

Cet article est disponible en ligne à l'adresse :

[http://www.cairn.info/article.php?ID\\_REVUE=CDLE&ID\\_NUMPUBLIE=CDLE\\_023&ID\\_ARTICLE=CDLE\\_023\\_0103](http://www.cairn.info/article.php?ID_REVUE=CDLE&ID_NUMPUBLIE=CDLE_023&ID_ARTICLE=CDLE_023_0103)

---

## Systèmes d'observation de classes et prise en compte de la complexité des événements scolaires

par Philippe DESSUS

| Université de Picardie Jules Vernes | Carrefours de l'éducation

2007/1 - n° 23

ISSN 1262-3490 | pages 103 à 117

---

Pour citer cet article :

— Dessus P., Systèmes d'observation de classes et prise en compte de la complexité des événements scolaires, Carrefours de l'éducation 2007/1, n° 23, p. 103-117.

---

Distribution électronique Cairn pour Université de Picardie Jules Vernes.

© Université de Picardie Jules Vernes. Tous droits réservés pour tous pays.

La reproduction ou représentation de cet article, notamment par photocopie, n'est autorisée que dans les limites des conditions générales d'utilisation du site ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Toute autre reproduction ou représentation, en tout ou partie, sous quelque forme et de quelque manière que ce soit, est interdite sauf accord préalable et écrit de l'éditeur, en dehors des cas prévus par la législation en vigueur en France. Il est précisé que son stockage dans une base de données est également interdit.

# Systemes d'observation de classes et prise en compte de la complexité des événements scolaires

▲ **Philippe Dessus**

*Philippe.Dessus@upmf-grenoble.fr*



On peut noter, dans la recherche en éducation francophone, un regain de l'usage des systèmes d'observation de classe (e.g., Berdot, Blanchard-Laville, & Chaussecourte, 2003; Leutenegger, 2001). Ces systèmes, élaborés pour la plupart dans les années 1950, peuvent être classés en trois grandes catégories (Fraser, 1998): 1° celle utilisant des observateurs entraînés codant les événements perçus (parfois nommées mesures de faible inférence); 2° celle recueillant par questionnaires les avis des enseignants et élèves de la classe; 3° celle utilisant des méthodes ethnographiques. Les données sur lesquelles s'appuient les deux premières méthodes sont quantitatives, la dernière utilisant des données qualitatives. Historiquement, les secondes méthodes se sont d'abord développées, suivies, dès les années 1950, par les premières, censées pouvoir mieux évaluer les aspects psycho-sociaux. Plus récemment, les méthodes ethnographiques ont supplanté les deux autres, en plaçant le texte (narration de protagonistes, discours sur l'action) en position d'être interprété. Cette position, en rendant l'action fixée, autonome et contextualisée autorise une interprétation estimée plus riche et ouverte par les chercheurs se réclamant de ce courant. Toutes ces méthodes ont bien évidemment des avantages et des inconvénients. La première fait moins appel à la subjectivité, mais ne tient pas compte des représentations des différents protagonistes et *vice versa* pour la seconde. La troisième méthode ne donne des interprétations valides qu'à un niveau individuel, rendant les généralisations très hasardeuses.

Quels que soient les intérêts de chacune de ces méthodes, il nous semble intéressant de reprendre l'étude de ces systèmes, à la lumière des résultats récents de la recherche sur l'enseignement. En effet, lorsqu'ils ont été conçus, l'enseignement était principalement vu comme un système d'entrée- (ensei-

gnement) sortie (résultats d'apprentissage) et les interactions à l'intérieur d'un tel système relativement déterministes. Plus récemment, on a pu considérer la classe comme un environnement dynamique complexe (e.g., Rogalski, 2003), ce qui devrait changer les manières de l'observer et de l'analyser. L'objet de cet article n'est pas de faire une revue de la question exhaustive de toutes les méthodes d'observation d'environnements scolaires (e.g., Medley & Mitzel, 1963; Postic & de Ketele, 1988; Simon & Boyer, 1970), mais de passer plus précisément en revue la recherche concernant le deuxième type de méthodes. Dans une première section, nous définissons les principaux termes liés à cette problématique, pour exposer ensuite, dans les deux sections suivantes, les différents systèmes, et décrire les étapes par lesquelles les systèmes d'observation passent d'une saisie d'événements par un observateur à un système de codes interprétable. Dans deux dernières sections, nous reprendrons les caractéristiques d'un environnement dynamique et présenterons un système d'observation et d'analyse qui tient compte de certaines de ces dernières.

### **Définitions: observation et grilles d'observation**

Pour commencer, il est important de noter que les premières méthodes d'observation directes (i.e., dans lesquelles l'observateur est directement confronté à l'événement, voir Kohn & Nègre, 1991) systématiques utilisées dans le domaine de l'éducation ou la psychologie ont pour origine l'éthologie, c'est pourquoi la définition donnée par Vauclair (1984), un spécialiste de la cognition des animaux, convient également à notre propos :

L'observation [...] désigne tout recueil de données établi à partir de la description du comportement spontané des animaux dans leur milieu naturel. » (id., p. 123, c'est l'auteur qui souligne).

