



**HAL**  
open science

# Des enseignants multimodaux en devenir : l'articulation de la recherche et de la formation

Marion Tellier

► **To cite this version:**

Marion Tellier. Des enseignants multimodaux en devenir : l'articulation de la recherche et de la formation. Cahiers de l'ASDIFLE, 2021, Multimodalité et multisupports pour l'enseignement-apprentissage des langues étrangères, 31, pp.59-78. hal-03516210

**HAL Id: hal-03516210**

**<https://hal.science/hal-03516210v1>**

Submitted on 19 Jan 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.

CECI EST LE MANUSCRIT AUTEUR DE LA PUBLICATION SUIVANTE :

TELLIER, M. (2021). Des enseignants multimodaux en devenir : l'articulation de la recherche et de la formation. *Cahiers de l'ASDIFLE*, 31,59-78.

## **Des enseignants multimodaux en devenir : l'articulation de la recherche et de la formation**

Marion TELLIER  
Aix Marseille Univ, CNRS, LPL, Aix-en-Provence, France

### **Résumé**

Le concept de multimodalité est aujourd'hui très présent dans de nombreux champs scientifiques et avec des acceptions parfois assez diverses. Dans une première partie, je présente deux visions de la multimodalité fréquemment utilisées en didactique des langues qui sont relativement complémentaires mais qui trouvent leur ancrage dans des cadrages théoriques différents. Ces deux perspectives théoriques sont illustrées par des analyses qualitatives de corpus écologiques de cours de FLE. Dans une seconde partie, j'interroge la formation des enseignants et la prise en compte de la multimodalité dans leur agir professoral. Comment la compétence multimodale se développe-t-elle ? Ce développement peut-il être accompagné en formation ? Peut-on apprendre à optimiser l'usage des ressources multimodales en formation ? J'apporte des premiers éléments de réponses en présentant deux projets de recherche en cours basés sur des modules de formation : le projet *Corps en Formation* et le projet *Vapvisio*.

*Mots clés : multimodalité / geste pédagogique / compétence techno-sémio-pédagogique / formation / visioconférence*

---

### **Introduction**

Le concept de multimodalité est aujourd'hui très présent dans de nombreux champs scientifiques et avec des acceptions parfois assez diverses. Je réduirai ici volontairement la définition de ce concept à deux visions fréquemment utilisées en didactique des langues qui sont relativement complémentaires mais qui trouvent leur ancrage dans des cadrages théoriques différents. Je commencerai donc, dans une première partie, par présenter le concept de multimodalité à partir de différents ancrages théoriques en faisant état de mon propre cheminement de pensée et de recherche à travers cette notion. Ce parcours définitoire sera donc forcément très subjectif. J'illustrerai ces définitions par des exemples mettant en avant la multimodalité de l'agir professoral (Cicurel, 2011) enseignant. Mon itinéraire de recherche étant plutôt axé sur la gestuelle des enseignants de langue, les exemples donnés reflèteront ce positionnement. Dans une seconde partie, j'interrogerai la formation des enseignants. Comment la compétence multimodale se développe-t-elle ? Ce développement peut-il être accompagné en formation ? Peut-on apprendre à optimiser l'usage des ressources multimodales en formation ? J'apporterai des premiers éléments de réponses en présentant deux projets de recherche en cours basés sur des modules de formation : le projet *Corps en Formation* et le projet *Vapvisio*.

## MULTIMODALITÉ : PARCOURS DÉFINITOIRE

Quand on parle de multimodalité et de didactique des langues, deux grandes définitions co-existent. Elles sont principalement ancrées dans deux cadrages théoriques que je propose de synthétiser et d'illustrer ici.

### Multimodalité de la parole

La première acception du terme « multimodalité » est liée aux *Gesture studies*, le domaine de recherche qui s'intéresse à l'étude de la gestuelle co-verbale. Les chercheurs dans ce domaine considèrent que le geste et la parole appartiennent à un seul et même système (Kendon, 2004 ; McNeill, 1992) et fonctionnent donc ensemble pour co-construire le sens (Azaoui & Tellier, 2020). Le verbal n'est donc pas dominant mais constitue une des modalités de la parole. En cela, la définition de Colletta est d'une grande clarté :

*« La parole est multimodale. [...] [Elle] est un comportement qui associe des sons articulés à des mouvements corporels, et lorsque vous percevez une conduite langagière, vous percevez celle-ci dans les différentes modalités sensorielles dont vous êtes équipés. Autrement dit, non seulement vous reconnaissez les mots prononcés, mais vous les entendez avec leur musique (ce qu'on appelle la prosodie), et vous percevez en même temps les gestes, les mimiques, les regards... l'ensemble des mouvements corporels produits par le locuteur. » (2004, p.15).*

Dans cette définition, ancrée dans la linguistique, la multimodalité est donc la combinaison des différentes modalités qui composent la parole à savoir, le verbal, la prosodie et les aspects posturo-mimo-gestuels. Rapportés à l'étude de l'enseignement/apprentissage des langues, deux angles d'étude sont possibles : l'analyse du discours multimodal de l'enseignant (Azaoui, 2014 ; Holt, 2020 ; Matsumoto & Dobs, 2017 ; Sato, 2020 ; Tellier, 2014, 2016), à titre d'exemples) et l'analyse du discours multimodal de l'apprenant (Gullberg, 2006 ; Kang, 2020 ; Sharifzadeh & Sarré, 2020 ; Stam, 2015). Cet article se focalisera sur la multimodalité du discours de l'enseignant, qui est tout particulièrement à l'œuvre dans les interactions didactiques et les pratiques de transmission (Azaoui, 2019).

Prenons l'exemple suivant de cette enseignante de FLE qui est face à un groupe d'un niveau A2 et commence son cours en faisant un rappel du contenu de la séance précédente que certains ont manquée.

Elle dit :

*« on a travaillé sur l'invitation comment proposer à quelqu'un une sortie et comment accepter l'invitation et comment refuser l'invitation hum »*

Cette courte transcription ne nous donne à voir que la modalité verbale du discours et si on s'arrête à cet aspect, on ne peut pas saisir toute la richesse multimodale de la parole enseignante.

