



**HAL**  
open science

## La médiation multimodale du savoir. Leçons tirées de pratiques explicatives en café scientifique

Claire Polo, Jean-Marc Colletta

► **To cite this version:**

Claire Polo, Jean-Marc Colletta. La médiation multimodale du savoir. Leçons tirées de pratiques explicatives en café scientifique. *Éducation & Didactique*, 2020, 14-1, pp.9-29. 10.4000/educationdidactique.5292 . hal-03583600

**HAL Id: hal-03583600**

**<https://hal.science/hal-03583600>**

Submitted on 3 Mar 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# **La médiation multimodale du savoir : leçons tirées de pratiques explicatives en café scientifique.**

**Claire Polo & Jean-Marc Colletta**

## **Résumé**

Cette étude présente une analyse linguistique des pratiques explicatives de lycéennes jouant le rôle d'animatrices pour des collégiens dans un café scientifique. Nous analysons leur utilisation des ressources verbales, gestuelles, et écrites dans leur travail de médiation. Les résultats mettent en évidence le rôle crucial des gestes référentiels et de la focalisation attentionnelle dans le travail de construction du savoir (*médiation conceptuelle*), et de mise en relation des différents modes sémiotiques représentant une même information (*médiation sémiotique*). Nous proposons de modéliser cette activité didactique de *guidage* des élèves comme une performance de médiation multimodale du savoir.

## **Abstract**

### **The Multimodal Mediation of Knowledge : Lessons based on the Analysis of Explanations during a Scientific Café.**

This study consists of a linguistic analysis of explanations made by female student instructors for younger students during a scientific *café*. We analyze how they use verbal, gestural and written semiotic resources to mediate knowledge. Our results show that referential gesturing and attention focus play a great role in both constructing knowledge (*conceptual mediation*) and relating diverse semiotic representation of such knowledge (*semiotic mediation*). We provide a model for such didactical activity of *guidance* as a multimodal performance of knowledge mediation.

**Mots-clefs** : Construction référentielle, Gestes, Guidage, Médiation, Multimodalité.

## Introduction

Les récentes recherches en éducation ont mis l'accent sur le rôle de la communication multimodale dans l'enseignement, comme un moyen de mettre à disposition des élèves plusieurs représentations du savoir ciblé. La dimension multimodale de l'activité d'enseignement a d'abord été étudiée sur le plan de la diversité des supports de représentation externes des concepts comme les images, les graphiques, les formules mathématiques, etc. (Duval, 1995). Cette perspective a été élargie pour prendre également en compte les ressources communicatives constituant des « ensembles sémiotiques » (*semiotic bundle*) (Arzarello, 2004 ; Flevaris et Perry, 2001 ; Lund, Bécu-Robinault, 2009 ; Bécu-Robinault, Lund, 2012). Des travaux se sont ainsi focalisés sur les ressources communicatives en face-à-face, notamment la parole et la gestuelle (Alibali, et al., 2014 ; Goldin-Meadow, 2004 ; Singer et Goldin-Meadow, 2005 ; Valenzeno, et al., 2003). D'autre part, les études portant sur le développement du langage et des compétences discursives ont montré que l'emploi de ressources multimodales se complexifie avec l'âge (Colletta, 2009 ; Colletta & Pellenq, 2009 ; Graziano, 2010), et dépendent du contexte interactionnel (Mazur-Palandre, Colletta, Lund, 2014).

Dans le contexte éducatif semi-formel étudié, à savoir un café scientifique sur l'eau potable, des élèves de niveau lycée se retrouvent en situation d'expliquer les réponses attendues à un quizz à des collégiens. Cet article décrit comment ces jeunes, ayant reçu une journée de formation à l'animation, s'acquittent d'une telle tâche, en situation écologique, à partir d'enregistrements audiovisuels d'interactions authentiques. Le corpus analysé regroupe des extraits de séances menées en France et aux USA. A partir de leur analyse multimodale, nous nous attachons à comprendre comment ces animateurs utilisent les ressources communicatives à leur disposition pour réaliser une médiation des savoirs en jeu auprès des élèves plus jeunes.

## Ancrages théoriques

### *La communication multimodale au cœur de la cognition*

L'approche des interactions verbales qui guide nos recherches est fondée sur une vision complexe et systémique de la communication humaine, dans laquelle le sens est co-construit par l'ensemble des personnes présentes lors d'un échange (voire par les acteurs non-humains, les objets, les configurations spatiales, etc.), et mérite d'être étudié de leur point de vue, dans une perspective « éémique » (e. g. Mondada, 2005). Un tel ancrage théorique rompt, depuis plusieurs décennies, avec la conception simpliste, télégraphique, de la communication, où les rôles d'émetteur et de récepteur sont clairement distincts et étanches et où un silence, par exemple, correspond à une absence de message, et « ne veut rien dire » (Winkin, 2000). Il implique également de considérer les multiples façons dont on peut exprimer et comprendre quelque chose, dans les interactions en face-à-face (voire en ligne), au-delà du seul langage verbal. Plus précisément, notre travail de linguistique cognitive se situe dans la lignée de la version « forte » de la recherche sur la multimodalité, qui considère que les différents éléments composant le message à travers diverses modalités communicatives participent à la construction et l'organisation de la pensée elle-même, et non seulement à son expression (Kita & Ozyurek, 2003, McNeill, 1992, 2000, Kendon, 2004). Le terme de « multimodalité » est fondé sur la métaphore d'un message communicatif perçu comme composé par plusieurs ensembles d'informations circulant sous des formats sémiotiques, ou « modes » divers, constituant une signification globale. Cette définition a donné lieu à

un certain nombre de débats sur ce qui peut être considéré comme mode, la nature spécifique de chacun de ces modes, l'éventuelle adéquation préférentielle entre certains modes et certains contenus informationnels, et les relations entre les modes. Dans les interactions en face-à-face, la parole joue généralement le rôle de mode-pivot, les autres modes étant souvent qualifiés de « co-verbaux ». Ici, notre analyse multimodale est basée sur l'étude de trois modes : la parole, les gestes, et les données écrites projetées sur un écran tout au long des séquences explicatives. Bien entendu, nous sommes conscients que plusieurs types d'informations sont susceptibles d'être véhiculées par chacun de ces trois modes. Ainsi, même le mode de référence, la parole, peut fournir à la fois la signification propre à l'emploi des mots prononcés, mais également des éléments signifiés par une certaine prosodie. Au sein des « gestes », des phénomènes communicatifs assez divers peuvent également être étudiés, allant de l'analyse des regards à la proxémie, ou se limitant à son acception restreinte, c'est-à-dire des mouvements des mains devant le locuteur. Enfin, les données écrites étudiées présentent également une diversité sémiotique, car elles incluent à la fois du texte et des tableaux. Nous précisons plus loin ce qui, pour ces trois modes, a été analysé pour cette étude, en adéquation avec nos questions de recherche. Nous estimons que ces ressources sémiotiques peuvent *a priori* servir chacun des trois buts typiques d'une interaction : a) coordonner la communication, b) signifier une appartenance sociale et c) élaborer des structures sémantiques via la construction référentielle. Notre perspective didactique centrée sur les savoirs nous conduit à nous concentrer sur la dimension référentielle des séquences étudiées (c).

Nous avons donc choisi de focaliser notre étude sur les gestes ayant une fonction référentielle : le pointage et les gestes représentationnels. Les gestes représentationnels ont pour caractéristique d'exploiter les « images » (Goldin-Meadow, 2004, p. 314), en fournissant des représentations incarnées dans le corps du référent. Ils peuvent prendre la forme de gestes des mains décrivant un objet physique concret en traçant dans l'espace sa forme et sa dimension, ou en le localisant (entièrement ou par parties) dans l'espace gestuel, ou encore en y dessinant sa trajectoire. Les gestes représentationnels peuvent également faire intervenir d'autres parties du corps, voire le corps entier, pour mimer des actions correspondant à un référent animé (Colletta, et al., 2010 ; Cosnier et Vaysse, 1997 ; Kendon, 2004 ; McNeill, 1992 ; Streeck, 2009). Enfin, d'autres gestes représentationnels fournissent des métaphores visuelles de concepts abstraits, comme, par exemple, le geste de donner exprime l'idée de partage (Calbris, 2011 ; Cienki et Muller, 2008 ; McNeill, 1992). Les gestes représentationnels peuvent être organisés en séquences sous forme de « phrases gestuelles », ce qui attribue une signification spécifique, à la pertinence locale et contextuelle, à la répétition de certains gestes ou à l'utilisation d'une certaine section de l'espace gestuel. Quant aux gestes de pointage, ils sont définis comme des mouvements corporels projetant un vecteur dans la direction d'un lieu ou d'un objet pertinents dans le contexte d'énonciation en cours (Enfield, et al., 2007, p. 1724). On distingue le pointage concret, dirigé vers des référents physiquement présents dans l'espace interactionnel, et le pointage abstrait, dirigé vers des zones vides de l'espace gestuel auxquelles une signification particulière a été conférée contextuellement. Le pointage concret apparaît tôt dans le développement du langage (Colletta et Pellenq, 2009). Dans notre étude, tous les gestes de pointage concrets sont dirigés vers l'écran où sont projetés les supports didactiques, ayant alors une fonction univoque de référence à ces informations écrites. Le pointage abstrait, comme tous les gestes de l'abstrait, apparaît plus tardivement dans le développement (Colletta & Pellenq, 2009).