Mais cette simple définition contient un présupposé à expliciter : que l'observateur n'intervient pas, ou le moins possible, dans la situation observée (observation naturaliste), ce qui n'est pas nécessairement le cas dans la recherche en éducation. Toutefois, l'inclusion de l'observateur dans la situation d'enseignement observée est nécessaire, car ce dernier est animé d'intentions particulières. Bunge (1984, p. 47), par exemple, souligne que l'observation

[est] une perception *préméditée* et *éclairée* : préméditée ou délibérée car elle est faite dans un but bien défini ; éclairée car elle est guidée, d'une façon ou d'une autre, par un corps de connaissances. (ibid., c'est l'auteur qui souligne)

L'observateur, de plus, est en général muni d'une « grille d'observation », que de Ketele (1987) définit ainsi :

Une grille d'observation est un système d'observation : systématique, attributive, allospective [pour observer les autres], visant à recueillir des faits et non des représentations, menée par un ou plusieurs observateurs indépendants et dans laquelle les procédures de sélection, de provocation,

d'enregistrement et de codage des « attributs » à observer sont déterminées le plus rigoureusement possible. (*id.*, p. 127)

Nous ne pouvons qu'être en accord avec cette définition très formelle : la rigueur est toujours à mettre de son côté quand il s'agit de méthode et il vaut mieux essayer de tenter de recueillir des faits plutôt que des représentations, même si cela risque parfois de ne rester qu'une tentative (que faire des représentations de l'observateur?). Nous pouvons noter que cette définition laisse de côté les éléments les plus importants d'une grille : que veut dire observer « systématiquement » une situation? Le plus gênant, dans cette définition, est qu'elle laisse le but de l'observation en dehors du propos, comme si l'on pouvait observer une situation sans but particulier. Nous voyons donc que les très nombreux instruments d'observation de situations d'enseignement répondent à autant de manières d'envisager le but de cette activité. Simon et Boyer (1970) font une hypothèse d'école judicieuse : tous les systèmes d'observation se répartissent le long d'un *continuum* entre, d'un côté, un système idéal qui permettrait de recueillir toutes les catégories de comportements possibles et, de l'autre côté, un système fruste avec seulement deux catégories, « quelqu'un parle » et « personne ne parle ».

Même avec une grille simplifiée à deux catégories, il reste deux problèmes importants : tout d'abord, déterminer précisément les critères permettant de classer les différents événements observés dans l'une ou l'autre des catégories ; et, avant même cela, déterminer quand on passe d'un événement à l'autre. S'il est aisé, en effet, de repérer des objets dans une classe, il est plus difficile de repérer des événements. Une définition hiérarchique des différents événements survenant en classe ne précise pas les limites entre chaque entité : comment savoir exactement où se termine une séquence? Il est possible de se fier, dans l'enseignement secondaire, aux changements horaires de cours, mais un enseignant ne peut-il pas continuer une séquence une prochaine fois? Nous reviendrons sur ces points plus loin mais, auparavant, nous passons en revue les systèmes d'observation existants.

### **Les différents systèmes d'observation : de catégories, signes ou échelles**

Les différents systèmes d'observation de situations scolaires sont traditionnellement classés selon qu'ils utilisent un système de codage en catégories, signes ou échelles. Mettre en place un *système de catégories*,

c'est limiter son observation à un aspect du comportement de la classe, déterminer une unité de comportement, et construire un ensemble fini de catégories dans lesquelles chaque unité peut être classée dans une et une seule de ces dernières. (Medley & Mitzel, 1963, p. 298)

Fonder ses observations sur un *système de signes*, à l'inverse, c'est définir *a priori* une liste d'événements spécifiques (actions, incidents, etc.) pouvant ou non surve-

nir pendant un intervalle de temps, et enregistrer leur fréquence dans cet intervalle. Le système de catégories vise à l'exhaustivité, ou plutôt à la complétude – tout événement est censé prendre place dans une catégorie, formant ensemble un tout cohérent. Un système de signes est censé décrire précisément un ensemble de comportements, et les comportements non observés sont laissés de côté. Les systèmes de signes ont été plutôt utilisés dans les études visant à rendre compte de l'efficacité de l'enseignant ; les systèmes de catégories dans celles visant à rendre compte du climat de la classe. Comme le notent Medley et Mitzel (1963), l'activité d'un observateur dans l'un ou l'autre des systèmes est radicalement différent. Dans un système de catégories, il s'agit d'attribuer une *tonalité* (i.e., un climat) à un épisode, presque indépendamment des nombreux événements s'y déroulant ; dans un système de signes, à l'inverse, un observateur sera vigilant à un éventail bien plus large d'événements, mais ne recueillera que ceux qui sont dans sa grille.

Décrivons l'un des plus célèbres systèmes, la grille de Flanders (1976), FIAC (pour *Flanders Interaction Analysis Categories*, ou Catégories d'analyse de l'interaction de Flanders). Cette grille est fondée sur l'observation de catégories d'événements survenant dans une classe, et son but initial est d'évaluer le niveau de directivité de l'enseignant (autoritaire vs. démocratique). L'observateur code, par intervalle de temps (e.g., toutes les trois secondes), toutes les occurrences d'un événement, qu'il détermine au sein d'une liste de dix événements, sept prenant l'enseignant pour objet (1. Accepte les sentiments, 2. Louanges ou encouragements, 3. Accepte ou utilise les idées des élèves, 4. Pose des questions, 5. Exposé, 6. Donner des directives, 7. Critiquer ou justifier l'autorité), deux centrés sur les élèves (8. Réponse de l'élève, 9. L'élève parle), et un dernier codant le silence ou le chahut. Il faut noter que l'observateur n'a pas à se préoccuper de segmenter ces événements en entités plus petites. Vu l'intervalle de temps très réduit, le même événement peut être observé et noté de nombreuses fois. Ce qui fait la spécificité de cette méthode est le type de lien de cause à effet qui est inféré entre deux catégories successives observées. Chaque catégorie est ensuite consignée dans un tableau à deux dimensions, de 10 par 10, dans lequel est reportée chaque paire contiguë d'événements. Ainsi, une partie de l'information initialement saisie est perdue (notamment la possibilité de retrouver des segments de séances), mais ce type de réduction permet de catégoriser les différentes cooccurrences d'événements.