Si on s'intéresse à présent à un aspect de la modalité prosodique, la pause<sup>1</sup>, nous pouvons ajouter dans la transcription entre parenthèses, les différentes pauses marquées par l'enseignante dans son discours (la durée est ici en millisecondes) :

*« on a travaillé (410) sur l'invitation (760) comment proposer à quelqu'un (210) une sortie (1120) et comment accepter l'invitation (380) et comment refuser l'invitation (565) hum »*

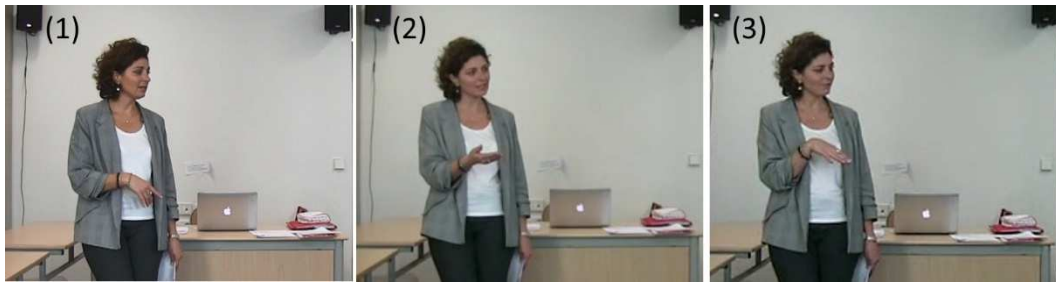
On constate alors que son discours est découpé en différentes unités inter-pausales basées sur la syntaxe de la phrase. Ainsi, elle isole certains compléments et certains groupes verbaux pour faciliter

---

<sup>1</sup> Bien sûr, limiter la prosodie au seul phénomène des pauses est très réducteur. J'ai fait ce choix pour simplifier l'exemplification.

le découpage de la chaîne parlée par les apprenants. Ces pauses sont caractéristiques du *teacher talk* (parler simplifié de l'enseignant) et facilitent l'accès au sens (Chaudron, 1982 ; Owen, 1996).

Ajoutons à présent la gestuelle (Figure 1)



« [on a travaillé] [sur l'invitation (1)] [comment] [proposer à quelqu'un (2)] [une sortie] [et comment accepter l'invitation 3] [et comment refuser l'invitation] hum »<sup>2</sup>  
 Figure 1 : Quelques gestes illustrant le discours pédagogique (Corpus CD, 2018)

Toutes les modalités s'articulent pour co-construire le discours pédagogique, les gestes de pointage dans différents espaces renforcent la segmentation des constituants de la phrase déjà marquée par les pauses. La Figure 2 est une capture d'écran faite à partir du logiciel *Elan* (Wittenburg *et al.*, 2006), un logiciel d'annotation multimodale. Les trois pistes en bas de l'image se lisent comme une partition de musique : de gauche à droite pour comprendre le déroulé temporel et verticalement pour saisir la co-occurrence des modalités. La première piste est une transcription du verbal de l'enseignante, la deuxième montre la présence des gestes et leurs différentes dimensions selon McNeill (1992). Ainsi, dans cet extrait, même si le verbal est ponctué de pauses, les gestes, eux, sont constants (les étiquettes indiquant les gestes se touchent ce qui signifie qu'il n'y a pas de moment sans geste). En regardant l'enregistrement vidéo de plus près cependant, on constate que les gestes marquent aussi des pauses (identifiables dans la transcription par les annotations bleues « Hold »). Les *hold* sont des tenues des gestes, qui sont comme suspendus dans leur mouvement, avant la rétraction finale (McNeill, 1992). Ce très court extrait montre la structure du discours de l'enseignant dans sa dimension multimodale.

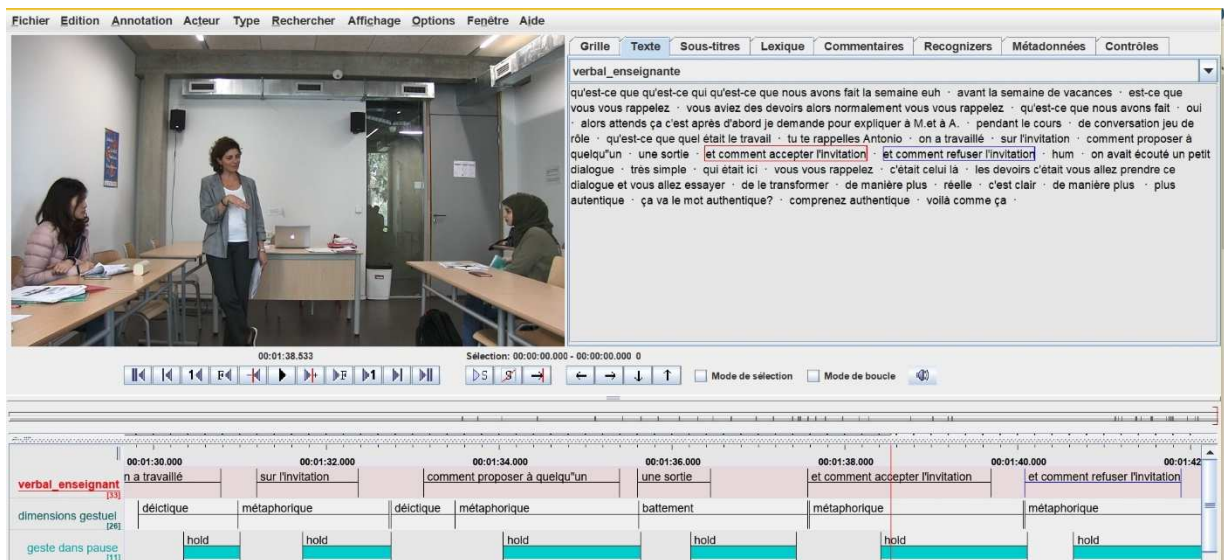


Figure 2 : Orchestration multimodale du discours enseignant (Corpus CD, 2018)

<sup>2</sup> Les parties entre crochets montrent les éléments verbaux accompagnés de gestes, chaque nouvelle paire de crochets indique la présence d'un nouveau geste. Afin de ne pas alourdir la présentation, seuls trois gestes sont présentés sur les photos ici (voir les numéros correspondants).

## Multimodalité initiée par les outils et les dispositifs

L'émergence du numérique dans les pratiques pédagogiques a fait apparaître de nouveaux supports et une redéfinition de la notion de multimodalité en didactique des langues. Les nouveaux outils du numérique (ordinateur, tablette, smartphone) facilitent à la fois l'utilisation en classe de ressources audio-visuelles diverses mais également la constitution, par les apprenants, de productions audio-visuelles en langue cible. De même, l'enseignement à distance, qu'il soit synchrone ou asynchrone, s'appuie sur des plateformes pédagogiques ou des outils de visioconférence combinant vidéo, son, image et texte. Les chercheurs qui s'intéressent à la multimodalité liée au numérique s'inscrivent le plus souvent dans le cadre théorique de la sémiotique sociale (à l'instar de Cappellini & Combe, 2017 ; Guichon, 2013 ; Hampel & Stickler, 2015 ; Holt, 2020 ; Trévisiol-Okamura & Marquilló-Larruy, 2015 parmi d'autres). Sagnier (2018) définit la multimodalité comme impliquant le recours massif et conjoint à divers modes d'expression (verbal, visuel, audio, tactile, etc.) dans la communication. Selon Jewitt (2009) une interaction langagière en ligne, *via* visioconférence, est multimodale dans la mesure où elle permet de combiner de l'image (et donc des gestes et des mimiques faciales), du son et du clavardage. Dans l'ouvrage de Develotte, Kern et Lamy « Décrire la conversation en ligne » (2011), les auteurs distinguent la conversation en ligne *faiblement multimodale* (en s'appuyant sur la notion de densité multimodale établie par Norris, 2004) limitée à l'écrit avec émoticône, au choix de la casse ou de la couleur de la conversation en ligne *fortement multimodale* : avec du graphisme, de l'audio et de la vidéo.

