



HAL
open science

The Multimodal Mediation of Knowledge: instructors' explanations in a scientific café

Claire Polo, Jean-Marc Colletta

► **To cite this version:**

Claire Polo, Jean-Marc Colletta. The Multimodal Mediation of Knowledge: instructors' explanations in a scientific café. *Multimodal Communication*, 2019, 8 (2), 10.1515/mc-2018-0009 . hal-03520314

HAL Id: hal-03520314

<https://hal.science/hal-03520314>

Submitted on 21 Feb 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.

The Multimodal Mediation of Knowledge: instructors' explanations in a scientific *café*.

Claire POLO & Jean-Marc COLLETTA

Abstract

In this paper, we present an in depth study of the interplay between various semiotic modes in the specific instructional setting of a scientific *café*. We analyzed the multimodal performance of five female instructors delivering a monologue explanation during instruction on the following dimensions: speech, gesture (hand gestures, head orientation and gaze) and use of written didactical material. Results first point out the crucial role played by referential hand gesture together with gaze-body behavior both in representing new concepts (*conceptual mediation*) and in building bridges between information displayed in several modes (*semiotic mediation*). They also show cross-individual differences in instructors' multimodal performance, that we propose to interpret as three diverse modes of mediating knowledge, *guiding* being the only one providing both conceptual and semiotic mediation.

1. Introduction

Recent literature emphasizes the role of multimodality in teaching, but how do instructors actually use and combine diverse semiotic modes in authentic face-to-face interactions with learners? This exploratory study offers insights into the different ways novice instructors multimodally mediate knowledge when using slides in a one to many lecturing scenario.

The multimodal dimension of teaching was first studied as using complementary external representations of concepts like pictures, diagrams, algebraic formula, and such (Duval, 1995). This perspective was enlarged to encompass oral communicative resources within what Arzarello calls the 'semiotic bundle' (Arzarello, 2004, Flevares & Perry, 2001, Lund, Bécu-Robinault, 2009, Bécu-Robinault, Lund, 2012). Some studies focused on face-to-face resources, notably speech and gesture (Alibali et al., 2014, Goldin-Meadow, 2004, Singer & Goldin-Meadow, 2005, Valenzeno, Alibali, Klatzky, 2003). Works on discursive skills also show that the use of multimodal resources gets more complex as a child grows (e. g. Colletta, 2009, Colletta & Pellenq, 2009, Graziano, 2010), and may vary depending on the interactional setting (Mazur-Palandre, Colletta, Lund, 2014).

A common instructional practice consists in asking some tutor students to explain something to their peers, or to younger students. Nevertheless, how they manage to do so has still not been precisely studied. Our paper explores this issue on the basis of videotaped explanations elaborated by senior students (here after named 'instructors') leading a scientific *café* about drinking water management, for junior students (here after named 'students'), in France and the USA. How do they make use of diverse semiotic resources to mediate the knowledge that they have to explain? As interactional linguists, we investigated this issue with an inductive approach, starting from *data sessions* and then defining our analytical categories according to two methodological principles: their data adequacy (ability to actually describe the observed phenomena) and their interpretative power (ability to make sense of the observed phenomena in the perspective of the participants) (e. g. Mondada, 2005).

2. State of the art: multimodality at the heart of cognitive processes

In opposition to a 'weak' approach of multimodality which considers non-verbal aspects of communication as mere supports to the verbal production, our work is in line with the 'strong' tradition of research on multimodality, in which these elements are recognized as playing a key role in the construction and organization of thought itself (Kita & Ozyurek, 2003, McNeill, 1992, 2000, Kendon, 2004). The term 'multimodality' is based on a metaphor of the communicative message as consisting of a combination of various information slots embedded in a variety of semiotic modes, which, together achieve a global meaning. This definition raises several issues about the specific nature of each mode in terms of information packing and expression, and whether specific multimodal combinations may display affordances to the construction of specific semantic structures (Gibson, 1979, Gerwing & Allison, 2009). The metaphor of the 'multimodal ensemble' also led some researchers to question the relations between the modes, for instance by studying their degree of interdependency, or their degree of redundancy (e. g. Bavelas, Beavin, Chovil, Lawrie, Wade, 1992, Goldin-Meadow, 2003). At the end of the day, the definition of 'mode' is still problematic, since distinction between modes can be done in various ways and at diverse scales. This leads Norris to define a mode as a '*system of mediated action that comes about through concrete lower-level actions that social actors take in the world*' (2013: 155). Moreover, what counts for a 'mode' varies from one research community to another, each one generally focusing only on a few semiotic communicative resources: picture-and-text, speech-graphs-and-inscriptions, speech-and-instrumental actions, etc. (Kress, 2010: 87). Multimodal studies on teaching discourse mostly focused, for instance, on the bimodal combinations of text and other external representations (Duval, 1995) or on speech and gesture (Alibali, et al., 2014, Goldin-Meadow, 2004, Singer & Goldin-Meadow, 2005, Valenzeno, Alibali, Klatzky, 2003). Still, the verbal mode very often seems to serve as a reference in comparison to which other modes' characteristics and functions are defined. Indeed, at least in research on face-to-face interactions, speech seems to be the pivotal mode that cannot be put aside, and, when it is studied together with other semiotic resources, they are called 'nonverbal' or 'coverbal'. Matters of synchronicity of information delivered through multiple modes are also generally addressed through a speech-centered perspective, questioning whether other semiotic communicative actions are aligned with speech. In this paper, our multimodal analysis is based on the study of three modes: speech (what you hear), gesture (what you see) and written data projected on a screen (what you read). The gestural mode is here understood in the extensive meaning of any body movement, *de facto* including three different types of lower-level actions in the present analysis: hand gestures, gaze and head orientation.

Our 'strong' conception of multimodality implies that any of these semiotic elements may serve one of the three key interactional goals of a) coordinating the communication, b) displaying social affiliation, and c) building semantic structures through the building of reference. For instance, interactive and pragmatic gestures are generally used to coordinate the communication and display social affiliation (a and b). Our interest in how the instructors use multimodal resources to explain concepts to younger students led us to mainly focus on the referential dimension of explanations (c). The literature distinguishes primary referential construction, corresponding to "*new information*" (e. g. Chui, 2005) from secondary referential construction, which consists of meta-elements added to previously brought information, to precise or disambiguate its meaning (Chui, 2008, Emmorey & Casey, 2001, Gerwing & Allison, 2009, Holler & Beattie, 2003). For the speech-gesture system, it has been shown that any mode can contribute to both the

primary and the secondary building of reference: gestures can contribute the same meaning as the words, or they can “*refine, qualify, or make more restricted the meaning of the words... and sometimes even provide aspects of reference that are not present at all in the verbal component*” (Kendon, 2004: 161).

