

Pour citer cet article : Dechamboux, L. (2020, prépublication). L'enquête formative : une proposition de modélisation de l'activité évaluative lors de séances de production écrite au CP en France. *La Revue LEeE*, 2. <http://revue.leeeonline/index.php/info/article/view/70>

L'ENQUÊTE FORMATIVE

Une proposition de modélisation de l'activité évaluative lors de séances de production écrite au CP en France

Lionel DECHAMBOUX

Version de la prépublication : juillet 2020
Évaluation ouverte et collaborative

Résumé

Cette contribution, issue d'une recherche doctorale, discute tout d'abord la question de l'évaluation formative. Devant la difficulté à l'appréhender empiriquement, elle lui préfère, dans un premier temps, la notion d'activité évaluative. Dans une perspective située de l'évaluation (Mottier Lopez, 2008), l'activité évaluative sera alors définie conceptuellement par une unité nommée valuation (Dewey, 2011), articulant une appréciation immédiate et un jugement évaluatif. Défini dans ses composantes en s'inspirant des préoccupations des théories évaluatives (notamment les rapports entre référé, référents, jugements évaluatifs) et du Cours d'action (Theureau, 2004), ce continuum émerge d'instant en instant et constitue un flux ininterrompu. Ce flux de valuations sera mis à contribution pour comprendre l'activité de trois enseignantes auprès de trois élèves en difficulté dans leurs classes lors de séances de production écrite au Cours Préparatoire (CP), en France. Les séances ont été filmées et transcrites intégralement. L'enseignante a été ensuite autoconfrontée à sa propre activité. Ces deux éléments (transcription et autoconfrontation) permettent de construire des données dont l'analyse sera ici illustrée au travers d'un exemple emblématique. En conclusion, une modélisation de l'activité évaluative sera présentée. Cette modélisation est nommée *enquête formative* en référence à l'enquête de Dewey (1993), où l'enseignant, face à une situation troublée puis problématique, met en œuvre des moyens pour revenir à une situation de production. Cette enquête est *in fine* prétendue formative puisqu'elle engage la participation de l'élève dans une construction de signification partagée.

Mots-clés : évaluation formative, activité évaluative, évaluation située, valuation, enquête

Abstract

This contribution, based on a doctoral research, first discusses the issue of formative assessment. Faced with the difficulty of apprehending it empirically, it prefers, initially, the notion of assessment activity. In a situated assessment perspective (Mottier Lopez, 2008), assessment activity will be conceptually defined by a unit called a valuation (Dewey, 2011), linking appraisal and evaluative judgement. Defined in its components by drawing on the concerns of evaluative theories (notably the relationships between referee, referents, evaluative judgements) and the Course of Action (Thereau, 2004), this continuum emerges from moment to moment and constitutes an uninterrupted flow. This flow of valuations will be used to understand the activity of three teachers with three students in difficulty in their classes during writing sessions at the Cours Préparatoire (CP), in France. The sessions were filmed and transcribed in full. The teacher was then confronted with her own activity. These two elements (transcription and self-confrontation) allow the construction of data, the analysis of which will be illustrated here through an emblematic example. In conclusion, a model of the assessment activity will be presented. This model is called a *formative inquiry*, in reference to Dewey's inquiry (1993), in which the teacher, faced with a disturbed and then problematic situation, implements means to return to a writing situation. This inquiry is ultimately claimed to be formative since it involves the student's participation in a construction of shared meaning.

Keywords: formative assessment, assessment activity, situated assessment, valuation, inquiry

1. Introduction

Quelques notions, concepts traversent parfois le ciel des sciences de l'éducation et l'éclairent d'une lueur éclatante. Cet éclat peut tenir tout à la fois de leur puissance propre mais également à l'enthousiasme de ses promoteur·rices qui, munis de réflecteurs, diffuseront cette lumière en dehors de la sphère strictement académique pour éclairer les acteur·rices de l'éducation que sont les enseignant·es. Celles-ci ou ceux-ci, dès lors, pourront en faire bénéficier leurs élèves. Cette description, dont les lecteur·rices se demanderont, à juste titre, si elle est portée par un idéalisme naïf ou par une ironie malvenue, pourrait être utilisée à propos de l'évaluation formative.

Si nous reprenons notre catéchisme, il en est ainsi et comme tout un chacun, nous savons *bien* que l'évaluation formative, conceptualisée par Bloom (1969) à la suite des travaux de Scriven (1967) à propos de l'évaluation de programme, *garantit* de meilleurs apprentissages pour les élèves (e.g. Black & Wiliam, 1998 ; Hattie, 2009). Mais au-delà de l'enthousiasme que nous avons connu en tant qu'enseignant et que nous connaissons en tant que formateur d'enseignant·es, nous avons pu, à l'occasion d'une recherche doctorale¹, avoir le loisir de nous poser (entre autres) deux questions à la suite de ce récit mythique : l'évaluation formative existait-elle avant Bloom ? Et finalement, comment reconnaît-on une évaluation formative ?

Cette dernière question est devenue cruciale au moment de réfléchir à la mise en œuvre concrète de notre dispositif de recherche afin de lui garantir un minimum de tenue scientifique. Comment caractériser empiriquement ce qui relevait de l'évaluation formative, que nous

¹ Sous la direction des professeur·es Lucie Mottier Lopez et Germain Poizat.



souhaitions étudier dans les classes de Cours Préparatoires (CP) en France qui nous avaient ouvert leurs portes lors de séances de production écrite ? Notre réflexion a abouti à un choix, discutable, que nous souhaitons exposer dans ses grandes lignes et discuter ici.

Ainsi, la première partie de cette contribution s'attachera à présenter un peu plus profondément les arguments qui nous ont conduit à choisir cette option que nous assumons aujourd'hui. Nous donnerons à voir ensuite les aspects contextuel et méthodologique de cette recherche avant de proposer une analyse emblématique du travail que nous avons réalisé sur les données recueillies. Nous terminerons en présentant la modélisation à laquelle nous avons abouti, celle de l'enquête formative qui, montre, d'une certaine manière, la fécondité de notre orientation, sans toutefois mettre un point final à notre questionnement.

2. De l'évaluation formative à l'activité évaluative

Nous relatons ici rapidement notre parcours réflexif à propos de l'évaluation formative, des problèmes rencontrés en essayant de la définir. En effet, nous souhaitons la circonscrire de manière très précise afin de pouvoir l'observer de manière stable sur nos terrains de recherche, sans négliger certaines de ses occurrences. En effet, l'une de nos questions de recherche visait justement à étudier les relations chronologiques entre les différentes évaluations formatives. Ce parcours nous a conduit finalement à abandonner notre prétention à l'observer *a priori* pour scruter une activité évaluative que nous pourrions, *a posteriori* affubler éventuellement du qualificatif de *formatif*.

2.2 Une évaluation formative en constante redéfinition

2.2.1 De la conception de Bloom à une conception élargie

Au commencement, Bloom (1969) avait conceptualisé une évaluation formative qui s'incarnait dans un dispositif nettement séparé de l'enseignement et de l'apprentissage. Il s'agissait d'un processus ordonné de trois phases s'intercalant entre chaque séquence d'enseignement et permettant, ou non, son incrémentation. Ainsi, l'enseignant·e délivrait à ses élèves un test visant à démontrer la maîtrise de l'apprentissage juste effectué. Ce test procurait à l'enseignant·e un feedback qui, selon sa teneur, permettait d'avancer sur les contenus enseignés ou de proposer aux élèves une remédiation.

Cette conception originale de l'évaluation formative sera étendue ensuite à d'autres pratiques plus informelles mais qui poursuivent une finalité identique : soutenir les apprentissages des élèves sur la base d'une prise d'information préalable. Ce constat permet à Allal et Mottier Lopez (2005) d'évoquer l'émergence d'une conception *élargie* de l'évaluation formative aujourd'hui largement reconnue dans ce champ de recherche et au-delà. Cette dernière conception s'intègre à toutes les situations d'enseignement. Elle tire parti de divers moyens permettant de recueillir des informations pouvant établir la nécessité d'une remédiation. Le feedback et la remédiation prennent le nom de *régulation*, notion qui connaîtra des développements importants et même une émancipation par rapport l'évaluation formative (Mottier Lopez, 2012).

2.2.2 Un problème de définition opérationnelle

Cet élargissement est certes bienvenu : d'autres pratiques que ces petits tests non notés participent certainement à cette finalité de soutien des apprentissages des élèves. Mais ces



pratiques informelles sont bien plus complexes à identifier. Nous pensons ici particulièrement à toutes les pratiques qui engendrent des régulations interactives immédiates (Allal, 1988 ; Mottier Lopez, 2012) et qui sont sans doute le cas le plus informel de ces pratiques informelles. Comment en effet les repérer ? Lesquelles inclure et lesquelles exclure de la catégorie « évaluation formative » ? Sur la base de quels critères ? Est-ce en raison d'un « effet » (positif) sur les apprentissages des élèves ? Ces « effets » positifs sont tout d'abord soumis à caution (Dunn & Mulvenon, 2006 à propos de Black & Wiliam, 1998). D'autre part, est-ce réellement possible de déterminer que c'est bien cette pratique-là qui a eu cet effet-là ? Et l'absence d'effet disqualifie-t-il la pratique candidate au titre d'évaluation formative ? Pour ne donner qu'un exemple, un·e enseignant·e peut prélever des informations dans une situation pour décider de l'opportunité de remédier à d'éventuelles difficultés d'apprentissage. Or, il arrive parfois (et heureusement) que des difficultés ne soient pas présentes. La pratique précédente n'était pour autant pas une évaluation formative ?

Autre option, pourrions-nous les identifier sur la base des déclarations des enseignant·es ? C'est ce que choisit adroitement de faire Mottier Lopez et al. (2010) dans une dynamique de recherche collaborative. Mais lorsque notre préoccupation est d'étudier les relations entre différentes évaluations formatives, cette modalité n'est pas satisfaisante. Sans mettre en cause la sincérité des acteur·rices, une reconstruction *a posteriori* n'est pas à exclure lors d'un entretien post-leçon, celle-ci pouvant négliger des éléments pourtant importants pour comprendre cette logique formative.

Une observation, même instrumentée et fine, pourrait également s'avérer insuffisante pour la raison discutée plus haut : certaines pratiques sont *invisibles* et, de plus, n'engagent pas de régulation.

La solution proposée par Black et Wiliam (2009) ne nous paraît pas acceptable dans une perspective compréhensive : ils proposent sept stratégies permettant d'adoubier une pratique du titre d'évaluation formative. Ces stratégies, non dénuées d'intérêts, s'avèrent cependant à la fois très générales et trop contraignantes. Très générales pour notre recherche car ne permettant pas réellement de discriminer ce qui relève de l'évaluation formative du « reste » et trop contraignantes car excluant par exemple une pratique non-verbale pouvant pourtant soutenir un apprentissage. Le problème de l'identification de l'évaluation formative en classe est donc loin d'être simple, pour ne pas dire quasiment insoluble.

2.3 Une alternative : l'activité évaluative

2.3.1 Abandon momentané de l'évaluation formative

C'est bien la question du « reste » qui demande à être examinée. Que reste-t-il en effet lorsque nous avons soustrait aux pratiques enseignantes tout ce qui vise à soutenir les apprentissages des élèves, même de manière informelle, sans présager d'un « effet » positif ou négatif en termes d'apprentissage (et en se prémunissant donc d'une logique prescriptive) ? La réponse à cette question serait bien injuste et cruelle pour ce qui resterait et que nous pourrions appeler *enseignement* !