Les *échelles d'évaluation* sont souvent présentées à part dans la littérature (Evertson & Green, 1986), alors qu'elles ne sont, tout compte fait, que des échelles de catégories avec un type de codage ordinal ou d'intervalle (de Ketele, 1987). Il s'agit de noter, selon une échelle, le niveau observé en classe d'un certain nombre de caractéristiques de l'enseignant, en relation avec son efficacité. La fidélité inter-observateurs de ce type d'échelles a bien sûr été critiquée. D'autres échelles ont pallié en partie la subjectivité du jugement, en se centrant sur le comportement de l'enseignant, l'observateur jugeant, non pas le climat de la classe, mais le comportement de l'enseignant ou des élèves (Wragg, 1994). Toutefois, comme l'indique

ce dernier, de telles échelles sont plus utiles à l'observateur témoin de nombreuses situations d'enseignement différentes, qu'au chercheur. Elles nécessitent de ce dernier un haut niveau d'interprétation, contrairement aux systèmes de catégories. De nombreux autres systèmes, souvent très sophistiqués, ont été élaborés à partir de ces précurseurs. Quels sont leurs points communs, leurs principales étapes de traitement ?

### **Comment fonctionnent les systèmes d'observation ?**

Un système d'observation permet à un observateur de transformer certaines données qu'il perçoit de la classe en informations pouvant être traitées et de caractériser les événements observés. Pour Kohn et Nègre (1991), mettre en œuvre une observation, c'est passer par deux étapes : *digitaliser*, c'est-à-dire « [...] standardiser les données et [...] déterminer des unités de compte. » (*id.*, p. 151) et *inférer*, de la manière la plus valide possible, une signification de ces données, c'est-à-dire extraire les seuls éléments signifiants et les regrouper par similarité. Quel que soit le courant théorique de leur concepteur, nous montrerons ici que les grilles sont censées aider l'observateur à opérer, dans l'ordre :

– Une *segmentation spatiale et temporelle* des événements. Les événements se déroulant continuellement dans l'environnement de classe, il est nécessaire de déterminer à la fois une unité de temps et d'espace permettant de se les représenter de manière discrète.

– Un *filtrage*, soit une discrimination entre les événements pertinents et non pertinents ; ce filtrage se réalise souvent par la détermination *a priori* d'un ensemble de catégories, dans lesquelles chaque événement sera classé. Ces deux premières activités sont souvent regroupées sous le terme « échantillonnage », mais il importe, à notre avis, de les distinguer. Si la segmentation nécessite l'interprétation de l'intention des protagonistes de l'événement observé, le filtrage implique une position épistémologique à propos des catégories conçues par le chercheur (*i.e.*, lui permettant de trier entre ce qui est à considérer de la situation et ce qui ne l'est pas).

– Une manière de *coder*, pour les consigner, les données perçues par l'observateur comme pertinentes. Cela se réalise notamment par l'utilisation d'un système de notation particulier à la grille, que des systèmes informatiques peuvent aider. Notons que ce codage est une *traduction*, c'est-à-dire qu'il introduit nécessairement une nouvelle distorsion entre les données filtrées et les données notées.

– *Réaliser une réduction des données*. Même si les trois opérations précédentes ont déjà sensiblement réduit les données parvenues à l'observateur, il est souvent nécessaire d'opérer une nouvelle réduction des données consignées, souvent – mais pas nécessairement – par le biais d'une méthode statistique. Le but premier de cette réduction est de permettre la mise en valeur de rapports de causalité entre données.

Exposons maintenant un bref état de la question selon les étapes ci-dessus.

## **Segmentation des données**

Tout enseignant, et tout observateur, segmente nécessairement les données se présentant à lui. C'est à notre avis un des problèmes les plus importants – pourtant assez peu mentionné dans la littérature – auquel est confronté un observateur de situations. Il doit rendre discret un processus continu en repérant les ruptures. Si certaines ruptures sont aisées à repérer (passage à une autre leçon, moment de flottement, etc.), le repérage de la plupart d'entre elles nécessite une bonne connaissance de ce qu'est le métier d'enseigner. Ces ruptures peuvent être caractérisées de plusieurs manières. La plus simple (mais aussi la plus arbitraire) est de les segmenter temporellement. L'observateur se donne un tempo interne (*i.e.*, une valeur de durée au-delà de laquelle il peut noter un nouvel événement), ce qui détermine de façon nette le niveau de grain des événements observés. Par exemple, toutes les trois secondes, l'observateur scrutera l'environnement à la recherche de l'événement en cours et le notera. Il est aussi possible de récupérer un échantillon temporel (récupérer les *n* minutes après le début de chaque séance) ou spatial (s'intéresser à tel ou tel élève) des événements scolaires. Il est aussi possible d'être attentif à tous les indices du discours de l'enseignant laissant supposer que ce dernier va passer à autre chose (*e.g.*, « alors », « bon », « donc », « nous avons vu que... »).

## **Filtrage des données**

Il s'agit ensuite de discriminer les données pertinentes des autres, ce qui dépend du but du chercheur, et fait intervenir des catégories d'événements. Wragg (1994, p. 20) en donne une liste, qui recoupe l'ancienne catégorisation de Simon et Boyer (1970) : caractéristiques personnelles, interaction verbale et non verbale, gestion de la classe, habiletés professionnelles, matériel d'enseignement, aspects affectifs, cognitifs, aspects sociologiques. Cette liste, un peu fourre-tout, montre que les buts d'observation d'une classe peuvent être divers et que les catégories se recourent largement : certaines interactions verbales peuvent aussi être rangées dans la catégorie « gestion de la classe ».