Our interest in the building of reference corresponding to the concepts explained in the situation studied led us to do what some authors call ‘a semantic feature analysis’: “*A semantic feature analysis is the systematic evaluation of how gestures and words represent meaning*” (Gerwing & Allison, 2009: 5). As a result, we focused on students’ gestures playing a referential function, i.e. representational and pointing gestures. Representational gestures directly “*exploit imagery*” (Goldin-Meadow, 2004: 314), and provide embodied images of the referent in a diversity of ways. Some representational hand gestures depict a concrete physical referent by drawing in space its imaginary shape and dimension, or by locating parts of it or tracing its trajectory in an imaginary space, while others use a part of the body, or the whole body to mime actions from an animated referent (Colletta, Capirci, Cristilli, Goldin-Meadow, Guidetti & Levine, 2010, Cosnier & Vaysse, 1997, Kendon, 2004, McNeill, 1992, Streeck, 2009). Other representational gestures help express ideas by providing visual metaphors for abstract concepts (e.g. the gripping / giving / trenching /chasing gestures to express respectively founding / sharing / dividing / chasing) (Calbris, 2011, Cienki & Muller, 2008, McNeill, 1992). Representational gestures can be organized sequentially, in a way that gives a specific meaning to the repetition of similar gestures, or to the use of a given section of the gestural space. Pointing gestures are defined as “*a communicative bodily movement which projects a vector whose direction is determined, in the context, by the conceived spatial location, relative to the person performing the gesture, of a place or thing relevant to the current utterance*” (Enfield, Kita, de Ruiter, 2007: 1724). Among pointing gestures, a distinction is needed between concrete pointing, directed to a person or to an object physically present and abstract pointing, directed to an empty section of the gestural space. Concrete pointing appears early in language development (Colletta & Pellenq, 2009) and has mostly a referential function (Enfield, Kita, de Ruiter, 2007). When concrete pointing is directed to a person, it can serve a different, interactive function (Bavelas et al., 1992). In the context of the present paper, all the concrete pointing gestures studied are directed to the screen where written data is projected, thus clearly referring to referential information provided on the slide. Abstract pointing belongs to the category of abstract deictics (McNeill, 1992), and, as all the abstract gestures, it appears later in language development (e. g. Colletta & Pellenq, 2009). Abstract pointing can be used in combination with representational gestures in gestural sequences. Pointing gestures are then directed to sections of the gestural space with specific attributed meaning and serve a referential function.

In instructional settings, multimodality and gesture have been studied with the idea that varying the modes of representation of concepts improves the teaching-and-learning process (Alibali et al., 2014, Goldin-Meadow, 2004, Singer & Goldin-Meadow, 2005, Arzarello, 2004, Flevaris & Perry, 2001, Lund, Bécu-Robinault, 2009, Bécu-Robinault, Lund, 2012, Valenzeno, Alibali, Klatzky, 2003). Among the consensual results, it is well established that as teachers get experienced, or trained, they tend to make a greater use of multimodal combinations, at least in terms of speech-gesture system (e. g. Neill & Caswell, 1993). Empirical studies also brought growing evidence that gestures improve the efficiency of teaching discourse (i. g. Alibali, Young, Crooks, Yeo, Wolfgram, Ledesma, Nathan, Church, Knuth, 2013, Cook, Duffy, Fenn, 2013, Perry, Berch & Singleton, 1995, Singer & Goldin-Meadow, 2005, Tellier, 2008, Valenzeno, Alibali, Klatzky, 2003).

Therefore, educational settings seem to be particularly appropriate places for the development of rich multimodal communication. The explanatory sequences studied in this paper are made by instructors with no experience in teaching and limited content knowledge about the topic. Through an empirical study, we try to identify whether some specific multimodal combinations are associated with deeper explanations in terms of the amount and nature (conceptual deepening or semiotic bridging) of the displayed information. We then propose a model of knowledge mediation as a multimodal performance.

3. Context & Method

3.1 Instructional context and data set

We investigated an extracurricular activity held at school in the USA and France, led by instructors aged 17-18, for students aged 12-15. Based on an international project of environmental education, this scientific *café* aimed at helping the students explore a socio-scientific issue (here drinking water management) and debate about it as future citizens. The senior students were volunteers, and they were trained for 7 hours in order to become 'instructors'. The *cafés* lasted between 90 and 120 minutes and alternated class discussion, group discussion, group voting and individual voting, on the basis on a multiple-choice questionnaire. The whole activity was divided into three topical sections: 1) water resources; 2) human uses of water; 3) water management, distribution and pricing. Some questions, called 'knowledge questions' (KQ), provided students with basic information, while others introduced debates about socio-scientific issues. To choose an answer about a KQ, the students used an electronic voting system. The percentages of votes for each option was then displayed on the screen, and the right answer turned green. In the next step, the instructor explained why this answer was considered true, on the basis of a slide called 'information desk' (consisting of written data and information source). This is the specific task that we studied, on the basis of an audiovisual corpus of the *cafés*, which were fully recorded.

We chose to focus on two KQ. The first one (here after referred to as Q1) deals with how much water is needed to make products of common use, or the 'virtual water' embedded in such products. The second one (here after referred to as Q2) consists of assertions about the price of drinking water in different places of Manila in comparison with its price in developed countries. The slides corresponding to Q1 and Q2, for the American and the French version of the activity, are reproduced in appendices 1-4.

All the activity was videotaped with several cameras, one of them recording a global view of the classroom combined with the sound from the instructor's microphone. This is the track that we used for the analyses presented here. Out of the 6 fully recorded *cafés* (3 in each country), one was left aside for this analysis because the instructor moved out of the camera frame while explaining the selected questions. Therefore, our final data set included the explanations made by 5 people, 3 American and 2 French female instructors (renamed for this research Cathy, Marlene and Iris, and Océane and Sylvie).

In the instructional sequence, the phase of explanation of a KQ is considered as interactive: the instructor must explain the right answer, answer students' questions and doubts, and make sure that they have well understood the involved concepts. Nevertheless, the time dedicated to this dialogue varied a lot. In order to work on comparable data, we decided to limit our analysis to the monologue part of this phase, when the instructor explains the right answer to the audience. Table 1 presents the data

collection. The 10 phases of ‘information desk’ lasted over than 18 minutes, while the monologue explanations totaled half of this.

| | Instructor | Question | Time spent on the ‘information desk’ slide | Time of explanation |
|--------------|---------------------|----------------------|--|---------------------|
| USA | Cathy | 1 | 1 min 26 | 50.1 sec |
| | | 2 | 2 min 16 | 1 min 25 |
| | Marlene | 1 | 1 min 46 | 45.6 sec |
| | | 2 | 2 min 40.4 | 1 min 34.4 |
| | Iris | 1 | 26 sec | 26 sec |
| | | 2 | 1 min 02 | 48.7 sec |
| FRANCE | Sylvie | 1 | 2 min 53.4 | 44.8 sec |
| | | 2 | 1 min 48 | 48.3 sec |
| | Océane | 1 | 1 min 30 | 1 min 01.6 |
| | | 2 | 2 min 24 | 1 min 30.8 |
| Total | 10 questions | 18 min 12 sec | 9 min 55.2 sec | |

Table 1. Overview of the video data.

3.2 Methodological path

The whole explanatory task can be considered as supplementing the ‘information desk’ slide in order to help the students understand the target knowledge. We defined such activity as mediation. Three research questions guided our work:

- 1) How do instructors make use of the available multimodal resources to mediate the target knowledge?
- 2) What are the effects of the multimodal combinations that they use regarding the information provided to the students?
- 3) Can we determine a general model of the different multimodal practices used to mediate knowledge?

We studied instructors’ multimodal mediation practice regarding three types of semiotic resources: speech, gestures, and the data written on the slide. The boundaries of what constitutes a mode in terms of channels providing different information pieces of the global message can be discussed. Even in speech, the information contained in words can be completed or disambiguated thanks to elements of prosody. The ‘gesture’ label may also refer to a large range of phenomena, from the study of gaze to kinesics and proxemics, including the restrictive use of the term ‘gestures’ corresponding to hand movements in an individual gestural space. The third type of semiotic objects that we study here, data to be visualized on the screen, also presents a diversity of representational forms, including both text and tables. For Q1, the written data source consisted both in text and a table (cf. appendix 2), while it was only text for Q2, (cf. appendix 4).