Et donc, inversement, tout ne serait qu'évaluation formative ? S'il nous a paru plus prudent de suspendre au moins momentanément le qualificatif de formatif, les précédentes considérations nous ont conduit à affirmer que, oui, tout est évaluation... Mais précisons le fond de notre pensée. Tout d'abord, pour « scandaleuse » qu'elle puisse paraître et même si nous la reprenons d'une manière quelque peu différente, cette idée n'est pas implicitement absente de la littérature lorsque Black et Wiliam (1998) déclarent : « Indeed, it is clear that



instruction and formative assessment are indivisible ” (p. 146). Et en pionnière, n'est-ce pas ce que suggère subtilement Allal (1988) avec la notion d'*évaluation continue* ?

Mais cette prise de position pose évidemment question. Les enseignant·es nomment-elles ou ils *a priori* ainsi l'ensemble de leur activité quotidienne ? Sans doute que non. Aussi nous proposons le raisonnement suivant en faisant appel à la notion d'*activité évaluative* (Figari & Achouche, 2001 ; Tourmen, 2009 ; Mottier Lopez, 2016). Cette notion met initialement en avant la nécessité de s'intéresser au moins tout autant au processus qui donne lieu à l'évaluation qu'au produit de l'évaluation. C'est justement le cœur de nos préoccupations de chercheur. Nous avons donc pris le parti de porter notre attention sur l'activité évaluative des enseignant·es avant de discuter *in fine* de son caractère formatif. Ceci étant dit, distinguer ce qui constitue, cette fois-ci, une activité évaluative du reste, n'est pas davantage résolu. Le problème reste donc entier ! Ou presque, c'est ce que nous montrons dans la partie suivante.

2.3.2 Une conception particulière de l'activité évaluative

Ainsi, nous avons envisagé non pas l'activité évaluative comme quelque chose que n'importe qui pourrait observer dans n'importe quelle classe et distincte d'autres moments d'enseignement mais bien plutôt comme le regard que pourrait porter un chercheur ou une chercheuse appartenant au champ de l'évaluation des apprentissages sur l'activité d'enseignant·es (ou d'élèves). Il s'agit, pour le dire autrement, de *voir* l'activité enseignante *comme* une activité évaluative. Et il ne s'agit plus de discriminer cela du « reste » puisque c'est précisément ce regard (du chercheur ou de la chercheuse) qui interprète *tout* ce qu'il voit comme relevant d'une activité évaluative. Il ne reste rien d'autre. Évidemment cela engage également une conception particulière de ce qu'est l'activité évaluative, ce que nous développons maintenant.

En préambule, nous dirons que notre positionnement se situe dans le *paradigme situé de l'évaluation* au sens de Mottier Lopez (2008). L'évaluation est considérée comme inscrite dans une situation en tout point singulière. Elle en porte les marques, elle procède d'un jugement dans un environnement, d'une construction de signification située.

Mais comment se construit ce jugement, cette signification ? De Ketele (2012), cherchant au travers des différentes conceptions de l'évaluation quelques principes permettant l'unification de ce domaine de recherche, les trouve dans « la confrontation d'un référent à un référé » (p. 206). Le référé désigne « ce qui est constaté ou appréhendé de manière immédiate » (Lesne, 1984, p. 132), « la partie de la réalité choisie » (Figari, 1994, p. 44). Le référent est « "*ce par rapport à quoi*" le jugement de valeur est porté ou est susceptible d'être porté » (Barbier, 1985, p. 72). Nè rompant pas avec cette tendance, c'est bien ce que nous observerons chez les enseignant·es. Mais précisons davantage nos hypothèses conceptuelles sur ce qui est constitutif de cette activité évaluative. L'ensemble des référents, entendu comme un agrégat de normes, forme le référentiel de l'évaluation. Sa nature hétérogène nous conduit à parler de *multiréférentiel*, suivant en cela la notion de *multiréférentialité* (Mottier Lopez, 2007), notion qui montre sa fécondité de travaux en travaux (Mottier Lopez, 2009, 2016 ; Mottier Lopez & Dechamboux, 2017) même si c'est parfois dans une formulation différente (Marcel & Murillo, 2013). Nous chercherons donc à mettre en évidence les référés, les multiréférentiels mobilisés par les enseignant·es.

Cette confrontation référé/multiréférentiel permet l'émergence d'un jugement évaluatif. Ce jugement peut prendre de nombreuses formes. Certaines sont classiques et attendues : un feedback verbal et une prise de décision. Cela correspond en effet à la manière dont en parle Allal (2008) dans sa définition de l'évaluation. Mais Barbier (1985) notait déjà que le jugement évaluatif prenait des formes parfois plus discrètes et ne se manifestait qu'au travers de ses effets ou de son énoncé. Nous pouvons donc reconnaître le jugement via un seul feedback



ou une prise de décision ou... rien du tout si ce jugement ne conduit pas à leur nécessité... Mais soulignons bien ici que ces considérations seront radicalisées. En effet, soucieux de suivre l'activité évaluative d'instant en instant pour en comprendre la logique, et la considérant comme formant un flux ininterrompu, nous désignerons aussi comme jugement tout acte même celui qui peut paraître le plus simple : un geste, un regard, un silence...

Mais cette confrontation entre référé et multiréférentiel donnant ce jugement ne s'effectue pas *in abstracto*. Elle est située. Or la situation n'est pas véritablement extérieure à l'évaluateur-riche, elle est construite par elle ou lui en fonction d'éléments « hors du corps » de l'évaluateur-riche qu'il ou elle sélectionne en fonction d'un *engagement* dans la situation et d'*attentes* particulières qu'il nous faut aussi interroger.

Enfin, le jugement actualise par ailleurs le multiréférentiel. Il fait expérience et peut potentiellement affecter le cours de l'activité évaluative. Nous avons donc besoin de questionner cette dernière composante pour caractériser l'activité évaluative et pour comprendre, par exemple, ses variations face à un « même » référé dans des situations différentes.

C'est donc ce complexe composé de six composantes – engagement / attentes / multiréférentiel / référé / jugement évaluatif / actualisation du multiréférentiel – qui forme l'unité de base de l'activité évaluative. Notre projet était donc de le renseigner d'instant en instant pour chacun des enseignant-es de cette recherche, afin d'en dégager la logique. Nous avons appelé ce complexe, une *valuation*, en référence aux travaux de Dewey (2011).

2.3.3 Quelques précisions sur les valuations

Dewey (2011) met, entre autres, en évidence par ce concept qu'une évaluation n'existe pas de manière isolée et que celle-ci s'ancre dans une appréciation immédiate. Nous reprenons ce fait fondamental en prétendant qu'il se reproduit à chaque instant de l'activité évaluative : une évaluation ne peut se former sans une appréciation immédiate préalable. Ou encore, une appréciation immédiate se prolonge dans une évaluation qui lui fait suite. Cette appréciation immédiate, que Dewey (2011) nomme *de facto valuing*, est, pour nous, représentée par ces engagements, ces attentes qui permettent l'émergence d'un référé confronté à un multiréférentiel singulier et lié à la situation, comme un instantané d'un flux multiréférentiel. Le référé est le signe de cette appréciation immédiate. Dit autrement, ce qu'un-e évaluateur-riche repère dans une situation comme significatif pour son évaluation révèle cette appréciation immédiate de la situation. Mais, sur un continuum, ce référé, confronté à ce multiréférentiel produit une action, et donc pour nous, un jugement évaluatif. Ce jugement évaluatif est quant à lui le signe d'une évaluation. Il matérialise le fait que l'évaluateur-riche ne s'est pas arrêté-e à la seule appréciation immédiate mais qu'elle ou il a élaboré un jugement sur la base de cette appréhension initiale. L'ensemble - de facto valuing et évaluation selon les termes de Dewey (2011) - forme donc notre unité, la valuation. L'activité évaluative, pour revenir à elle, est alors conçue comme une succession ininterrompue de valuations dont notre question de recherche vise à comprendre la chronologie et à comprendre également² la logique interne, au sein de chaque valuation et plus précisément entre les différentes composantes des valuations³.

Nous devons enfin expliciter un élément qui aura sans doute déjà été repéré par des lecteur-rices familier-es de ces approches. Le recours à la valuation de Dewey (2011) scelle finalement l'hybridation que nous avons opérée entre les apports du champ de recherche sur l'évaluation des apprentissages dans lequel nous nous inscrivons et les réflexions stimulantes

² Ces deux aspects sont évidemment intimement liés même si la formulation de ces questions les sépare.

³ Dans notre étude, nous avons surtout pu étudier les relations entre les référés, les jugements évaluatifs, le multiréférentiel et son actualisation.



menées dans le champ de l'analyse de l'activité et plus particulièrement celles issues du programme de recherche Cours d'action (Poizat, 2006 ; Theureau, 2004, 2006). Nous avons trouvé dans cette approche des éléments tout à fait compatibles et féconds avec notre conception de l'évaluation située. En effet, le cours d'action s'inscrit dans le paradigme de l'enaction (Varela, Thompson & Rosch, 1993) qui est une radicalisation des approches situées dans laquelle nous nous retrouvons. Il conçoit son objet théorique, le cours d'action, comme la part significative pour les acteur·rices de leur activité pouvant être vue comme une concaténation de signes. Ces signes se déploient selon une logique inspirée par les travaux de Peirce (1978) et que nous reprenons à notre compte en les déclinant avec les composantes propres à l'évaluation au travers des valuations. Tous ces éléments nous ont permis de conceptualiser l'activité évaluative telle que nous l'utilisons dans cette recherche⁴.

Ce flux continu, ce *cours de valuations* que nous souhaitons construire évitait ainsi l'écueil rencontré en amont à propos d'une évaluation formative insaisissable sans être dans un rapport limitatif, prescriptif à son égard. Il nous restait donc à l'alimenter avec des données recueillies dans des classes afin de répondre aux questions de recherche juste évoquées.

3. Terrain de recherche et cadre méthodologique

Dans cette partie, nous répondons à deux préoccupations. La première est de présenter le terrain de recherche dans lequel nous avons eu la chance de travailler, la seconde d'exposer la manière dont nous avons construit nos données afin de nourrir les composantes de l'activité évaluative et les analyser.

3.2 Trois classes de Cours Préparatoire (CP) lors de séances de production écrite

3.2.1 La production écrite au CP

Même si nous assumons une approche transversale de l'évaluation, l'objet de savoir n'est pas pour autant absent de notre recherche. Nous nous intéressons ainsi à la production écrite dans des classes où les élèves sont engagé·es dans l'apprentissage initial de la lecture. Il s'agit de la classe de CP en France. Ce choix n'est pas le fruit du hasard. La production écrite est un exercice complexe (Simard, Dufays, Dolz, & Garcia-Debanc, 2010) et pose problème aux enseignant·es et aux élèves en France (rapport du Conseil National d'Évaluation Scolaire (CNESCO), 2018⁵). Pourtant, l'enjeu de cet apprentissage est immense. En tant que tel bien sûr mais de manière encore plus aigüe au CP au regard de l'apprentissage de la lecture. En effet, même si la production écrite est une activité d'apprentissage plus complexe que la lecture, la production écrite est particulièrement propice pour comprendre ce qui est en jeu dans la lecture (Aeby Daghé & Dolz, 2007), étant donnée la relative symétrie existant entre ces deux activités. La production écrite peut rendre explicite les processus devant être mobilisés lors de la lecture, qui reste une activité opaque aux élèves car souvent silencieuse et toujours invisible. Cette activité mérite donc des recherches variées pour en comprendre les ressorts et notamment comment les enseignant·es soutiennent les apprentissages de leurs élèves lors de ces activités.

⁴ L'espace limité de cette publication ne nous permet pas d'en dire davantage. Nous renvoyons les lecteur·rices intéressé·es aux ouvrages de Theureau (2004, 2006) pour des informations de première main sur le cours d'action et à notre manuscrit de thèse (Dechamboux, 2018) pour connaître notre argumentation complète sur cette hybridation avec l'évaluation et sa viabilité.

⁵ <https://www.cnESCO.fr/fr/ecriture-et-rediger/> site consulté le 05.06.2020.