## **Les unités de codage**

L'étape suivante consiste à consigner de manière stable les événements préalablement observés et filtrés. Il est possible de classer tout événement survenant dans une classe suivant une hiérarchie à deux dimensions. Une hiérarchie « partie-de », partonomique, et une hiérarchie « sorte-de » taxonomique (Zacks & Tversky, 2001). La première est une hiérarchie temporelle, la seconde est fonctionnelle. Voici différents événements scolaires selon ces deux hiérarchies. La hiérarchie « partie-de » comprendrait les éléments suivants (il existe d'autres niveaux de description, *e.g.*, le trimestre, mais ils ont un rôle plus institutionnel que lié au contenu, il est donc difficile de les insérer comme éléments de taxonomie) :

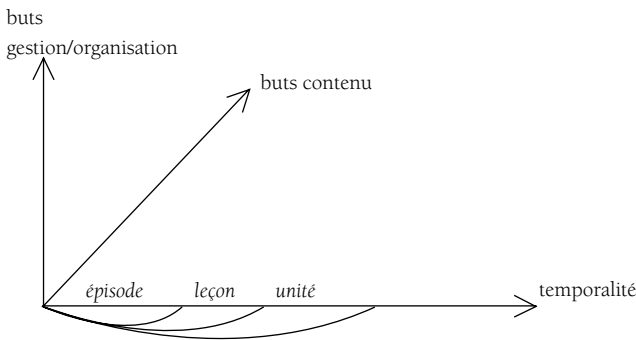
– *Épisode* : c'est la plus petite séquence d'observation pouvant être reliée à un but d'enseignement ou d'apprentissage. On peut également la définir selon la tâche remplie par les élèves, sachant que les tâches contenues dans un seul épisode peuvent être multiples ;

– *Leçon* : comprend plusieurs épisodes contigus dans le temps (une leçon comprend plusieurs épisodes), et est parfois nommée séance.

– *Unité* : c'est la plus large séquence d'observation reliée à un but d'enseignement (e. g, enseigner les événements de la Révolution Française). Une unité, parfois appelée séquence, comprend donc plusieurs leçons.

Passons maintenant à la hiérarchie « sorte de ». D'un point de vue fonctionnel, l'activité de l'enseignant concourt à remplir deux types de buts (Leinhardt & Greeno, 1986). Des buts liés à l'organisation (préparation) et la gestion en direct de la classe, et des buts liés au contenu enseigné. Ainsi, l'enseignant est face à une double hiérarchie de type « sorte-de » : l'une en rapport avec le contenu et l'autre en rapport avec la manière de gérer/organiser la classe (voir figure 1).

**Figure 1 – Hiérarchies des événements d'une situation d'enseignement.**



Ces hiérarchies peuvent être utilisées pour coder ces différents événements survenant dans un environnement scolaire. Les manières de coder un événement sont classiques et bien documentées dans la littérature, et comprennent au moins les catégories suivantes (Zacks & Tversky, 2001) : espace, temps, intentionnalité (ou motivations), causalité, protagonistes et/ou objets. Il s'agit bien de comprendre que ces unités ne sont présentes que par commodité. Il serait tout à fait possible que l'observateur note ce qu'il observe, en direct et en langage naturel, mais cela donnerait des protocoles bien trop variables pour être analysés. Le choix des unités de codage consiste donc bien en une première réduction de dimension de la réalité observée.

## Réduction

Le but de la quasi-totalité des grilles d'observation n'est pas tant de décrire les différents événements survenant dans l'environnement, mais plutôt de faire des liens entre eux afin d'y attribuer un sens. Ainsi, pour reprendre l'exemple de la grille de Flanders, il existe de nombreux indices calculables à partir des données recueillies. Certains sont simples, comme le pourcentage de parole de l'enseignant ou des élèves. D'autres sont plus sophistiqués, comme le pourcentage de réponses de l'enseignant (*Teacher response ratio*) rendant compte de la tendance de l'enseignant à réagir aux idées et sentiments des élèves, ou le pourcentage d'événements de la croix du contenu (*Content cross ratio*) exprimant le degré de centration de l'enseignant sur le contenu enseigné. Cette croix du contenu désigne les lignes et les colonnes des catégories 4 et 5 qui, étant centrales, forment une croix dans le tableau d'analyse.

Il faut noter que, comme ce tableau d'analyse, à deux dimensions, reporte chaque événement et son suivant comme autant de paires, les indices calculés à partir de ces données présupposent que le premier événement cause le second, *du seul fait qu'ils sont successifs*. De nombreuses grilles et systèmes d'observation reprennent ainsi le présupposé de Flanders (1976) : un événement donné est la cause du suivant, qui est en retour la conséquence du précédent. Ce présupposé, directement lié au but général de la FIAC, qui était de déterminer le degré de liberté laissé aux élèves par l'enseignant, a pu s'avérer gênant lorsqu'il a été reconduit sans examen particulier dans d'autres systèmes d'observation. Il est important de souligner cet *a priori*, à notre avis insuffisamment relevé dans la littérature. Flanders avait pourtant ce souci à l'esprit, car l'usage de sa grille ne se résume pas à repérer des *paires* d'événements survenant souvent, mais aussi et surtout à repérer leurs *successions* fréquentes. Par exemple, Flanders (1967) a mis au jour des *patterns* de comportements de l'enseignant qu'il analyse en cycles de paires d'interaction. Certaines méthodes statistiques ont été également utilisées à ces fins de réduction de l'information recueillie. Souvent fondées sur des analyses factorielles, elles permettent de réduire les dimensions dans lesquelles sont consignés les événements. Par exemple, la méthode d'analyse hiérarchique en grappes (*clusters*), que Swank *et al.* (1989) ont utilisée, sans succès, pour trouver des correspondances entre des observations à un niveau qu'ils appellent « macro » (interactions de l'enseignant au groupe d'élèves) et des observations à un niveau « micro » (interactions de l'enseignant à un élève).