Lorsque l'on enseigne avec le numérique et notamment à distance, on doit donc jongler avec les modalités à disposition et les potentialités offertes par les outils, ce que l'on appelle les affordances. Prenons la même enseignante que dans l'exemple précédent. Comme une grande partie des enseignants de langues, elle a dû, lors de la crise sanitaire liée au Covid 19, s'initier à l'enseignement en ligne, notamment *via* la visioconférence. Le passage à distance nécessite de se construire une présence corporelle à l'écran et de questionner le rôle et la place du corps et des sens dans ce contexte (Guichon & Tellier, 2017 ; Cappellini & Combe, 2017). En ce qui concerne le corps de l'enseignante, la mise en scène de celui-ci requiert une certaine adaptation au dispositif de la visioconférence. La visibilité globale de l'enseignant est essentielle en ligne pour rassurer l'apprenant sur l'attention que l'enseignant lui porte et pour renforcer la dimension psychoaffective (Guichon, 2017). Le cadrage de l'image de l'enseignant est également fondamental pour rendre visible sa gestuelle pédagogique (Holt *et al.* 2015). La Figure 3 montre des images d'une même enseignante à 15 jours d'écart. Lors de sa première séance en ligne (image a), elle était novice dans l'enseignement à distance (alors que très expérimentée en présentiel), on voit qu'elle adopte un cadrage resserré qui rend invisibles ses gestes manuels. Cependant, après 15 jours d'enseignement en ligne, on voit (image b) qu'elle a élargi son cadrage ce qui lui permet de rendre visible sa gestuelle (ici elle explique « attraper » et « contact »). Elle développe progressivement ses compétences techno-sémio-pédagogiques (Guichon, 2012 ; Cappellini & Combe, 2017) et apprend à positionner son corps à l'écran afin d'en optimiser l'usage pédagogique. Les compétences techno-sémio-pédagogiques regroupent le volet technique (savoir utiliser les outils de la communication numérique), le volet sémiotique (savoir choisir les modes appropriés ou la bonne interface en fonction de l'activité pédagogique) et le volet pédagogique (savoir gérer les activités d'apprentissage, respecter sa planification ou l'adapter, réguler les échanges et évaluer).

Dans les images c et d, elle montre son manuel et indique précisément un exercice à faire, tout en contrôlant la visibilité de son corps et de son geste de pointage. En effet, l'orientation de son regard vers le bas de l'écran révèle qu'elle est en train de contrôler son cadrage à l'aide du retour caméra de son image.



Figure 3 : Construire une présence corporelle en ligne (Corpus Cours confinés, 2020)

Enfin, pour renforcer l'attention portée à la parole des apprenants, elle fait preuve d'une « *écoute manifeste* » (Guichon, 2017, p. 40) par la production de nombreux hochements de tête. Elle a également recours à une posture d'écoute exagérée (corps penché vers l'écran, mains sur ses écouteurs), ce que Guichon appelle « *tendre une oreille métaphorique* » (2017, p. 41) (Figure 4).



Figure 4 : Posture d'écoute renforcée (Corpus Cours confinés, 2020)

Mais la multimodalité dans l'usage du numérique ne se résume pas à savoir cadrer son corps à l'écran, cela concerne aussi le recours aux autres affordances (potentialités de l'outils) telles que le partage de documents, le clavardage, le recours à d'autres supports et applications.

Azaoui (2019) évoque même d'un « *multi-agenda augmenté* » en se basant sur la notion de multi-agenda de Bucheton & Soulé (2009) pour évoquer l'intégration des outils numériques aux pratiques pédagogiques et montrer qu'elles peuvent parfois alourdir les actions de l'enseignant. L'usage optimisé des ressources multimodales ne va pas de soi et peut nécessiter une formation pour les enseignants.

### **FORMER DES ENSEIGNANTS MULTIMODAUX**

Que l'on soit en présentiel ou en ligne, l'usage de la multimodalité n'est pas une évidence chez les enseignants de langues novices (voir Tellier & Yerian, 2018 et Tellier, à paraître b). Les projets de recherche portant sur la formation des enseignants et l'effet de ces formations sur leur pratique multimodale sont encore peu nombreux. Je présenterai ici deux projets de recherche : le premier portant sur l'étude de la multimodalité de parole (Corps en formation) et le deuxième concernant l'enseignement en ligne et l'usage des ressources multimodales propres au numérique (VAPVISIO). Ce sont des projets en cours dont ne seront présentés ici que des résultats qualitatifs à titre d'exemple.

#### **Projet « Corps en Formation »**

Nous ne sommes pas tous égaux dans la maîtrise du corps et de la voix. Certes, certains jeunes

enseignants sont spontanément à l'aise face à une classe et savent attirer l'attention sur eux et accepter les regards des apprenants sur leur corps, mais tous les professeurs ne sont pas prêts à jouer leur rôle d'enseignant et à maîtriser les outils de transmission que sont le corps et la voix. Enseigner est un métier. Cela implique des savoirs disciplinaires et des savoir-faire pédagogiques. Le contrôle de son corps et de sa voix permet la mise en œuvre de ces savoir-faire et n'a rien de naturel pour reprendre les propos de Pujade-Renaud : « *L'utilisation de l'espace et d'une force de dissuasion, la maîtrise du regard et de la voix, le travail de la tension et de la mise en scène ne sont peut-être pas aussi 'naturels' qu'on l'affirme parfois.* » (1983/2005, p. 53).

