Le rôle des gestes dans les explications lexicales par visioconférence

The role of gestures in videoconferenced lexical explanation sequences

Benjamin Holt

Édition électronique

URL : <http://journals.openedition.org/tipa/3458>

DOI : 10.4000/tipa.3458

ISSN : 2264-7082

Éditeur

Laboratoire Parole et Langage

Référence électronique

Benjamin Holt, « Le rôle des gestes dans les explications lexicales par visioconférence », *TIPA. Travaux interdisciplinaires sur la parole et le langage* [En ligne], 36 | 2020, mis en ligne le 01 juin 2020, consulté le 27 janvier 2021. URL : <http://journals.openedition.org/tipa/3458> ; DOI : <https://doi.org/10.4000/tipa.3458>

Ce document a été généré automatiquement le 27 janvier 2021.

La revue *TIPA. Travaux interdisciplinaires sur la parole et le langage* est mise à disposition selon les termes de la licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International.

Le rôle des gestes dans les explications lexicales par visioconférence

The role of gestures in videoconferenced lexical explanation sequences

Benjamin Holt

Introduction

- 1 Les gestes que l'on fait quand on parle, autrement dit les gestes coverbaux, sont spontanés, fortement liés au langage et participent à la communication (Goldin-Meadow, 2003 ; Kendon, 2004 ; McNeill, 1992). Ils permettent à celui qui parle d'organiser ses pensées et à trouver ses mots, tout en aidant le destinataire à décoder le sens de l'énoncé (Stam & McCafferty, 2008). Indissociables et complémentaires, les gestes et la parole doivent être pris en compte ensemble pour pouvoir appréhender la totalité du message émis par un interlocuteur. Bien que la parole puisse être comprise sans gestes, au téléphone par exemple, les sujets saisissent plus d'informations lorsqu'ils voient les gestes et entendent la parole en même temps (Beattie & Shovelton, 1999). Si les gestes et la parole d'un interlocuteur sont contradictoires, le destinataire aura plus de difficulté à comprendre le message (Kelly, Özyürek, & Maris, 2010). L'hypothèse selon laquelle les gestes sont allocentrés, c'est-à-dire faits pour l'autre, a été confirmée par Alibali, Heath, & Myers (2001) qui ont trouvé que les sujets produisent plus de gestes représentationnels quand l'interlocuteur est visible. Ceci est vrai non seulement en présentiel mais aussi en visioconférence. A cet effet, Mol, Krahmer, Maes, & Swerts (2011) ont trouvé que les sujets produisent plus de gestes représentationnels en visioconférence quand la webcam est allumée et quand ils savent qu'ils sont vus par un interlocuteur.
- 2 En contexte pédagogique, nous nous intéressons aux gestes produits par les enseignants lorsqu'ils enseignent la langue étrangère (Azaoui, 2019 ; Sime, 2006, 2008 ; Tellier, 2008a ; Tellier & Cadet, 2014). Il s'agit d'étudier ce que Tellier (2008a) nomme les

gestes pédagogiques. Ces gestes sont plus conscientisés que les gestes coverbaux sans visée didactique et servent à faciliter l'accès au sens dans une langue étrangère. Les gestes manuels et les expressions faciales sont utilisés par les enseignants pour remplir différentes fonctions pédagogiques.