Maintenant que nous avons décrit les principales étapes du processus d'observation impliquées dans ces systèmes, nous allons montrer que certaines caractéristiques des classes, récemment mises au jour dans la littérature, sont peu ou mal prises en compte par ces étapes.

## **Les classes vues comme des environnements dynamiques complexes**

Flanders (1991) s'était lui-même rendu compte que, au vu de la complexité des interactions dans la classe, il était nécessaire d'assouplir la manière de les consigner, afin de rendre compte de la flexibilité des conduites de l'enseignant. Cette flexibilité joue sur la manière de réagir à des événements, et correspond à certaines caractéristiques des systèmes complexes, récemment mises au jour (e.g., Hoc, 1996). Ces caractéristiques, malgré certains avis précurseurs (Doyle, 1986), ont été largement ignorées dans la recherche en éducation (voir toutefois Rogalski, 2003), et en particulier pour la conception de nouveaux systèmes d'observation de classes.

– *Évolution dans le temps.* Les classes sont des systèmes dynamiques, évoluant dans le temps en partie sans l'intervention de leurs superviseurs, les enseignants. Cela a une influence sur le codage des événements se déroulant dans la classe. En effet, l'environnement va pouvoir évoluer sans l'intervention du superviseur, qui peut décider ainsi de *ne pas* décider, et espérer ainsi que la classe retrouve un fonctionnement normal sans intervention. À l'inverse, et contrairement à un environnement statique, une intervention de l'enseignant ne sera pas nécessairement suivie d'effets. Il est donc nécessaire de disposer d'un codage précis pour ces deux types d'issues.

– *Variables cruciales et leur évolution.* Certaines des variables de l'environnement sont cruciales, c'est-à-dire que leur évolution influe sur le comportement général de ce dernier, et doit donc être contrôlée. Pour un environnement scolaire, nous pouvons dire que ces variables évoluent de manière continue plutôt que discrète, avec une pression temporelle pouvant être importante. Cette caractéristique rend problématiques à la fois la segmentation des événements, mais aussi l'option, souvent choisie, de relever un événement pour un intervalle de temps fixe. Ce dernier point peut rendre incomplète la saisie des données du point de vue de l'observateur, et peut nécessiter un recodage *a posteriori* de ces derniers.

– *Lien de causalité entre variables.* Il est probable que toutes les variables cruciales recueillies n'influent pas de la même manière sur le système-classe. Certaines peuvent avoir un rôle plus important que d'autres et, de plus, le lien de causalité entre deux variables peut être complexe. Pour certaines, il peut être direct (e.g., le volume de la parole influe directement sur le niveau de bruit de la classe), pour d'autres, des variables intermédiaires, et parfois difficilement accessibles, peuvent également jouer un rôle (e.g., le niveau d'attention des élèves dépend du contenu du cours, mais aussi du moment de la journée), rendant difficile la mise au jour de la chaîne causale. Cette caractéristique influe à la fois sur la nature des unités de codage choisies (pour observer la causalité d'une variable, il faut déjà qu'elle ait été sélectionnée) et sur la nature de la réduction (une réduction trop importante peut brouiller la mise en évidence de causalités).

– *Comportement du système-classe*. Enfin, les variables cruciales sont plus ou moins opaques, c'est-à-dire que leur accès est plus ou moins aisé. Il est notable que certaines des variables les plus importantes du point de vue de l'enseignant (e.g., le niveau de compréhension des élèves) sont également celles les plus difficilement accessibles, pour l'enseignant ou l'observateur. Cette caractéristique influe sur le fonctionnement général de la grille d'observation, car les événements dans un système opaque seront plus difficilement segmentables, filtrables et codables. Et les solutions à ce problème ne sont guère évidentes. Il est probable que les récentes avancées dans le domaine de l'enseignement à distance en mettent en œuvre certaines.

La principale conséquence de ces différentes caractéristiques est la difficulté, pour l'observateur ou même pour l'enseignant, d'évaluer la causalité des événements qui s'y déroulent. Les systèmes classiques mettent en œuvre, nous l'avons vu, une « causalité limitée », alors que les environnements dynamiques nécessitent la prise en compte d'une causalité plus étendue. En effet, dans de tels environnements, les événements qui s'y déroulent sont polysémiques (un même événement peut avoir plusieurs sens ou causes) et peuvent avoir des homonymes (deux événements différents peuvent avoir le même sens ou la même cause). Nous décrivons maintenant un système de notre conception palliant ce dernier problème.