3.2.2 Trois classes de CP en France

Nous avons eu le bonheur de travailler dans quatre classes de CP en France mais nos analyses n'ont été formalisées de manière aboutie uniquement dans trois de ces classes en raison de la quantité de données construites. Celles-ci s'avéraient suffisantes pour les analyses effectuées. Nous discuterons donc seulement de ces trois terrains de recherche.

Ces classes étaient toutes situées, en France, dans le même secteur géographique (Haute Savoie) et présentaient une population relativement comparable. Les enseignantes, Alice, Gaby et Suzie⁶, étaient toutes chevronnées, avec une expérience d'enseignement supérieure à 15 ans.

Nous avons assisté à six séances de production écrite dans chacune de ces classes et nous avons effectué des entretiens *ante* et *post* avec chaque enseignante. Les séances observées ont été réparties régulièrement entre octobre et mars de l'année scolaire 2013-2014.

Notre dispositif de récolte de données ne nous permettait pas de focaliser notre attention sur les interactions entre chaque enseignante et l'ensemble des élèves de leur classe. Nous avons donc demandé aux enseignantes concernées, lors de l'entretien *ante* ayant eu lieu en début d'année scolaire (fin du mois de septembre) de nous proposer trois de leurs élèves pour lesquels elles pressentaient des difficultés dans leur apprentissage de la production écrite. Ce qu'elles firent. Nous avons donc pu mettre en œuvre la construction de nos données que nous présentons ci-dessous.

3.3 Construction des données

3.3.1 Observation en classe et autoconfrontation des enseignantes

Nous avons répliqué le protocole suivant sur l'ensemble des 18 séances observées. Nous avons filmé toutes les séances. Nos questions de recherche s'intéressant à un grain fin de l'activité enseignante, à l'intime des interactions qu'elles avaient avec leurs élèves, nous n'avons donc pas filmé en plan d'ensemble mais au contraire en plan rapproché sur les trois élèves choisis par chacune des enseignantes pour capter au plus près les interactions enseignante-élève. Nous pourrions parler de centration sur les régulations interactives immédiates (Mottier Lopez, 2012).

Suite à chacune de ces séances, nous avons autoconfronté les enseignantes à leur propre activité. Nous leur avons montré le film de la séance et leur avons demandé de décrire le plus spontanément ce qu'elles faisaient, percevaient, ressentaient d'instant en instant en évitant de tomber dans le commentaire ou l'analyse. Ces éléments méritent quelques explications complémentaires. En effet, l'activité évaluative, à l'instar de l'activité, n'est pas forcément observable de manière extérieure à l'acteur·rice considéré·e. Une part parfois conséquente de celle-ci se déroule alors que, pour un·e observateur·rice même attentif·ve, rien ne transparait. Pourtant, pour construire cette activité évaluative, il nous fallait avoir accès à ce qui était significatif pour les acteur·rices dans la situation d'instant en instant et ces informations sont susceptibles de ne pas être restituées de manière fidèle lors d'un entretien classique post-leçon. Ceci n'est finalement possible qu'en apportant du crédit à une hypothèse empruntée au Cours d'action, celle de la conscience préreflexive et à un dispositif permettant d'en faire le rappel, l'autoconfrontation. La conscience préreflexive est une forme de présence à soi qui accompagne toute activité humaine (Theureau, 2006) offrant « un accès non illusoire bien que limité à l'activité humaine comme enaction que ne permet pas la conscience

⁶ Les prénoms ont été anonymisés.



réflexive » (Theureau, 2014, p. 4). L'activité de l'acteur·rice donne lieu à chaque instant à expérience. Ce flux d'expérience *est* la conscience préreflexive. A l'aide du visionnage du film de l'activité qu'il nous intéresse d'explorer et d'un accompagnement du chercheur ou de la chercheuse, il est possible de « replonger » l'acteur·rice dans ce flux d'expérience passé afin qu'elle ou il puisse le verbaliser en référence à cette situation passée et non celle, actuelle, réunissant acteur·rice et chercheur ou chercheuse. Cette remise en situation dynamique est évidemment délicate à mener mais évite également des projections trop importantes du chercheur ou de la chercheuse à propos de l'activité observée. Ces autoconfrontations ont été toutes filmées.

Enfin, à la fin de chacune des séances de production écrite, nous procédions à une dernière prise de données : nous photocopions également les textes écrits par les trois élèves.

3.3.2 Un protocole à deux volets

L'étape suivante a consisté en l'élaboration d'un protocole à deux volets permettant de construire les valuations. Pour le premier volet, nous avons procédé à une description intégrale et systématique des actions observées sur l'enregistrement vidéo et des verbalisations de l'enseignante et de ses élèves pendant la séance de production écrite. Ce volet restitue ce qui se passe à l'« extérieur » du corps de l'actrice autoconfrontée.

Le second volet est renseigné à l'aide de la transcription intégrale des verbalisations des actrices lors des autoconfrontations, synchronisée avec le contenu du premier volet. Ces transcriptions constituent la source principale et directrice de la part intrinsèque de nos données (des inférences restent nécessaires pour compléter ces verbalisations même si nous les avons limitées au maximum).

3.3.3 Construction du flux de valuations

A l'aide du protocole ainsi constitué, nous avons pu élaborer le flux de valuations représentant l'activité évaluative de ces enseignantes d'instant en instant lors de ces séances de production écrite. Dans un tableur, nous avons ainsi établi les valuations se succédant pendant la séance en renseignant pour chacune d'elles et autant que possible les six composantes discutées dans notre partie conceptuelle (engagement / attentes / multiréférentiel / référé / jugement évaluatif / actualisation du multiréférentiel). Un extrait de ce travail est visible dans le tableau 1, il présente neuf valuations. Sur l'ensemble des séances analysées, nous avons construit environ 4400 valuations. C'est donc le résultat de cette dernière étape, ces cours de valuations, que nous avons analysé afin de répondre à nos questions de recherche.

Tableau 1

Valuations déployées lors du 20^{ème} épisode interactif de la 3^{ème} séance de production écrite dans la classe de Gaby

Episode Interactif	Elève	Engagement E	Attentes A	Multiréférentiel M	Référé R	Jugement évaluatif J	actualisation du Multiréférentiel aM
20	L	Permettre aux élèves de poursuivre la tâche	L ait écrit "des"	Le son /g/ n'a pas été travaillé en classe		Demande à L s'il a trouvé le "des"	
			L ne soit pas capable d'écrire /gi/ GUI pour le moment	Le son /an/ n'a pas été travaillé en classe	L montre sur son cahier	Validant sa réponse, lui demande "des quoi"	



			L écrit "des guirlandes" en utilisant le cahier blanc	Guirlande est un long mot avec trop de sons inconnus	L lit "/ki/"	pense que L a essayé d'écrire "guirlande" tout seul avec le G de G	
			L ne passe pas du temps à écrire "guirlande", ce qui n'est pas l'objectif de la séance	Gloria et Guillaume sont des prénoms de la classe qui peuvent servir à écrire le son [g]	L a mis un G	Lui demande s'il sait écrire "guirlande" tout seul	L a envie de se détacher de ce cahier blanc ce qui est également positif
						Choisit de lui faire trouver plutôt que de lui faire décomposer	
						Lui demande où il pourrait trouver ce mot	
					L répond que non		
					L prend le cahier blanc	Valide	
					L ouvre à la page des prénoms	Lui demande s'il croit trouver guirlande dans cette page	
					L répond que non	Lui disant qu'il est perdu ce matin, lui demande de chercher la page des décorations	
					L cherche dans son cahier		

Note. Un niveau de gris représente une valuation au niveau des jugements évaluatifs. L'élève est désigné par la lettre L.

Dans le tableau, des cases vides subsistent. Elles manifestent le fait que nous n'avons parfois pas voulu aller trop loin dans les inférences et rester au plus près du vécu des acteur·rices. Cette décision s'anticipe dès l'autoconfrontation et se matérialise à cet instant par une attitude du chercheur ou de la chercheuse très en retrait, intervenant surtout pour éviter que les acteur·rices ne soient trop dans l'analyse, tentante mais contreproductive ici, de leur propre activité.

3.3.4 Méthode d'analyse

Une fois ces tableaux contenant toutes les valuations construites, nous sommes passé à la phase d'analyse à proprement parler. Notre objectif était donc d'élucider autant que possible la logique à l'œuvre entre les différentes composantes des valuations et ensuite entre les valuations. Cette phase d'analyse a nécessité plusieurs « passages » sur les données construites afin de stabiliser les catégories permettant de qualifier les données entrant dans les différentes composantes des valuations qu'étaient le référé, le multiréféréntiel, le jugement



évaluatif et l'actualisation du multiréférentiel⁷. Nous avons, par anticipation, construit un certain nombre de catégories puis, nous avons complété cela avec des catégories émergentes au cours de nos analyses.

Nous avons évidemment puisé dans la didactique du français et notamment dans les typologies d'erreurs (e.g., Catach, 1980 ; Groupe EVA, 1991 ; Simard, 1992, 1999 ; Simard, Dufays, Dolz & Garcia-Debanc, 2010 ; Thevenaz-Christen & Leutenegger, 2007) pour catégoriser les référés sélectionnés par les enseignantes. Nous avons progressivement affiné nos choix et nous avons finalement fonctionné avec les catégories présentées dans le tableau 2. Chaque catégorie est associée à un code couleur qui nous a été utile lors du passage à la représentation graphique de nos données (voir [annexe A](#) pour un exemple).

Tableau 2

Catégories de référés issus de la didactique du français

Codes couleur	Catégories	Indications, précisions
	Aspect sémantique	Lié au sens relativement à la situation de référence.
	Aspect pragmatique	Lié à la situation de production et notamment aux contraintes de la tâche, de la consigne
	Aspect syntaxique	Lié à la syntaxe de la phrase
	Aspect phonologique	Lié à un choix erroné de la correspondance grapho-phonétique avec altération de la prononciation
Le	Orthographe lexicale Morphographie lexicale	Encodage erroné du mot (choix des correspondances grapho-phonétiques inexact d'un point de vue orthographique) et/ou morphogramme lexical absent ou erroné
Gr	Morphographie grammaticale	Morphogramme grammatical absent ou erroné
	Segmentation des mots	Segmentation erronée des mots
	Aspect typographique	Erreur dans le choix de l'écriture en cursif / script, suivi de la ligne, mise en page
	Majuscule et ponctuation (code textuel)	Choix erroné ou absence de majuscule, de ponctuation

Mais ces catégories n'ont pas épuisé l'ensemble des référés identifiés. Il a fallu adjoindre à ces catégories, issues de la didactique, des catégories émergentes liées notamment au comportement des élèves. Nous reviendrons sur cet aspect dans la partie suivante de cet article.

Concernant les jugements évaluatifs opérés par l'enseignante, nous avons anticipé trois catégories, à la croisée entre les opérations nécessaires à la production écrite (Mas, 1991) et les opérations de régulation métacognitive (Allal, 2007) : *anticiper*, *contrôler*, *ajuster*. Mais comme pour les référés, ces catégories initiales seront insuffisantes.

Enfin, sur la question du multiréférentiel et de son actualisation (les catégories sont identiques dans l'une ou l'autre de ces composantes), nous avons abandonné la catégorie relative au référentiel institutionnel, celui-ci n'était jamais invoqué lors des autoconfrontations. Cependant, nous ne dirons pas qu'il est absent des situations puisque le fait même que cette activité de production écrite existe en classe montre son influence structurante. Ceci étant dit, les catégories finalement retenues sont au nombre de sept :

- *aide disponible* : l'enseignante se réfère à l'une des aides matérielles disponibles en classe,

⁷ Nous n'avons pas étendu ce travail aux composantes de l'engagement et des préoccupations. Il nous paraissait moins pertinent de le faire à ce moment-là de notre recherche. La modélisation présentant la généralisation de nos analyses rend compte finalement de ces deux composantes, même si c'est d'une manière générale.