- 3 Les gestes pédagogiques sont utiles en particulier lors des explications lexicales. Lazaraton (2004 : 89) remarque que *one sort of information that is commonly conveyed in the ESL classroom is the meaning of vocabulary words, which is perhaps the prototypical language use situation requiring the explicit deployment of gestures for teaching purposes*.¹ Les gestes iconiques se prêtent bien à l'explication lexicale car ils renforcent le lien entre la forme phonologique d'un mot et son signifié (Kelly, McDevitt, & Esch, 2009). Lorsqu'un mot est associé à un geste, il est mieux mémorisé (Macedonia, 2014 ; Macedonia & von Kriegstein, 2012 ; Quinn-Allen, 1995 ; Tellier, 2008b), surtout lorsque l'apprenant reproduit le geste, ce qui s'appelle l'*enactment effect* (Engelkamp & Krumnacker, 1980 cité dans Macedonia & Klimesch, 2014) ou bien un *subject-performed task* (Cohen, 1981). Ce n'est pas le simple fait de bouger le corps qui renforce l'encodage de mots, les gestes doivent avoir une valeur représentationnelle, un certain degré d'iconicité. Macedonia, Müller, & Friederici (2011) ont ainsi démontré que les gestes iconiques renforcent la mémorisation de mots mieux que les gestes dépourvus de sens. L'effet positif des gestes sur la mémorisation d'items lexicaux a été observé non seulement en laboratoire mais aussi en salle de classe (Macedonia & Klimesch, 2014), pour les phrases entières et pour les mots abstraits (Macedonia & Knösche, 2011). Dans ces expériences, la reproduction d'un geste par l'apprenant a eu un meilleur effet sur la mémorisation que le visionnage d'un geste sans reproduction kinésique. Une étude récente a toutefois démontré que le fait de visionner un geste sans le reproduire a aussi un effet positif sur la mémorisation, car cela active des réseaux sensorimoteurs (Macedonia, Repetto, Ischebeck, & Mueller, 2019). En bref, les gestes, quand ils sont vus par l'apprenant et surtout quand ils sont reproduits, facilitent grandement l'apprentissage d'items lexicaux en langue étrangère.
- 4 Le présent article se donne pour objectif d'analyser le déploiement de gestes pédagogiques par des apprentis-enseignants qui enseignent le français par visioconférence poste à poste². La visioconférence est un contexte propice à la négociation du sens (Nicolaev, 2010 ; Sarré, 2011 ; Van der Zwaard & Bannink, 2019), et c'est à travers la négociation que la langue est apprise (Ellis, Tanaka, & Yamazaki, 1994 ; Hatch, 1978 ; Long, 1981, 1983 ; Newton, 1995 ; Pica, 1994 ; Pica, Young, & Doughty, 1987 ; Porquier & Py, 2008 ; Swain, 1985, 1995). De plus, que l'échange ait lieu en présentiel (de Pietro, Matthey, & Py, 1989 ; Fasel Lauzon, 2014) ou en ligne (Blake, 2000 ; McDonough & Sunitham, 2009 ; Nicolaev, 2010 ; Renner, 2017 ; Smith, 2003), la plupart des séquences de négociation du sens sont déclenchées par l'incompréhension ou la recherche d'un item lexical.
- 5 A partir d'un corpus d'interactions pédagogiques recueilli dans le cadre d'un projet qui s'inscrit dans la lignée du Français en Première Ligne (FIL) (Develotte, Guichon, & Kern, 2008), nous proposons une étude de l'apport des gestes pédagogiques aux séquences d'explication lexicale. Cette étude va nous permettre de mettre à jour les fonctions des gestes pédagogiques dans ce contexte bien spécifique qui est celui des explications lexicales par visioconférence. Dans un premier temps, nous dresserons un état de l'art sur l'enseignement par visioconférence, l'explication lexicale en langue étrangère, et la multimodalité. Puis, nous décrirons la façon dont nous avons repéré les séquences d'explication lexicale et les avons annotées. Enfin, à l'aide de quelques exemples, nous

mettrons à jour certaines fonctions des gestes pédagogiques lors des séquences d'explication lexicale par visioconférence.

Visioconférence, multimodalité et gestes

Enseignement par visioconférence : quelles spécificités ?