La question principale du projet de recherche *Corps en Formation* était : « Peut-on, avec un module de formation, permettre l'optimisation de son corps et de sa voix en tant qu'outils pédagogiques ? » Pour cette étude, un module pour la formation initiale d'enseignants de français langue étrangère (FLE) a été conçu. Il était composé de la manière suivante : 3h théoriques sur la gestuelle et la voix (rapport à la parole, aspects pédagogiques, aspects culturels) et 6h de formation pratique : 2h d'entraînement à l'explication lexicale et à la délivrance de consignes (de manière libre sans guidage particulier), 1h de travail sur le corps et l'expression corporelle, 1h de travail sur la voix et la posture, 1h d'autoconfrontation collective (à partir d'extraits vidéo issus des 2 premières séances d'entraînement) et enfin, 1h de travail sur l'explication lexicale et la délivrance de consignes (cette fois en intégrant les aspects abordés dans la formation). Le dispositif s'est étalé sur 6 semaines. Il a été intégralement filmé par deux caméras offrant des points de vue différents, à l'automne 2017. Le dispositif est présenté en détails dans Tellier (à paraître a).

Les grands principes établis en formation pour l'utilisation du corps étaient les suivants :

- « - **critère de visibilité** : faire en sorte que la gestuelle et le visage soient visibles par les apprenants en évitant de parler face au tableau et en se positionnant au centre de la salle et, si possible, debout. Éviter les gestes parasites, comme les gestes nerveux (grattements, tics...), qui peuvent gêner la communication. Libérer les mains pour faciliter la production gestuelle ;
- **critère d'iconicité** : lorsque la compréhension des apprenants est difficile, illustrer le contenu du discours avec son corps en ayant recours à des mimes, des gestes iconiques mais aussi des postures et des mimiques faciales ;
- **critère de pertinence** : éviter la surcharge en produisant trop de gestes ou des gestes moins ciblés. Il ne faut pas chercher à illustrer tous les mots, mais plutôt à marquer les mots-clés afin de faciliter une compréhension globale chez l'apprenant ;
- **critère de bienveillance** : adopter une attitude bienveillante en réduisant l'espace proxémique entre les apprenants et l'enseignant, travailler le regard pour bien intégrer tous les participants dans l'échange et utiliser des signaux de feedback positifs comme le sourire et les hochements de tête pour encourager les apprenants ;
- **critère de constance** : utiliser certains gestes de façon routinisée pour créer un code gestuel commun dans la classe. » (Tellier & Yerian, 2018)

Quant à la voix, l'accent a été mis à la fois sur l'aspect préventif pour éviter le forçage vocal et les complications médicales sur les cordes vocales, ainsi que les postures et techniques pour faciliter la projection de la voix. Les aspects prosodiques pour jouer sur le rythme, l'intensité et les pauses dans le discours pédagogique sont également pris en compte (Guimbretière, 2014).

Lors des deux premières séances du module consacrées à l'explication lexicale et à l'émission de consignes, les futurs enseignants n'avaient reçu aucun conseil particulier et se sont lancés sans filet dans les différentes activités proposées (par exemple, tirer au sort un mot à expliquer devant ses camarades en faisant comme si l'on s'adressait à des apprenants de FLE, être interrogé par ses camarades sur la différence entre deux mots en français, donner des consignes d'activité, etc.).

Si certains étudiants étaient à l'aise avec leur gestuelle, beaucoup présentaient des difficultés et semblaient embarrassés avec leur corps. En analysant ces deux premières séances, 3 difficultés

majeures liées au corps ont été mises au jour : 1) la posture (certains tournaient le dos à une partie de la classe ou se tenaient debout dans une posture instable) ; 2) le regard (certains ne regardaient pas toute la classe, voire regardaient dans le vague ou leur document, le cas échéant) et 3) l'absence de recours à une gestuelle facilitante (certains produisaient des gestes nerveux, d'autres des gestes imprécis, voire pas de geste) (ces résultats sont présentés en détail dans Tellier & Yerian 2018).

À titre d'exemple, je vais me concentrer sur ce dernier point : la gestuelle illustrative.

Alice est étudiante en master 1, c'est une jeune femme très timide. En début de module (c'est-à-dire avant la formation), elle doit expliquer ce qu'est une *girafe* comme si elle se trouvait devant des apprenants de FLE. Les captures d'écran ci-dessous (Figure 5) montrent les gestes qu'elle produit en disant « un très long cou ». On voit qu'elle fait un geste de la main gauche qui n'est pas iconique, c'est-à-dire qu'il n'illustre pas des propriétés visuelles du référent. Son geste est ce que l'on dénomme un métaphorique (McNeill, 1992). Il représente abstraitement le mot qu'il accompagne, comme si elle tenait le concept dans sa main. Dans une situation exolingue, en classe de FLE par exemple, ce geste n'aiderait pas un apprenant à comprendre son explication verbale car il n'est pas assez illustratif (Sime, 2008 ; Sueyoshi & Hardison, 2005 ; Tellier, 2008).



Figure 5 : « *c'est un animal qui vit en Afrique et qui a un très long cou* » (Corpus Corps en Formation, 2017)

Lors de l'autoconfrontation collective (Figure 6), un montage vidéo est présenté aux étudiants, présentant chaque participant en train d'expliquer un mot. Ils doivent dire ce qui est positif dans leur façon d'expliquer et ce qui est à améliorer. Les deux formatrices complètent leur analyse. Lorsque vient le tour d'Alice, elle dit de manière presque inaudible : « *ben moi c'est une catastrophe c'est pas articulé c'est pas du tout fort et la posture est très statique pas de geste voilà* ». Lorsqu'on le lui demande, elle dit qu'elle ne voit rien de positif. L'emploi de termes très négatifs comme « catastrophe » et de tournures impersonnelles (elle ne dit jamais « je ») et une description par la négative (« c'est pas » ou « pas » utilisé 3 fois) montrent son malaise face à sa performance. Ses camarades et ses formatrices s'emploient alors à valoriser ce qui peut l'être et ce sur quoi elle devrait travailler.



Figure 6 : Autoconfrontation collective (Corpus Corps en Formation, 2017)

Ainsi, à l'issue du module de formation, elle doit à nouveau expliquer girafe (Figure 7). Son



explication verbale est exactement la même mais on constate que sa gestuelle a changé<sup>3</sup>. En effet, elle produit à présent un geste iconique pour illustrer « très long cou ». Elle positionne ses deux mains, l'une au niveau de la poitrine et l'autre au niveau du cou puis elle les écarte, la première descendant vers le ventre et la seconde montant au-dessus de la tête, pour montrer un cou qui s'étire. Dans cette deuxième explication, non seulement Alice a eu recours à un geste plus illustratif mais également à un geste plus grand. En effet, dans la première explication, son espace gestuel était réduit à la partie centre, c'est-à-dire au niveau de son ventre alors que dans la deuxième explication, elle explore la périphérie de son espace gestuel (McNeill, 1992). Elle remplit donc le critère d'iconicité et de visibilité énoncés ci-dessus, mais également de pertinence puisque c'est bien le long cou qui est prototypique de la girafe et qu'il convient d'illustrer.