### **Un outil pour analyser les observations de situations d'enseignement**

Nous avons contribué à ce thème de recherche en concevant un système informatique réalisant une analyse de contenu d'observations de situations d'enseignement recueillies par un observateur situé au fond de la classe (Allègre & Dessus, 2003; Dessus & Allègre, 2004; Dessus, Allègre, & Maurice, 2005). Le but principal de ce système est de rendre compte de la signification complexe des événements scolaires. Si la méthode de codage est classique (utilisation de raccourcis-clavier pour réduire la durée de saisie des événements), la méthode de réduction utilisée est plus originale. Il s'agit d'utiliser l'analyse sémantique latente (LSA, pour *Latent Semantic Analysis*, Landauer & Dumais, 1997), une méthode statistique d'analyse de données textuelles proche d'une analyse factorielle, pour permettre un repérage des régularités des événements observés en classe. La similarité des événements (ou des épisodes) peut être calculée, en partant du principe que des événements similaires se déroulent dans des épisodes similaires, et *vice versa*. Cette définition récursive est appliquée par la méthode d'analyse, et permet, sans établir *a priori* aucune correspondance entre leçons, de classer les différents épisodes selon leur but (i.e., exposé de la consigne, résoudre le problème, corriger le problème). Cette démarche est issue du modèle de la perception d'événements de Zacks et Tversky (2001), qui signalent que :

[L]es unités d'événement peuvent être caractérisées selon leurs propriétés qualitatives ou statisti-

quement, selon le lieu où ces derniers surviennent. Selon cet aspect, ils ressemblent aux unités de parole tels que les phonèmes, syllabes, ou mots. (*id.*, p. 12)

Selon ce modèle, la cooccurrence d'événements n'est pas arbitraire, pour les deux raisons suivantes : – un événement donné peut appartenir à des plans plus larges, ayant une cause, ou un but, partagés avec d'autres événements, cooccurents ou non ; – les protagonistes d'un événement agissent généralement en fonction d'un but. Ces buts et plans peuvent s'exprimer dans un objet de connaissance particulier, le schéma d'événements, ou script, structure cognitive prédisant ou activant des événements sociaux typiques, et ayant une structure hiérarchique partonomique (voir section « *Les unités de codage* »).

Notre travail (voir une description complète dans Dessus, Allègre & Maurice, 2005) part de l'idée que l'enseignant et les élèves, dans une classe, effectuent des unités d'action, des « coups », dans un contexte précis (*i.e.*, un épisode de la résolution d'un problème de mathématiques). Le postulat que nous faisons est qu'il est possible d'analyser, avec LSA, les cooccurrences de ces « coups » suivant les différents contextes dans lesquels ils surviennent, et que LSA sera capable de rendre compte de la signification de chaque « coup », en fonction, d'une part, de son insertion dans tel ou tel épisode et, d'autre part, de ses voisins au sein de l'épisode. LSA va nous permettre de repérer des *patterns* à deux niveaux :

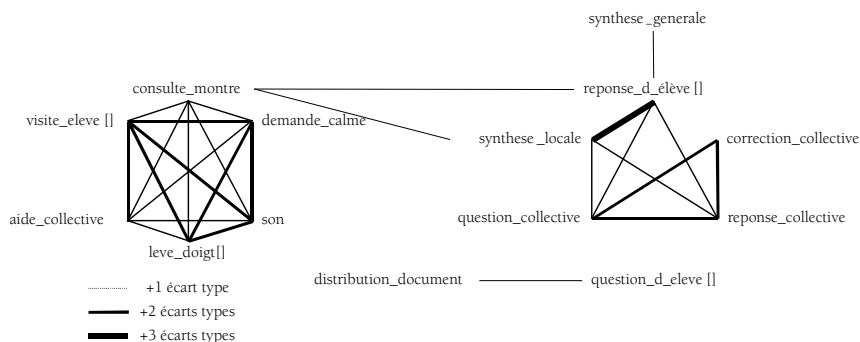
– *Au niveau des épisodes*. LSA permet de comparer directement la proximité de deux épisodes. Nous pourrions ainsi répondre aux questions suivantes : qu'est-ce qui rend deux épisodes proches ? Le fait qu'ils soient dans la même séance ? Qu'ils soient à propos du même objet (*i.e.*, objectif général) ? Nous faisons l'hypothèse que les épisodes ayant un même objet seront plus proches entre eux que les épisodes ayant un autre objet, cette discrimination se faisant uniquement à partir des « coups » qui les composent. Pour vérifier cette hypothèse, nous comparerons systématiquement chaque épisode avec tous les autres, et vérifierons si, au-delà d'un certain seuil (correspondant à la moyenne des proximités interépisodes augmentée d'un écart type), les épisodes les plus proches sont bien de même objet.

– *Au niveau des « coups »*. Nous pourrions déterminer si des « coups », ou événements, sont d'autant plus proches (toujours au-dessus d'un seuil fixé comme précédemment) qu'ils apparaissent dans le même épisode, ou bien dans des épisodes différents, mais d'objet similaire. Par exemple, existe-t-il un (ou des) événement(s) dont la moyenne de proximité avec les autres événements est plus grande que les autres ? Ainsi, on pourra montrer que certains événements, à première vue semblables, sont évalués comme dissemblables, parce qu'ils sont présents dans des épisodes différents. Nous pourrions ainsi construire une carte des événements qui prenne en compte leur réalisation dans les séances observées.

Nous avons récupéré la description 41 leçons de mathématiques d'un même enseignant, dans une même classe (Maurice & Allègre, 2002). Ce corpus comprend environ 2 300 événements, réalisés au sein de 123 épisodes, pouvant être de trois

types (Énoncé du problème, Recherche individuelle ou en groupes de sa solution, Correction du problème). Il a été traité par LSA, qui a projeté les différentes données (événements et épisodes) dans un espace multidimensionnel. Les résultats montrent que notre système est bien capable de discriminer les épisodes en analysant les *patterns* des événements qui les composent. La quasi-totalité des proximités interépisodes supérieures au seuil fixé (soit 0,59) concerne bien des épisodes de même objet. Les seules exceptions concernent les différents épisodes comprenant un seul événement, temporisation, soit trois épisodes sur 3400 comparaisons. LSA arrive donc à discriminer les épisodes ayant le même objet. Cependant, des épisodes d'objet semblable peuvent être évalués comme très proches, alors qu'ils comportent des événements différents. Passons enfin aux proximités entre événements. En comparant les événements entre eux, et filtrant les proximités au-delà d'un seuil arbitraire, nous avons construit la figure 1 ci-dessous, dans laquelle on voit se détacher trois ensembles d'événements interconnectés.