- 6 L'enseignement par visioconférence comporte de nombreux avantages pédagogiques. Des apprenants de langue n'ont pas besoin de se déplacer pour découvrir la langue et la culture d'un pays étranger et de futurs enseignants peuvent gagner de l'expérience pédagogique en restant chez eux (Develotte et al., 2008 ; O'Dowd, 2000, 2016). Le temps de prise de parole chez les apprenants est plus important que dans une salle de classe classique puisque ceux-ci sont souvent seuls ou en binôme face à un enseignant. Cela permet aux apprenants de bénéficier de plus d'input, modifié sur mesure pour l'apprenant dans l'interaction (Long, 1981, 1983), et de plus de production d'output provoqué par l'enseignant (Swain, 1995, 2000). Les apprenants jouissent de la possibilité de pouvoir demander à tout moment à l'enseignant une explication qui sera *a priori* adaptée au niveau langagier de l'apprenant et qui pourra être négociée, modifiée, et reformulée si besoin. Un autre avantage lié à ce contexte (du moins pour le F1L) concerne l'écart d'âge faible entre les participants qui fait que les conversations ont un ton plus informel. Les relations interpersonnelles sont ainsi plus intimes avec un degré de connivence accru (Develotte & Drissi, 2013 ; Guichon, 2017a ; Guichon & Drissi, 2008).
- 7 Alors que certains des avantages susmentionnés pourraient être attribués aussi à l'enseignement en présentiel, notamment aux cours particuliers où l'apprenant est seul face à l'enseignant, il reste néanmoins quelques particularités propres à la communication médiée par ordinateur (désormais CMO) qui méritent d'être abordées. La visioconférence offre des possibilités riches de communication par rapport à d'autres moyens de communication à distance plus rudimentaires, mais tout type de CMO provoque des ruptures par rapport à l'enseignement classique en salle de classe, notamment des ruptures spatiales, médiatiques, et temporelles (Develotte & Mangenot, 2007). La visioconférence est un outil de médiation qui ne restitue pas parfaitement le son et l'image de l'autre : images floues, mouvements saccadés, désynchronisation entre parole et image, et coupures intempestives constituent autant de phénomènes auxquels doivent faire face les utilisateurs de l'outil (Azaoui, 2017). Ces particularités ont un impact sur la communication, et donc sur la gestion pédagogique et sémiotique nécessaire.
- 8 Une maîtrise de la technologie est donc nécessaire afin que la focalisation sur les opérations techniques ne prime pas sur la focalisation sur l'interaction pédagogique (Guichon & Drissi, 2008). La gestion de plusieurs outils à la fois nécessite une certaine *polyfocalisation* de l'attention (Jones, 2004). A cette polyfocalisation sur les outils techniques s'ajoute la dimension langagière de l'échange. A l'instar de Bange (1992) qui parle de *bifocalisation* sur l'activité et sur la langue, nous pouvons parler alors de *trifocalisation* : focalisation sur l'activité en cours, focalisation sur la langue, et focalisation sur la manipulation de l'outil technique. Il revient donc aux apprentis-enseignants de maîtriser des compétences techno-sémio-pédagogiques (Cappellini & Combe, 2017) qui incluent la capacité à gérer les côtés organisationnel, pédagogique, socio-affectif et technique de l'échange en ligne. Nous verrons aussi que cette

compétence inclut la capacité à utiliser ses gestes à des fins pédagogiques et de les rendre visibles, particulièrement lors des séquences d'explication lexicale.

Connaître un item lexical en langue étrangère

- 9 Les procédés d'apprentissage du lexique en langue étrangère changent selon le niveau de l'apprenant. Les apprenants les plus avancés peuvent apprendre du lexique en lisant ou en écoutant, en s'appuyant sur leurs connaissances préalables afin d'inférer le sens des mots inconnus. Ce mode d'apprentissage, dit incident (Grossmann, 2011), est la façon dont les locuteurs apprennent la plupart des mots qu'ils connaissent dans leur langue maternelle. En langue étrangère, l'apprentissage incident est compliqué au début car pour comprendre un texte il faut connaître à peu près 95% des mots, soit environ 3 000 à 4 000 mots (Nation, 2013). A ce niveau, un mot sur vingt sera inconnu. Plus difficile encore, pour être en mesure d'inférer le sens des mots inconnus, il faut comprendre 98% des mots d'un texte, ce qui nécessite des connaissances lexicales encore plus riches, de l'ordre de 6 000 à 9 000 mots pour un texte écrit (Huckin & Coady, 1999) et 5 000 à 7 000 mots pour un discours oral (Schmitt, 2008). En dessous de ce seuil, l'apprentissage incident et autonome de nouveaux mots est problématique. C'est pour cette raison que l'apprentissage explicite du lexique est à privilégier pour les niveaux débutant et intermédiaire (Folse, 2004).
- 10 Toute explication lexicale a pour objectif de transférer et de faire converger des connaissances, et nos analyses didactiques ont pour objectif d'étudier ce transfert de connaissance. Afin de mener de telles analyses, il faut aborder le concept de *connaissance* d'un item lexical. Certains auteurs adoptent la conception classique : signifiant et signifié (Van der Linden, 2006), tandis que d'autres conçoivent la connaissance d'un item lexical par degrés de familiarité (Waring, 2002). Les connaissances productives et réceptives sont souvent distinguées (Tréville & Duquette, 1996). D'autres encore définissent différentes dimensions de la connaissances d'un item lexical (García, 2008 ; Laufer, 1994).
- 11 Nation (2013) résume la conception de la connaissance d'un mot et regroupe les dimensions en trois grandes catégories : la forme, le sens et l'usage. La forme englobe celles qui sont orale et écrite ainsi que leurs composants ; le sens englobe les concepts et référents du mot ; enfin l'usage comprend les fonctions grammaticales, les collocations, et les contraintes liées à l'usage comme le registre et la fréquence. Chacune de ces facettes est subdivisée encore entre connaissance réceptive et connaissance productive. Cette classification nous semble utile dans la mesure où elle permet d'identifier assez clairement la/les visée(s) d'une séquence d'explication lexicale. Dans le cadre de cette étude, nous nous limiterons à l'intention du locuteur puisque nous ne disposons pas de moyens pour mesurer l'acquisition. En nous basant sur ce modèle, nous allons évaluer la contribution des gestes aux séquences d'explication lexicale.