The overall explanation performance of the instructors was studied through three analytical steps, focusing on 1) verbal explanation of the slide content, 2) gestural behavior, and 3) identification of relevant cross-individual differences among the empirically observed multimodal combinations. First, we focused on the speech dimension, and fully transcribed the instructors’ speech (for transcription conventions cf. appendix 5). We then identified the parts of the transcripts that corresponded to a literal re-use of the words of the ‘information desk’ slide. Examples of such transcripts are provided for the case studies detailed in section 4. Most of the instructors’ speech was not pure repetitions but actually embedded innovative content. These innovative

elements were classified in three categories: *rewording*, *elucidation* and *addition*. By *rewording*, the instructor does not really provide new information, but she rephrases the information of the slide with her own words. Other parts of the transcripts correspond to *elucidation*: the instructor then relates pieces of information that are presented separately in the slide, or produces inferences based on the content of the slides. When the instructors verbally provided totally new information that was not in the slide, it was classified as *addition*. *Additions* mostly consisted in giving examples, providing reasons or anticipating implications and consequences of the ideas presented on the slide. Table 2 summarizes this classification of these verbal elements and provides an authentic example for each category. In our data, we found a total of 32 instances of *rewording*, 23 of *elucidation*, and 36 of *addition*.

| Type of mediating verbal element | Rewording | Elucidation | Addition |
|---|---|--|--|
| Definition | No new information. Rephrasing the information presented on the slide in the instructor's own words. | Relating two pieces of information from the slide. OR (Interpretative) inferences based on the slide informational content. | Providing new pieces of information that cannot be inferred from the slide. An addition may consist of an example, justification, implication, etc. of a fact mentioned on the slide. |
| Examples from explanations about Q2 (appendix 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> |
| Initial text provided visually (slide reproduced in appendix 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> |

Table 2. Typology of mediating elements in instructors' speech.

Our second analytical step focused on gestures. What we coded as gestures included the 3 typical phases of preparation, stroke and retraction (Kendon, 1980), sometimes with a holding phase. When a gesture is hold and then serves as a basis for another gesture, then the new one is counted as another gesture, even if the first one has not finished yet. Such gestural analysis was conducted using the ELAN software. First, the attention focus of each leader all along his/her explanation was coded, on the basis of head and gaze orientation. This led us to observe an attention focus pattern distributed between the screen and the audience (this result is detailed in 4.1). For each time slot corresponding to a mediating verbal element (91 in total), we identified co-occurring representational and pointing gestures. For representational gestures and abstract pointing, a referent was systematically associated to each gesture. A table visualization of each excerpt was made in order to put gestural information in perspective with the attention focus of the leader and her co-occurring speech. Examples of this visualization are given in section 4. This analytical step made it possible to identify information gesturally added by the instructors in their explanations.

Our last analytical phase consisted in stepping away from this accumulation of qualitative results, trying to identify tendencies among the explanatory practices analyzed. We empirically characterized instructional explanatory practice as

combinations of two co-verbal gestural micro-activities: representational gesturing associated to focusing attention to the audience and concrete pointing associated to focusing attention to the screen. As a combination of multimodal resources, we could then define the instructors' instructional activity as enriching the building of reference by both adding new information (*conceptual mediation*) and establishing bridges between different semiotic representations of the involved concepts (*semiotic mediation*).

4. Results

Our first general result was to observe a clear pattern of attention focus alternation between the screen and the audience (4.1). Through a first case study we explore what typically occurs when the leaders are facing the students (4.2), making use of representational gestures to complement their speech. We then turn to another case study (4.3), corresponding to a transitional phase in the focus alternation pattern, in which both speech, gestures and head orientation serve as bridges to mediate concepts between diverse semiotic modes. Finally, we propose a multimodal model characterizing the explanatory practice of instructors as providing the students with more or less conceptual and/or semiotic mediation of knowledge (4.4).

4.1 Paying attention to both the students and the data to explain: attention focus alternation

A first striking result is the observation that all the instructors alternate phases of attention focus directed to the screen and phases of attention focus directed to the students, the attention focus being defined according to the head and gaze orientation. Figures 1, 2, 3, 4 and 5, based on screen captures of ELAN time lines show this pattern respectively for each instructor. Moments spent on the 'information desk' slide which do not correspond to monologue explanation were not coded and appear in black. The moments of attention focus directed to the students are highlighted in grey, the rest of time corresponding to an attention focused on the screen.

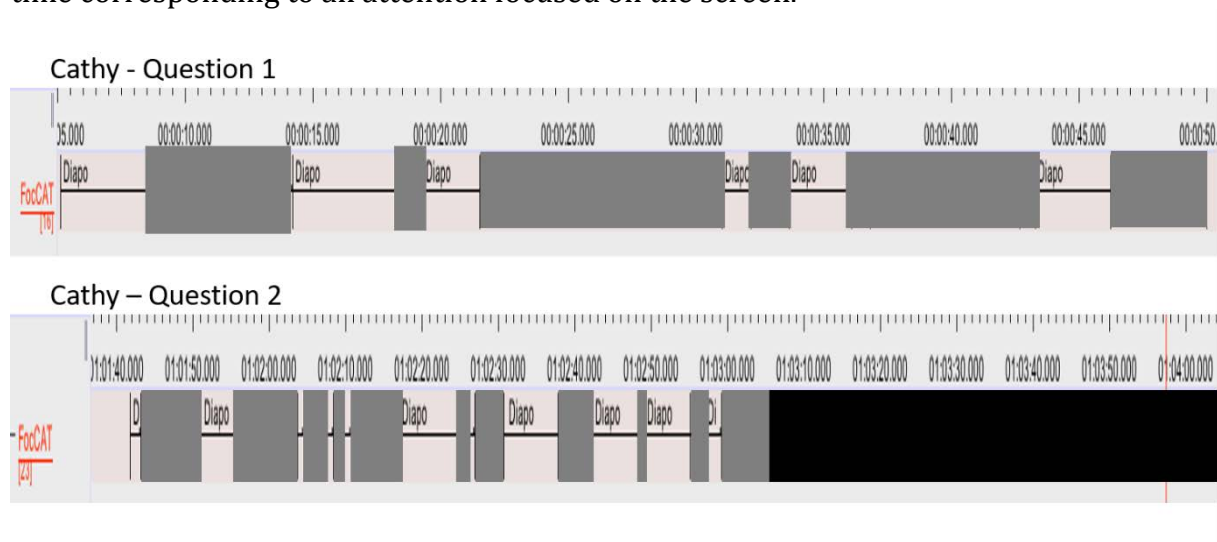
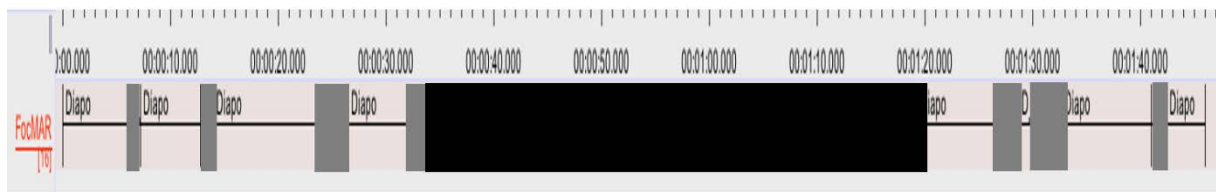


Figure 1. Attention focus during monologue explanation: Cathy.