- *travail et progression de la classe* : l'enseignante se réfère à l'avancée du travail en classe, à l'état de la progression au moment de la séance,
- *norme de la classe* : l'enseignante se réfère à des habitudes spécifiques à la classe,
- *connaissance de l'élève* : l'enseignante se réfère à sa connaissance de l'élève avec lequel elle interagit,
- *savoir didactique* : l'enseignante se réfère à un savoir issu d'une théorie didactique relative à la production écrite, d'une manière théorique, générale, sans relation explicite avec la classe ou avec son expérience,
- *savoir professionnel* : l'enseignante se réfère à un savoir dont l'énonciation montre qu'il a été soutenu, confirmé par l'expérience professionnelle de l'enseignante,
- *contrainte de la tâche et de la situation* : l'enseignante se réfère à la tâche, à ses contraintes.

Signalons que les quatre premières catégories ont été anticipées grâce aux travaux de Mottier Lopez et Allal (2008) à propos de la multiréférentialité.

Une fois ces composantes catégorisées, nous avons procédé à une première analyse très descriptive à propos de chacune de ces catégories. Si nécessaire que fut sans doute cette étape, elle n'a pas été très fructueuse en termes de résultat, c'est pour cette raison que nous l'évoquons seulement dans cette partie méthodologique. Il nous fallait une réflexion sans nul doute plus interactionniste entre les différentes composantes mais rendue difficile en raison de la quantité de données qualitatives à traiter. Nous avons trouvé une issue en passant à une représentation graphique des cours de valuations (cf. [annexe A](#)). La création de ces diagrammes a permis d'obtenir une vue panoptique de l'activité évaluative et de pouvoir ensuite dégager des *séquences* parmi ce flux de valuations, c'est-à-dire, en suivant Theureau (2004), des groupes rassemblant plusieurs valuations formant une cohérence thématique. Nous détaillons tous ces éléments dans les parties suivantes.

4. Analyses et résultats

Le volume des données était important. Il est donc impensable de vouloir en présenter ici l'analyse qualitative effectuée de manière exhaustive. Nous faisons donc le choix d'en montrer des brides à propos des catégories émergentes et un exemple emblématique d'une séquence prélevée dans un cours de valuations.

4.2 Catégories émergentes

4.2.1 Au niveau des référés, premières analyses et résultats

Nous avons analysé des référés correspondant aux catégories anticipées à ce niveau, qui étaient d'ordre didactique. Ceux-ci étaient présents en abondance. Mais leur abondance est à discuter ici pour éviter d'associer trop hâtivement référé et erreur dans la production écrite de l'élève, c'est l'objet de ce premier temps, et référé et référé didactique de manière exclusive, ce sera l'objet d'une autre partie, après un détour nécessaire à la compréhension de la logique de l'activité évaluative, traitant des jugements évaluatifs analysés.



L'erreur est un élément déviant de la norme du français écrit, repérable par un·e observateur·rice extérieur·e à la situation. Le référé, quand il est didactique, est surtout un élément considéré comme acceptable ou inacceptable à ce moment de l'année dans cette classe-là. Il est tout à fait fréquent que des erreurs ne soient pas relevées en tant que référés par l'enseignante. Cela ne signifie pas son incompetence à « corriger » mais bien davantage son adéquation à la situation. Par exemple, dans la classe de Gaby à laquelle nous nous référerons principalement, la correction orthographique est exigée pour des mots présents dans les aides disponibles en classe (affichage, répertoire, etc.). Pour les autres, la correction phonologique est tout à fait suffisante. Ainsi, dans la première séance observée, Neil⁸ écrit le mot « copain » *copin. Gaby ne le relèvera même pas en tant que référé alors qu'il s'agit de toute évidence d'une erreur d'orthographe lexicale. Cela s'explique sans aucun doute par le fait que ce mot n'était pas présent dans les aides et que sa prononciation n'est pas altérée. Ce qui est donc une erreur ne l'est que pour quelqu'un d'extérieur. Cette distinction est cruciale car elle rend compte de la logique située, spatialement et temporellement, des valuations.

4.2.2 Au niveau des jugements évaluatifs

Les catégories *anticiper*, *contrôler* et *ajuster* furent opérationnelles mais au prix d'un ajustement conséquent qui nous a cependant permis d'entrer dans la logique de l'activité évaluative. Nos analyses ont pu mettre en évidence que des jugements se référaient effectivement à des opérations d'anticipation, de contrôle ou d'ajustement mais qu'ils étaient aussi inscrits dans deux temps distincts mais potentiellement articulés. Le premier est un temps d'*exploration* pendant lequel l'enseignante explore visiblement la situation pour apprécier si l'élève a bien anticipé ou contrôlé son écrit. Le cas échéant, l'activité évaluative se dirige en général vers d'autres élèves. Dans le cas contraire, remarquable notamment lorsque l'enseignante identifie un *référé didactique inadéquat*, elle entre dans un second temps pendant lequel les jugements sont des jugements de *soutien à une opération d'anticipation* ou de *contrôle* jusqu'à ce que l'élève ait accompli cette opération de manière correcte pour l'enseignante.

Par exemple, pendant la deuxième séance dans la classe de Gaby, elle *demande* à Léonard « Alors, qu'est-ce que tu veux me raconter là ? ». Il s'agit pour nous d'un *jugement évaluatif* visant à explorer si Léonard a accompli *une opération d'anticipation* (de ce qu'il souhaite écrire). Léonard répond « La tortue touche les fleurs ». Ce référé est considéré par Gaby comme incorrect car la phrase proposée par Léonard ne correspond pas à l'image qu'il devait décrire (consigne de la tâche). Ce référé est un pivot qui fait entrer Gaby dans ce deuxième temps, de *soutien*, initié par le jugement évaluatif suivant : « dit⁹ à Léonard : "c'est pas des fleurs, elle mange quoi la tortue ?" ». Ici, dans cette nouvelle phase, elle soutient donc Léonard dans son opération d'anticipation.

Autre exemple, lors de la cinquième séance, Gaby *pointe* le début de la phrase écrite par Gloria, ce qui est pour nous un *jugement évaluatif relatif à une opération de contrôle* mais dans un temps *exploratoire*. Elle repère que Gloria a écrit le mot « jardin » *gardin. Ce référé fait basculer Gaby du temps exploratoire à un temps de soutien à une opération de contrôle puisque son jugement évaluatif suivant sera « lui rappelant son prénom, lui demande ce que font un G et un A ensemble » par lequel elle aide donc Gloria à contrôler son écrit.

⁸ Les prénoms des élèves sont anonymisés. Les prénoms autres que Suzie, Alice ou Gaby désignent ces élèves dans la suite de l'article.

⁹ Dans nos données, nous avons formulé les jugements évaluatifs sans sujet et du point de vue de l'enseignante qui le vit.



Enfin, et concernant l'opération d'ajustement, celle-ci arrive en général après une opération de contrôle ayant montré sa nécessité (temps exploratoire) puis sa réalisation (temps de soutien). Ainsi, les jugements évaluatifs relatifs à une opération d'ajustement ont pris principalement la forme d'une validation de la proposition de correction émise ou effectuée par l'élève à l'issue d'un temps de soutien au contrôle. Elle signe d'ailleurs généralement la fin de ce que nous avons pu identifier comme *séquence* dans un cours de valuations. A titre d'exemple, Gaby, dans la 5^{ème} séance, se rend compte que Neil a mis un accent circonflexe sur le premier E du mot « fenêtre ». Ce référé orthographique (orthographe lexicale) inadéquat enclenche une phase de soutien au contrôle de ce mot. Finalement, Neil comprend son erreur et la corrige avec l'aide de Gaby. L'enseignante effectue le jugement évaluatif suivant : « validant, lui demande de corriger l'autre mot fenêtre ». Ce jugement est relatif à l'opération d'ajustement (elle valide l'ajustement proposé par Neil). Par contre, contrairement aux autres discutés ici (anticiper et contrôler), nous ne trouvons que rarement les deux temps d'exploration et de soutien. Cela s'explique par le fait que ce qui est à ajuster a été bien souvent établi juste avant et que l'enseignante considère que l'élève, dans ces conditions, est capable d'ajuster seul, sans soutien.

Nous avons pu noter que ces deux temps (exploration et soutien) structuraient fortement l'activité évaluative et donnaient la « clé » de sa compréhension en tant qu'*enquête* (au sens de Dewey, 1938) sur l'ensemble des valuations analysées, ce que nous détaillerons ci-dessous. Quant aux jugements évaluatifs nombreux que nous avons identifiés mais qui ne « rentraient » pas de manière satisfaisante dans les trois catégories anticipées, nous avons alors pu les spécifier avec d'autres catégories qu'il serait trop long de détailler ici. Mais pour le dire rapidement, il s'agit de jugements évaluatifs qui consistent à souvent *étayer* les élèves dans leurs opérations d'anticipation ou de contrôle. Il s'agit d'étayages « classiques » décrits par les fonctions d'étayage de Bruner (1983) mais aussi de questions, de questions où certains éléments sont soulignés par l'intonation, dans la formulation, des renvois en miroir de la réponse précédente de l'élève, d'appel à l'aide d'un·e pair·e, etc. Bref, nous avons identifié un panel très riche de jugements qui s'adaptent de manière très fine aux réponses (verbales, comportementales) de l'élève et un art subtil d'amener vers la réponse attendue sans jamais pratiquement la donner explicitement à l'élève. Nous reverrons cela et en comprendrons la motivation lors de notre exemple emblématique.

4.2.3 D'autres considérations à propos des référés

Tout d'abord, précisons que les référés didactiques sont finalement de deux types. Soit le référé correspond à ce qui était attendu par l'enseignante et l'activité évaluative ne se poursuit pas par des valuations de soutien à une opération d'anticipation ou de contrôle. Soit le référé ne correspond pas à ce qui pouvait être attendu et il est, nous l'avons déjà dit, à l'initiation de valuations visant à soutenir l'anticipation ou le contrôle. Mais si ces *référés didactiques inadéquats* sont systématiques à cet endroit, nous avons pu constater clairement qu'ils n'initiaient pratiquement jamais une *séquence* de valuations. D'autres référés s'enchaînent et se déterminent progressivement pour arriver au référé didactique. Il s'agit d'un point crucial quant à la compréhension de l'activité évaluative.

Si nous cherchons en amont de ces *référés didactiques (adéquats ou inadéquats)*, nous constatons que juste avant un référé didactique qui initiera (ou non car adéquat) un jugement évaluatif visant à soutenir une opération de contrôle, existent des référés que nous avons nommés *référés trace d'un écrit*. Il s'agit de traces de lettres, de phrases, etc. mais sans qu'ils ne soient encore saisi avec leur signification « classique ». Si nous nous risquons au vocabulaire saussurien, ce serait comme si le signifiant existait mais sans son signifié¹⁰. C'est

¹⁰ Il s'agit d'une image limitée car évidemment s'il n'a pas son signifié « habituel », ce signifiant est bien doté d'une signification que nous cherchons justement à faire comprendre.



un signe, un référé donc, mais qui n'a pas de sens autre que de montrer sa présence. Il acquiert au moment où il devient un référé didactique cette signification habituelle, ce signifié. Avant, ce n'est que la *trace d'un écrit*, une forme noire sur un fond blanc. Par exemple, Gaby dans la quatrième séance constate que « Gloria semble avoir terminé sa phrase (sans arriver à la lire) ». Suzie voit que Casey, son élève, « fait des lignes et des lignes ». Dans les valuations qui suivront, ces lignes, ces phrases, ces mots, etc. seront véritablement lues et donneront des référés didactiques. Mais pas à cet instant.