Figure 7 : « une girafe c'est un animal qui vit en Afrique et qui a un très long cou » (Corpus Corps en Formation, 2017)

Les données issues de ce module sont encore en cours d'analyse, cependant on constate déjà l'intérêt de faire conscientiser les futurs enseignants sur la place de leur corps en formation, ainsi que sur son potentiel pour l'accès au sens, mais également l'animation du cours (Holt, 2020 ; Tellier, 2008, 2016). L'aspect vocal n'a pas été évoqué ici faute de place, mais les premières analyses qualitatives montrent que les futurs enseignants parlent plus fort et ralentissent leur débit à l'issue du module.

### Projet VAPVISIO

Le projet ANR VAPVISIO<sup>4</sup> a pour objectif d'identifier les compétences techno-sémio-pédagogiques (Guichon, 2012, voir la définition dans la partie 1 du présent article) nécessaires pour l'enseignement des langues par visioconférence. Ce travail a également pour but d'organiser ces compétences dans un référentiel, pour une mise à disposition en formation de formateurs (Cappellini, 2020). Dans ce projet, des étudiants de master 1 en didactique du FLE suivent dans leur formation un module de 26h intitulé « Concevoir et tutorer un cours de langue en ligne », co-animé par Christelle Combe et Marco Cappellini. Dans ce cours, ils apprennent à enseigner à la fois en mode asynchrone et synchrone. Ils complètent la formation théorique et méthodologique par une mise en pratique en dispensant quelques séances de tutorat hebdomadaire en ligne en visioconférence avec des apprenants de FLE américains (voir Cappellini & Hsu, 2020 pour un descriptif du dispositif). Dans le cadre du projet de recherche, les séances de visioconférence ont été enregistrées et pour quelques étudiants de master, nous avons également enregistré leurs mouvements oculaires pour voir quels étaient leurs foyers d'attention sur l'écran (ces données ne seront pas analysées ici, mais voir l'article de Marco Cappellini dans ce numéro, qui se base sur ce projet). En outre, les étudiants étaient invités à participer à des entretiens

<sup>3</sup> Malheureusement le cadrage lors de ce moment du module a été trop large et la qualité de l'image n'est pas très bonne.

<sup>4</sup> VAPVISIO : Vers une approche comparative de l'apprentissage/enseignement des langues étrangères par visioconférence pour développer les compétences techno-sémio-pédagogiques d'enseignants en formation. Cette recherche a bénéficié du soutien de l'Agence Nationale de la Recherche et rentre dans le projet ANR JCJC VAPVISIO <https://anr.fr/Projet-ANR-18-CE28-0011>.

d'autoconfrontation individuels avec un de leurs formateurs après les séances de télécollaboration, afin d'analyser comment se développait leur compétence techno-sémio-pédagogique.

À titre d'exemple, nous pouvons présenter ici le cas de Carole, qui n'a quasiment pas d'expérience pédagogique au moment de l'étude et très peu de pratique de la visioconférence<sup>5</sup>. Une particularité de la plateforme *Adobe Connect* utilisée dans ce projet est que l'affichage, la taille et la disposition des vidéos, documents partagés, et autres fenêtres choisies par l'animateur sont imposées aux autres participants. C'est-à-dire que tous voient la même chose (contrairement à certaines plateformes comme *Zoom* où chacun est libre d'agrandir et d'afficher ou non les fenêtres vidéo par exemple). Par ailleurs, plusieurs dispositions sont disponibles, notamment un mode *présentateur* qui facilite les présentations plus magistrales et un mode *discussion* qui est plus pertinent pour les interactions et le partage de documents. Tout cela relève des ressources multimodales typiques des échanges en ligne. Apprendre à manier ces différents affichages et à les utiliser à bon escient sollicite tous les volets de la compétence techno-sémio-pédagogique. Le volet technique pour savoir naviguer à travers les modes d'affichage de la plateforme pédagogique, le volet sémiotique pour optimiser les affordances de l'outil et rendre visible ce qui est nécessaire à l'échange pédagogique et le volet pédagogique pour servir la planification et faciliter l'interaction didactique et les apprentissages. La maîtrise de tous ces aspects au service de l'agir professoral (Cicurel, 2011) rappelle la notion de *multi-agenda augmenté* d'Azaoui (2019) évoquée plus haut.

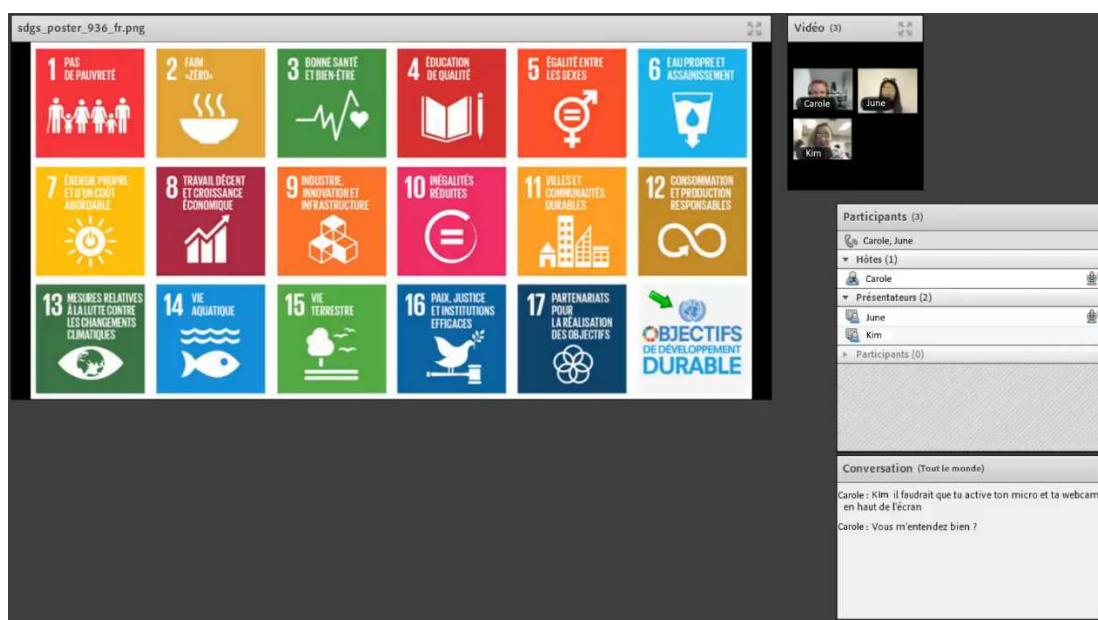


Figure 8 : Séquence de compréhension d'un document en ligne (séance 1 de Carole) (Corpus Vapvisio, 2019)

On note une certaine évolution dans la pratique multimodale de Carole. Lors de sa première séance en visioconférence, elle affiche un document iconographique sur lequel elle pose des questions de compréhension globale assez classiques pour l'accès au sens d'un document dans un cours de langue (Figure 8).