### ***Interactions didactiques et médiation du savoir***

Alors que la conversation ordinaire est principalement à finalité interne, visant avant tout la construction et le maintien des relations sociales (but indiqué « b » plus haut), les interactions didactiques, sont, elles à finalité externe, leur raison d'être étant attachée à la (re/co)construction de savoirs (sous-domaine du but indiqué « c » plus haut). En effet, toute situation à finalité didactique est conçue comme fondamentalement constituée par un « système didactique » tripolaire, correspondant aux interactions entre enseignant, élèves et savoirs (Brousseau, 1998). La dissymétrie entre élèves et enseignant, vis-à-vis du savoir, constitue la raison d'être de ce système : sans intervention de ce dernier, qui détient déjà le savoir en jeu, le milieu seul ne permet pas aux élèves de l'acquérir. L'enseignant doit organiser le milieu pour qu'un environnement adidactique devienne une situation didactique, favorable à l'acquisition du nouveau savoir en jeu. De plus, l'ensemble des « routines communicatives » (Gumperz & Hymes, 1972, p. 17) portant sur les conduites attendues des élèves et de l'enseignant vis-à-vis du savoir forment le contrat didactique. Si elles sont susceptibles d'être en permanence négociées pour partie dans les interactions en classe, des éléments contextuels participent à la définition de ces règles, comme le cadre institutionnel posé par les programmes et les rythmes d'enseignement (Bouchard & Traverso, 2006, p. 189), ou la culture scolaire.

En particulier, la dimension multimodale des interactions didactiques commence à être étudiée dans l'idée que faire varier les modes de représentations des concepts faciliterait le processus d'enseignement-apprentissage (Alibali et al., 2014 ; Goldin-Meadow, 2004 ; Singer et Goldin-Meadow, 2005 ; Arzarello, 2004 ; Flevares et Perry, 2001 ; Lund, Bécu-Robinault, 2009 ; Bécu-Robinault, Lund, 2012 ; Valenzeno et al., 2003). Il est en tout cas établi que les enseignants expérimentés ont des pratiques multimodales plus riches, tout du moins en terme de combinaison gestes-parole (Caswell et Neill, 1993). Plusieurs études empiriques donnent à penser que les gestes rendent le discours d'enseignement plus efficace (Alibali, et al., 2013 ; Cook et al., 2013 ; Perry et al., 1995 ; Singer et Goldin-Meadow, 2005 ; Tellier, 2008 ; Valenzeno et al., 2003).

Mais le contexte étudié ici, s'il est bien didactique au sens où il vise des apprentissages, s'éloigne à plusieurs titres du système didactique tripolaire classique. Alors que la didactique s'est historiquement construite sur l'étude des enjeux soulevés par des connaissances disciplinaires, les objectifs pédagogiques portent plutôt sur des compétences transversales (cf. annexe 1). De plus, la figure de l'enseignant a été volontairement « sortie » de la situation, même si elle est évidemment organisée selon une intention didactique. Même si le café scientifique a lieu en classe, ce choix vise à favoriser l'émergence d'un contrat didactique détaché de la pression liée au programme scolaire et tel que personne, dans la situation, n'est dépositaire du savoir. Ainsi, les animateurs n'ont pas, contrairement à l'enseignant, un rôle d'expertise ni d'évaluation. Néanmoins, ils sont amenés à développer des séquences explicatives à propos des savoirs mobilisés dans le café scientifique. L'explication, en tant que genre discursif, existe bien entendu également dans d'autres contextes, et notamment dans les contextes didactiques plus formels, qu'il s'agisse d'une séquence explicative de l'enseignant envers l'ensemble de la classe, ou en aparté pendant un exercice, ou encore d'une explication réalisée entre pairs engagés dans une tâche collaborative. Dans le cas présent, les jeunes qui animent le café ont pour support un « point info » projeté au tableau, justifiant la réponse attendue, qu'ils cherchent à éclairer à travers leurs explications, qui, si elles sont supposées prendre appui sur cette diapositive, s'en tiennent rarement à leur stricte explicitation et véhiculent souvent d'autres éléments de savoir. Pour mener à bien leur tâche, ces novices utilisent spontanément les ressources multimodales à leur disposition, dans un travail relevant de ce que nous appelons la « médiation du savoir ». Dans une telle situation didactique, la posture d'animation consiste ainsi à faciliter et accompagner des personnes

dans la compréhension d'un matériau que l'on s'est préalablement approprié, produit par un autre, plus expert que soi. Le parti pris d'un tel dispositif pédagogique repose sur le processus d'identification des élèves à l'animateur, qui peut jouer ce rôle d'intermédiaire grâce à sa proximité sociale avec eux (âge, statut). Tout l'enjeu de cet article est de documenter finement comment se réalise concrètement cette médiation du savoir auprès des élèves plus jeunes, à travers la caractérisation de la performance explicative multimodale associée. Plus largement, une telle étude peut être riche d'enseignements sur la construction multimodale de la référence, dans toute communication interpersonnelle authentique à visée explicative. Avant de préciser plus avant nos questions de recherche, il est nécessaire de détailler davantage en quoi consiste la situation didactique étudiée.

## **Contexte didactique et orientations méthodologiques**

### ***Un corpus d'enregistrements audiovisuels de cafés scientifiques***

Notre corpus est tiré d'enregistrements de cafés scientifiques menés aux USA et en France, animés par des élèves de 17-18 ans pour des élèves de 12-15. Elle s'inscrit dans un projet international d'éducation à l'environnement qui a donné lieu à une adaptation du modèle du café scientifique à destination de la jeunesse, sous la forme d'un dispositif dénommé *YouTalk*. Le *YouTalk* vise à familiariser les jeunes avec les concepts de base leur permettant, ensuite, de débattre d'une controverse socioscientifique (ici la gestion de l'eau potable). Ce dispositif a été conçu dans un partenariat, en France, entre une équipe de recherche en didactique des sciences et une association d'éducation populaire. Les établissements scolaires, en France et aux USA, ont principalement été sollicités pour accueillir une formation à l'animation pour les élèves plus âgés, puis quelques cafés scientifiques ponctuels. Les objectifs pédagogiques revendiqués par le projet, sur lesquels l'ensemble de ces acteurs s'accordaient, mettaient l'accent sur le travail de compétences transversales, où l'éducation scientifique est principalement mobilisée au titre d'éléments littéraciques indispensables à la formation du citoyen (cf. annexe 1).

L'ensemble du café est animé par des élèves de niveau lycée volontaires et ayant reçu une formation spécifique de 7 heures, les adultes parfois présents au cours du café gardant un rôle d'observation. Chaque café a duré environ 90 minutes aux USA, et 120 minutes en France. Il comporte des temps de discussion en classe entière, de discussion en groupe de 4 élèves, des phases de vote individuel à l'aide d'un dispositif électronique de sondage en temps réel (*Activote*), des phases de vote de groupe à propos de problèmes socio-scientifiques, à l'aide d'un carton levé. La séance commence par une grande question, qui donne l'orientation générale du café, à propos de laquelle les opinions initiales sont recueillies, et qui est longuement discutée en fin d'activité : « De quoi dépendra l'accès à l'eau potable à l'avenir ? ». Le café est structuré par un questionnaire à choix multiple, organisé en trois cycles thématiques conçus pour étayer la réflexion au sujet de cette grande question : 1) l'état des ressources en eau ; 2) les usages de l'eau ; 3) la gestion de l'eau potable. Certaines questions, dites « questions de connaissance » (QC), fonctionnent comme des devinettes permettant d'acquérir des savoirs de base sur le sujet, tandis que d'autres ouvrent des débats d'opinion. Chaque cycle thématique comprend 2 (USA) ou 3 (France) QC puis une question de débat. Les participants votent électroniquement à propos des QC. Les résultats des votes sont immédiatement affichés, et la bonne réponse apparaît en vert. Les animateurs doivent alors expliquer pourquoi cette réponse était attendue, à l'aide d'une diapositive projetée à l'écran nommée « point

info », qui mentionne les sources des informations fournies. C'est cette tâche spécifique que nous étudions dans cet article.

Toute l'activité a fait l'objet d'un enregistrement audiovisuel complexe, comprenant des vues de deux groupes d'élèves, un enregistrement de l'écran d'ordinateur projetant le diaporama, et une vue globale de la classe. C'est ce dernier film qui a principalement été utilisé pour l'analyse présentée ici. Six cafés ont été entièrement enregistrés (3 dans chaque pays), mais l'un d'eux a été mis de côté car l'animateur sortait trop du champ de la caméra lors de ses explications.

### *Enjeux didactiques soulevés par les séquences explicatives sélectionnées*

Nous nous sommes concentrés sur deux QC impliquant des informations de statuts épistémiques différents. Notre corpus d'analyse comprend donc 2 séquences pour 5 sujets, tous féminins, 3 américaines et deux françaises (ci-après nommées respectivement Cathy, Marlene, Iris, Océane et Sylvie). Ces séquences ont été conçues comme interactives, les animatrices pouvant être amenées à interagir avec les élèves et, dans ces données, le temps dédié au dialogue était très variable. A des fins de comparaison, nous avons donc décidé de limiter notre propos aux parties monologiques des explications réalisées lors de ces « points infos ». Le tableau 1 récapitule les données analysées. L'ensemble des 10 séquences compte plus de 18 minutes, et les parties monologiques comptent pour près de 10 minutes.