Ce ne sont cependant pas encore les référés initiant les séquences de valuations. Il faut aller encore plus en amont pour les identifier. Que ce soit lors d'un temps d'exploration à visée d'anticipation ou de contrôle, les premiers référés sont quasiment systématiquement des référés liés au comportement des élèves. L'intégration de ces référés à l'activité évaluative est un résultat important selon nous car il montre que la seule focalisation sur le didactique et son apparition ne sont pas suffisantes pour appréhender l'activité évaluative des enseignant·es. Mais commençons par des référés comportementaux n'initiant pas de temps exploratoire. Ils sont rares dans notre corpus mais leur évocation nous paraît importante, là encore dans une visée de compréhension. Ils sont homogènes dans les trois classes et relèvent d'un travail « normal » d'un ou de plusieurs élèves. Par exemple, dans la première séance chez Gaby, elle perçoit « l'attitude travailleuse de Neil », dans la troisième séance, elle retient que « Gloria est dans le travail, elle cherche ». Alice, dans sa classe, repère lors de la troisième séance « pas trop d'agitation visuelle et sonore, pas de bruit de fond ». Ces référés qui ne déclenchent pas de séquence particulière ont été nommés *référés comportementaux de production conforme*. Pour le dire simplement, les élèves travaillent normalement, tout va bien...

A l'inverse, nous avons pu analyser trois autres types de référés pouvant donc être au déclenchement d'une séquence de valuations. Les deux premiers, très nombreux dans notre corpus, sont des *référés comportementaux d'arrêt de production conforme* ou *non conforme*. Ils seront généralement suivis d'autres référés (dont des référés didactiques) et seront le marqueur du début d'une séquence de valuations. Parmi les référés comportementaux d'arrêt de production conforme, nous avons par exemple « Neil demande comment écrire "des" » dans la troisième séance chez Gaby. Nous avons aussi « Lemmy lève la main » dans la troisième séance chez Alice cette fois-ci. Parmi les référés comportementaux d'arrêt de production non conforme, nous notons « Neil perturbe Léonard » dans cette même troisième séance dans la classe de Gaby ou encore « Andy est retourné, il n'est pas dans son travail » dans la première séance chez Suzie.

Les troisièmes de ce groupe sont des *référés de production non conforme*. Nous en avons repéré quelques-uns comme lorsque Gaby repère le « temps mis par Gloria pour écrire la date » dans la première séance ou que, toujours dans la même classe, « Neil vocalise les sons qu'il est en train d'encoder ». Au-delà des différences, ce qui relie ces trois types de référés est qu'ils initient donc une séquence de valuations que nous spécifierons tout de suite après dans sa globalité. Ils signifient, pour le dire simplement, que les élèves ne travaillent pas ou travaillent de manière étonnante, ce qui motive une intervention de la part de l'enseignante...

4.3 Un exemple emblématique de séquence de valuations : l'enquête formative

A cet endroit, nous pensons avoir donné suffisamment d'éléments qui permettront aux lecteur·rices de comprendre la logique que nous avons pu ainsi dégager progressivement de notre corpus, moyennant quelques ajouts supplémentaires au cours de cette dernière analyse



présentée. Nous prenons ici un exemple de séquence de valuations, que nous proposons tout d'abord dans un tableau, puis que nous analysons en mobilisant le concept de l'enquête de Dewey (1993)¹¹ qui nous a apparu tout à fait congruent pour comprendre l'activité évaluative enseignante et que nous reformulerons en conclusion. De manière très superficielle ici, l'enquête est un processus qui permet à l'être humain, que ce soit lors d'une activité quotidienne ou hautement scientifique, de résoudre les problèmes qui se posent à lui et troublent sa tranquillité¹². Au départ, la situation dans laquelle il se trouve est sans dissonance, les habitudes permettent à l'être humain de s'y mouvoir sans penser. A un moment, une difficulté surgit et la situation, jusque-là *apaisée, déterminée* selon le vocabulaire de Dewey, devient *troublée*. Face à ce trouble, l'être humain a un objectif, retrouver une situation apaisée. Il observe la situation troublée, l'organise selon ce qu'il connaît et ne connaît pas. Il peut alors formuler une hypothèse lui permettant de comprendre l'origine de ce trouble. La situation est désormais *problématique*. En fonction de l'hypothèse choisie, l'acteur·rice met en œuvre des moyens cohérents avec l'hypothèse, le but recherché, pour revenir à une situation apaisée, pour qu'elle soit à nouveau *déterminée*. Il s'agit non pas d'un retour au point de départ mais une arrivée dans une situation, nouvelle, enrichie par l'expérience de cette enquête, mais apaisée, comme avant l'enquête. Voyons maintenant comment l'enquête permet de comprendre l'activité évaluative de ces enseignantes lors de séances de production écrite.

Nous nous retrouvons donc lors de la quatrième séance observée dans la classe de Gaby. Pendant cette séance, les élèves doivent décrire trois images séquentielles. Nous focalisons dès lors notre attention sur une séquence durant laquelle Gaby est en interaction avec Léonard. Cette séquence court sur une durée cumulée de 1 minute et 40 secondes (mais est entrecoupée par d'autres séquences dont nous ferons ici abstraction). Le tableau 3 présente les premières valuations construites à son propos.

Tableau 3

Valuations de la séquence emblématique de l'enquête formative

Episode	Elève(s) concerné(s)	N° valuation	Multiréférentiel (M)	Référés (R)	Jugements évaluatifs (J)	Actualisation du Multiréférentiel (aM)
G_S4_6	L	1	L a une histoire à décrire différente de N	Temps important passé sans avoir vu L	Lui demande de dire toute sa phrase	
		2	Le mot ski est écrit au tableau	L répond "le garçon fait du ski"		
		3		La phrase a un sens		
		4		La phrase débute par une majuscule		
		5		Le K de ski est mal formé		

¹¹ Nous devrions peut-être avoir présenté « classiquement » ce concept dans la partie conceptuelle de cet article mais nous privilégions ici une restitution au plus près (autant que possible) de la logique de notre recherche dans son déroulement. Le concept de l'enquête n'a pas été formulé de manière initiale (contrairement aux valuations qui ont structuré la construction de nos données). Sa pertinence et sa mobilisation a surgi au cours de nos analyses, suite à ces premières spécifications de référés et de jugements évaluatifs, dans une abduction dont le secret nous échappe encore en dépit de son actuelle évidence...

¹² Cette tranquillité peut se troubler à l'insu de l'acteur·rice ou de son plein grès comme dans la recherche scientifique.



		6		La phrase n'a pas de point		
		7		Fait est écrit "fé"	Demande à L si c'est bien	A ce niveau de l'année, c'est plutôt encourageant
		8		Absence de réaction de L	Dit à L que sa phrase décrit la 1ère image	
		9			Demande à L si cette phrase est finie	
		10		L rajoute tout de suite un point	Valide sa proposition	
		11	L est performant en déchiffrage		Isolant avec ses doigts la lettre « é », demande à L ce que ça fait	
		12	Beaucoup de sons commencent à avoir été étudiés en classe	L répond [fe]	Invalidant sa proposition, lui demande de rester sur le seul son montré	
		13	Certaines graphies de son se confondent facilement comme [u] et [ɔ], le [e] et le [ɛ] sur le plan visuel	L répond [ɛ]	Invalide sa proposition	
		14			Montrant les affichages des mots-référents, lui demande de regarder le sens de l'accent	
		15		L remarque que ce n'est pas le même	Validant sa proposition, lui demande ce que ça fait	
		16		Absence de réaction de L	Demande à L si l'accent dans ce sens fait [e] ou [ɛ]	
		17		Absence de réaction de L	Demande à L de regarder les affichages	
		18		L dit que ça fait [e]	Valide en disant qu'il s'agit du [e] de étoile	
		19			Répète la phrase de L en insistant sur le [fe]	



		20			Lui dit qu'elle veut [fɛ]	
		21			Lui dit qu'elle va l'aider et qu'il s'agit du [ɛ] de balai	
		22			Lui demande de chercher dans les affichages	

D'une situation apaisée à une situation troublée

Le référent initial est : « temps important passé sans avoir vu Léonard ». Cela nous indique qu'au-delà d'une activité évaluative conduite par des troubles de la situation « subis » par l'enseignante, il existe donc une modalité d'enquête où l'enseignante choisit de troubler d'elle-même la situation, ce qui constitue un stade supérieur d'enquête selon Dewey (Frega, 2006). Il y aurait donc une sorte de veille qui s'effectue lorsque la situation est calme. Ce constat est pour nous positif et témoigne d'une professionnalité de l'enseignante qui n'attend pas mais est, dans une certaine mesure, à l'initiative. Cependant et très concrètement, nous n'avons que très exceptionnellement observé de telle initiation tant les situations se troublent « spontanément », tant la gestion de ces enquêtes laisse peu de place à des enquêtes instituées volontairement. Pour revenir à ce référent initial, nous l'avons codé comme *référé comportemental de production non conforme* car révélant une forme de surprise de Gaby de ne pas avoir vu Léonard plus souvent (en raison de l'habitude qu'elle a d'interagir fréquemment avec lui à cause de ses difficultés en production écrite).

D'une situation troublée à une situation problématique

Gaby demande alors à Léonard de lui lire sa phrase, ce qui est typique d'une exploration d'un écrit en vue d'un contrôle. Nous savons grâce au Multiréférentiel que Léonard a une histoire à décrire différente de celle de Neil (qui est assis à côté) et que le mot « ski » (thème de son histoire) est écrit au tableau noir. La phrase écrite par Léonard est : *Le garçon fé du ski. Mais ce que nous observons comme référents dans l'activité évaluative de Gaby est assez exceptionnel¹³. Nous avons ici accès à un nombre important de référents. Nous disposons ici non seulement de référents problématiques mais également de référents didactiques adéquats¹⁴ :

- la phrase a un sens (référé didactique de type sémantique adéquat, codé en rose et positif >0¹⁵),
- la phrase débute par une majuscule (référé didactique de type code textuel adéquat, codé en vert et positif >0),
- le K de ski est mal formé (référé didactique de type typographique inadéquat, codé en marron et négatif <0),
- la phrase n'a pas de point (référé didactique de type typographique inadéquat, codé en vert et négatif <0),

¹³ Grâce, rappelons-le, à l'explicitation de Gaby lors de l'autoconfrontation, ce qui nous permet retenir tous ces référents.

¹⁴ Ce qui nous permet de connaître les critères de réussite de la tâche demandée.

¹⁵ Tous ces codes font référence au diagramme que nous avons construit pour cette séquence et qui est disponible dans l'[annexe A](#). Nous ne le commenterons pas plus, faute d'espace.



- le mot 'fait' est encodé [fe] (référé didactique de type orthographique lexical¹⁶ inadéquat, codé en orange et négatif <0).

Nous avons ici, au travers de ces référés, un compte-rendu détaillé des différentes recherches réalisées par Gaby lors de ce temps exploratoire. Il rend explicite *l'organisation de la situation troublée aboutissant à la situation problématique* : Léonard ne peut plus produire¹⁷ en raison de trois éléments correspondant aux référés didactiques inadéquats.

Précisons également que nous observons dans ces premières valuations, l'actualisation du Multiréférentiel : « à ce niveau de l'année, c'est encourageant ». Un *référé potentiel relatif à la connaissance de l'élève* est ainsi créé.

D'une situation problématique à la résolution de la situation à nouveau déterminée ?

La situation étant désormais problématique, examinons la manière dont Gaby engage sa résolution au travers de ce que nous avons identifié plus haut comme le *temps de soutien lié à une opération de contrôle*.