**Figure 2 – Relations de proximité des différents événements. Un lien signifie « les deux événements reliés entretiennent l'un l'autre une proximité supérieure à 0,43 » et son épaisseur est proportionnelle à la proximité (voir cartouche).**



Cette figure représente un réseau de proximité des événements du corpus : on peut constater que la proximité interévénements n'est pas directement reliée à la personne qui exerce les actions (voir notamment le lien entre *distribution\_document* et *question\_d\_élève*), mais est liée aux cooccurrences d'événements dans les épisodes. On peut dégager trois grands ensembles d'événements, ayant en commun des liens aux événements *consulte\_montre*, *reponse\_d\_élève* et *synthèse\_locale*. Un premier ensemble a trait aux événements survenant dans les épisodes liés à la Recherche de la solution à un problème (*son*, *lève\_doigt*, *visite\_élève*, *aide\_collective*, *demande\_calme*). Le deuxième ensemble, lui, est aux épisodes de Correction (*correction\_collective*, *réponse\_collective*, *question\_collective*, *synthèse\_locale*). Enfin, un troisième groupe de deux événements amène à consi-

dérer comme proches `distribution_document` et `question_d_élève`, sans doute parce que ces deux événements sont cooccurrents dans les épisodes Énoncés, et uniquement dans ces derniers, ce qui explique l'absence de liens avec les autres événements. Au vu de cette répartition en trois groupes, nous pouvons noter que les événements scolaires observés sont très typés, c'est-à-dire qu'ils s'inscrivent dans un épisode à la finalité bien établie (Énoncé, Recherche, ou Correction) et que peu d'événements (si ce n'est `consulte_montre`, `réponse_d_élève` et `synthèse_locale`) sont menés indifféremment dans plusieurs épisodes.

Si cette méthode ne peut s'adapter à toutes les caractéristiques de l'environnement de classe listées dans la section précédente, nous avons montré qu'une méthode de traitement statistique permettait de mettre au jour des liens de proximité, et donc vraisemblablement de causalité, interévénements et interépisodes, et cela pour des événements non consécutifs.

## **Discussion**

Dans cet article, nous avons décrit le processus de traitement de données d'observation d'environnements scolaires, et montré que certaines des caractéristiques de ces derniers, récemment mises au jour, ne sont pas toujours prises en compte. Les techniques des méthodes d'observation de classe ont relativement peu évolué, même si leur informatisation a apporté quelques améliorations. Par exemple, l'aide informatique à la saisie et au codage des événements observés (Kahng & Iwata, 1998), et aussi l'utilisation de méthodes statistiques pour réduire les données. Le récent regain d'utilisation de ces méthodes dans la recherche en éducation francophone nous a amené à la fois à exposer leurs principales fonctions, mais aussi à montrer leurs limites et quelques possibilités d'amélioration. La quasi-totalité des systèmes d'observation reprend le principe de la FIAC (Flanders, 1976), qui part du principe que les interactions entre enseignant et élèves *s'enchaînent*, c'est-à-dire qu'elles ont une forte dépendance deux à deux. Même si des méthodes plus récentes ont été élaborées (e.g., la grille d'organisation et de gestion pédagogique, Altet *et al.*, 1996), elles restent très inspirées de la FIAC. Ces différentes méthodes ont deux inconvénients, elles ne considèrent pas que les différents événements observés se déroulent dans des intervalles de temps pouvant se chevaucher et, de plus, elles ne supposent pas que les actions de l'enseignant non contiguës peuvent être causalement reliées, ce qui empêche l'analyse d'éventuelles régularités de l'action de l'enseignant, à un niveau plus large de description.

Les recherches futures sur les systèmes d'observation visent à combiner les mesures de haut et de bas niveau d'inférence, les observations et les performances des élèves, les observations et les questionnaires, ainsi que les modélisations statistiques. Enfin, l'une des questions importantes auxquelles aura à répondre la recherche en éducation est la manière dont ces méthodes tiennent compte du caractère dynamique d'environnements d'enseignement/apprentissage. En effet,

les événements scolaires surviennent de manière souvent simultanée et rapide, peuvent parfois difficilement être prédits, et le comportement de l'enseignant est parfois improvisé. Leur analyse en termes de causalité est de plus délicate à réaliser. Ces caractéristiques rendent complexe l'implantation de logiciels d'aide à l'observation de classes, et la recherche à venir devra en tenir compte.

Philippe Dessus, laboratoire des sciences de l'éducation & IUFM (E.A. 602), université de Grenoble.