	Animatrice	Question	Durée passée sur le "point info"	Durée de monologue explicatif
USA	Cathy	1	1 min 26	50.1 s
		2	2 min 16	1 min 25
	Marlene	1	1 min 46	45.6 s
		2	2 min 40.4	1 min 34.4
	Iris	1	26 s	26 s
		2	1 min 02	48.7 s
FRANCE	Sylvie	1	2 min 53.4	44.8
		2	1 min 48	48.3 s
	Océane	1	1 min 30	1 min 01.6
		2	2 min 24	1 min 30.8
<b>Total</b>		<b>10 questions</b>	<b>18 min 12 s</b>	<b>9 min 55.2 s</b>

**Tableau 1. Corpus d'analyse**

Les deux QC concernées présentent chacune des enjeux didactiques spécifiques quant aux objets de savoir mobilisés. Les diapositives correspondant à ces questions sont reproduites en annexes 2 à 5. La première (désormais Q1) traite de la quantité d'eau nécessaire pour fabriquer des produits de consommation courante. Il s'agit d'un raisonnement typique de la doxa de l'éducation environnementale, mais qui est largement contre-intuitif, et éloigné du savoir scolaire traditionnel. En effet, dans une optique d'éducation à l'environnement, l'enjeu est de faire prendre conscience aux élèves de l'empreinte écologique, ici en termes de consommation d'eau, de quelques-unes de leurs pratiques quotidiennes (habitudes alimentaires, utilisation de produits électroniques). L'objectif n'est évidemment pas qu'ils retiennent les chiffres associés à la production d'un kilogramme de chacun des éléments considérés, mais qu'ils réalisent les deux points suivants : 1) les processus de production peuvent varier (notamment selon les pays) ; 2)

ils sont, de ce fait, susceptibles d'avoir une empreinte écologique plus ou moins importante ; 3) des tendances globales apparaissent tout de même, notamment le coût important en eau de la consommation de bœuf. Cette question vient préparer les élèves au débat sur la question finale du cycle thématique, concernant les efforts qu'ils seraient prêts à consentir pour préserver les ressources en eau, qui mentionne le végétarisme.

La seconde (Q2) compare le prix de l'eau potable à différents endroits de Manille et dans des pays développés. Cette question fait davantage appel à des concepts courants de coût et de prix, dont les participants ont déjà une représentation, même si elle ne correspond pas nécessairement à leur définition économique. Ici encore, l'intention didactique ne porte pas sur l'apprentissage des données chiffrées, mais sur la compréhension des quelques mécanismes suivants : 1) le prix payé pour l'eau potable n'est pas égal à ce qu'elle coûte (ce qui suppose d'acquérir une distinction conceptuelle claire entre « coût » et « prix ») ; 2) il peut être inférieur, notamment dès lors que l'alimentation en eau potable est mise en œuvre dans le cadre d'un service public, organisant une collectivisation et un subventionnement des coûts de réseau ; 3) il peut être supérieur, notamment dès lors que la vente de l'eau est réalisée par l'initiative privée individualisée, à l'image du « porteur d'eau », qui a besoin de faire un bénéfice pour vivre de cette activité ; 4) actuellement, ce sont les personnes les plus pauvres du monde qui paient l'eau le plus cher. Ces éléments visent à constituer des bases de réflexions pour la question de débat qui clôt ce cycle thématique, portant sur la façon dont il serait souhaitable que le prix de l'eau soit déterminé.

### *Questions de recherche et démarche analytique*

Notre étude visait à comprendre comment les animatrices utilisent spontanément les ressources multimodales à leur disposition dans ces séquences explicatives, pour caractériser leur travail de médiation des savoirs en tant que performance communicative et pratique didactique. Dans cette optique, nous nous sommes focalisés sur quatre questions de recherche :

- 1) Comment caractériser la contribution de leur parole à ce travail de médiation ?
- 2) Quelles informations viennent l'enrichir par le recours à d'autres modes sémiotiques (gestes, focalisation attentionnelle, supports didactiques écrits) ?
- 3) Comment l'ensemble de cette performance multimodale fonctionne-elle de façon cohérente et globale pour fournir une explication des concepts accessible aux élèves plus jeunes ?
- 4) Au-delà de cette médiation conceptuelle, les animatrices accompagnent-elles les collégiens dans le passage d'un registre sémiotique à l'autre, sous la forme de ce qu'on pourrait appeler une « médiation sémiotique » ?

La performance des animatrices a été analysée selon trois étapes : a) la médiation verbale ; b) le comportement gestuel, et c) la comparaison des performances multimodales observées, en terme de médiation conceptuelle et de médiation sémiotique, par le recours à des cas d'étude adressant les questions 3 et 4.

La première phase, centrée sur le matériau verbal, s'est appuyée sur une transcription complète des dires des animatrices, selon les conventions définies en annexe 6. Nous avons ensuite identifié les éléments parlés correspondant à une répétition littérale des termes écrits sur la diapositive (soulignés dans les transcriptions). Généralement, ils consistaient en quelques mots. Pour l'une des animatrices, cependant, toute une phrase voire plusieurs phrases étaient répétées, révélant une stratégie de simple lecture de la diapositive. Mais la plupart des discours verbaux comprenaient de véritables apports contribuant à la médiation du

contenu informationnel de la diapositive. Ces éléments de médiation verbale ont donné lieu à une classification.

La seconde phase analytique, centrée sur les gestes des animatrices, a été réalisée à l'aide du logiciel ELAN, où extraits vidéo et transcriptions ont été importés. Nous avons d'abord codé leur focus attentionnel, sur la base de l'orientation du regard et de la tête. Nous avons ensuite identifié et analysé les gestes représentationnels et les pointages réalisés lors des paroles correspondant à de la médiation verbale. Un référent a été associé à chaque geste représentationnel ou de pointage abstrait. Une visualisation sous forme de tableau a ensuite été réalisée pour chaque extrait (91 au total), afin de mettre en perspective les informations attentionnelles, gestuelles et verbales co-occurentes. Cette étape analytique nous a permis de repérer les éléments de médiation inédits, provenant uniquement de l'information gestuelle. Parfois, les informations de médiation passaient à la fois par la parole et par le geste. Souvent, les gestes apportaient des éléments spécifiques qui n'étaient pas véhiculés par la parole. Ainsi, des extraits classés comme simples « reformulation » à partir du matériau verbal consistaient en réalité à un travail de médiation plus profond, des apports informationnels plus importants étant réalisés en parallèle au plan gestuel. Cela nous a conduit à modéliser les apports co-verbaux comme l'alternance de deux attitudes typiques : une gestualité représentationnelle face au public et un pointage concret face à l'écran de projection.

Enfin, nous avons cherché à identifier, en dernière phase analytique, des tendances récurrentes dans les performances multimodales étudiées. Cette activité de médiation peut ainsi être décrite comme l'enrichissement de la construction référentielle à la fois par l'ajout d'informations plus ou moins inédites (médiation conceptuelle), et par l'établissement de ponts entre différentes représentations sémiotiques des concepts en jeu (médiation sémiotique). Les différences et similitudes inter-individuelles observées nous ont conduit à nous interroger sur les effets didactiques de ces performances communicatives, et à proposer, pour discussion, un modèle de la médiation multimodale du savoir.

## **Résultats et études de cas**

### ***Paroles médiatrices : classification des apports verbaux***

La première étape analytique a été réalisée par comparaison des transcriptions des paroles des animatrices avec les diapositives « point info » qu'elle avaient pour tâche d'expliquer. Leur discours verbal a ainsi été segmenté en unités sémantiques classés en 3 catégories : *reformulation*, *élucidation* et *ajout*. Lors d'une *reformulation*, l'animatrice n'apporte pas réellement de nouvelle information, mais redit l'information présente avec ses propres mots. Lors d'une *élucidation*, l'animatrice met en relation des informations présentes séparément sur la diapositive. Enfin, lorsque le discours verbal des animatrices mobilise des informations absentes de la diapositive, nous parlons d'*ajouts*. Ces apports consistaient principalement en des exemples, raisons, causes ou implications et conséquences d'informations présentes sur la diapositive. *Le tableau 2* résume cette classification, en fournissant un exemple authentique pour chaque catégorie analytique. Dans notre corpus, nous avons identifié au total 32 *reformulations*, 23 *élucidations* et 36 *ajouts*.

Type d'apport verbal	Reformulation	Elucidation	Ajout
<b>Définition</b>	Pas d'apport informationnel  Reformulation des informations présentes sur la diapositive, dans des termes propres.	Mise en relation d'informations présentes à différents endroits de la diapositive  OU  Inférences réalisées à partir du contenu de la diapositive, éventuellement avec une interprétation	Apport de nouvelles informations non présentes sur la diapositive.  <i>Exemples</i> <i>Justifications, causes, raisons.</i> <i>Conséquences</i> <i>Etc.</i>
<b>Exemples tirés de Q2</b> (annexe 3)	<i>so the way they get their water is they have to buy it from er: tank trucks and street vendors\</i>	<i>the reason that it's more expensive is because they don't have access to: like a network of: of plumbing\ and su- and supplies like that\</i>	<i>like in kenosha it doesn't matter if you're richer or if you don't have as much money you guys are all paying the same for your water\</i>
<b>Texte de la diapositive</b> (reproduite en annexe 4)	<i>people living in shantytowns in the developing world buy their drinking water at the highest price in the world, from tank trucks and street vendors</i>	<i>in cities, there can be very sharp differences in prices depending on whether one is connected to local water networks or not</i>	<i>in general, once one is connected to local water networks, household users all pay the same price, regardless of whether they are rich or poor</i>