Modéliser une séquence d'explication lexicale

- 12 Depuis que Schegloff, Jefferson, & Sacks (1977) se sont intéressés aux réparations langagières liées au code, plusieurs modèles ont été élaborés pour rendre compte des différentes étapes suivies par les participants d'une interaction qui négocient de

manière collaborative le sens d'un item langagier. L'un des modèles les plus connus est celui de Varonis & Gass (1985) à quatre étapes qui commence lorsqu'un item langagier, le *déclencheur* (1), devient *indicateur* (2) d'un problème, et se termine lorsque celui qui n'a pas compris donne sa *réaction* (4) à la *réponse* (3) de l'explicateur. Ce modèle, très polyvalent, permet de suivre du début à la fin une séquence de négociation du sens pour faire état des techniques interactionnelles, pédagogiques, sémiotiques et multimodales déployées par les participants. Ces séquences de négociation, qui arrivent souvent lors des échanges pédagogiques par visioconférence (Nicolaev, 2010 ; Sarré, 2011 ; Van der Zwaard & Bannink, 2019), sont propices à l'acquisition langagière car elles privilégient la focalisation sur la forme (Long, 1991). Nous nous sommes appuyé sur ce modèle dans Holt, Tellier, & Guichon (2015) pour mesurer la production gestuelle des apprentis-enseignants lors des séquences d'incompréhension.

- 13 Il existe un modèle plus spécifiquement conçu pour les séquences d'explication lexicale, c'est-à-dire les séquences de négociation du sens qui s'imposent lorsque c'est un item lexical qui fait obstacle à la communication réceptive ou productive. La structure canonique que nous empruntons a été modélisée par Gülich (1990), De Gaulmyn (1991), et Lüdi (1991) et reprise par Fasel Lauzon (2008, 2014). Ce modèle est constitué de trois étapes :

L'ouverture : l'item lexical non compris est identifié, repéré, localisé, mis en exergue, et problématisé ;

Le noyau explicatif : l'item lexical non compris est expliqué, le plus souvent par l'enseignant, mais souvent en collaboration avec les apprenants ;

La clôture : l'enseignant ratifie et évalue la compréhension de l'apprenant.

- 14 Nous avons suivi ce modèle canonique pour diviser les séquences d'explication lexicale de notre corpus en phases. Lors de chaque séquence que nous avons repérée, des ressources sémiotiques et multimodales, y compris des gestes, sont déployées par les participants pour mener à bien la séquence d'explication lexicale.

Une perspective multimodale

- 15 Souvent, quand les séquences d'explication sont analysées, les ressources sémiotiques³ autres que le langage verbal oral se voient accorder une place accessoire. Or, lors d'une séquence de négociation du sens, les indices multimodaux comme la prosodie, les gestes, les expressions faciales, et les changements de posture jouent un rôle non seulement dans l'indication d'une incompréhension (Develotte & Drissi, 2013 ; Nicolaev, 2010) mais aussi dans la partie explicative de la séquence. L'importance du « non-verbal » lors des séquences d'explication a été évoquée par De Gaulmyn (1991 : 292) :

Il faudrait y ajouter le non-verbal. Parmi les indices non-verbaux, les gestes manuels ont une part importante, les locuteurs montrent les objets, les manipulent quand ils sont sous leurs yeux, ils figurent, miment, représentent spatialement les objets et les activités invisibles.

- 16 Une approche multimodale est donc nécessaire afin d'appréhender la complexité des séquences d'explication lexicale en contexte de visioconférence. Des analyses

précédentes réalisées en adoptant une telle approche (Holt & Tellier, 2017 ; Wigham, 2017) ont mis en évidence la variété de modes sémiotiques utilisés par les participants.

