lors d'un calcul, quand il y a des prioritées. »

II – Choix incorrect / Justification incorrecte : explicatif et formel
L'élève a recours à la causalité en commençant son discours justificatif par « car ». Son discours relève de l'explication avec l'utilisation d'arguments formels erronés.
Exemple :« car on multiplie de gauche à droite »

Interprétation de l'analyse d'un point de vue didactique

L'étude empirique que nous venons de développer, montre que les élèves utilisent de façon récurrente des structures linguistiques formelles pour exprimer leurs justifications en langage mathurel et que l'usage de ces formes révèle des niveaux d'avancée différents dans la pensée algébrique. Il s'agit d'interpréter ces résultats d'un point de vue didactique.

Concernant l'égalité $a^3a^2=a^5$
Le fait que cette égalité vraie soit proche d'une règle souvent vue en cours par les élèves, joue un rôle essentiel. Cela décourage chez de nombreux élèves l'usage du registre argumentatif. Seuls ceux ayant a priori atteint un niveau avancé dans la compétence algébrique produisent une justification de niveau conceptuel : les élèves concernés ont répondu correctement à l'ensemble des questions et ont passé l'ensemble du test avec un certain succès. Inversement, l'usage du registre descriptif est favorisé : il s'appuie sur le contexte représenté par l'égalité donnée pour le rapprocher de la règle « bien connue » : $a^ma^n=a^{m+n}$. Pour les mêmes raisons, les élèves ayant fait un choix incorrect, reconnaissant néanmoins une règle vue dans le cours, ont souvent recours à un discours de type légal.

Concernant l'égalité $a^2=2a$
À l'inverse de l'égalité précédente, le fait que l'égalité soit fausse et non « semi-calquée » sur une règle d'algèbre du cours, fait que le registre descriptif est relativement peu utilisé. On note également de ce fait une quasi-absence du registre légal en ce qui concerne les justifications de réponses incorrectes qui restent dans un mode explicatif. En revanche, dans ce cas de figure, un nombre plus important d'élèves a recours à un registre de type argumentatif. Ces derniers justifient leur choix en mettant en relation les deux membres de l'égalité en spécifiant le type de relation : l'opposition, ou à travers l'emploi d'une conjonction de coordination.

Concernant l'égalité $2a^2=(2a)^2$
Le cas de cette égalité est à rapprocher du cas précédent, à savoir la mise en relation des deux membres de l'égalité fausse. Néanmoins, concernant les élèves ayant apporté une réponse incorrecte, la justification produite s'inscrit à nouveau parfois dans le registre du légal. Ce recours au légal n'est plus sollicité par la forme de l'égalité (qui ne réfère pas à une règle vue en cours) mais plutôt par un élément précis de son contenu, les parenthèses. La parenthèse a une fonction et c'est cette fonction qui est explicitée dans le registre légal. D'une part les parenthèses semblent attirer l'attention, comme seul élément variant d'un membre de l'égalité à l'autre. D'autre part, les élèves tentent de rappeler le rôle de ce symbole, souvent mis en avant ou explicité par les professeurs en cours de mathématiques.

Résultats
L'interprétation de nos analyses nous a permis de mettre en relation d'une part l'usage de structures de phrase complexes et de marqueurs linguistiques et, d'autre part, des caractéristiques de la tâche considérée (par exemple la structure syntaxique marquant l'opposition pour le cas d'égalités fausses). Elle nous a permis de dégager des variables relatives aux égalités données qui nous paraissent influencer le type de discours justificatif produit par les élèves :
- égalité vraie / fausse,
- complexité des expressions algébriques (qui peuvent être développées, réduites, etc.),
- présence dans l'égalité d'éléments variants/invariants (d'un membre à l'autre),
- présence dans l'égalité d'éléments spécifiques des cours d'algèbre (tels les parenthèses, ou la proximité apparente de règles algébriques usuelles, etc.).

Perspectives

Cette étude exploratoire ouvre des perspectives selon trois directions.
Tout d'abord, il s'agit de tester certaines des hypothèses *a priori* sur lesquelles nous nous sommes appuyées tant d'un point de vue didactique que linguistique : nous souhaitons confronter nos suppositions sur l'avancée en algèbre d'élèves à partir de l'analyse linguistique de leurs discours, aux profils cognitifs en algèbre établis par le logiciel Pépite (à partir de l'analyse de leurs réponses à l'ensemble du test). Une confrontation systématique entre réussite au test, profils cognitifs et analyse linguistique des productions d'élèves permettra d'attester de la pertinence de notre étude.