Elle demande tout d'abord à Léonard « si c'est bien ». Elle accomplit ici une demande rare dans notre corpus : une incitation à une autoévaluation (jugement codé AE). Nous pensons qu'il s'agit là d'une part d'annoncer l'initiation de ce temps de soutien au contrôle. D'autre part, cette question modélise le début de ce qu'elle souhaite voir apparaître chez Léonard, une compétence à se contrôler seul. Une opération de contrôle devrait débiter par une question de ce type destinée à « troubler soi-même » la situation. Et si ce trouble se confirme, cela incite à débiter sa problématisation pour éliminer le trouble. Evidemment cette question convie Léonard à partager le constat problématique de Gaby accessible *via* ses précédents référés. Nous avons pu montrer dans notre corpus (et comme nous le verrons ici) que ce qui caractérise les relations entre référés et jugements évaluatifs lors de ces phases de soutien est la construction d'une signification partagée avec l'élève à propos de ce que l'enseignante a jugé problématique dans la situation et les moyens de la résoudre. Un jeu tout à fait subtil se dessine au cours duquel, lorsque le jugement évaluatif permet à l'élève de construire une signification partagée avec son enseignante (ce qu'elle peut constater au travers du référé suivant), elle proposera un étayage plus « léger » voire la conclusion de ce temps de soutien à partir du moment où elle considère qu'il ou elle est en capacité d'ajuster seul. Au contraire, si le jugement évaluatif n'a pas permis à l'élève de progresser vers la construction d'une signification partagée, alors l'enseignante engagera un jugement évaluatif plus étayant et ce, jusqu'à ce que l'élève construise cette signification. Autrement dit, nous assistons à une adaptation de l'étayage très fine, d'instant en instant. Nous parlerons dès lors de *jugements et de référés de signification partagée* se répondant de valuations en valuations. Or ici. Léonard ne réagit pas à la demande d'autoévaluation de Gaby. Ce référé de signification partagée est donc codé négativement. Les moyens à employer pour la suite de l'enquête devront être ajustés en conséquence pour assurer une résolution selon la fin-en-vue visée.

Gaby explicite alors que la phrase écrite par Léonard décrit effectivement la première image. Ce jugement évaluatif est à mettre en relation avec le référé « la phrase a un sens » (codé Exp). Il rend visible une partie de l'organisation de la situation vers laquelle Gaby souhaite entraîner¹⁸ Léonard.

¹⁶ Ce mot étant plutôt simple et courant, Gaby exiger la correction orthographique plutôt que phonologique comme nous le constaterons, ce qui, rétroactivement dans notre élaboration de cette valuation, nous permet de le coder ainsi.

¹⁷ Même si en l'espèce, c'est Gaby qui stoppe sa production.

¹⁸ Entraîner dans le double sens de ce terme : faire exercer et amener à.



Elle demande ensuite à Léonard « si cette phrase est finie » (Q). Cette question est pour nous un jugement évaluatif destiné à faire prendre conscience à Léonard que, bien que sa phrase soit terminée du point de vue sémantique, il manque le point final. Elle est inachevée sur un aspect du code textuel. C'est ce que comprend également Léonard puisqu'il rajoute aussitôt un point. Par cette question, Gaby modifie l'environnement en permettant de focaliser l'attention de Léonard sur la fin de la phrase sans lui en dire plus. Cette modification est suffisante pour que Léonard problématise (partiellement) la situation comme Gaby (i.e., il reconnaît qu'il manque un point) et qu'il mette en œuvre un moyen pour rétablir la situation (ajout du point). Ce référent de signification partagée est donc positif sur le plan du code textuel (code vert, >0). La situation, globalement, n'est pas encore rétablie pour autant mais elle progresse vers cette fin-en-vue. Gaby valide cette action de Léonard (V).

Gaby s'attèle alors à la résolution d'une autre partie du « problème » : l'encodage du mot « fait ». Elle mobilise au sein du Multiréférentiel plusieurs référents importants pour la compréhension des jugements suivants : « Léonard est performant en déchiffrement » (référent relatif à la connaissance de l'élève), « beaucoup de sons commencent à avoir été étudiés en classe » (référent relatif à la progression de la classe), « certaines graphies de sons se confondent facilement comme le [u] et le [ɔ] ou le [e] et le [ɛ] sur le plan visuel » (référent relatif à un savoir professionnel). Par cette entremise, nous pouvons inférer l'hypothèse de Gaby quant à l'origine de cette erreur. Léonard a essayé d'encoder le mot 'fait' tout seul. Il n'a pas confondu les phonèmes [e] et [ɛ] (ce qui est cohérent avec ses performances en déchiffrement) mais, plutôt, au moment de les encoder, les graphèmes correspondants (ce qui est cohérent avec le savoir professionnel mobilisé).

Sur cette base qui délimite finalement l'orientation de son soutien, ou mieux, maintenant, les moyens à employer pour rétablir la situation apaisée, Gaby organise alors une séquence de jugements successifs pour amener progressivement Léonard à la prise de conscience de son erreur. Elle isole avec ses doigts la lettre 'E accent aigu' sur le cahier de Léonard en lui demandant quel son cela produit. Nous entendons ce jugement comme un pointage (Po). Léonard répond « [ɛ] ». Ce référent de signification partagée indique à Gaby que Léonard, non seulement ne lit pas vraiment ce que Gaby lui désigne mais qu'il reste sur sa confusion (il lie le graphème « é » avec le phonème [ɛ]). Invalidant sa proposition (InV), elle lui demande de focaliser sur la seule graphie désignée. Il s'agit du même pointage que le précédent mais accentué par le commentaire de Gaby (codé Po+). La relation entre le jugement de la valuation (n) et le référent de la valuation (n+1) étant défavorable, Gaby accentue alors le soutien apporté par le jugement (n+1) relativement au soutien procuré par le jugement (n)¹⁹ et ce afin d'aboutir à une situation relativement partagée. Ce sont ces adaptations subtiles des jugements évaluatifs qui sont frappantes à ce grain d'analyse. Elles sont éminemment articulées avec le référent de la valuation en cours « résultant » du jugement évaluatif précédent.

Léonard répond alors « [ɛ] ». Cette réponse valide en quelque sorte l'hypothèse de Gaby sur une confusion visuelle. Mais Léonard ne réalise toujours pas, malgré cette modification de l'environnement, qu'il utilise le mauvais graphème. La situation reste donc problématique pour Léonard (référent codé <0). Gaby invalide (InV) puis tournant les yeux vers les affichages au-dessus du tableau noir, lui dit : « non / regarde le sens de l'accent ». Nous avons codé ce jugement C sur fond vert. Il s'agit d'une incitation à comparer (C) la production effective et le modèle (code vert pour désigner les aides matérielles). Nous pouvons à nouveau apprécier la conséquence de la relation jugement (n) / référent (n+1) négative : l'enseignante accentue d'un cran son étayage (nous surcodons également ce jugement avec un +). Léonard observe puis remarque « ah c'est pas le même » (>0). La demande de comparaison a porté ses fruits et Gaby infère donc que la situation est en voie d'être problématisée par Léonard sur ce point au moins (le jugement précédent était donc bien ajusté). Validant sa remarque (V), elle lui

¹⁹ Dont la relative inefficacité a donc été « démontrée » à Gaby par le référent (n+1).



demande « ça fait quoi ça ? » (Q). Léonard ne réagit pas (<0). Ce référé de signification partagée négatif indique à Gaby que la problématisation ne se poursuit pas. Elle doit encore accentuer la modification de l'environnement de Léonard en espérant que celle-ci transforme sa situation : elle précise ce qui doit être comparé en demandant « Le E avec l'accent dans ce sens ça fait [e] ou ça fait [ɛ] ? ». Nous codons ce jugement évaluatif C+ afin de montrer que la nouvelle demande de comparaison est considérée plus étayante relativement au jugement précédent (une simple question). Malgré cela, Léonard ne réagit pas (<0). Gaby montre alors les affichages avec sa main en lui disant « regarde là-haut ». Elle oriente donc davantage Léonard sur l'objet de la comparaison, ce que nous codons donc C+ sur fond vert pour témoigner de cette accentuation à l'aide des outils disponibles du fait de la relation négative (n) / (n+1)²⁰. Léonard répond alors « [e] ». A cet instant, Gaby pense sans doute que la problématisation avance chez cet élève. Il sait désormais qu'il a encodé le phonème [ɛ] avec la graphie qui encode le [e] (>0). Elle valide sa réponse en précisant qu'il s'agit bien du [e] de « étoile » (le mot référent de cette relation graphème-phonème).

C'est un pas important mais insuffisant pour que la situation soit rétablie de manière satisfaisante. En effet, plusieurs graphies sont disponibles pour encoder le phonème [ɛ] (et disponibles à ce moment de l'année dans cette classe). Le mot « fait » étant courant, Gaby exige une correction à un niveau non pas phonologique mais orthographique. Si plusieurs graphèmes sont connus dans cette classe pour encoder le son [ɛ], rien ne permet pour autant à Léonard de savoir lequel choisir. Il s'agit d'un savoir à restituer que nous ne pouvons donc restituer si nous ne le connaissons pas, ce qui est visiblement le cas de Léonard (qui lui s'était orienté vers un « è », même s'il l'a confondu avec le « é »). Gaby en est consciente et devient à cet instant beaucoup plus directive. C'est en effet le seul moyen qui permette d'avancer face à cette part arbitraire de la langue pour reprendre une expression saussurienne. Elle valide tout d'abord le choix de Léonard (V) : « C'est le [e] de étoile / donc ça fait [fe] ». Elle poursuit en lui formulant la phrase avec l'encodage erroné : « le garçon [fe] du ski » (M : Miroir²¹). Elle lui explicite ce qu'elle attend : « moi je veux [fɛ] alors je vais t'aider c'est le [ɛ] de balai » (Exp).

A l'issue de ces jugements de signification partagée permettant de pallier à ce que Léonard ne peut savoir (i.e., il faut utiliser ici la graphie du mot référent « balai »), elle remet Léonard dans une position d'acteur avec ce jugement codé comme une injonction (I) : « d'accord / tu cherches pour écrire avec le [ɛ] de balai [*regardant au tableau, sans doute dans la direction de ce mot référent*] / ils sont là-haut tous les mots tu les cherches [*montrant avec son doigt l'ensemble des affichages*] ». Nous avons codé le dernier jugement (I) en rouge comme enjoignant Léonard à effectuer une opération d'ajustement. Gaby suspend ainsi l'enquête pensant sans doute que la signification est cette fois-ci partagée de manière optimale et que la situation se rétablira grâce à Léonard qui a tous les éléments pour reprendre sa production. Elle s'attend donc à ce que Léonard ajuste en conséquence son écrit. Elle part ensuite aider d'autres élèves puis revient ensuite voir Léonard un peu plus tard²² avec cette attente (composante non représentée). Nous présentons tout d'abord le tableau des valuations (Tableau 4) puis l'analysons.

Tableau 4

Valuations de la séquence emblématique de l'enquête formative (suite et fin)

²⁰ Nous n'avons cependant pas mis deux signes + pour ne pas alourdir les diagrammes à l'excès.

²¹ Il s'agit d'une forme de jugement où l'enseignante soumet la réponse que l'élève a formulé à ce même élève (en accentuant ou non certains traits) afin qu'il se rende compte (en général) de son erreur.

²² Par souci d'exactitude, nous signalons qu'elle revient en fait le voir aussi avant mais pour une autre enquête. C'est aussi un résultat de notre thèse, l'*enchâssement des enquêtes* entre différents élèves mais concernant également le même élève.