- 17 Selon une perspective multimodale sur l'interaction (Jewitt, 2014), le langage verbal fait partie d'un ensemble multimodal et ne prime pas sur les autres modes. Chaque mode réalise une tâche communicative différente, et ces tâches peuvent changer d'un moment à un autre. Les acteurs communiquent en sélectionnant et en configurant les modes. Par exemple, un apprenti-enseignant qui explique un item lexical par visioconférence peut passer d'une explication orale à une explication écrite, en ayant recours à l'envoi d'une photo pour illustrer son propos. La partie orale de son explication peut être émaillée de gestes. Face à un choix de modes, un interlocuteur va sélectionner le mode le plus adapté ou les modes les plus adaptés pour faire passer le message (Baldry & Thibault, 2006). Ces choix seront influencés par le profil interactif de chacun (Develotte, Guichon, & Vincent, 2010) et par les ressources disponibles. Parfois un mode sera plus porteur de sens qu'un autre, ce que Norris (2004) appelle la densité modale. Les gestes constituent un mode parmi d'autres, et nous nous intéressons aux moments où les gestes sont employés par les apprentis-enseignants pour être porteurs de sens à différentes phases des séquences d'explication lexicale.
- 18 Le champ de la webcam offre des informations riches : le visage de l'interlocuteur, ses expressions faciales, ses habits, ses postures, et la salle derrière (Guichon, 2013). Les gestes sont visibles aussi, à condition qu'ils soient produits dans le champ de la webcam. Nous verrons plus bas que la médiation a un effet considérable sur la production gestuelle. Puisque les gestes sont souvent produits hors champ (Guichon & Wigham, 2016), les mimiques faciales prennent une place importante et jouent un rôle compensatoire (Cosnier & Develotte, 2011).

Le rôle des gestes pédagogiques

- 19 Les gestes pédagogiques ont déjà été étudiés en salle de classe. A partir d'observations des apprenants, Sime (2006, 2008) s'est intéressée aux fonctions pédagogiques des gestes produits par des enseignants d'anglais. Elle a trouvé que les apprenants faisaient attention aux gestes quand ceux-ci avaient un effet sur l'apprentissage. Notamment, les apprenants prêtaient attention aux gestes en situation de détresse langagière quand la compréhension orale était difficile. Cette auteure décline trois fonctions des gestes produits par les enseignants : une fonction cognitive, une fonction émotionnelle/ affective, et une fonction organisationnelle. Concernant la fonction émotionnelle, une posture relâchée, un regard maintenu, et la présence de beaucoup de gestes manuels et d'expressions faciales servent à encourager l'apprenant et à lui donner des retours correctifs. Pour ce qui est de la fonction organisationnelle, l'usage de gestes emblématiques (Ekman & Friesen, 1969) et déictiques sert à gérer les tâches et les tours de parole. Les fonctions cognitives des gestes accomplissent trois sous-fonctions : améliorer la compréhension, faciliter le processus d'apprentissage, et indiquer une réaction vis-à-vis de l'*output* de l'apprenant. Des entretiens avec les élèves ont soulevé l'habileté des gestes à marquer et à expliquer des mots importants, et à aider l'apprenant à trouver le mot recherché.
- 20 Tellier (2008a) répertorie aussi trois fonctions des gestes pédagogiques : informer, animer, et évaluer. Cette auteure explique qu'il y a trois types de gestes pour informer : les gestes d'information grammaticale, les gestes d'information lexicale, et les gestes

d'information phonologique et phonétique. Les gestes d'animation servent entre autres à passer d'une activité à une autre, donner des consignes, donner la parole à quelqu'un, interroger, faire répéter, etc. Enfin, les gestes d'évaluation servent à évaluer la production langagière de l'apprenant.

- 21 Ces fonctions peuvent être appliquées à des séquences d'explication lexicale. Les gestes d'information servent à expliquer une facette de la connaissance d'un item lexical (la forme, le sens, ou l'usage). Les gestes d'animation sont déployés lors du passage d'une phase d'explication lexicale à une autre, ou quand l'enseignant cherche à impliquer l'apprenant dans la séquence explicative. On retrouve les gestes d'évaluation suite à une erreur lexicale d'un apprenant, ce qui déclenche une séquence d'explication lexicale, ou à la fin d'une séquence d'explication lorsque l'enseignant évalue la compréhension dont fait preuve l'apprenant.

Questions de recherche

- 22 Compte tenu de l'importance de la forme, du sens, et de l'usage des items lexicaux (Nation, 2013), et de l'importance de l'implication des apprenants dans les séquences d'explication lexicale (Fasel Lauzon, 2014 ; Laufer & Hulstijn, 2001), nous nous efforçons de répondre aux questions suivantes : comment les gestes produits par visioconférence contribuent-ils à la focalisation sur la forme, sur le sens, et sur l'usage des items lexicaux ? Comment les gestes sont-ils utilisés pour impliquer les apprenants dans les séquences d'explication lexicale ?