Dans l'échange transcrit ci-dessous (Figure 9), elle demande qui a produit ce document (celui utilisé pour la compréhension écrite, voir Figure 8). Au début de l'interaction, la qualité sonore de la visioconférence est médiocre et Carole n'entend pas la réponse faite par Kim, son apprenante. Une fois la réponse répétée par Kim, Carole ratifie (« *c'est ça* ») et montre où se trouvait la réponse sur le document (= le logo des Nations Unies).

<sup>5</sup> Les corpus ont été enregistrés au printemps 2019, avant la crise sanitaire liée au Covid 19 et donc avant que la visioconférence dans l'enseignement universitaire devienne d'usage courant.

Carole	et + est-ce que vous avez une idée de qui euh a fait ce qui a créé ce document (0.47)
Kim	hum + les + euh + nations (2.40)
Carole	j'ai pas entendu j- + désolée (1.97) vous vous m'avez dit j'ai pardon (3.95) j'ai j'ai pas du tout entendu votre réponse je suis désolée (1.94) vous m'avez dit que vous sa- vous avez vu qui c'était qui euh
Kim	ah hum c'est euh (0.44)
Carole	qui avait <u>fait le document</u>
Kim	<u>oui euh</u> + les nations unies hum + c'est le document et euh j'ai vu euh ce document + <u>hum</u>
Carole	<u>c'est ça</u> (0.97)
Kim	dans un autre classe euh + on + pour des {incompréhensible} (0.36) hum (0.22) aussi (1.20)
Carole	hum (1.06) mais on fait si on regarde là c'est normalement vous voyez mon curseur en vert (0.31)
Kim	hum
Carole	si vous regardez vous voyez que c'est le logo <u>des Nations Unies</u>
Kim	<u>hum hum</u> (0.38)
Carole	et en fait ils ont travaillé avec euh l'UNESCO (0.92) pour euh pour faire euh ce document

Figure 9 : Transcription de séquence de compréhension d'un document en ligne

Dans un cours de langue en présentiel, l'enseignant aurait pointé du doigt la partie du document concernée, mais en visioconférence, si les gestes pédagogiques sont possibles, un pointage concret de cette sorte est impossible. Carole substitue alors cette fonction par le recours à une flèche verte (qu'elle nomme curseur) et qu'elle déplace pour montrer le logo. L'utilisation de ce curseur permet d'établir une attention conjointe : « *mais on fait si on regarde là c'est normalement vous voyez mon curseur en vert* ». Cette petite flèche devient donc un geste déictique numérique.

Si lors de cette première séance, Carole gère bien l'usage du curseur, les autres modalités sont, en revanche, sous-exploitées. Comme on peut le voir dans la Figure 8, la vidéo est de taille très réduite ce qui ne permet pas bien aux apprenants de voir leur enseignante (une partie du visage de Carole est même dissimulée derrière l'étiquette de son nom), ni ses gestes si elle en utilise. La modalité écrite est également peu sollicitée, la partie « conversation » en bas à droite n'a servi qu'à vérifier le bon fonctionnement du matériel audio-visuel en début de séance et n'a pas été utilisée à des fins pédagogiques. Lors de son entretien d'autoconfrontation sur cette séance, elle évoque très peu l'usage de la multimodalité dans sa pratique à part quelques références à ses gestes et à l'agrandissement d'un document partagé.

Lors de sa 3<sup>e</sup> séance, l'utilisation de la plateforme a relativement évolué. Carole a choisi cette fois le mode *discussion* de la plateforme, qui lui permet d'agencer différentes fenêtres et d'en régler la taille. La Figure 10 montre la densité multimodale (Norris, 2004) de cette interaction didactique. Carole a disposé les fenêtres vidéo en haut de l'écran et les a suffisamment agrandies pour être visible lorsqu'elle produit des gestes pédagogiques (ici un geste iconique dessinant un carré pour expliquer ce qu'est un *flyer*). En bas à gauche, elle a affiché le document produit par son apprenante sur lequel elles sont en train de discuter (une charte pour réduire les déchets lors de l'organisation de manifestations scientifiques), à côté il y a une image qui illustre le thème du document (la réduction des déchets), et elle a également disposé en bas à droite une fenêtre « notes de discussion » qui lui sert de tableau pour synthétiser les idées et le vocabulaire nouveau.

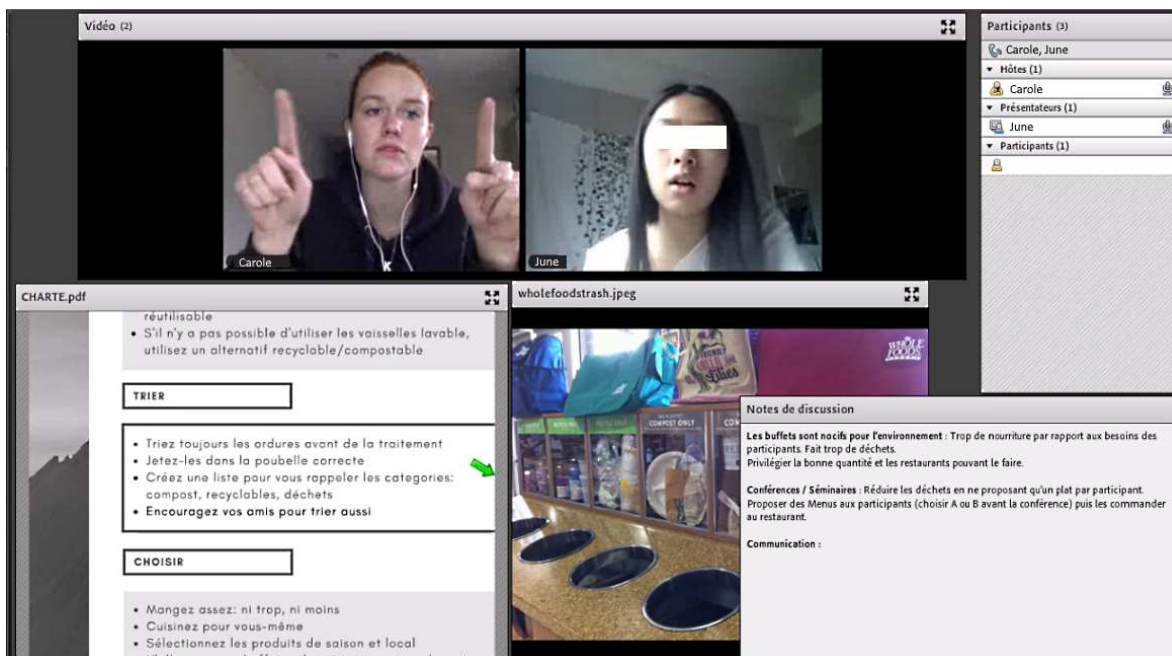


Figure 10 : Densité multimodale (3<sup>e</sup> séance de Carole) (Corpus Vapvisio, 2019)