Nous envisageons également de tester la typologie élaborée sur un autre corpus de données. Il s'agit de proposer un test comportant des questions relatives au même type de tâches (*valider ou invalider l'égalité de deux expressions algébriques*), élaborées en modulant les variables dégagées par l'étude (égalité vraie/fausse, complexité des expressions algébriques, etc.).
Enfin il s'agit d'améliorer le système de diagnostic automatique de Pépite. Sur la base de nos premiers résultats, nous pouvons envisager d'implémenter un système capable premièrement d'évaluer le caractère correct/incorrect des justifications produites en langage mathurel par les élèves. Deuxièmement en étudiant plus en profondeur la justifications à l'aide des critères linguistiques dégagés (marqueurs et structures grammaticales) il est possible d'assigner un niveau de rationalité algébrique pour compléter le profil cognitif existant.

Références

V. Aleven, O. Popescu, K. Koedinger, 2002 Pilot-Testing a Tutorial Dialog System That Supports Self-Explanation, Proceedings of ITS'2002, Biarritz (France), 5-8 juin 2002. Cerri S., Gouardéres G., Paraguaçu F. (eds.), Springer-Verlag, p. 344-354.

I. Arroyo, R. Conejo, E. Guzman, B. P. Woolf, 2001, An Adaptive Web-Based Component for Cognitive Ability Estimation, in J.D. Moore et al. (Eds.), Artificial Intelligence in Education, IOS Press, 2001, 456-466.

M. Artigue, T. Assude, B. Grugeon, A. Lenfant, 2001 Teaching and Learning Algebra : approaching complexity trough complementary perspectives, In Helen Chick, Kaye Stacey, Jill Vincent, John Vincent (Eds), The future of the Teaching and Learning of Algebra, Proceedings of 12 th ICMI Study Conference, The University of Melbourne, Australia, December 9-14, 2001, 21-32.

J. L. Austin, 1962, How to do the things with words. Cambridge, Cambridge University Press.

O. Ducrot, 1984, Le dire et le dit, Paris, Minuit

É. Delozanne, D. Prévit, B. Grugeon, P. Jacoboni, 2003, Supporting teachers when diagnosing their students in algebra, Workshop Advanced Technologies for Mathematics Education, Supplementary Proceedings of Artificial Intelligence in Education, Sydney, July 2003, 461-470.

B. Grugeon, 1995, Etude des rapports institutionnels et des rapports personnels des élèves à l'algèbre élémentaire dans la transition entre deux cycles d'enseignement : BEP et Première G, thèse de doctorat, Université Paris VII.

S. Jean, E. Delozanne, P. Jacoboni, B. Grugeon, 1999, A diagnostic based on a qualitative model of competence in elementary algebra, in S. Lajoie, M. Vivet, AI&ED'99, IOS Press, Amsterdam, , Le Mans 491-498

N. Hefferman, K. Koedinger, 2002, An Intelligent Tutoring System Incorporating a Model of an Experienced Human Tutor, Proceedings of ITS'2002, Biarritz (France), 5-8 juin 2002. Cerri S., Gouardéres G., Paraguaçu F. (eds.), Springer-Verlag, p. 596-608.

J. Kay, 2000, Stereotypes, Student Models and scrutability, in G. Gauthier, C. Frasson, K. VanLehn (eds.), Intelligent Tutoring Systems, 5th International Conference, ITS 2000, Montreal, Canada, 19-30.

C. Kieran, 1992, The Learning and Teaching of School Algebra, Handbook of research on Mathematics Teaching and Learning, Douglas Grouws Ed. Macmillan publishing company.

Pépite, 2004 : http://pepite.univ-lemans.fr

D. Prévit, E. Delozanne, B. Grugeon, Modélisation cognitive en algèbre élémentaire, Proceedings of TICE 2004 « Technologies de l'Information et de la Connaissance dans l'Enseignement Supérieur et l'Industrie », (ce volume).

D. Rasseneur, E. Delozanne, P. Jacoboni, B. Grugeon,2002 Learning with virtual agents: Competition and Cooperation in AMICO, Proceedings of ITS'2002, Biarritz (France),. Cerri S., Gouardéres G., Paraguaçu F. (eds.), Springer-Verlag, p. 61-70.

C. P. Rosé, A. Roque, D. Bhembe, K. VanLehn, 2003a, A Hybrid Text Classification Approach for Analysis of Student Essays, Proceedings of the HLT-NAACL 03 Workshop on Educational Applications of NLP.

C. P. Rosé, A. Roque, D. Bhembe, K. VanLehn, 2003b, Overcoming the Knowledge Engineering Bottleneck for Understanding Student Language Input, in Ulrich Hoppe, Felisa Verdejo, Judy Kay (eds.) Proceedings of Artificial Intelligence in Education, Sydney, IOS Press, Amsterdam, 315-322.

J. R. Searle, 1969, Speech Acts, An essay in the Philosophy of Language, Cambridge, CUP.