Méthodologie

Le contexte

- 23 La présente étude s'appuie sur le corpus ISMAEL (Guichon & Tellier, 2017) qui a été recueilli lors du premier semestre de l'année universitaire 2013-2014 à l'Université Lumière Lyon 2. Ce projet, cogéré par le laboratoire ICAR de l'Université de Lyon sous la direction de Nicolas Guichon, et par Dublin City University sous la direction de Françoise Blin, a réuni 12 futurs enseignants de français langue étrangère à Lyon et 18 apprenants à Dublin. Les apprentis-enseignants étaient inscrits dans un Master 2 FLE professionnalisant et les apprenants étaient inscrits en licence dans une école de commerce et devaient apprendre le français afin d'effectuer un stage l'année suivante en France. Ils avaient le niveau B1-B2 selon le CECRL (Conseil de l'Europe, 2001)⁴ et interagissaient une fois par semaine pendant 25 à 40 minutes sur des thèmes divers ayant trait au monde du commerce (les droits du travail, les expériences professionnelles, la préparation d'un stage, la gestion d'un projet, l'entretien d'embauche). Les séances ont été préparées de manière collaborative par les apprentis-enseignants. La plateforme de visioconférence utilisée pour l'étude s'appelle *Visu* et a été développée par l'Université de Lyon pour l'enseignement des langues (Guichon, Bétrancourt, & Prié, 2012).

Le corpus d'étude

- 24 D'autres types de données ont été recueillis en plus des captures d'écran dynamiques : les plans des cours, les bilans correctifs créés par les apprentis-enseignants, des enregistrements vidéo des interactions par une caméra hors-champ, des post-interviews avec les apprentis-enseignants et les apprenants, et des enregistrements des débriefings qui se sont tenus le lendemain avec le formateur. Avant de procéder aux enregistrements, tous les participants ont signé un formulaire de consentement pour participer à l'étude et pour donner aux chercheurs le droit de diffuser leur image et productions langagières. Certains participants n'ont pas souhaité participer au projet, et d'autres n'ont pas souhaité que leur image soit diffusée. Les visages de ces derniers ont alors été floutés. En fonction des réponses aux formulaires de consentement, 7 apprentis-enseignants et 12 apprenants ont été retenus pour l'étude. Des pseudonymes ont été attribués à tous les participants.
- 25 Ces données ont fait l'objet de sélection, de structuration et de partage afin de mériter le titre de « corpus » (Chanier & Ciekanski, 2010). Comme Rivens Mompean & Guichon (2009), nous travaillons principalement sur des captures d'écran. En collaborant avec la Cellule Corpus Complexes (CCC)⁵ du laboratoire ICAR pour les données en utilisant divers logiciels de traitement vidéo, nous avons reconstruit 28 séances pour un total de 15 heures et demie d'interaction.
- 26 Contrairement à d'autres corpus qui sont constitués d'interactions visiophoniques provoquées uniquement pour la recherche (Develotte, Kern, & Lamy, 2011), le présent corpus est plus écologique (Tellier, 2013) dans le sens où aucun chercheur n'est vraiment intervenu sur les consignes. Chaque apprenti-enseignant était libre de réinterpréter la tâche et de conduire l'interaction à sa propre façon, et la plupart des séquences d'explication lexicale n'ont pas été planifiées. Les interactions auraient eu lieu même sans enregistrement, et l'avantage de l'enregistrement par capture d'écran dynamique est que *la webcam est peu intrusive, les utilisateurs oubliant rapidement qu'un logiciel garde une trace de leur activité, et cette technique ne dénature pas l'activité des sujets comme cela est davantage le cas avec la présence d'une caméra* (Guichon, 2013 : 105). Nous considérons donc que comme une salle de classe classique, l'environnement virtuel dans lequel ont eu lieu les interactions visiophoniques est un véritable '*observatoire*' qui *permet de découvrir de quelle manière les données langagières sont maniées par les uns et les autres* (Cicurel, 2011 : 25).

Repérage des séquences d'explication lexicale

- 27 Nous avons défini une séquence d'explication lexicale comme un échange où une ou plusieurs facettes de la connaissance d'un item lexical sont abordées, questionnées, traitées, discutées, ou expliquées par les participants. La séquence d'explication lexicale peut résulter ou bien d'un problème de réception où un item lexical n'est pas compris par l'apprenant, ou bien d'un problème de production où un apprenant a une lacune lexicale et cherche le *mot juste* (Grossmann, 2011). Nous considérons que les séquences d'explication lexicale monogérées où l'apprenant ne parle pas doivent être prises en compte au même titre que les séquences d'explication lexicale qui sont dialogiques ou polygérées. Nous nous alignons sur les propos de Gülich (1991 : 334) qui argumente que :

Même si l'un des participants n'y contribue pas verbalement ou s'il s'agit du tour de parole d'un seul locuteur, le partenaire coopère à son accomplissement par le fait même que le locuteur s'adresse à lui ; par son existence même, l'interlocuteur coopère à la production discursive du locuteur, qui ne peut pas parler sans faire des hypothèses sur l'autre.