**Tableau 2. Catégories analytiques pour l'étude des apports verbaux à la médiation**

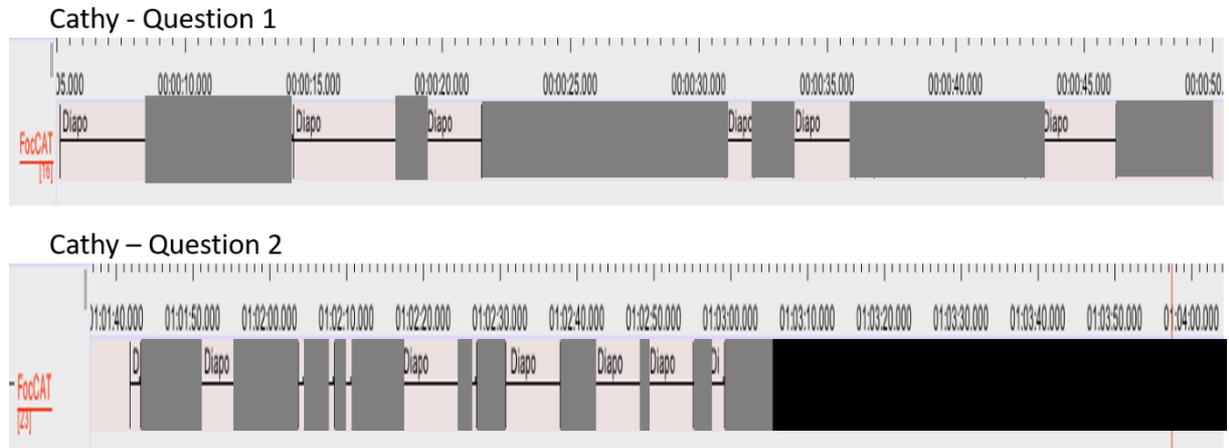
Lors des phases analytiques suivantes, incluant des aspects co-verbaux de ces séquences explicatives, un premier résultat a consisté à mettre en évidence un schéma d'alternance de focalisation attentionnelle entre la diapositive projetée et les élèves, associée à une gestualité référentielle spécifique. Après avoir présenté ces tendances générales, nous détaillons un cas illustrant ce qui arrive typiquement lors de la focalisation « élèves », où les animatrices utilisent souvent des gestes représentationnels en complément de leur parole. Un second cas correspondant à une phase de transition entre deux focalisations attentionnelles différentes, donne à voir comment la parole, les gestes et l'orientation de la tête participent d'une médiation entre différentes représentations sémiotiques d'un même concept. Finalement, nous proposons un modèle caractérisant ces pratiques de médiation pour leurs apports didactiques multimodaux.

### *Contributions co-verbales à la médiation des savoirs : tendances générales*

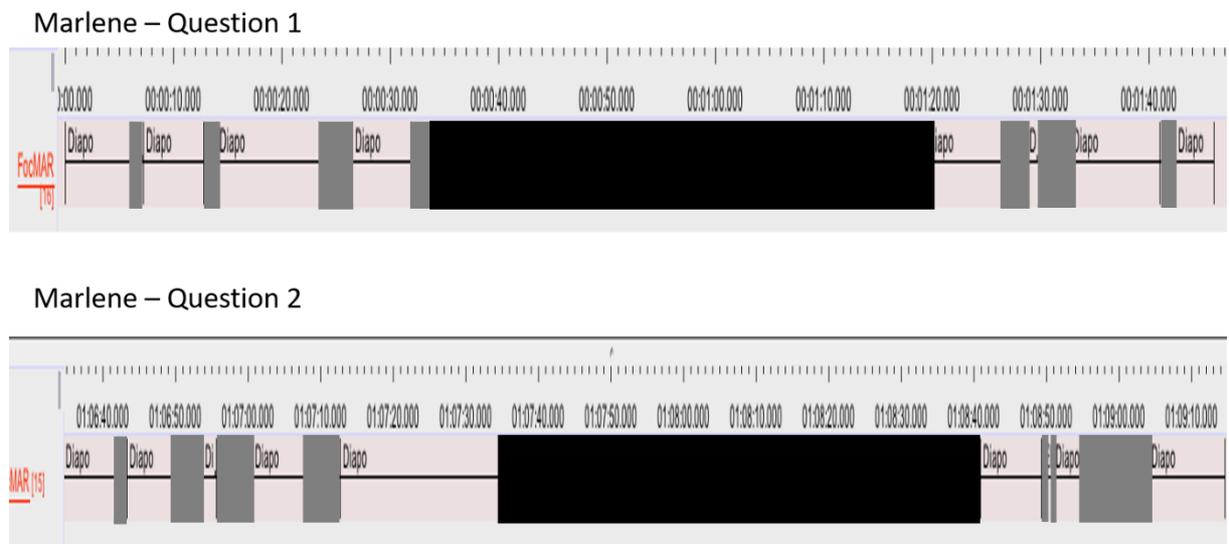
Avant d'en venir à l'analyse gestuelle au sens strict, nous nous sommes intéressés à la direction générale de l'attention des animatrices, repérée par l'orientation de leur regard et de leur tête, parfois suivis par le buste ou l'ensemble du corps. Or les animatrices ont également tendance à varier énormément le type de gestualité référentielle qu'elles produisent selon leur focalisation attentionnelle : gestes représentationnels et médiation conceptuelle face aux élèves contre pointage concret et médiation sémiotique face à la diapositive.

*Partager son attention : un schéma d'alternance élèves / diapositive*

Il est frappant de remarquer que tous les sujets étudiés partagent un schéma d'alternance attentionnelle entre le support didactique projeté et les élèves. Les figures 1 à 5, construites à partir de captures d'écran des analyses réalisées dans ELAN, donnent à voir cette tendance pour chacune des animatrices. Les moments de dialogue n'ont pas été codés [en noir]. Les moments de focalisation attentionnelle vers les élèves sont grisés.

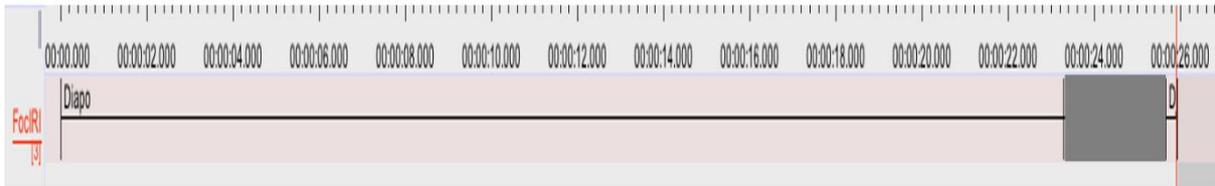


**Figure 1. Focalisation attentionnelle en explication monologuée : Cathy.**



**Figure 2. Focalisation attentionnelle en explication monologuée : Marlène.**

Iris – Question 1



Iris – Question 2

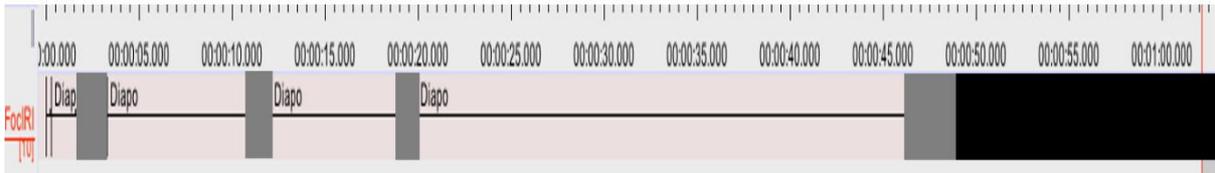
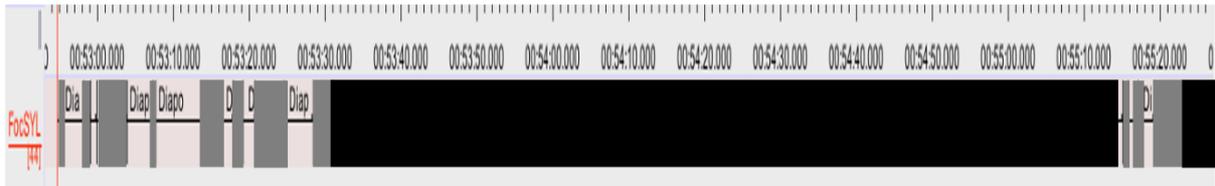


Figure 3. Focalisation attentionnelle en explication monologuée : Iris.

Sylvie – Question 1



Sylvie – Question 2

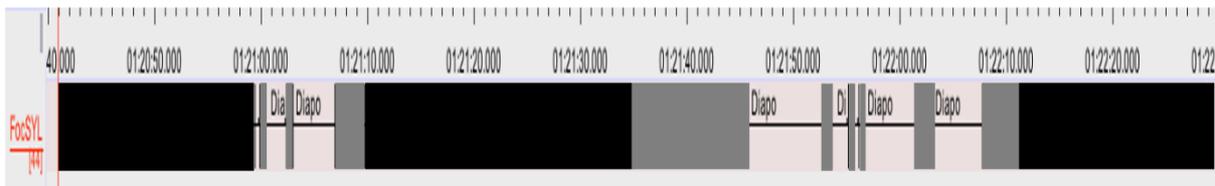


Figure 4. Focalisation attentionnelle en explication monologuée : Sylvie.

### Océane – Question 1



### Océane – Question 2



**Figure 5. Focalisation attentionnelle en explication monologuée : Océane.**

Comme on peut le constater sur ces figures, le temps de chaque épisode de focalisation attentionnelle varie beaucoup. Pour la seule Q2, l'épisode de focalisation attentionnelle « élèves » le plus long dure 18 s. (Océane), et le plus court 0,5 s (Sylvie). Ainsi, on ne peut pas établir de durée typique pour un épisode de focalisation attentionnelle, à partir de ces données. Cependant, on peut s'intéresser aux parts d'attention « élèves » et « diapositive » respectives dans le temps d'une séquence explicative. Cette information est fournie dans le tableau 3, qui présente également la part de gestes représentationnels et de pointage sur la durée de l'explication. Ces données témoignent de l'importance des gestes dans leur performance, les animatrices fournissant en moyenne, toutes les 10 secondes - mis à part le cas exceptionnellement pauvre en gestes, d'Iris - entre 2,4 et 6,3 gestes représentationnels ainsi que 1,2 à 6,2 gestes de pointage.