Lors de son entretien d'autoconfrontation sur cette séance (Figure 11), son formateur l'interroge sur cette disposition :

Formateur	et concernant la la disposition spatiale euh comment elle comment elle a été conçue ?
Carole	euh ben au début comme j'avais pas la charte et puis j'ai mis en espace euh en discussion je crois et pas en en présentation comme on avait fait les autres fois pour euh puisque pour moi là c'était plus un échange sur la charte et pas et pas vraiment une séance de tutorat sur le document
Formateur	ouais
Carole	donc euh pour qu'on puisse se voir puisque bon c'est c'est important pour lire la charte aussi et d coup j'ai essayé de manier puisque j'ai pas pu le faire avant le début
Formateur	oui parce qu'il y avait eu un problème et y avait pas non plus d'accord et euh question bête pourquoi les visages en haut plutôt qu'en bas ou
Carole	ben parce que euh
Formateur	s'il y a une raison
Carole	pour qu'on ait l'impression qu'on se regarde il faut regarder la caméra
Formateur	hum hum
Carole	donc pour regarder la caméra c'est plus simple euh d'avoir les le visage juste en dessous de ma caméra centrale donc c'est ça
Formateur	d'accord c'était fait exprès
Carole	hum
Formateur	d'accord ok

Figure 11 : Extrait de l'entretien d'autoconfrontation de la 3<sup>e</sup> séance

Carole explique qu'elle a choisi le mode « *discussion* » volontairement pour faciliter l'interaction « *j'ai mis en espace euh en discussion je crois et pas en en présentation comme on avait fait les autres fois pour euh puisque pour moi là c'était plus un échange sur la charte* ». Afin de bien voir la charte sur laquelle elles échangent, l'affichage à l'écran lui semble nécessaire pour renforcer l'attention conjointe sur ce document. Le formateur l'interroge ensuite sur le choix fait d'avoir placé les vidéos en haut. On sent dans son discours qu'il ne croit pas à un choix délibéré « *question bête* », « *s'il y a une raison* ». Et pourtant, la décision est bien réfléchie et montre le développement de la compétence techno-sémio-pédagogique de Carole. En effet, elle sait que la prise de vue de la caméra se fait en

haut de l'écran et que pour donner l'impression de regarder son interlocuteur, c'est là qu'il faut diriger le regard (volet technique). Elle positionne donc les fenêtres vidéo en haut de l'écran pour rapprocher son regard de la caméra lorsqu'elle regarde son apprenante (volet sémiotique). Elle sait que le regard dans l'interaction didactique est important pour engager l'apprenant dans l'interaction et pour montrer l'attention qu'on lui porte (volet pédagogique). Lors de cette 3<sup>e</sup> séance en visioconférence, Carole montre donc une meilleure maîtrise de l'outil.

## Conclusion

Cet article a présenté deux projets de recherche qui portent sur la multimodalité dans la formation des enseignants de langue, et notamment de français langue étrangère. Si dans le cadre de ces projets, on voit que les futurs enseignants optimisent l'usage des ressources multimodales qu'ils ont à leur disposition (celles de la parole et/ou celles de l'environnement numérique qu'ils utilisent), rien ne permet d'affirmer que cette modification soit durable. En effet, il serait pertinent de suivre ces étudiants dans leur pratique d'enseignement et voir si, en dehors du cadre du projet de recherche et de la formation, ils mettent également en pratique les principes sur lesquels leur attention a été attirée. Car l'objectif de la formation à l'optimisation des ressources multimodales n'est pas de façonner les pratiques enseignantes en leur imposant un modèle multimodal à suivre mais bien de leur faire prendre conscience des ressources à leur disposition et de leur potentiel pédagogique (Tellier et Cadet, 2014 ; Tellier & Yerian, 2018). Chacun ensuite doit explorer le style qui lui convient pour développer son propre agir professoral multimodal.

L'auteure remercie les autres participants au projet *VAPVISIO* (Marco Cappellini, Christelle Combe, Isabelle Cros et Benjamin Holt) et *Corps en formation* (Keli Yerian), ainsi les membres du CEP du Laboratoire Parole et Langage qui ont contribué au recueil de données et traitement du corpus (Carine André, Cyril Deniaud, Alain Ghio et Antonio Serrato).

VAPVISIO a bénéficié du soutien de l'Agence Nationale de la Recherche et rentre dans le projet ANR JCJC <https://anr.fr/Projet-ANR-18-CE28-0011>.

## RÉFÉRENCES BIBLIOGRAPHIQUES

- AZAOUI, B.** *Coconstruction de normes scolaires et contextes d'enseignement. Une étude multimodale de l'agir professoral*, Thèse de doctorat non publiée, Université Paul Valéry, 2014.
- AZAOUI, B.** Multimodalité, transmodalité et intermodalité : considérations épistémologiques et didactiques. *Revue de recherches en littérature médiatique multimodale*, n°10, 2019.
- AZAOUI, B., & TELLIER, M.** Comment le corps coconstruit les discours et le sens. *TIPA. Travaux interdisciplinaires sur la parole et le langage*, n°36, 2020.
- BUCHETON, D., & SOULÉ, Y.** Les gestes professionnels et le jeu des postures de l'enseignant dans la classe : Un multi-agenda de préoccupations enchâssées. *Education didactique*, Vol. 3(3), pp. 29-48, 2009.
- CAPPELLINI, M.** Télécollaboration et formation de formateurs en langues au tutorat en ligne – Un état de l'art. *Alsic. Apprentissage des Langues et Systèmes d'Information et de Communication*, Vol. 23, n° 1, 2020.
- CAPPELLINI, M., & COMBE, C.** Analyser des compétences techno-sémio-pédagogiques d'apprentis tuteurs dans différents environnements numériques : Résultats d'une étude exploratoire. *Alsic. Apprentissage des Langues et Systèmes d'Information et de Communication*, Vol. 20, n° 3, 2017.
- CAPPELLINI, M., & HSU, Y.-Y.** When future teachers meet real learners through