Marlene – Question 1



Marlene – Question 2

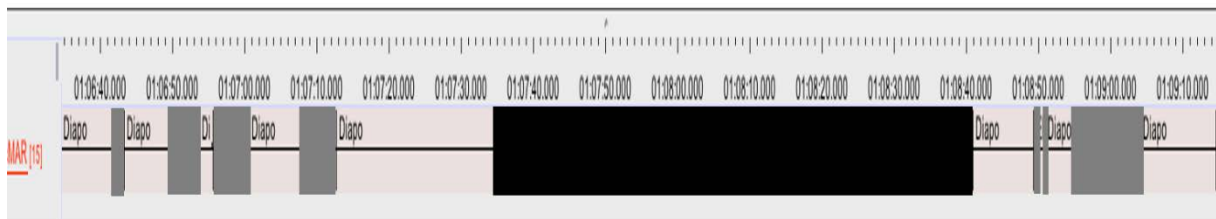
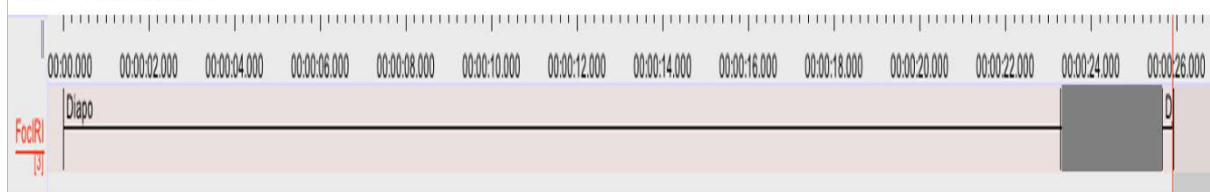


Figure 2. Attention focus during monologue explanation: Marlene.

Iris – Question 1



Iris – Question 2

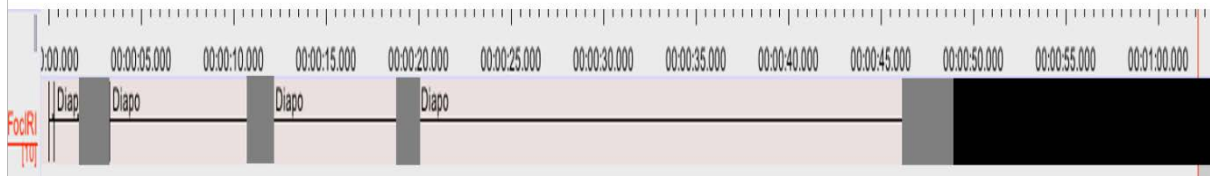


Figure 3. Attention focus during monologue explanation: Iris.

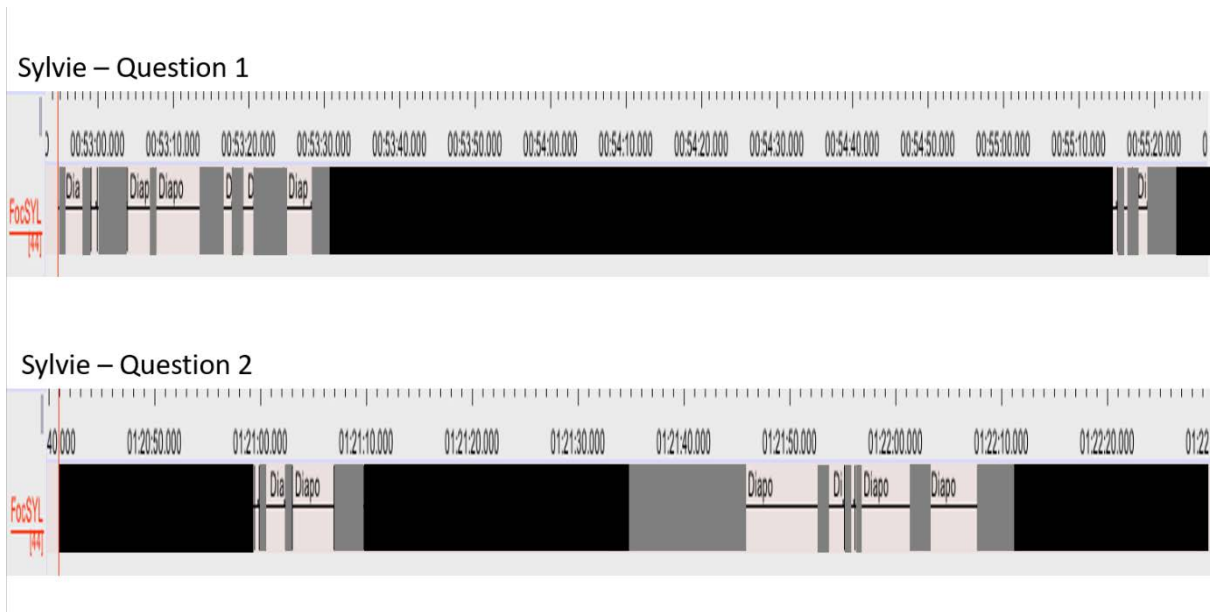


Figure 4. Attention focus during monologue explanation: Sylvie.

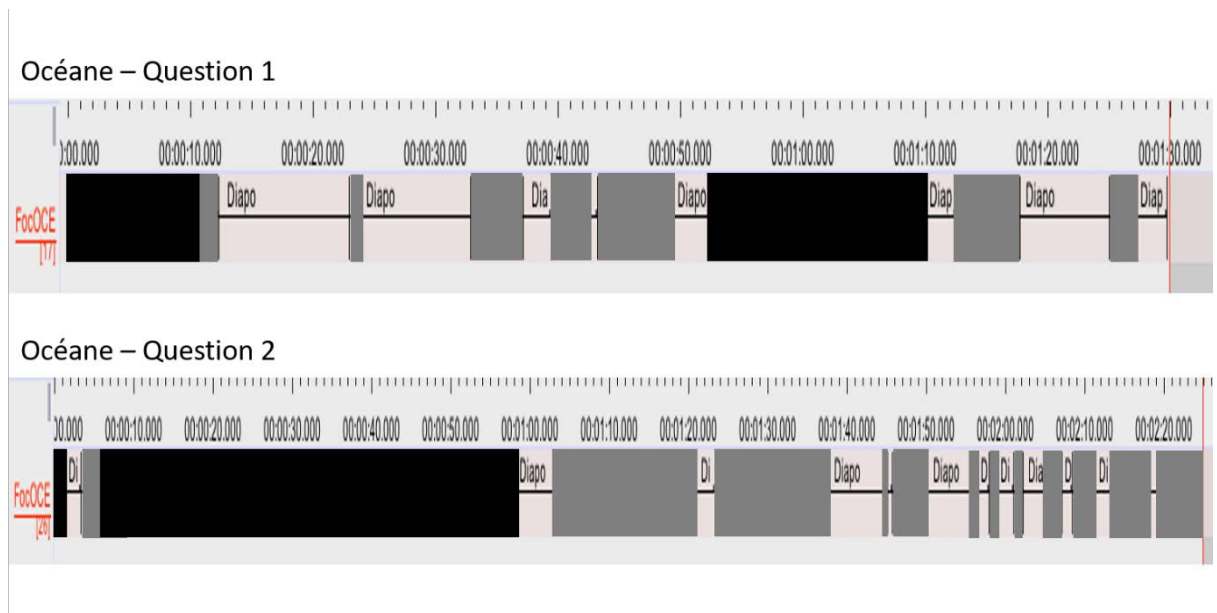


Figure 5. Attention focus during monologue explanation: Océane.