Anim.	Q.	Durée monologuée	Fraction temporelle vers les élèves	Fraction temporelle avec des gestes repr.	Nombre de gestes repr. (/10 sec)	Fraction temporelle vers la diapo.	Fraction temporelle avec des gestes de pointage	Nombre de pointages (/10 sec)	
USA	CAT	1	50.1 sec	55,5 %	27,0 %	12 (2.4)	44,5 %	38,9 %	19 (3.8)
		2	1 min 25	53,5 %	30,1 %	24 (2.8)	46,5 %	29,7 %	25 (2.9)
	MAR	1	45.6 sec	23,0 %	27,9 %	20 (4.4)	87,0 %	14,9 %	9 (2.0)
		2	1 min 34.4	27,6 %	<b>45,8 %</b>	59 (6.3)	72,4 %	9,8 %	11 (1.2)
IRI	1	26 sec	<b>9,1 %</b>	28,4 %	6 (2.3)	<b>90,9 %</b>	<b>0,0 %</b>	<b>0 (0.0)</b>	
	2	48.7 sec	<b>14,3 %</b>	<b>12,9 %</b>	<b>5 (1.0)</b>	<b>85,7 %</b>	<b>1,8 %</b>	<b>1 (0.2)</b>	
FR.	SYL	1	44.8 sec	48,6 %	25,3 %	17 (3.8)	51,4 %	22,3 %	6 (1.3)
		2	48.3 sec	44,8 %	19,6 %	20 (4.1)	55,2 %	<b>57,6 %</b>	<b>30 (6.2)</b>
	OCE	1	1 min 01.6	39,0 %	<b>22,4 %</b>	<b>26 (4.2)</b>	61,0 %	<b>42,3 %</b>	<b>19 (3.1)</b>
		2	1 min 30.8	64,1 %	<b>43,8 %</b>	<b>42 (4.6)</b>	35,9 %	30,6 %	19 (2.1)

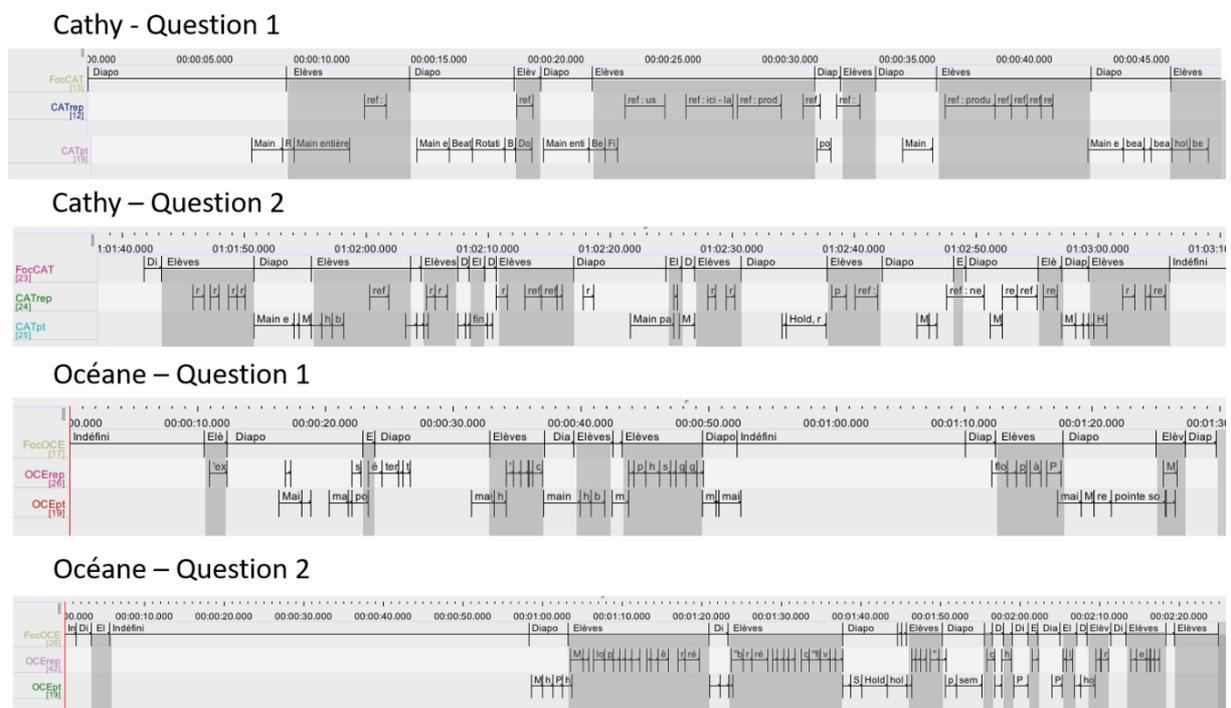
**Tableau 3. Fractions temporelles de focalisation attentionnelle vers les élèves et vers la diapositive, et proportions de gestes représentationnels et de pointage concret.**

## Une gestualité référentielle associée à la focalisation attentionnelle

Le tableau 3 met aussi en évidence des différences inter-individuelles. Iris regarde peu les participants (moins de 10% du temps pour Q1 et de 14% pour Q2), et focalise surtout son attention sur la diapositive. C'est aussi elle qui produit le moins de gestes, ne faisant même aucun pointage pour Q1, et ne produisant des gestes représentationnels que sur 12,9% du temps pour Q2 (soit 1 geste en moyenne toutes les 10 secondes).

Marlene passe un quart à un tiers de son temps face aux élèves, tandis que Cathy et Sylvie y passent environ la moitié de leur temps. La pratique d'Océane varie davantage, passant environ 40% du temps face public pour Q1, contre plus de 64% pour Q2. D'ailleurs, la fraction de temps passée à faire des gestes représentationnels est 2 fois plus élevée pour Q2 que pour Q1 (respectivement 43,8% et 22,4%), même si son taux gestuel reste similaire (4,6 contre 4,2 gestes pour 10 secondes), du fait que ses gestes durent plus longtemps dans le second cas.

Nous avons ainsi observé une tendance à l'association *attention focalisée sur l'écran et pointage concret* d'une part et *attention vers les élèves et gestualité représentationnelle* d'autre part. Ce schéma de comportement est particulièrement visible chez Océane et Cathy. La figure 6 présente les captures d'écran ELAN correspondantes.



**Figure 6. Corrélation entre focalisation attentionnelle et type de gestes référentiels : Cathy et Océane.**

La première ligne code la focalisation attentionnelle, grisée lorsqu'elle est face aux élèves. La seconde code les gestes représentationnels, et la troisième le pointage concret. La plupart des gestes représentationnels sont réalisés lors d'une focalisation vers les élèves (zone grisée), et la plupart des pointages concrets en focalisation vers la diapositive (hors zone grisée). Le tableau 4 reprend ces données sous forme de nombres de gestes. Sur 44 gestes de pointage, Cathy n'en fait que 14 face aux élèves, et Océane 6 sur ses 38 gestes de pointage. L'association gestes représentationnels et attention « élèves » semble encore plus systématique : Cathy produit seulement 3 de ses 36

représentationnels face à la diapositive, et Océane 5/68. Quelle est la fonction de cette combinaison multimodale parole, gestes représentationnels et attention « élèves » ? Nous abordons maintenant cette question via une étude de cas.

	<b>Animatrice</b>	<b>Gestes représentationnels</b>	<b>Gestes de pointage</b>	<b>Total</b>
<b>Attention vers les élèves</b>	Cathy	33	14	47
	Océane	63	6	69
<b>Attention vers la diapositive</b>	Cathy	3	30	33
	Océane	5	32	37
<b>Total</b>	Cathy	36	44	80
	Océane	68	38	106
	Total	104	82	186

**Tableau 4. Nombre de gestes représentationnels et de pointage selon la focalisation attentionnelle : Cathy et Océane.**

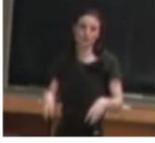
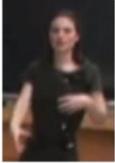
*Cas 1 : médiation conceptuelle par la gestualité représentationnelle face aux élèves*

Nous nous attachons ici à la performance explicative de Cathy à propos de Q1. Si elle répète plusieurs termes écrits sur la diapositive, elle ne se contente pas de la lire, et réalise de nombreux ajouts verbaux. Deux segments de son discours ont ainsi été identifiés comme des reformulations, deux comme des élucidations, et trois comme des apports. Ils sont présentés dans le tableau 5, les termes repris de la diapositive étant soulignés (mots exacts et expressions lexicales).

Reformulations	Elucidations	Ajouts
<i>but worldwide (...) all those coffee beans being produced and grown in other countries takes an immense amount of water\</i>	<i>so: (0.9) the: w- w- of course the water footprint how much is used pound for pound varies between countries\</i>	<i>because most of our: er: coffee beans and everything are not grown in the us\ so then the production em: as far as we would see it here is lower/</i>
<i>em are they any other option (...) er maybe you don't know how necessarily the water is used</i>	<i>so: em: (0.5) as you can see in the us: it's very: very small compared to: other countries:\ and em the world average\</i>	<i>is what you have to keep in mind em: (0.5) is that</i>
		<i>that you guys: strongly thought of that (...) and wanna ask about it/</i>

**Tableau 5. Caractérisation des apports verbaux – Cathy, Q1.**

Pour illustrer notre propos, nous nous concentrons maintenant sur le premier segment, selon un ordre chronologique, identifié comme un *ajout* (encadré dans le tableau 5). Cathy ne se contente pas d'apporter une information inédite verbalement, elle produit dans le même temps 6 gestes complétant ce contenu informationnel. Ses 5 premiers gestes sont représentationnels, et produits face aux élèves. Son dernier geste consiste à pointer la diapositive avec sa tête, en lui faisant face, et intervient sur la fin de la clause verbale « *is lower* ». Le tableau 6 décrit chacun de ces gestes, en perspective avec le matériau verbal co-produit, qui apparaît en gras. Un référent a également été attribué à chaque geste ou composant de geste.