- 28 Afin de repérer les séquences d'explication lexicale telles que nous les avons définies, nous avons minutieusement visualisé notre corpus d'étude. Au total, 295 séquences d'explication lexicale ont été annotées dont 44% sont liées à des problèmes de compréhension et 53% sont liées à des problèmes de production. Les 2% des séquences restantes sont à la fois des problèmes de compréhension et de production. Concernant la gestion de l'équilibre parfois délicat entre focalisation sur la communication et focalisation sur la forme (Guichon *et al.*, 2012), le repérage montre que 103 séquences ont été déclenchées par l'apprenant, 133 ont été déclenchées par l'enseignant, et 59 ont été déclenchées par les deux à la fois. Les tableaux ci-dessous offrent une vision globale des séquences.

Figure 1 : Type de panne lexicale

Type de panne lexicale	Nb.	%
Problème de compréhension	131	44,4
Problème de production	157	53,2
Les deux à la fois	7	2,4
Total	295	100

Figure 2 : Participant initiant la séquence

Initiateur de la séquence	Nb.	%
Apprenant	103	34,9
Enseignant	133	45,1
Les deux à la fois	59	20
Total	295	100

Annotation des données

- 29 Nous avons utilisé le logiciel ELAN (Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006) pour annoter la parole orale et les phénomènes multimodaux. À l'aide d'une équipe de transpositeurs, la parole orale de l'intégralité du corpus a été transcrite. Nous avons choisi d'adopter la convention ICOR (Groupe ICOR, 2007) car elle est polyvalente et facilite le partage du corpus avec d'autres chercheurs possédant des préoccupations scientifiques diverses et complémentaires (Guichon, 2017b).

- 30 Après avoir repéré les séquences d'explication lexicale, nous les avons délimitées et les avons divisées en phases suivant le modèle canonique évoqué plus haut (ouverture, noyau et clôture). Pour chaque séquence, nous avons annoté les postures, les regards, les expressions faciales et les gestes produits par les apprentis-enseignants. Ultérieurement, nous avons comptabilisé les stratégies déployées par les apprentis-enseignants pour se focaliser sur la forme, sur le sens, sur l'usage des items lexicaux, et pour impliquer les apprenants dans les séquences d'explication. Notre façon de représenter nos annotations sur une page sera expliquée plus bas.
- 31 L'annotation des gestes produits en visioconférence n'est pas la même chose que l'annotation des gestes produits dans une salle de classe. En effet, la production de gestes en visioconférence est particulière car la webcam a un effet sur la visibilité du corps de l'interlocuteur. Même si on voit les mêmes gestes en face à face et sur un écran, les gestes sont perçus différemment et l'effet de la médiation peut en réduire l'efficacité communicative (Beattie & Shovelton, 1999). Le positionnement de la webcam, la taille de l'écran, la résolution de la vidéo, et le nombre d'images par seconde impactent la production et la réception des gestes en visioconférence.
- 32 Si un sujet est assis face à une caméra ou à un interlocuteur, on peut conceptualiser l'espace gestuel comme un espace couvrant plusieurs zones concentriques du centre du corps jusqu'aux extrémités (McNeill, 1992). Or, cela n'est pas envisageable en visioconférence en raison du champ réduit où ne sont visibles que la tête et les épaules du sujet (Cosnier & Develotte, 2011 ; De Chanay, 2011). Alors que l'espace gestuel de McNeill lui permet de distinguer et de décrire avec précision les tendances des différentes dimensions gestuelles qu'il décrit, les particularités de notre corpus nous obligent à nous contenter d'une classification binaire des gestes. Au lieu de dire que le geste est produit dans le centre, dans la périphérie, à gauche, ou à droite du champ de la webcam, il semble plus pertinent de dire que le geste est produit ou bien « dans le champ », ou bien « hors champ ». Une telle classification a permis à Guichon & Wigham (2016) de constater, à l'aide d'une caméra extérieure, que la plupart des gestes ont été produits hors champ. Bien que cette classification réduise la précision des analyses, elle a toutefois un avantage : on peut faire l'hypothèse que les gestes qui sont produits dans le champ ont probablement une plus grande intention communicative. Les participants, du fait qu'ils voient en miroir le retour de leur propre image, peuvent surveiller leurs propres gestes pour s'assurer qu'ils sont visibles.
- 33 Concernant l'annotation des mouvements, les gestes sont souvent annotés par des *unités gestuelles* (Kendon, 2004 ; McNeill, 1992). Afin d'annoter des gestes de cette façon, deux conditions au moins doivent être remplies. D'abord, on doit voir la position de repos des mains afin de pouvoir délimiter les unités gestuelles et ensuite, l'enregistrement vidéo doit être assez fluide pour pouvoir apercevoir les noyaux et délimiter les phases gestuelles. Les enregistrements vidéo dont nous disposons ne remplissent malheureusement pas ces deux conditions. Concernant la position de repos, on ne voit que la tête et les épaules des interactants la plupart du temps. Quand un geste est produit, on ne voit pas la position de repos, à moins que les mains ne partent de la tête de l'apprenti-enseignant. Quand un geste est produit totalement ou partiellement en dehors du champ de la webcam, soit il n'y a rien qui apparaît à l'image, soit on ne voit que le bout des doigts et le geste est non identifiable. Enfin, même si le noyau de l'unité gestuelle est visible, on peut l'identifier mais on ne reconnaît pas toujours le nombre de gestes en raison du caractère saccadé des