Still, although the alternating pattern was found for the five instructors, the time dedicated to each single episode directed to one or the other varies a lot. Both occurring during Q2, the longest student-focused episode's duration is 18s. and shows in Océane's instruction, and the shortest student-focused episode's duration is 0.5 s. and shows in Sylvie's instruction. While only two other episodes last more than 10 s., there are 20 episodes lasting 1 s. or less. No typical duration can be established for a single episode of attention focus. Nevertheless, it is worth mentioning the general time proportion of the explanation directed to the screen vs focused on the students. This information is provided in table 3. Table 3 also presents the time proportion of each explanation corresponding to representational and pointing gestures. The instructors' behavior is especially rich in referential gestures, since, summing up pointing and representational gestures, the instructors (except Iris, who gestures very little) spent from 42% to 77% of their time providing referential information gesturally.

| | Leader | Question | Time of monological explanation | Student-focused time | Proportion of time with repr. gestures | Number of repr. gestures (rate/10 sec) | Proportion of time with pointing gestures | Number of pointing gestures (rate/10 sec) |
|-----|--------|----------|---------------------------------|----------------------|--|--|---|---|
| USA | CAT | 1 | 50.1 sec | 55,5 % | 27,0 % | 12 (2.4) | 38,9 % | 19 (3.8) |
| | | 2 | 1 min 25 | 53,5 % | 30,1 % | 24 (2.8) | 29,7 % | 25 (2.9) |
| | MAR | 1 | 45.6 sec | 23,0 % | 27,9 % | 20 (4.4) | 14,9 % | 9 (2.0) |
| | | 2 | 1 min 34.4 | 27,6 % | 45,8 % | 59 () | 9,8 % | 11 (1.2) |
| | IRI | 1 | 26 sec | 9,1 % | 28,4 % | 6 (2.3) | 0,0 % | 0 (0.0) |
| | | 2 | 48.7 sec | 14,3 % | 12,9 % | 5 (1.0) | 1,8 % | 1 (0.2) |
| FR. | SYL | 1 | 44.8 sec | 48,6 % | 25,3 % | 17 (3.8) | 22,3 % | 6 (1.3) |
| | | 2 | 48.3 sec | 44,8 % | 19,6 % | 20 (4.1) | 57,6 % | 30 (6.2) |
| | OCE | 1 | 1 min 01.6 | 39,0 % | 22,4 % | 26 (4.2) | 42,3 % | 19 (3.1) |
| | | 2 | 1 min 30.8 | 64,1 % | 43,8 % | 42 (4.6) | 30,6 % | 19 (2.1) |

Table 3. Time proportions of student-focused attention, representational gesturing and concrete pointing in explanations and associated gesture rates.

In table 3, clear cross-individual variation appears. Iris spent relatively little time facing the students (less than 10% for Q1 and less than 14% for Q2). She is also the one producing the least gestures, holding the record of the smallest proportion of pointing time (0%, for Q1), pointing gesture rate (0.2 for Q2), and the smallest proportion of time with representational gestures (12,9%, corresponding to the smallest rate of representational gestures too, 1 per 10 s., for Q2). Her poor gesture production goes along with a predominant attention focus toward the screen.

Marlene spent one quarter to one third of her explanation time facing the students (23% for Q1 27,6% for Q2). Two instructors used this student focus about half the time, a little bit more for Cathy (54,5% in average), and a little bit less for Sylvie (46,7% in average). Océane showed greater variation, spending about 40% of her explanation time facing the students during Q1, and more than 64% during Q2. It is interesting to note that the proportion of time when she is producing representational gestures in Q2 is also about twice as large as it is for Q1 (respectively 43,8% and 22,4%), even though her representational gestures rate is only a little bigger (4.6 vs 4.2 gestures per 10 s.). As she also does a lot of pointing gestures at the screen during Q1 (spending more than 43% of the time pointing, with more than 3 gestures per 10 s), we can hypothesize that her short representational gestures are meant to help the students do semiotic transfers (from verbo-gestural concepts to written ideas), while the longer representational gestures produced during Q2 might rather foster conceptual deepening.

We also found a tendency to associate attention focusing on the screen and concrete pointing on the one hand, and focusing on the audience and representational gesturing, on the other hand. This is very clear for the explanations made by Océane and Cathy. Figure 6 presents the corresponding ELAN screen captures.

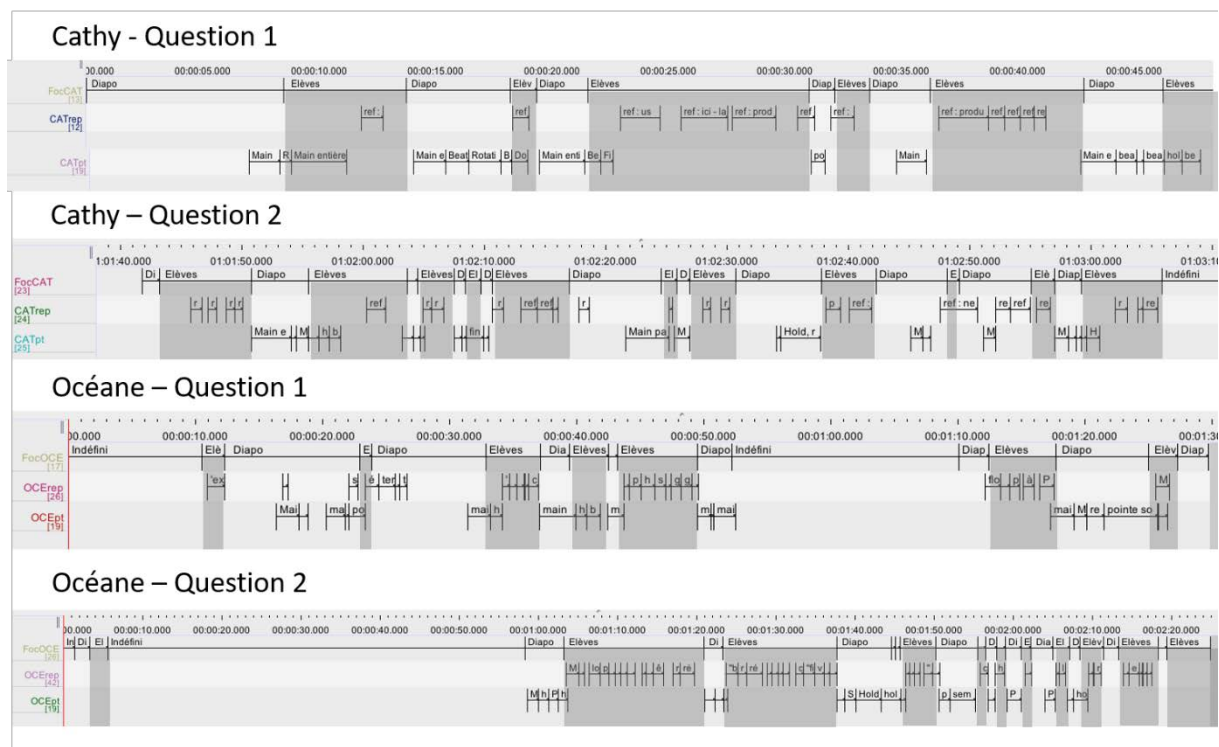


Figure 6. Correlation between the attention focus and the type of gesture – Cathy and Océane (ELAN screen captures).

The first line indicates the attention focus, marked in grey when directed to the students. In the second line were coded the representational gestures, and, in the third, the concrete pointing gestures. Most of the instructors' representational gestures occurred when they were facing the students (grey zone), while most of their concrete pointing is produced facing the screen. Out of a total of 44 pointing gestures, Cathy only made 14 of them facing the students. Océane only made 6 of her 38 pointing gestures facing the students. The correlation between attention to the students and representational gestures looks even stronger than the one between attention to the screen and concrete pointing gestures. Cathy made 36 representational gestures in total, and only produced 3 of them facing the screen. Océane made a total of 68 representational gestures, and only 5 of them were produced facing the screen, 3 of which were enumeration gestures. What does this multimodal combination of speech and representational gesturing facing the audience bring to this explanatory task? The next subsection (4.2) presents a case study that is typical of the use of such resources by the instructors for their explanations.