#### BIBLIOGRAPHIE

- ALLÈGRE (E.) & DESSUS (P.). Un système d'observation et d'analyse en direct de séances d'enseignement, In J. M. C. BASTIEN (Ed.), *Actes des 2<sup>e</sup> journées d'étude en psychologie ergonomique (EPIQUE 2003)* (p. 85 à 90), Roquencourt, INRIA, 2003.
- ALTET (M.), BRESSOUX (P.), BRU (M.), & LAMBERT (C.). Étude exploratoire des pratiques d'enseignement en classe de CE2, *Les dossiers d'Éducation et Formations*, 70, 1996.
- BERDOT (P.), BLANCHARD-LAVILLE (C.), & CHAUSSECOURTE (P.). Analyse clinique, in C. Blanchard-Laville (éd.), *Une séance de cours ordinaire* (p. 159-198), Paris: L'Harmattan, 2003.
- BUNGE (M.). L'observation, in M.-P. Michiels-Philippe (éd.). *L'observation* (p. 47-59), Neuchâtel: Delachaux et Niestlé, 1984.
- CHI (M. T. H.). Quantifying qualitative analyses of verbal data: A practical guide, *The Journal of the Learning Sciences*, 6 (3), 271-315, 1997.
- DE KETELE (J.-M.). *Méthodologie de l'observation*, Bruxelles: De Boeck, 1987.
- DESSUS (P.), & ALLÈGRE (É.). *Quelles actions d'élèves (de l'enseignant) déclenchent quelles actions de l'enseignant (des élèves)?* Communication au 5<sup>e</sup> congrès de l'AECESE, Paris, 2004.
- DESSUS (P.), ALLÈGRE (É.), & MAURICE (J.-J.). Enseignement en tant que supervision d'un environnement dynamique, *L'Année de la recherche en sciences de l'éducation*, 149-162, 2005.
- DOYLE (W.). Classroom organization and management, in M. C. Wittrock (Ed.), *Handbook of Research on Teaching* (3<sup>e</sup> éd., p. 392 à 431), New York: McMillan, 1986.
- EVERTSON (C. M.), & Green (J. L.). Observation as inquiry and method, in M. C. Wittrock (éd.), *Handbook of Research on Teaching* (3<sup>e</sup> éd., p. 162-213), New York: McMillan, 1986.
- FLANDERS (N. A.). Interaction models of critical teaching behaviors, in E. J. Amidon & J. B. Hough (Eds.), *Interaction Analysis: Theory, Research, and Applications* (p. 360 à 374), Reading, Addison-Wesley, 1967.
- FLANDERS (N. A.). Analyse de l'interaction et formation, in A. Morrison & D. McIntyre (éds.), *Psychologie sociale de l'enseignement* (vol. 1, p. 57 à 69), Paris: Dunod, 1976.
- FLANDERS (N. A.). Human interaction models of teaching, in K. Marjoribanks (Ed.), *The foundations of students' learning* (p. 109-126), Oxford: Pergamon Press, 1991.
- FRASER (B. J.). Classroom environment instruments: Development, validity and applications, *Learning Environments Research*, 1 (1), 7-33, 1998.
- HOC (J.-M.). *Supervision et contrôle de processus*, Grenoble: PUG, 1996.
- KAHNG (S. W.), & IWATA (B. A.). Computerized systems for collecting real-time observational data, *Journal of Applied Behavior Analysis*, 31, 253-261, 1998.
- KOHN (R. C.), & NÈGRE (P.). *Les voies de l'observation*, Paris: Nathan, 1991.
- LANDAUER (T. K.), & DUMAIS (S. T.). A solution to Plato's problem: the Latent Semantic Analysis theory of acquisition, induction and representation of knowledge, *Psychological Review*, 104 (2), 211-240, 1997.
- LEINHARDT (G.), & GREENO (J. G.). The cognitive skill of teaching, *Journal of Educational Psychology*, 78 (2), 75-95, 1986.

LEUTENEGER (F.), Un atelier de mathématiques : pratiques enseignantes, *Les Dossiers des sciences de l'éducation*, 5, 79-95, 2001.

MAURICE (J.-J.), & ALLÈGRE (É.). Invariance temporelle des pratiques enseignantes : le temps donné aux élèves pour chercher, *Revue française de pédagogie*, 138, 115-124, 2002.

MEDLEY (D. M.), & MITZEL (H. E.). Measuring classroom behavior by systematic observation, in N. L. GAGE (Ed.), *Handbook of Research on Teaching* (1<sup>re</sup> éd., p. 247-328), Chicago : Rand McNally, 1963.

POSTIC (M.), & DE KETELE (J.-M.). *Observer les situations éducatives*, Paris : PUF, 1988.

QUESADA (J.), KINTSCH (W.), & GOMEZ (E.). Complex problem-solving : A field in search of a definition? *Theoretical Issues in Ergonomics Science*, 6 (1), 5-33, 2005.

ROGALSKI (J.). Y a-t-il un pilote dans la classe ? Une analyse de l'activité de l'enseignant comme gestion d'un environnement dynamique ouvert, *Recherche en didactique des mathématiques*, 23 (3), 343-388, 2003.

SIMON (A.), & BOYER (E. G.). *Mirrors of Behavior II : An Anthology of Observation Instruments*, Philadelphie : Classroom Interaction Newsletter, 1970.

SWANK (P. R.), TAYLOR (R. D.), BRADY (M. P.), & FREIBERG (H. J.). Sensitivity of classroom observation systems : Measuring teacher effectiveness, *Journal of Experimental Education*, 57 (2), 171-186, 1989.

VAUCLAIR (J.). L'observation en éthologie, in M. P. MICHIELS-PHILIPPE (Ed.), *L'observation* (p. 123-136), Neuchâtel : Delachaux et Niestlé, 1984.

WRAGG (E. C.). *An Introduction to Classroom Observation*, Londres : Routledge, 1994.

ZACKS (J. M.), & TVERSKY (B.). Event structure in perception and conception, *Psychological Bulletin*, 127 (1), 3-21, 2001.