Episode	Elève(s) concerné(s)	N° valuation	Multiréférentiel (M)	Référés (R)	Jugements évaluatifs (J)	Actualisation du Multiréférentiel (aM)
G_S4_15	L	1	L a été informé que [fɛ] s'écrit avec le [ɛ] de balai (G_S4_6)	L a écrit [fɛ] FEST	Demande à L s'il se rappelle de quel [ɛ] ils avaient parlé	L a oublié ce qu'ils avaient décidé
		2	L n'est pas très concentré en classe	Absence de réaction de L	Reformule la question	
		3	L oublie régulièrement ce qu'il doit faire	L regarde le tableau et dit "balai"	Valide sa réponse	
		4			Lui demande qu'elle est le [ɛ] de balai	
		5		L répond "A / I"	Valide sa réponse	
		6			Lui demande si c'est ce qu'il avait mis	
		7		L répond que non	Demande à L de corriger	
		8		L corrige		

Gaby arrive donc auprès de Léonard sans être « attirée » par un référent comportemental, ce qui nous paraît compréhensible puisque, pour elle, l'enquête précédente concernant Léonard n'est pas refermée (ou en tout cas, elle n'en est pas certaine). Elle n'a pas « besoin » de ce type de référent pour revenir s'intéresser à la production de Léonard. Son attente est cependant qu'il ait corrigé le mot « fait », elle veut s'en assurer avant de clore l'enquête. Gaby mobilise au sein de son Multiréférentiel trois référents : « Léonard a été informé que [fɛ] s'écrit avec le [ɛ] de balai », « Léonard n'est pas très concentré en classe », « Léonard oublie régulièrement ce qu'il doit faire », trois référents relatifs à sa connaissance de l'élève. Les deux derniers référents lui apportent une explication à la surprise générée par l'inadéquation entre son attente, le premier référent et le référent didactique relatif à l'orthographe lexicale problématique (code orange, <0) : Léonard a écrit « fest ». Ce nouveau référent actualise le Multiréférentiel en renforçant le référent concernant les oublis de Léonard : « Léonard a oublié ce qu'ils avaient décidé ». La signification n'est pas ou plus partagée. Ce que nous appelons plus haut la phase de soutien à une opération de contrôle doit reprendre.

Gaby lui pose alors une question (jugement codé Q) : « tu te rappelles ce que je t'avais dit c'était quel [ɛ] ? ». Léonard ne réagit pas (<0). Elle pose alors la même question en la précisant davantage (Q+) en raison de la relation défavorable jugement (n) / référent (n+1) : « tu te rappelles ce que je t'avais dit le [ɛ] de fait ? ». Léonard regarde alors le tableau noir et dit « balai ». Ce regard et cette réponse constituent donc un référent de signification partagée positif (>0). Gaby valide (V). La situation est à nouveau similaire à la fin de l'épisode précédent.

C'est à la fois positif mais insuffisant puisque ce degré de signification partagée n'a pas suffi pour que Léonard ajuste correctement son écrit. Gaby doit aller encore plus loin dans la tentative de transformation de la situation de Léonard. Elle pose alors une question : « balai / alors tu l'as vu balai ? / c'est quoi le [ɛ] de balai ? » (Q). La réponse de Léonard « A / I » constitue alors un référent clairement positif (>0) signifiant la fin prochaine de l'enquête et le retour à une situation apaisée. Gaby demande à Léonard de comparer sa production et cette



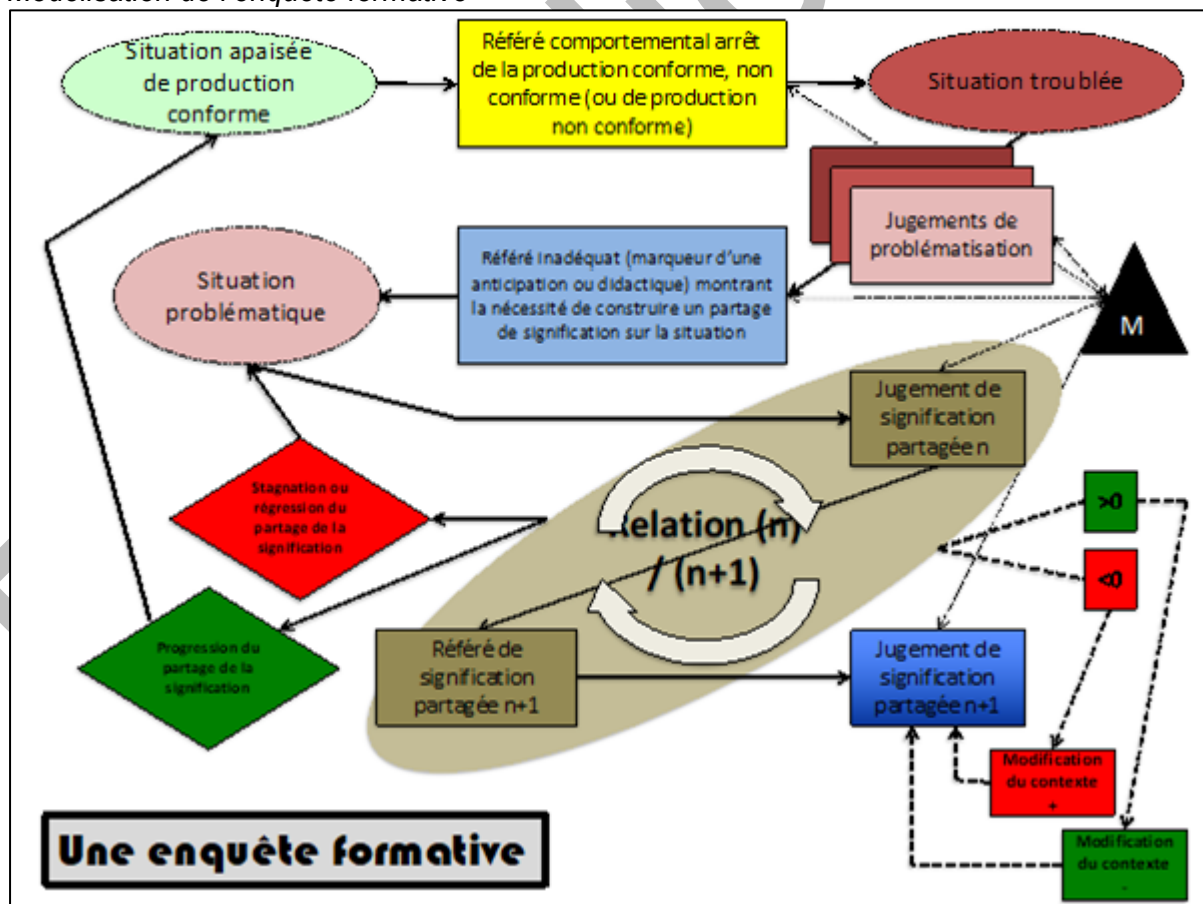
relation grapho-phonétique : « Un A et un I c'est ce que tu m'as mis ? » (Comparaison, C). Ce à quoi Léonard répond par la négative, ce qui est pour Gaby un référé définitivement positif (>0). Elle lui demande alors de corriger, dernier jugement de cette enquête sous la forme d'une injonction (I). Léonard corrige (référé montrant à Gaby non pas ce qu'il fait précisément mais qu'il est en train de le faire) : la situation est à nouveau apaisée pour Gaby, cette enquête est terminée.

Cet exemple, s'il nous paraît particulièrement explicite, est emblématique de l'analyse des données éprouvée sur l'ensemble de notre corpus. Cette congruence récurrente nous permet de proposer une modélisation de l'activité évaluative que nous détaillons ci-après, en guise de conclusion.

5. Conclusion : l'enquête formative, une modélisation de l'activité évaluative des enseignantes

Comme dit à l'instant, ce travail nous a permis de dégager une modélisation de l'activité évaluative des enseignant-es lors de situation de production écrite, en classe de CP. Nous proposons ci-dessous cette modélisation et la discutons brièvement ensuite.

Figure 1
Modélisation de l'enquête formative



Tout d'abord, le triangle noir en interrelation avec tous les référents et tous les jugements de l'enquête symbolise le rôle et la place du *Multiréférentiel*. Il alimente (mobilisation de référents)

et se nourrit (actualisation du Multiréférentiel) lors des enquêtes. C'est par lui également qu'est permise la continuité au sein de l'enquête et entre les enquêtes, qu'elles soient dans la même séance ou situées lors de séances différentes.

Partons maintenant de la *situation apaisée de production conforme* (ellipse verte en haut à gauche). Ici, tout est déterminé : la classe, les élèves, un élève, sont en situation de production selon les us et coutumes de la classe. Cette situation est le fond à partir duquel peut se déclencher une enquête. Un *référé comportemental de production non conforme, d'arrêt de la production conforme* ou *non conforme* peut ainsi venir troubler la situation apaisée. La situation n'est plus déterminée. Face au trouble émergeant chez l'enseignant·e, une enquête se met en mouvement avec la visée de revenir à une situation apaisée en employant des moyens conformes à une valeur centrale : une reprise de la production certes, mais cette enquête doit être l'occasion de permettre une *émancipation* de l'élève. Il ne s'agit pas seulement de rétablir la situation de production. Cette enquête doit également permettre à l'élève de *pratiquer* l'enquête en compagnie de son enseignant·e pour pouvoir s'en détacher un jour.

Le début de l'enquête est le premier pas vers sa résolution comme nous le dit Dewey (1993) et il consiste à *organiser la situation* avec ce que nous en savons, ce que nous en ignorons afin de faire une hypothèse plausible sur les moyens à employer pour atteindre la fin-en-vue. Ce premier pas, le temps de *problématisation* (que nous avons précédemment nommé temps d'*exploration*) est constitué de jugements, les *jugements de problématisation*, qui cherchent à identifier la cause hypothétique du trouble. La réponse à cette question se précipite²³ généralement dans un *référé marqueur d'une anticipation* ou dans un *référé didactique*, l'un ou l'autre étant considérés comme *inadéquat*. Elle ou il sait pourquoi la situation est troublée (ou en fait l'hypothèse tout du moins : cette organisation pouvant se reconfigurer en cours d'enquête si l'hypothèse initiale n'est pas correcte, c'est-à-dire si elle ne permet pas de déployer des moyens modifiant la situation selon la fin-en-vue). La situation est désormais *problématique*.

Un deuxième temps de l'enquête peut alors démarrer, le temps de *résolution* (ou temps de *soutien* dans le début de nos analyses). Il est ici un temps de *construction de signification partagée, à propos de l'identification de la situation troublée et des moyens de sa résolution, avec l'élève*. C'est cette caractéristique qui lui permet de prétendre non pas au seul libellé d'*enquête* mais au titre d'*enquête formative*, d'argumenter *a posteriori* pour une « fonction formative » de cette activité évaluative. La résolution de l'enquête, une fois la situation problématique, ne se fait jamais par le biais des jugements de l'enseignant·e corrigeant elle-même ou lui-même directement les erreurs de son élève. Elle ou il l'accompagne pour qu'elle ou il soit capable d'effectuer cette correction. Rappelons que l'élève a elle ou lui aussi construit une signification, il s'agit de contribuer à la faire converger vers celle de l'enseignant·e et non de lui en assigner une alors qu'elle ou il en serait dépourvu.

Le moyen à disposition de l'enseignant·e pour aboutir à cette construction de signification partagée est la *modification de l'environnement* par des jugements divers (appelés *jugements de signification partagée*). En effet, l'enseignant·e ne peut modifier directement la situation de l'élève. Elle ou il a sa propre situation, l'élève à la sienne, toutes les deux construites sur le fond d'un même environnement mais vécu différemment. La seule chose qu'elles ou ils partagent est cet environnement. Ce temps de construction d'une signification partagée est fondé dans son évolution sur l'appréciation par l'enseignant·e d'une *relation* qui l'informe sur l'état de partage de la signification à propos de la situation et de sa résolution. Cette relation est constituée, dans ce temps, par le *rapport entre le jugement effectué à une valuation (n) et le référé émergent à la valuation suivante (n+1)*. Si ce rapport est positif, et/ou évolue

²³ Au sens chimique du terme.



positivement, c'est-à-dire si le référent montre un partage de signification progressant du côté de l'élève, les jugements de signification partagée (n+1) ne modifient pas davantage l'environnement. Si ce rapport est négatif, les jugements de signification partagée (n+1) s'attachent à modifier davantage l'environnement en espérant que cette modification fasse *signe* pour l'élève et transforme sa situation.

L'appréciation de ce rapport se renouvelle (flèches circulaires sur le schéma autour de (n) / (n+1)) et progresse ainsi jusqu'à ce qu'un référent construit par l'enseignant·e lui indique que la signification est partagée et que la situation pourra être à nouveau apaisée. C'est alors que l'enquête s'arrête tout en ayant permis la participation de l'élève au rétablissement de la situation. La situation est alors déterminée ou l'enseignante a de bonnes raisons de penser qu'elle est en passe de l'être grâce notamment à un ajustement de la production dans lequel l'élève peut s'engager. La production écrite peut reprendre !