4.2 Didactical use of representational gestures while facing the students

Here we focus on Cathy's explanation about Q1. Below is transcribed her corresponding speech (as for all the verbal transcripts reproduced in this paper, the parts of her speech corresponding to exact words or lexical expressions appearing on the slide are underlined):

so: (0.9) the: w- w- of course the water footprint how much is used pound for pound varies between countries\ 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\ 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/ but worldwide is what you have to keep in mind em: (0.5) is that all those coffee beans being produced and grown in other countries takes an immense amount of water\ em are they any other option that you guys: strongly thought of that er maybe you don't know how necessarily the water is used and wanna ask about it/

Even if Cathy reuses several terms from the slide, she is not reading and she also brings new vocabulary and information verbally. Seven parts of her speech were classified as innovative verbal elements, 2 of them consisting in rewording, 2 serving an elucidation function, and 3 identified as additions. These elements are presented in table 4.

| Rewording | Elucidation | Addition |
|---|---|---|
| <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 woud see it here is lower/</i> |
| <i>em are they any other option (...) 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> |

Table 4. Characterization of innovative elements in Cathy’s explanatory speech about Q1.

For the present case study, we turn to a micro-level analysis and focus on one segment of her speech: the first *addition* (framed in table 4). Even if Cathy then elaborates innovative information verbally, she also produces, meanwhile, 6 gestures, using them to bring meaning that is not embedded in her speech. Her first five gestures in this excerpt are representational. They are produced facing the student, and while saying what was transcribed from “*because*” to “*here*”. Her last gesture consists in pointing to the screen with her head, facing it, and occurs together with the final part of the transcript “*is lower*”. Table 5 provides a description of each of her representational gesture, together with the accompanying speech. The parts of speech uttered during the production of the gesture appear in bold, and so does the referent identified for each gesture or part of a gesture.

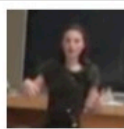
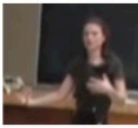
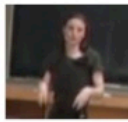
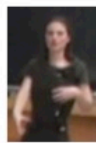
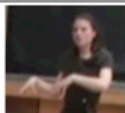
| Speech | Representational gesture | Picture |
|---|---|---|
| 1 <i>Because most of our: er: coffee beans</i> | Two hands : - parallel in conduit metaphor > quite a big thing/quantity - palms to herself > hers |  |
| 2 <i>and everything</i> | Keeping the parallel hands of the conduit metaphor > thing /quantity + abstract pointing on the right > another, different |  |
| 3 <i>are not grown in the us\ so then</i> | Hands pointing down, with beats on 'not', 'grown' and 'us' > ground, land, here |  |
| 4 <i>the production em: as far as we would see it</i> | Polysign mainly using the right hand - claw shape: to take - moving up: to grow - cyclic trajectory: process > to cultivate, to harvest |  |
| 5 <i>here</i> | Two hands pointing down > this land, the usa + on the right > getting closer to the slide |  |

Table 5. Case study 1: Cathy's speech and representational gestures.

Cathy first uses what McNeill (1992: 15) calls a 'content' metaphor to represent the coffee beans as quite a big thing or quantity. This first gesture is a polysign (1), in Calbris' sense: "*one gesture may (...) represent several notions simultaneously because more than one of its components has an analogical link and, thus, it contains more than one gestural sign. In this case it is a polysign gesture*" (Calbris, 2011: 28). Here, the shape of the hands specifies this thing with the palms directed to Cathy's body, indicating ownership, as her simultaneous use of the word 'our'. Then, keeping the parallelism between her hands, she produces an abstract pointing gesture to her right (2), providing an interpretation of her vague speech component 'and everything' corresponding to another thing of the same size as, or in the same quantity as, or as important as the 'coffee beans'. Here, her gesture mainly disambiguates the verbal referential construction. Her third gesture (3) consists in pointing down with her two hands, a general structure that she holds all along the utterance 'are not grown in the us', with emphasizing beats on 'not', 'grown' and 'us'. Doing so, she refers to the ground, specifying an aspect of her wording 'grown' as 'coming from the ground', but she also refers metaphorically to the land, this land, elaborating on the verbal referent 'the us' with an information that does not appear in her speech, meaning that here, we are in the US. The gesture that she makes as she is pronouncing the word 'production' (4) is emblematic of a polysign. The first dimension of this gesture, the shape of her right hand, is similar to a claw orientated to the ground, involving the idea of taking something. The second dimension, the direction of her hand, is globally moving up depicting the image of something that is growing from the ground. Third dimension: more specifically, her hand follows a cyclic trajectory, involving the other hand, in a commonplace metaphor for "a process". Through this three-dimensional gesture, Cathy specifies what she means by 'production', as cultivation or harvest. Her last representational gesture (5) consists in pointing down again while she utters 'here', referring to this land again. In the context of her previous speech and gesture, this one

can also be interpreted as referring to the US. Actually, a very interesting aspect of Cathy's representational gestures in this excerpt is that they function sequentially: the second one takes the meaning of 'another' because of a move on her side; the last pointing down takes the meaning of 'US' because a similar gesture was done earlier with a beat on 'the us'. Last but not least, this last representational gesture is produced on the right of Cathy's gestural space, as she is about to point at the screen, which is on her right. She does this concrete pointing saying 'is lower'. Somehow, this move to the right is a way to include the latest gesture in her global gestural sequence. She is literally 'bringing' the referent (US production) to the screen as a way to make sure that the audience understands what 'lower' refers to, and looks at the correct place in the slide (comparing it to the world average use of water for coffee production, cf. appendix 2).

4.3 Mediating concepts by establishing bridges between a variety of semiotic modes

This last aspect of our first case study is similar to what some instructors do more systematically in their explanations while transiting from one attention focus to another. The global pattern of focus alternation itself works as a way to establish bridges between the semiotic modes used on the slides (text, table) and the semiotic modes used in the oral explanations (speech and gesture). Nevertheless, in some occasions, the instructors more explicitly mediate this transfer from one mode to another. In this subsection, we provide an emblematic case of such *semiotic mediation*, based on an analysis of an excerpt of Marlene's explanation about Q1:

coffee: (0.6) requires: hum seventeen thousands/ (1.1) liters per kilogram\ (0.7) hum: (1.2) this is worldwide though\ this is no:t just the usa\ as you can see ours is: (0.5) significantly less\ hum: (0.8) it varies: depending on: (0.6) how you produce it/ and also where you're producing it\ so that might be why ours is lower\ (0.6) also: are you are you guys surprised that your second runner up was actually microchips\ (...) and also cattle\ you have to (0.8) cattle is your third one up\ you have to (1.5) the cattle has to drink\ it has to eat\ the plants\ you have to grow the plants with the water\ (0.6) so\ em: and this comes from: the: scientific journal of water resource management\ water footprint of nations\ (1.0) so that was just: some knowledge for you guys to use\ (2.6) 'kay

Verbally, Marlene reuses some vocabulary written on the slide, but she is not just reading it, except when she mentions the title of the scientific source used. Her explanation is actually quite rich in terms of innovative verbal elements: 5 parts of this transcript were identified as *rewording*, 4 as *elucidation*, 4 as *addition*. Such coding is presented in table 6.

| Rewording | Elucidation | Addition |
|---|---|--|
| <i>coffee: (0.6) requires: hum seventeen thousands/ (1.1) liters per kilogram\ (0.7)</i> | <i>ours is: (0.5) significantly less\</i> | <i>this is no:t just the usa\</i> |
| <i>hum: (1.2) this is worldwide though\</i> | <i>so that might be why ours is lower\</i> | <i>as you can see</i> |
| <i>hum: (0.8) it varies: depending on: (0.6) how you produce it/ and also where you're producing it\</i> | <i>(0.6) also: are you are you guys surprised that your second runner up was actually microchips\</i> | <i>you have to (0.8) (...) you have to (1.5) the cattle has to drink\ it has to eat\ the plants\ you have to grow the plants with the water\ (0.6) so\</i> |
| <i>and also cattle\</i> | <i>cattle is your third one up\</i> | <i>so that was just: some knowledge for you guys to use\ (2.6) 'kay\</i> |
| <i>em: and this comes from: the: scientific journal of water resource management\ water footprint of nations\</i> | | |

Table 6. Characterization of innovative elements in Marlene's explanatory speech about Q1.