- telecollaboration : An experiential approach to learn how to teach languages online. *Journal of Virtual Exchange*, 3, pp. 1-11, 2020.
- CHAUDRON, C.** Vocabulary Elaboration in Teachers' Speech to L2 Learners. *Studies in Second Language Acquisition*, 4(2), pp. 170-180, 1982.
- CICUREL, F.** *Les interactions dans l'enseignement des langues. Agir professoral et pratiques de classe.* Paris, Éditions Didier, 2011.
- COLLETTA, J.-M.** *Le développement de la parole chez l'enfant âgé de 6 à 11 ans.* Sprimont, Editions Mardaga, 2004.
- GUICHON, N.** *Vers l'intégration des Tic dans l'enseignement des langues.* Paris, Didier, 2012.
- GUICHON, N.** Une approche sémio-didactique de l'activité de l'enseignant de langue en ligne : Réflexions méthodologiques. *Éducation et didactique*, 7(1), pp. 101-116, 2013.
- GUICHON, N.** Construire une présence pédagogique en ligne. In Guichon, N. & M., Tellier (dir.). *Enseigner l'oral en ligne, une approche multimodale.* Paris, Didier, pp. 29-58, 2017.
- GUICHON, N. & TELLIER, M.** (dir.) *Enseigner à l'oral en ligne : une approche multimodale des interactions pédagogiques.* Paris, Didier, 2017.
- GUIMBRETIERE, E.** La voix de l'enseignant. In M. Tellier & L. Cadet (dir.) *Le corps et la voix de l'enseignant: théorie et pratique.* Paris, Éditions Maison des Langues, pp. 15-28, 2014.
- GULLBERG, M.** Handling Discourse: Gestures, Reference Tracking, and Communication Strategies in Early L2. *Language Learning*, 56(1), pp. 155-196, 2006.
- HAMPEL, R., & STICKLER, U.** (dir.). *Developing Online Language Teaching.* Palgrave Macmillan, New York, 2015.
- HOLT, B., TELLIER, M. & GUICHON, N.** The use of teaching gestures in an online multimodal environment: the case of incomprehension sequences. In Ferré, G. & M. Tutton (dir.) *Gesture and Speech in Interaction - 4th edition.* Nantes, 2-4 September 2015, pp. 149-154, 2015.
- HOLT, B.** Le rôle des gestes dans les explications lexicales par visioconférence. *TIPA. Travaux interdisciplinaires sur la parole et le langage*, n°36, 2020.
- KANG, S.-T.** Utilisation du geste illustratif par les apprenants coréens dans la classe de langue. *TIPA. Travaux interdisciplinaires sur la parole et le langage*, n°36, 2020.
- KENDON, A.** *Gesture : Visible Action as Utterance.* Cambridge University Press, 2004.
- MATSUMOTO, Y., & DOBS, A. M.** Pedagogical Gestures as Interactional Resources for Teaching and Learning Tense and Aspect in the ESL Grammar Classroom. *Language Learning*, n°67(1), pp. 7-42, 2017.
- MCNEILL, D.** *Hand and Mind : What Gestures Reveal About Thought.* The Chigago University Press, Chicago, 1992.
- NORRIS, S.** *Analyzing multimodal interaction: a methodological framework.* New York, London, Routledge, 2004.
- OWEN, D.** Do teachers modify their speech according to the proficiency of their students ? *ELTED*, n°2(1), pp. 31-51, 1996.
- PUJADE-RENAUD, C.** *Le corps de l'enseignant dans la classe.* Paris, L'Harmattan, 1983/2005.
- SATO, R.** Gestures in EFL classroom: Their relations with complexity, accuracy, and fluency in EFL teachers' L2 utterances. *System*, n°89, 2020.
- SHARIFZADEH, S., & SARRÉ, C.** Les processus inférentiels dans l'interaction multimodale en L2. *TIPA. Travaux interdisciplinaires sur la parole et le langage*, n°36, 2020.
- SIME.** "Because of Her Gesture, It's Very Easy to Understand"—Learners' Perceptions of Teachers' Gestures in the Foreign Language Class. In McCafferty, S. G. & Stam, G. (dir.). *Gesture: Second language acquisition and classroom research.* New York, Routledge, pp. 271-291, 2008.
- STAM, G.** Changes in Thinking for Speaking: A Longitudinal Case Study. *The Modern Language Journal*, n°99(S1), pp. 83-99, 2015.
- SUEYOSHI, A., & HARDISON, D.** The Role of Gestures and Facial Cues in Second Language Listening Comprehension. *Language Learning - LANG LEARN*, n°55, pp. 661-699, 2005.

- TELLIER, M.** Dire avec des gestes. *Le Français dans le monde. Recherches et applications*, pp. 40-50, 2008.
- TELLIER, M.** Donner du corps à son cours. In M. Tellier & L. Cadet (dir.), *Le corps et la voix de l'enseignant : Théorie et pratique*, Paris, Maison des Langues, pp. 101-114, 2014.
- TELLIER, M.** Prendre son cours à bras le corps. De l'articulation des modalités kinésiques avec la parole. *Recherches en didactique des langues et des cultures. Les cahiers de l'Acedle*, n°13, pp. 13-1, 2016.
- TELLIER, M.** Ceci est mon corps : le corps de l'enseignant dans la classe. In N. Olympio, F. Davin, M. Crocco, C. Hache, M. Cappellini (dir.). *Apprentissage et éducation : conditions, contextes et innovations pour la réussite scolaire, universitaire et professionnelle*. Aix en Provence, Presses Universitaire de Provence, à paraître (a).
- TELLIER, M.** L'émergence de la compétence techno-semio-pédagogique des futurs enseignants de FLE lors d'interactions didactiques en visioconférence. In A. Burrows, I. Cros, N. Kübler & G. Miras (dir.). *Entre présence et distance : Enseigner et apprendre les langues à l'université à l'ère numérique*. Paris, Presses de l'INALCO, à paraître (b).
- TELLIER, M. & CADET, L.** (dir.). *Le corps et la voix de l'enseignant : théorie et pratique*. Paris, Éditions Maison des Langues, 2014.
- TELLIER, M., & YERIAN, K. D.** Mettre du corps à l'ouvrage : Travailler sur la mise en scène du corps du jeune enseignant en formation universitaire. *Recherche et pratiques pédagogiques en langues de spécialité. Cahiers de l'Apliu*, Vol. 37 n°2, 2018.
- TRÉVISIOL-OKAMURA, P., & MARQUILLÓ-LARRUY, M.** La reformulation et multimodalité dans un cours en visioconférence. *Corela. Cognition, représentation, langage*, n°HS-18, 2015.
- WITTENBURG, P., BRUGMAN, H., RUSSEL, A., KLASSMANN, A., & SLOETJES, H.** *ELAN: A professional framework for multimodality research*, pp. 1556-1559, 2006.