Parole	Gestes représentationnels	Photo
1. <i>Because most of our: er: coffee beans</i>	Deux mains : - Parallèles en métaphore du contenant > <b>chose ou quantité assez grande</b> - Paumes vers elle > <b>sienne</b>	
2. <i>and everything</i>	Parallélisme de la métaphore du contenant maintenu > <b>chase/quantité</b> + pointage abstrait à droite > <b>autre, différente</b>	
3. <i>are not grown in the us\ so then</i>	Mains pointant vers le bas, battements sur « not », « grown » et « us » > <b>sol, terre, ici</b>	
4. <i>the production em: as far as we would see it</i>	Polysigne principalement avec la main droite - forme de griffe : <b>prendre</b> - mouvement vers le haut : <b>pousser</b> - trajectoire cyclique : <b>processus</b> > <b>cultiver, récolter</b>	
5. <i>here</i>	Deux mains pointant vers le bas > <b>cette terre, usa</b> + mouvement vers la droite > rapprochement de la diapositive	

**Tableau 6. Cas 1 – Discours et gestes représentationnels de Cathy.**

Cathy emploie d'abord ce que McNeill (1992, 15) appelle une métaphore « du contenant » pour représenter une relativement grande quantité ou chose (co-occurent avec la mention verbale des grains de café). Le premier geste est un polysigne (1), au sens de Calbris (2011, p. 28) : il représente simultanément plusieurs notions à travers plusieurs composants gestuels fournissant chacun un lien analogique à la fonction de signe. Ici, la forme de la main, paume vers Cathy, renvoie à l'idée de propriété, simultanément à l'emploi du mot « *our* » (notre). Gardant l'effet de parallélisme entre ses mains, elle produit ensuite un pointage abstrait vers la droite (2), venant spécifier l'information verbale vague « *and everything* » (et tout) comme une chose de même taille ou en même quantité, mais d'une autre nature. Le troisième geste (3) consiste à pointer vers le bas ses deux mains, une structure gestuelle qu'elle maintient tout au long de l'énoncé « *are not grown in the us* » (n'est pas cultivé aux Etats-Unis), avec des battements insistant sur « *not* », « *grown* » et « *us* ». Elle se réfère ainsi au sol, spécifiant l'usage de « *grown* » comme « *provenant du sol* », mais elle fait référence également, au plan métaphorique, à la terre, cette terre, complétant le référent verbal « *the us* », comme pour rappeler qu'ici nous sommes aux USA. Le geste co-réalisé avec le mot « *production* » (4) est emblématique du polysigne. Sa main droite forme une griffe orientée vers le sol, donnant à voir l'idée de prendre quelque chose ; son mouvement vers le haut fournit en même temps l'idée d'une plante qui pousse ; enfin, sa trajectoire cyclique, faisant intervenir l'autre main, est une métaphore courante de la notion de processus. Ce sont ces trois aspects de son geste qui permettent à Cathy de préciser ce qu'elle entend par « *production* » comme « *culture* » ou « *récolte* ». Son dernier geste

représentationnel (5) consiste en la répétition du pointage vers le bas lors de l'énoncé « *here* » (ici), faisant encore référence à « cette terre », qui, en contexte, renvoie aux Etats-Unis. Ainsi, la gestualité représentationnelle de Cathy fonctionne ici de façon séquentielle : le second geste prend le sens de « *another* » car elle déplace ses mains sur le côté ; le cinquième réfère aux USA car un geste similaire a été fait préalablement en prononçant « *the us* ». Cette gestualité représentationnelle permet à Cathy de préciser sa médiation verbale sur un élément de savoir visés par Q1, à savoir le fait que les processus de production peuvent varier selon les pays, qui apparaissent ici de façon métaphorique comme des « terres », présentes sous les pieds des participants, ou « ailleurs ». Cet univers métaphorique permet également de préparer la suite de son explication, ayant trait au second élément de savoir important dans cette question, la comparaison des empreintes écologiques des différents modes de production. En effet, l'image de la terre, uniquement mobilisée au plan gestuel, permet d'englober 3 aspects centraux des savoirs abordés ici : elle précise le processus de production dont il est question, en renvoyant à la culture, qui rend le lien avec la consommation d'eau plus évidente ; elle rend concrète la notion de pays comme étant le sol sur lequel on est ou un autre ; et elle ancre l'explication dans un univers métaphorique qui facilite la mobilisation de la notion analogique d'« empreinte en eau » (*water footprint*).

Enfin, il convient de préciser que Cathy produit son cinquième geste représentationnel sur sa droite, alors qu'elle s'apprête à pointer l'écran, également situé à sa droite. Ce pointage concret est co-réalisé avec l'énoncé « *is lower* » (est inférieure). Ce premier mouvement vers la droite lors du geste (5) peut ainsi être perçu comme une façon d'intégrer le pointage qui suit à la phrase gestuelle globale. Elle transporte littéralement le référent (la production américaine) vers la diapositive pour s'assurer que les élèves comprennent à quoi renvoie « *lower* ».

### ***Cas 2 : La médiation sémiotique : établir des ponts entre les modes de représentation***

Ce dernier point ressemble à une pratique plus systématique chez certaines animatrices, lorsqu'elles passent d'une focalisation attentionnelle à une autre. Le schéma d'alternance de focalisation attentionnelle lui-même peut d'ailleurs être compris comme jouant cette fonction de pont entre des représentations sémiotiques différentes d'une même information, verbo-gestuelle et écrite (texte et tableau). Cependant, les animatrices fournissent parfois un travail de médiation spécifique pour accompagner ce transfert informationnel d'un mode à l'autre. Nous présentons ici un cas emblématique de cette pratique, basé sur un extrait de la performance explicative de Marlene lors de la Q1. Globalement, Marlene reprend des termes de la diapositive, mais elle ne la lit franchement que lorsqu'elle mentionne la source scientifique des informations présentées. Cette séquence est assez riche en apports verbaux, comptant 5 reformulations, 4 élucidations, 4 ajouts.

Nous nous intéressons ici plus spécifiquement à la gestualité de Marlene lors du second segment, par ordre chronologique, identifié comme *élucidation*. Au plan verbal, elle établit une relation entre des éléments présentés séparément sur la diapositive, comparant la consommation d'eau pour la production de café aux USA et sa moyenne mondiale. Le tableau 7 détaille ses co-productions verbales et gestuelles, et indique la focalisation attentionnelle concomitante. Les segments verbaux correspondant aux gestes représentationnels sont en gras, et ceux co-produits avec des gestes de pointage sont soulignés.

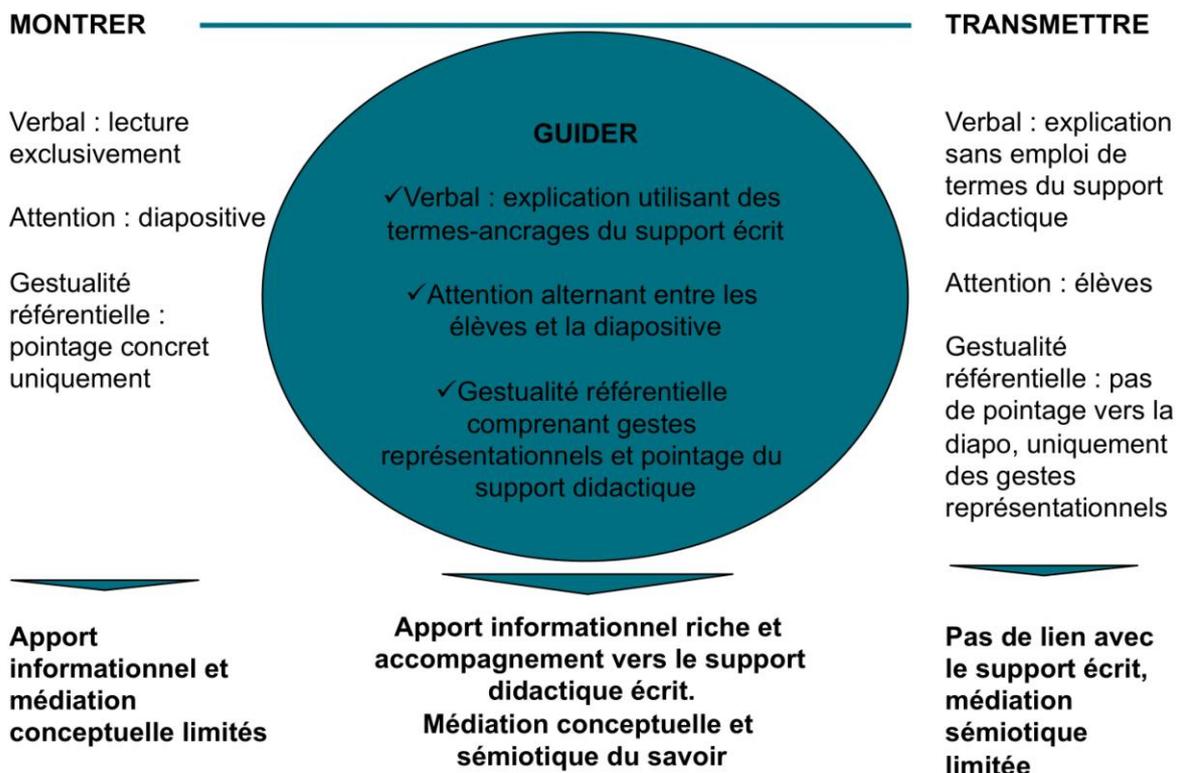
Parole	Attention	Geste	Photo
1. <i>so that</i>		Représentationnel, à deux mains : Pointage abstrait à droite > <b>causalité</b>	
2. <i>might be why</i>	Elèves	Trajectoire des mains évoluant en pointage concret vers la diapositive	
3. <i>ours is</i>		Pointage concret : Index droit pointant vers la diapositive, pour indiquer un point plus précis	
4. <i>lower</i>	Diapositive	Représentationnel, main droite : - vers le bas : <b>plus petit, inférieur</b> - forme de ligne : <b>niveau</b> > <b>Niveau bas, inférieur</b>	