We are now going to analyze Marlene's gesturing during the second excerpt chronologically identified as playing an elucidation function, framed in table 6. Verbally, she is establishing a relation between informational elements that are presented separately on the slide, comparing water consumption for coffee production in the US and worldwide. In table 7 is reproduced her speech with a description of her accompanying gestures and attention focus. The parts of speech uttered during the production of a representational gesture appear in bold, and so does the referent identified for each gesture or component of a gesture. The parts of her speech corresponding to a concrete pointing gesture are underlined.



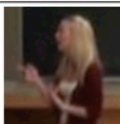

| Speech | Focus | Gesture | Picture |
|----------------------------|----------|--|---|
| 1 so that | | Representational, two-hand: abstract pointing to the right> causality |  |
| 2 <u>might be why</u> | Students | Hands' trajectory transformed into a concrete pointing at the slide |  |
| 3 <u>ours</u> <u>is</u> | | Concrete pointing: right hand index finger pointing at the slide, as referring to a precise point on the screen |  |
| 4 lower\ | Screen | Representational, right hand: - moving down: smaller - flat line shape: level > low level |  |

Table 7. Case study 2: Marlene's speech, attention focus and gestures.

With the pronoun *'that'*, Marlene is referring to what she has just said about the fact that the amount of water needed for coffee production varies depending on *'how you produce it/ and also where you're producing it'*, a *rewording* that she produced facing the students. In this short excerpt, Marlene is now establishing a causal link between this previous statement and some facts presented in the table projected on the screen. With a first two-hand move to the right (1), she signifies an implication of what she has previously explained. After this first stop, the two hands complete their trajectory up, transforming into a concrete pointing at the slide (2), while Marlene is still facing the students. She then turns her head and attention to the screen, and transforms her hands pointing into an index finger pointing (3), as referring to a precise part of the slide, actually verbally indicated (*'ours'*, stating for the water footprint of coffee produced in the USA, the last square of the second column of the table appearing on the slide). During the entire gestural sequence, Marlene is guiding the students to help them establish a link between the information given in her speech and the information presented on the slide. After clearly specifying the causal nature of this link, she gradually draws their attention to the relevant information in the table. When facing the slide, she keeps on working on this transition between semiotic modes. She uses another representational gesture addressed to the students, as if she was *'translating'* to them the information content of both the slide and her speech. She then moves her right hand down, in a flat line shape (4), referring to a lower level, as she is getting to the key point of her verbal sentence: *'lower'*.

4.4 Toward a multimodal description of knowledge mediation

Most of the instructors demonstrated their ability to develop rich explanatory discourse by taking advantage of a diversity of multimodal resources. Alternating their head and attention between the screen and the students, they mediated the concepts by providing innovative information in their speech and gesture. Such *conceptual mediation* was associated to a *semiotic mediation* of knowledge: by pointing at the screen, they established bridges between the oral and the written semiotic modes, and guided the students in their understanding of the slide. In this last subsection, we step away from micro-analysis and propose a general model to characterize the instructional behavior of instructors, as a combination of multimodal resources. We describe the most frequent explanatory practice as *'guiding'*. Guiding can be understood as a continuum of intermediate behaviors from the two other stereotypical practices of *'showing'* and *'transmitting'*. Figure 4 represents these three explanatory practices and their characteristics.

Figure 4. Model of the multimodal explanatory performance of instructors: showing, guiding and transmitting.

Showing consists of acting as if *'everything was on the slide'*, and as if there were no need for mediating the target knowledge through innovative information, neither verbally nor gesturally. It is associated with a verbal behavior limited to reading the slide, an attention focus only directed to the screen, and the absence of representational gestures. As this is a stereotype, no instructor in our data exactly behaved this way. Nevertheless, we have an example of an American leading student, Iris, whose instruction tended to be a showing one. For instance, see below the transcription of her explanation of 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 \

Most of her verbal production consisted in reading, and no innovative information was provided. Moreover, during all this explanation, her only representational gesturing activity was an enumeration, which did not bring much new information compared to her speech.

Transmitting stands at the other end of the spectrum of explanatory multimodal performance. It corresponds to behaving as if one could explain everything from scratch, only through oral discourse (including speech and gesture), without referring to the didactical written resources. Then, almost no word appearing on the slide is reused and much innovative information is also brought verbally and gesturally. The attention is only directed to the students, with no concrete pointing at the screen. In terms of instructional effects, the knowledge is extensively *conceptually mediated*, but no link is established with the written semiotic modes with which the students need to become familiar: there is no *semiotic mediation*. Of course, no instructor totally fell into this category either. But Océane, one of the French instructors, sometimes tended to be *transmitting*, in parts of her explanations. For instance, during more than 15 seconds and a half, while explaining Q1, she faced the students, without pointing at the screen. She then provided a lot of new vocabulary in her speech, only reusing 5 words from the text, even if she was rewording the content of the slide. Meanwhile, she also produced 8 representational gestures, specifying a lot the few key referents that she was considering.

Finally, most of instructors' authentic explanatory practices consisted in *guiding*. *Guiding* is characterized by alternating attention focus between the screen and the students, using both concrete pointing at the slide and representational gestures, together with a speech including lexical anchors from the written data. As an intermediate behavior between the two stereotypes of *showing* and *transmitting*, *guiding* may be embedded in a variety of communicative styles that all provide, to some extent two elements: 1) the necessary additional information to provide a *conceptual mediation* and 2) the bridging work to mediate their cognitive circulation between semiotic modes, and especially their understanding of the written didactical data, or *semiotic mediation*.

5. Conclusion

Novice instructors proved capable of mediating knowledge using a diversity of semiotic resources. Studying the relation between their speech, gesture, and use of the data written on the slide, we could specify different styles of multimodal explanatory performance. Except in one case, the instructors produced a speech that informed the information written on the slide, either through *rewording* it, *elucidating* it, or *adding* elements. Their attention focus defined by the gaze and head orientation, revealed a clear pattern of alternation between the students and the screen, for all the instructors. Through representational gestures and abstract pointing, the instructors highlighted parts of their speech, but also conveyed complementary information disambiguating their verbal expression, or specifying a referent. As in case study 1, they tended to perform such gestures facing the students, and developing complex gestural sequences. Concrete pointing always referred to the slide, either globally or to a precise part of it, and, on the contrary, tended to be produced facing the screen. Nevertheless, as in case study 2, some pointing gestures were produced facing the students, and some

representational gestures facing the screen, especially in transition phases just before or after a shift of attention focus. Such behavior establishes bridges between different representations of the concepts and may help the students understand the information written on the slide.