Ensuite, de nouvelles enquêtes pourront débiter. La situation est certes apaisée pour l'enseignant·e mais ce n'est plus la *même* situation. Les enquêtes futures pourront compter sur les savoirs construits au travers des enquêtes précédentes et qui ont actualisé le multiréférentiel.

Cette logique de l'enquête formative relève donc d'une inter-objectivation (Zask, 2004) avec l'élève. L'objet n'est véritablement identifié qu'à l'issue de l'enquête et son identification tient au fait que la production écrite a pu reprendre son cours (« la signification d'un objet n'est pas substantielle, mais fonctionnelle » (Zask, 2004, p. 4)). Or, un point crucial de l'enquête formative réside justement dans la participation de l'élève. En raison de sa qualité de participant à cette enquête nous pouvons donc nous demander ce qu'il en est, pour lui, de cette objectivation.

A cet égard, un « objet » n'est pas ce qui s'impose au même titre à tout le monde, mais ce qui sert de pivot à une pluralité d'expériences ultérieures. Plus la liberté et les opportunités d'expériences futures sont importantes, plus l'objet constitué est pertinent (p. 6). Si l'inter-objectivation dans l'enquête est un acquis (la production reprend grâce à la participation de l'élève), cette dernière citation suggère un critère d'évaluation du potentiel formatif de l'enquête. Nous avons certes montré que les enquêtes considérées impliquaient la participation (guidée, Mottier Lopez, 2015 ; Rogoff, 1990) de l'élève mais nous pouvons nous demander dans quelle mesure et à quelles conditions l'enquête permet-elle *véritablement* l'émancipation de l'élève (c'est ainsi que nous interprétons l'idée de liberté et d'expériences futures dans la dernière référence à Zask, 2004) ? Même si nous n'avons dans cette étude que peu d'éléments à apporter concrètement sur ce sujet, cette question reste d'importance et devrait être prolongée par d'autres enquêtes (de chercheurs et de chercheuses) notamment au travers de la notion d'*appropriation* (Theureau, 2011).

Références

- Aeby Daghé, S., & Dolz, J. (2010). Stratégies d'enseignement de la lecture lors de l'écriture de divers genres de textes. *Langage & pratiques*, 40, 44-52.
- Allal, L. (2008). Evaluation des apprentissages. In A. van Zanten (Ed.), *Dictionnaire de l'éducation* (pp. 311-314). Presses Universitaires de France.
- Allal, L. (2007). Régulation des apprentissages : Orientations conceptuelles pour la recherche et la pratique en éducation. In L. Allal & L. Mottier Lopez (Eds.), *Régulation des apprentissages en situation scolaire et en formation* (pp. 7-23). De Boeck.



- Allal, L. (1988). Vers un élargissement de la pédagogie de maîtrise : processus de régulation interactive, rétroactive et proactive. In M. Huberman (Ed.), *Assurer la réussite des apprentissages scolaires. Les propositions de la pédagogie de maîtrise* (pp. 86-126). Delachaux et Niestlé.
- Allal, L., & Mottier Lopez, L. (2005). Formative Assessment of Learning: A Review of Publications in French. In *Formative Assessment - Improving Learning in Secondary Classrooms* (pp. 241-264). OECD Publication.
- Barbier, J.-M. (1985). *L'évaluation en formation*. Paris : Presses Universitaires de France.
- Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139-149.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5-31.
- Bloom, B. S. (1969). Some theoretical issues relating to educational evaluation. In R. W. Tyler (Ed.), *Educational evaluation: new roles, new means: the 63rd yearbook of the National Society for the Study of Education (part II)* (Vol. 69(2), pp. 26-50). University of Chicago Press.
- Bruner, J. (1983). *Le développement de l'enfant : savoir-faire, savoir-dire*. PUF.
- Catach, N. (1980). *L'orthographe française. Traité théorique et pratique avec des travaux d'application et leurs corrigés*. Nathan.
- De Ketele, J.-M. (2012). A la recherche de paradigmes intégrateurs ou de paradigmes unificateurs. In L. Mottier Lopez & G. Figari. (Eds.), *Modélisations de l'évaluation en éducation – questionnements épistémologiques* (pp. 195-210). De Boeck.
- Dechamboux, L. (2018). *L'enquête formative dans le paradigme de l'évaluation située : une modélisation enactive de l'activité évaluative d'enseignantes lors de séances de production écrite au CP* [Thèse de doctorat]. Université de Genève.
- Dewey, J. (2011 / textes traduits par A. Bidet, L. Quéré, G. Truc). *La formation des valeurs*. La Découverte.
- Dewey, J. (1993). *Logique, la théorie de l'enquête*. Presses Universitaires de France. (Original publié en 1938)
- Dunn, K. E., & Mulvenon, S. W. (2009). A critical review of research on formative assessment: the limited scientific evidence of the impact of formative assessment in education. *Practical Assessment Research & Evaluation*, 14(7), 1-11.
- Figari, G. (1994). *Evaluer, quel référentiel ?* De Boeck-Wesmael.
- Figari, G., & Achouche, M. (Eds.). (2001). *L'activité évaluative réinterrogée*. De Boeck Supérieur.
- Frega, R. (2006). *Pensée, expérience, pratique. Essai sur la théorie du jugement de John Dewey*. L'Harmattan.
- Groupe EVA (1991). *Evaluer les écrits à l'école primaire*. Hachette.



- Hattie, J. A. C. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.
- Lesne, M. (1984). *Lire les pratiques de formation d'adultes – essai de construction théorique à l'usage des formateurs*. EDILIG.
- Marcel, J.-F., & Murillo, A. (2013). Analyse des pratiques d'évaluation. In J.-F. Marcel & H. Savy (Eds.), *Evaluons, évoluons. L'enseignement agricole en action* (pp. 137–152). Educagri Editions.
- Mas, M. (1991). Savoir écrire : c'est tout un système ! Essai d'analyse didactique du « savoir écrire » pour l'école élémentaire. *Repères, recherches en didactique du français langue maternelle*, 4, 23-34.
- Mottier Lopez, L. (2016). Une modélisation pour appréhender la référentialisation dans l'évaluation des apprentissages des élèves. In P. Detroz, M. Crahay & A. Fagnant (Eds.), *L'évaluation à la lumière des contextes et des disciplines* (pp. 169-192). De Boeck.
- Mottier Lopez, L. (2015). Evaluation-régulation interactive : étude des structures de participation guidée entre enseignant et élèves dans le problème mathématique "Enclos de la chèvre". *Mesure et évaluation en éducation*, 38(1), 89-120.
- Mottier Lopez, L. (2012). *La régulation des apprentissages en classe*. De Boeck.
- Mottier Lopez, L. (2009). La reconnaissance d'une professionnalité : un processus situé d'évaluation interprétative. In A. Jorro (Ed.), *La reconnaissance professionnelle : évaluer, valoriser, légitimer* (pp. 123-146). Presses Universitaires d'Ottawa.
- Mottier Lopez, L. (2008). *Apprentissage situé : la microculture de classe en mathématiques*. Peter Lang.
- Mottier Lopez, L. (2007). L'impact des communautés de pratique dans les entretiens de coévaluation. In A. Jorro (Ed.), *Evaluation et développement professionnel* (pp. 153-167). L'Harmattan.
- Mottier Lopez, L., & Allal, L. (2008). Le jugement professionnel en évaluation : un acte cognitif et une pratique sociale située. *Revue Suisse des sciences de l'éducation*, 3, 465-482.
- Mottier Lopez, L., Borloz, S., Grimm, K., Gros, B., Herbert, C., Methenitis, J., Payot, C., Pot, M., Schneuwly, J.-L., & Zbinden, S. (2010). Les interactions de la régulation entre l'enseignant et ses élèves, expérience d'une recherche collaborative. In L. Lafortune, S. Fréchette, N. Sorin, P.-A. Doudin & O. Albanese (Eds.), *Approches affectives, métacognitives et cognitives de la compréhension* (pp. 33-50). Presses de l'Université du Québec.
- Mottier Lopez, L., & Dechamboux L. (2017). D'un référentiel d'évaluation fixe à une co-constitution référentielle dynamique, ce que nous apprend le jugement situé de l'enseignant. *Contextes et Didactiques*, 9, 12-30. http://www.espe-guadeloupe.fr/wp-content/uploads/2017/07/2_ArticleMottierDechamboux_VFF.pdf
- Peirce, C. S. (1978). *Ecrits sur le signe* (G. Deledalle, trad.). Editions du Seuil.



- Poizat, G. (2006). *Analyse en ergonomie cognitive de l'activité collective en tennis de table. Contribution à la connaissance des interactions humaines* [Thèse de doctorat], Université de Rouen.
- Rogoff, B. (1990). *Apprenticeship in thinking*. Oxford University Press.
- Simard, C. (1999). L'annotation des textes d'élèves. *Québec français*, 115, 32–38.
- Simard, C. (1992). L'écriture et ses difficultés d'apprentissage. In R. Ouellet & L. Savard (Eds.), *Pour favoriser la réussite scolaire* (pp. 276-294). CEQ-Editions Saint-Martin.
- Simard, C., Dufays, J.-L., Dolz, J., & Garcia-Debanc, C. (2010). *Didactique du français langue première*. De Boeck.
- Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R. M. Gagné & M. Scriven (Eds.), *Perspectives of curriculum evaluation* (pp. 39-83). Rand Mc Nally.
- Theureau, J. (2014, septembre). *Compendium 2014 des notions & méthodes développées jusque là dans le cadre du programme de recherche empirique 'cours d'action', en appui à des cours et séminaires*, 18 Sept. – 30 Sept. Sao Carlos, Belo Horizonte, Brasilia, Rio de Janeiro (Brésil).
- Theureau, J. (2011, juin). *Appropriations 1, 2, 3 ou un exemple de pouvoir heuristique et de capacité de croissance d'un programme de recherche ou appropriation, incorporation et In-culturation*. Communication au colloque Journée Ergo-Idf, CNAM, Paris.
- Theureau, J. (2006). *Le cours d'action. Méthode développée*. Octarès.
- Theureau, J. (2004). *Le cours d'action. Méthode élémentaire*. Octarès.
- Thevenaz-Christen, T., & Leutenegger, F. (2007). Stratégies de régulation de la lecture et de l'écriture de consignes. *Langage & pratiques*, 40, 53-63.
- Tourmen, C. (2009). L'activité évaluative et la construction progressive du jugement. *Les dossiers des sciences de l'éducation*, 22, 101-119.
- Varela, F., Thompson, E., & Rosch, E. (1993). *L'inscription corporelle de l'esprit. Sciences cognitives et expérience humaine*. Le Seuil.
- Zask, J. (2004). L'enquête sociale comme inter-objectivation. In B. Karsenti & L. Quéré (Eds.), *La croyance et l'enquête. Aux sources du pragmatisme* (pp. 141-163). De Boeck.

Remerciements

L'auteur profite de cette publication pour remercier les professeur·es Germain Poizat et Lucie Mottier Lopez qui l'ont accompagné avec exigence et bienveillance dans cette recherche doctorale. C'est grâce à leur soutien sans faille que cette enquête fut aussi formative pour l'auteur...

Lionel Dechamboux

Chargé d'enseignement à l'Université de Genève (section des Sciences de l'Éducation), il appartient à l'équipe EReD (Evaluation, Régulation et Différenciation des apprentissages) sous la direction de la professeure Mottier Lopez. Ses intérêts



scientifiques sont liés à la régulation, la différenciation et plus précisément aux interactions des élèves avec leur environnement dans les dispositifs conçus pour soutenir leurs apprentissages.

Prépublication