**Tableau 7. Cas 2 – Discours verbal, focalisation attentionnelle et gestes de Marlene.**

En utilisant le pronom « *that* », Marlene fait référence à la reformulation qu'elle vient de réaliser face public sur le fait que la quantité d'eau requise pour la production de café varie selon comment et où il est produit (« *how you produce it/ and also where you're producing it* »), correspondant à l'un des trois éléments de savoir visés pour Q1. Dans ce court extrait, elle établit un lien entre cet énoncé antérieur et des données projetées sur l'écran. Via un premier mouvement des deux mains vers la droite (1), elle représente un lien de causalité avec ce qu'elle vient d'expliquer. Après un arrêt, ses deux mains poursuivent leur trajectoire vers le haut, et le geste se transforme en un pointage concret vers la diapositive (2), alors qu'elle regarde toujours les élèves. C'est alors qu'elle tourne sa tête vers l'écran, et fait simultanément évoluer son pointage en geste plus précis mobilisant l'index seul (3), pour indiquer un chiffre du tableau, vers lequel elle oriente également l'assistance verbalement (« *ours* » - la nôtre – faisant référence à l'eau consommée par unité de café produite aux Etats-Unis). L'ensemble de cette phrase gestuelle permet à Marlene d'aider son assistance à établir un lien entre les informations verbo-gestuelles de son discours et leur représentation sur la diapositive. Une fois qu'elle est face à l'écran, elle continue à travailler à cette médiation entre les modes sémiotiques, en utilisant un geste représentationnel (4) : elle déplace sa main droite, en forme de ligne, vers le bas, alors qu'elle atteint l'acmé de sa phrase verbale : « *lower* » (inférieure). Elle représente ainsi métaphoriquement un niveau plus bas, en même temps qu'elle pointe vers la donnée pertinente du tableau projeté.

## Discussion : modèle multimodal de l'explication et médiation double du savoir.

Dans notre corpus, la plupart des lycéennes-animatrices, pourtant inexpérimentées, ont développé des séquences explicatives riches en exploitant une diversité de ressources multimodales. Elles ont notamment fourni des informations additionnelles pour expliciter les savoirs présentés dans les diapositives via leur communication verbale et gestuelle, et accompagné leur public dans l'appréhension de la représentation écrite de ces concepts. Nous proposons ici une hypothèse théorique à partir de cette étude empirique, sous la forme d'un modèle général de performance multimodale lors d'explications en situation de médiation (cf. figure 7). La pratique la plus fréquente dans notre étude consistait à « guider » l'auditoire, un comportement communicationnel intermédiaire entre deux autres archétypes consistant à « montrer » ou à « transmettre ».



**Figure 7. Modèle multimodal de l'explication et prise en charge de la double nature conceptuelle et sémiotique de la médiation : montrer, guider ou transmettre.**

*Montrer* consiste à fournir un travail de médiation minimal, en considérant que les données projetées « parlent d'elles-mêmes ». Verbalement, la pratique typique associée à ce type de performance est la simple lecture de la diapositive. Au plan corporel, l'attention est alors seulement dirigée vers l'écran, et les gestes sont limités à du pointage concret. Dans notre étude, c'est Iris qui se rapproche le plus de ce pôle. Voyez ci-dessous son discours verbal pour Q1 :

*so: the water footprint of a product varies between countries to countries different due to different product practices and local climate.so as you can see:*

*coffee: was the leader followed by microchips beef em: rice wheat and apples*  
*and the source was: statistics from the water research management article*

Ainsi, il consiste principalement à lire la diapositive, et ne présente aucun *ajout* informationnel. De plus, la seule gestualité représentationnelle qu'elle produit est une énumération, ce qui relève plutôt de la structuration du discours que de la construction référentielle primaire, au sens de l'expression d'une nouvelle information. Ici, ce geste est redondant avec le classement exprimé verbalement.

*Transmettre* constitue le pôle opposé du continuum. Ce type de performance revient à tout expliquer par le seul discours verbo-gestuel face-à-face, comme s'il n'y avait pas de support didactique écrit. La plupart de la production verbale est alors inédite, et la gestualité, uniquement représentationnelle, est abondante. Les concepts sont expliqués avec force détail, mais la part du travail de médiation consistant à établir des ponts entre différents modes sémiotiques, et à accompagner les élèves vers la prise en main de l'écrit est inexistante. Sur le plan des effets didactiques, il y a alors un travail de médiation conceptuelle, mais sans médiation sémiotique. Sans correspondre totalement à cet archétype, Océane, dans notre étude, tend parfois à s'en rapprocher pendant des segments de ses séquences explicatives. Par exemple, pendant Q1, elle est restée face public pendant plus de 15 secondes, sans pointer l'écran, en introduisant verbalement beaucoup de vocabulaire inédit, et en produisant 8 gestes représentationnels.

La plupart des animatrices de notre étude, la plupart du temps, réalisent plutôt un travail de médiation complet consistant à *guider* les élèves vers les concepts apparaissant sur les diapositives, même si certaines mettent plutôt l'accent sur la médiation sémiotique (pont entre les modes de représentation) et d'autres sur la médiation conceptuelle (élaboration verbo-gestuelle des savoirs). Le travail didactique de médiation consistant à *guider* l'appréhension des savoirs peut alors être caractérisé comme une performance communicative présentant 3 traits : 1) l'alternance de focalisation attentionnelle support didactique/élèves ; 2) l'usage de gestes représentationnels comme de gestes de pointage concret ; 3) un discours verbal explicitant le support écrit tout en réutilisant des termes servant d'« ancrages » (*anchors*) à la médiation. Grâce à sa capacité à prendre en charge simultanément les deux volets du travail de médiation, conceptuel et sémiotique, nous faisons l'hypothèse que ce type de performance explicative multimodale est le mieux susceptible d'expliquer les savoirs aux élèves.

## **Conclusion**

Nous avons étudié les explications réalisées par cinq lycéennes-animatrices (3 américaines, 2 françaises) lors d'un café scientifique comme des performances multimodales de médiation du savoir. Plus spécifiquement, nous nous sommes intéressés aux relations entre leurs paroles, leurs gestes, et l'utilisation d'un support didactique écrit. A l'exception d'une animatrice américaine, toutes ont réalisé un véritable travail de médiation verbale des concepts en jeu, via des *reformulations*, des *élucidations* et des *ajouts* informationnels. Cette étude a permis de préciser la nature de l'activité de médiation comme présentant deux dimensions : une facette conceptuelle, consistant à élaborer de nouveaux éléments d'explication des savoirs au plan verbo-gestuel ; et une facette sémiotique, visant à établir des liens entre différents modes de représentation de ces concepts, et notamment entre leur représentation verbo-gestuelle et leur représentation écrite. La première dimension passe principalement par des *élucidations* et des *ajouts* verbaux, et par des gestes représentationnels. La seconde s'appuie sur un schéma d'alternance de focalisation attentionnelle entre les élèves et le support didactique, et l'intégration de gestes de pointage vers ce dernier au discours

verbo-gestuel explicatif. A partir de cette étude empirique, nous proposons de modéliser l'activité de médiation en tant que performance multimodale relevant d'une pratique communicationnelle de *guidage*, définie comme une position intermédiaire entre deux archétypes comportementaux : *montrer* et *transmettre*. Nous formulons l'hypothèse que *guider* les élèves constitue la forme la plus élaborée au plan didactique, car elle prend en charge les deux facettes, conceptuelle et sémiotique, de la médiation. Pour confirmer cette hypothèse, il serait nécessaire de compléter, dans un travail futur, cette analyse des pratiques des animatrices par une étude de leur réception chez les élèves.

Au plan théorique, nos résultats vont dans le sens de la vision des interactions didactiques comme des situations de communication complexes, multimodales, et remplissant une multiplicité de fonctions entrecroisées, en temps réel. Ainsi, un trait relevant habituellement d'une fonction gestuelle interactive, à savoir la focalisation attentionnelle vers les interlocuteurs, se révèle, dans la situation étudiée, étroitement liée à une fonction de construction référentielle assurée par la production de gestes représentationnels. Cette conclusion est cohérente avec des travaux antérieurs, comme l'analyse du discours enseignant par Azaoui (2014) comme caractérisé par une « ubiquité énonciative ». Au plan pratique, notre étude sur la performance multimodale de médiation peut informer la formation des professionnels de l'éducation (enseignants, médiateurs culturels ou scientifiques, animateurs) susceptibles de travailler dans des situations de communication similaires.