We finally propose a model of this instructional activity that includes three explanatory multimodal styles. On one end of the continuum, *showing* consists in minimal gestural and verbal activity. At the other end stands the *transmitting* style, characterized by a high level of *conceptual mediation* without *semiotic mediation*. Most of authentic studied instructional practices correspond to the third, intermediate style, *guiding*. *Guiding* is a richer instructional discourse in the sense that it provides both *conceptual* and *semiotic mediation* of knowledge, and we can hypothesize that it would turn out to be the most efficient multimodal explanatory performance. Nevertheless, further work is needed to specify this hypothesis, and notably the diverse forms of *guiding*. Future research should also analyze students' reception of the three styles of explanatory performance and their corresponding learning outcomes to test this hypothesis.

At a theoretical level, our results strengthen the vision of didactical interactions as complex communicative situations involving real-time management of multiple goals. While studying the building of reference in instructors' explanation, we revealed the importance of facing the students, an attention focus related to the production of rich gestural referential constructions. A typical 'interactive' feature (looking at the audience) thus appeared as strongly related to the 'cognitive' dimension of the building of reference. This is consistent with previous work, for instance Azaoui's analysis of language teaching as involving the skill of addressing several interlocutors and conversations simultaneously or 'enunciative ubiquity' (Azaoui, 2014). Our findings can inform the training of teachers and communicators likely to deal with similar settings to make them aware of their multimodal explanatory style and its implications for conceptual and semiotic mediation, so that they can adjust their behavior to their goals. Several studies held in instructional contexts already suggest that people can, in some conditions, learn to adapt their gestural activity (e. g. Alibali et al., 2013, Goldin-Meadow, 2004, Tellier & Cadet, 2014).

REFERENCES

- Alibali, M. W., Nathan, M. J., Wolfgram, M. S., Breckinridge Church, R., 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), pp. 65-100.
- Alibali, M. W., Young, A. G. Crooks, N. M., Yeo, A., Wolfgram, M. S., Ledesma, I. M., Nathan, M. J., Breckinridge Church, R., Knuth, E. J. (2013). Students learn more when their teacher has learned to gesture effectively. *Gesture* 1(2), pp. 210-233.
- Arzarello, F. (2004). Semiosis as a Multimodal Process. *Relime*, pp. 267-299.
- Azaoui, B. (2014). Analyse multimodale de l'agir professoral et degré de granularité de traitement. *Réflexions méthodologiques. Lidil*, 49, pp. 17-32.
- Bavelas, J. B., Chovil, N. Lawrie D. A., Wade, A. (1992). Interactive gestures. *Discourse Processes* 15 (4), pp. 469-489.
- Bécu-Robinault, K. & Lund, K. (2012). Modèles implicites dans les reformulations multimodales des enseignants. In: *Proceedings of the 7th ARDiST Conference*. Bordeaux, France, pp.383-389.
- 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.
- Chui, K. (2005). Topicality and Gesture in Chinese Conversational Discourse. *Language and Linguistics*, 6(4), pp. 635-654.
- Chui, K. (2008). Complementary Gestures and Information Types. *Language and Linguistics*, 9(1), pp. 1-22.
- 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), pp. 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*, ANR "Multimodality", ANR-05-BLANC-0178-01 & -02, Lidilem, University of Grenoble Alpes. <http://lidilem.u-grenoble3.fr/IMG/pdf/anrmultimodalityresearch-codingmanual.pdf>
- Colletta, J.-M. & Pellenq, C. (2009). The development of multimodal explanations in French children. In: M. A. Nippold & C. M. Scott, ed., *Expository Discourse in Children, Adolescents, and Adults. Development and Disorders*, Taylor & Francis, pp. 63-100.
- Cook, S., Wagner, R., Duffy, G., Fenn, K. M. (2013). Consolidation and transfer of learning after observing hand gesture. *Child Development*, 84(6), pp. 1863-1871.
- Cosnier, J. & J. Vaysse (1997). Sémiotique des gestes communicatifs. *Nouveaux Actes Sémiotiques*, 9, 52-53-54, pp.7-28.
- Duval, R. (1995) *Semiosis et Pensée Humaine*. Paris: Peterlang.
- Emmorey, K. & S. Casey (2001). Gesture, thought and spatial language. *Gesture* 1(1), pp. 35-50.
- Enfield, N. J., Kita, S., de Ruiter, J. P. (2007). Primary and secondary pragmatic functions of pointing gestures. *Journal of Pragmatics*, 39, pp. 1722-1741.
- Flevares, 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, pp. 330-345.
- Gerwing J. & Allison, M. (2009). The relationship between verbal and gestural contributions in conversation: a comparison of three methods. *Gesture* 9(3), pp. 312-336.
- Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Houghton Mifflin.
- Goldin-Meadow, S. (2004). Gesture's role in the learning process. *Theory into practice*, 43(4), pp. 314-321.
- Goldin-Meadow, S. (2003). *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, pp. 113-138.
- Holler, J. & Beattie G. (2003). How iconic gestures and speech interact in the representation of meaning: are both aspects really integral to the process? *Semiotica*, 146, pp. 81-116.
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge University Press.
- Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In: M. Ritchie Key, ed., *The relationship of verbal and non-verbal communication*, Mouton Publishers, pp. 207-227.
- 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), pp. 16-32.

- Kress, Gunther R. (2010). *Multimodality: A Social Semiotic Approach to Contemporary Communication*. Taylor & Francis.
- Lund, K., Bécu-Robinault, K. (2009). La reformulation multimodale et polysémiotique comme aide à la compréhension de la physique. In: A. Rabatel, ed., *Analyse sémiotique et didactique des reformulations*, Presses Universitaires de Franche-Comté, pp. 191-215.
- Mazur-Palandre, A., Colletta, J.-M., Lund, K. (2014). Context sensitive 'how' explanation in children's multimodal behavior, *Journal of Multimodal Communication Studies*, 2, pp. 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.
- Norris, S. (2013). What is a mode? Smell, olfactory perception, and the notion of mode in multimodal mediated theory. *Multimodal Communication*, 2(2), 155-169.
- 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, pp. 213-235.
- Singer, M. A. & S. Goldin-Meadow (2005). Children learn when their teacher's gestures and speech differ. *Psychological Science*, 16(2), pp. 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), pp. 219-235.
- Tellier, M. & Cadet, L. (2014). *Le corps et la voix de l'enseignant: théorie et pratique*. Paris: Editions Maison des langues.
- Valenzeno, L., Alibali, M. W., Klatzky, R. (2003). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology*, 28(2), pp. 187-204.

APPENDICES


Appendix 1. Slide corresponding to Q1*.

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 ? 

Appendix 2. 'Information desk' slide explaining the answer to 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 |
| 1 kg of coffee | 17,373 | 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/Reports/Hoekstra_and_Chapagain_2007.pdf) website at http://www.waterfootprint.org/Reports/Hoekstra_and_Chapagain_2007.pdf

__YOU TALK !__ got water ? 

Appendix 3. Slide corresponding to Q2*.

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/m³).
- E **All of these statements are true.**
- F None of these statements are true.

__YOU TALK !__ got water ? 

* For the slides of Q1 and Q2 (appendices 1 and 3), the expected correct answers appear in bold and are underlined.

Appendix 4. 'Information desk' slide explaining the answer to 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 ?

Appendix 5. Conventions used for verbal transcripts.

The conventions are adapted from the ICOR group norms, with a few specificities:

| | |
|--------|--|
| : | elongated sound |
| / or \ | rising or falling intonation |
| WORD | augmented volume |
| ‘ | non-standard elision |
| (2.1) | measured pauses (more than 0.5 seconds), s. |
| (...) | dialogue interruption during the explanation |