## REFERENCES

- Alibali, M. W., Nathan, M. J., Wolfgram, M. S., Church, R. B., Jacobs, S. A., Martinez, C. J., Knuth, E. J. (2014). How Teachers Link Ideas in Mathematics Instruction Using Speech and Gesture: A Corpus Analysis. *Cognition and Instruction*, 32(1), 65-100.
- Alibali, M. W., Young, A. G., Crooks, N. M., Yeo, A., Wolfgram, M. S., Ledesma, I. M., Nathan, M. J., Church, R. B., Knuth, E. J. (2013). Students learn more when their teacher has learned to gesture effectively. *Gesture* 1(2), 210–233.
- Arzarello, F. (2004). Semiosis as a Multimodal Process. *Relime*, 267-299.
- Azaoui, B. (2014). Analyse multimodale de l'agir professoral et degré de granularité de traitement. *Réflexions méthodologiques. Lidil* 49, 17-32.
- Bécu-Robinault, K. & Lund, K. (2012). *Modèles implicites dans les reformulations multimodales des enseignants*. Communication présentée au 7<sup>ème</sup> congrès de l'ARDiST, Bordeaux, France, 14-16 mars, 383-389.
- Bouchard, R. & Traverso, V. (2006). Objets écrits et processus d'inscription entre planification et émergence. Dans M.-C. Guernier & V. Durand-Guerrier (dir.), *Interactions verbales didactiques et apprentissages* (p. 185–219). Presses universitaires de Franche-Comté.
- Brousseau, G. (1998). *Théorie des situations didactiques : didactiques des mathématiques 1970-1990*. Grenoble: La Pensée Sauvage.
- Calbris, G. (2011). *Elements of Meaning in Gesture*. Amsterdam/Philadelphia, John Benjamins Publishing Company.
- Caswell, C., & Neill, S. (1993). *Body language for competent teachers*. Routledge.
- Cienki, A. & Müller, C. (2008). *Metaphor and Gesture*. John Amsterdam/Philadelphia, John Benjamins Publishing Company.
- Colletta, J.-M. (2009). Comparative analysis of children's narratives at different ages, a multimodal approach. *Gesture* 9(1), 61–96.
- Colletta, J.-M., Capirci, O., Cristilli, C., Goldin-Meadow, S., Guidetti, M., Levine, S. (2010). *Multimodal Data Transcription and Annotation with ELAN*. Coding Manual. Disponible sur <http://lidilem.u-grenoble3.fr/IMG/pdf/anrmultimodalityresearch-codingmanual.pdf>.

- Colletta, J.-M. & Pellenq, C. (2009). « The development of multimodal explanations in French children ». Dans M. A. Nippold, & C. M. Scott (dir.), *Expository Discourse in Children, Adolescents, and Adults. Development and Disorders* (p. 63-100). Taylor & Francis.
- Cook, S. W., Duffy, R. G., Fenn, K. M. (2013). Consolidation and transfer of learning after observing hand gesture. *Child Development* 84(6), 1863–1871.
- Cosnier, J. & Vaysse, J. (1997). Sémiotique des gestes communicatifs. *Nouveaux Actes Sémiotiques* 9(52-53-54), 7-28.
- Duval, R. (1995). *Semiosis et Pensée Humaine*. Paris, Peterlang.
- Enfield, N. J., Kita, S., de Ruiter, J. P. (2007). Primary and secondary pragmatic functions of pointing gestures. *Journal of Pragmatics* 39,f 1722–1741.
- Flevaris, L. M., & Perry, M. (2001). How many do you see? The use of nonspoken representations in first-grade mathematics lessons. *Journal of Educational Psychology* 93, 330–345.
- Goldin-Meadow, S. (2005). Gesture's role in the learning process. *Theory into practice* 43 (4), 314-321.
- Goldin-Meadow, S. (2004). *Hearing gesture: How our hands help us think*. Cambridge, MA, Harvard University Press.
- Graziano, M. (2010). Le développement des gestes pragmatiques et leur relation avec le développement de la compétence textuelle chez l'enfant âgé de 4 à 10 ans. *Lidil* 42, 113-138.
- Gumperz, J. J., & Hymes, D. H. (1972). *Directions in sociolinguistics*. Holt, Rinehart and Winston.
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge University Press.
- Kita, S. & Özyürek, A. (2003). What Does Cross-Linguistic Variation in Semantic Coordination of Speech and Gesture Reveal? Evidence for an Interface Representation of Spatial Thinking and Speaking. *Journal of Memory and Language* 48(1), 16-32.
- Lund, K. & Bécu-Robinault, K. (2009). La reformulation multimodale et polysémiotique comme aide à la compréhension de la physique. Dans A. Rabatel (dir.), *Analyse sémiotique et didactique des reformulations* (p. 191-215). Presses Universitaires de Franche-Comté.
- Mazur-Palandre, A., Colletta, J.-M., Lund, K. (2014). Context sensitive 'how' explanation in children's multimodal behavior. *Journal of Multimodal Communication Studies* 2,1-17.
- McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. University of Chicago Press.
- McNeill, D. (2000). *Language and Gesture*. Cambridge University Press.
- Mondada, L. (2005). *Chercheurs en interaction. Comment émergent les savoirs*. Lausanne : Presses Polytechniques et Universitaires Romandes.
- Perry, M., Berch, D. B., Singleton, J. L. (1995). Constructing shared understanding: The role of nonverbal input in learning contexts. *Journal of Contemporary Legal Issues* 6, 213-235.
- Singer, M. A. & Goldin-Meadow, S. (2005). Children learn when their teacher's gestures and speech differ. *Psychological Science* 16(2), 85-89.
- Streeck, J. (2009). *Gesturecraft. The manu-facture of meaning*. Amsterdam/Philadelphia, John Benjamins Publishing Company.
- Tellier, M. (2008). The effect of gestures on second language memorisation by young children. *Gesture* 8(2), 219-235.
- Valenzeno, L., Alibali, M. W. Klatzky, R. (2003). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology* 28(2), 187-204.
- Winkin, Y. (2000). *La nouvelle communication*. Paris : Seuil.

## ANNEXES

### Annexe 1. Objectifs pédagogiques du café scientifique *YouTalk*

1) Créer un climat de confiance, de respect et d'écoute mutuels qui permette à chaque élève de s'exprimer. On cherche pour ce faire à faciliter la formulation orale des idées, l'esprit de synthèse, la pertinence des interventions et la capacité d'écoute.

2) Inciter les élèves à distinguer, dans la perspective de concourir à renforcer leur esprit critique, les éléments de la discussion qui :

- ont trait à la science stabilisée : ceci correspond aux connaissances et au vocabulaire apportés afin d'établir une « base de discussion » ;
- relèvent de la science en cours, et sont, à ce titre, incertains ;
- procèdent d'un choix, d'une orientation politique ;
- établissent un lien de l'un à l'autre.

Cet objectif fait écho à une visée importante du socle commun, dans son pilier 3 : « faire comprendre aux élèves la distinction entre faits et hypothèses vérifiables d'une part, opinions et croyances d'autre part. »

3) S'entraîner à s'approprier un sujet nouveau et complexe en identifiant :

- les nécessités et contraintes à prendre en compte pour traiter la problématique ;
- les principaux points de désaccord dans le débat.

4) Leur permettre de se forger une opinion sur une question, a minima se sentir capable d'en avoir une (identifier les arguments en faveur ou à l'encontre de telle ou telle position et/ou les incertitudes à lever pour prendre position).

5) S'entraîner à travailler en groupe : coopérer, écouter les arguments des autres, intervenir avec pertinence dans la discussion.

6) S'entraîner à prendre la parole en public, et à s'exprimer clairement à l'oral.

Annexe 2. Diapositive correspondant à la question 1\*

**9. Among the following products, which requires the most water, pound for pound, for its production?**

- A Wheat
- B Beef
- C Rice
- D **Coffee**
- E Microchips
- F Apples

**\_\_YOU TALK !\_\_** got water ? 

\* Pour Q1 et Q2 (annexes 1 et 3), les réponses attendues sont soulignées et apparaissent en gras.

Annexe 3. Diapositive « Point information » expliquant la réponse à Q1

**Information Desk** 

Product	Water footprint world average (litres/kg)	Water footprint USA (litres/kg)
1 kg of wheat	1334	849
1 kg of beef	15,497	13,193
1 kg of rice	2291	1275
<b>1 kg of coffee</b>	<b>17,373</b>	4864
1 kg of microchips	16,000	
1 kg of apples	700	

***Pound per pound, worldwide, coffee production uses on average the most water.***

The water footprint of a product varies between countries due to different production practices and local climates.

**SOURCE:** Statistics taken from the scientific journal *Water Resource Management*, "Water footprints of nations: Water use by people as a function of their consumption pattern" by A. Y. Hoekstra and A. K. Chapagain, 2006. The article is available through the [waterfootprint.org](http://www.waterfootprint.org) website at [http://www.waterfootprint.org/Reports/Hoekstra\\_and\\_Chapagain\\_2007.pdf](http://www.waterfootprint.org/Reports/Hoekstra_and_Chapagain_2007.pdf)

**\_\_YOU TALK !\_\_** got water ? 

#### Annexe 4. Diapositive correspondant à la question 1\*

**12. Which of these statements is true for people living in a shantytown in Manila (Philippines) without indoor plumbing?**

- A** They pay 10 times more than someone in Manila who has indoor plumbing.
- B** They pay 4 times more than a middle-class resident of Manhattan
- C** They pay 4 times more than a very rich resident of Manhattan
- D** They pay \$3.00 per cubic meter (in Kenosha, the price is \$0.68/m<sup>3</sup>).
- E** **All of these statements are true.**
- F** None of these statements are true.

**\_\_YOUTALK !\_\_** got water ? 

#### Annexe 5. 'Diapositive « Point information » expliquant la réponse à Q2

**Information Desk** 

*All of these statements are true.*

Relatively little information is available on the household budgets of the poorest people on the planet. Nevertheless, the studies that are available tend to show that :

- people living in shantytowns in the developing world buy their drinking water at the highest price in the world, from tank trucks and from street vendors
- in cities, there can be very sharp differences in prices depending on whether one is connected to local water networks or not
- drinking water in shantytowns costs even more than it does in the world's richest countries, where people have greater access to local water networks.
- In general, once one is connected to local water networks, household users all pay the same price, regardless of whether they are rich or poor.

**SOURCE:** Kevin Watkins (dir.) United Nations Development Programme, « Beyond scarcity: Power, poverty and the global water crisis », *Human Development Report*, 2006

**\_\_YOUTALK !\_\_** got water ? 

#### Annexe 6. Conventions utilisées pour la transcription du discours verbal

Elles sont adaptées des normes proposées par le groupe ICOR :

- : son allongé
- / or \ intonation montante ou descendante
- WORD volume augmenté
- ‘ élision non-standard
- (2.1) pauses mesurées (durée supérieure à 0.5 secondes)
- (...) interruption de l'explication monologuée par une séquence de dialogue