enregistrements vidéo. La qualité de la vidéo est dégradée par rapport à une vidéo prise par une caméra classique en raison du débit de la connexion Internet et de la résolution de la webcam.⁶ Nous représentons donc les mouvements saillants par des flèches accompagnées d'une description textuelle (cf. exemples ci-dessous).

- 34 Concernant la classification des gestes, nous considérons que la classification par fonction pédagogique proposée par Sime (2006, 2008) et Tellier (2008a) est plus adaptée à l'analyse de séquences d'explication lexicale qu'une catégorisation par dimension à la McNeill.

Analyses

Rendre ses gestes visibles pour expliquer

- 35 La saillance d'un geste n'est qu'un seul critère d'analyse, mais elle prend tout son sens en contexte médié par visioconférence où la visibilité de la webcam est réduite et où le simple fait d'être visible est une mesure importante. Sur une sous-partie du présent corpus, nous avons étudié dans Holt, Tellier, & Guichon (2015) la production gestuelle de quatre apprentis-enseignants lors des séquences d'incompréhension. Une telle séquence commence lorsque l'un des interactants arrête le déroulement de la conversation pour se focaliser sur un élément langagier, et se termine lorsque l'apprenti-enseignant décide de la clôturer et de reprendre la conversation momentanément suspendue (Varonis & Gass, 1985). Les résultats de cette étude ont infirmé notre hypothèse de départ selon laquelle les apprentis-enseignants allaient produire davantage de gestes lors des séquences d'incompréhension.
- 36 Il est cependant raisonnable de postuler, malgré ces résultats, que les apprentis-enseignants adapteront leurs gestes en les simplifiant, en les caricaturant, en les maintenant plus longtemps, et en les rendant plus visibles, afin de pallier les contraintes de résolution et de fluidité liées à la visioconférence. L'exemple suivant illustre ce phénomène. Victor, apprenti-enseignant, parle des compétences requises pour un stage en France lorsque la compréhension du mot « lien » fait défaut à son apprenant.⁷

Exemple 1 : Visibilité croissante des gestes

Transcription de la parole			Image de Victor	Phén. observés
1	VIC	et et est-ce que ce projet pourrait euh:/ c- le projet dont tu me parles/ (0,4) est-ce que tu pourrais l'utiliser/ pour euh vendre euh ta candidature/ (0,3) pour le stage\ (0,3) ou y a pas de lien (3,3)		Aucun geste n'est produit
2	LIM	pardon/ euh répétez s'il vous pl[âit/]		

3	VIC	[est-ce] que tu arriverais à faire un lien / entre ton expérience (0,3) de: que tu me parles maintenant/ (0,3) et: le: stage (0,6) que: que je t'ai envoyé/ l'annonce de stage (0,7)		Regard pensif et geste iconique peu visible représentant un lien
		est-ce qu'il y a des liens possibles/ ou c'est du c'est c'est pas du tout pareil/ (2,4)		
4	LIM	hum: (0,4) qu'est-ce que c- qu'est/ (0,6) qu'est-ce que c'est euh des liens/ (1,1)		Geste emblématique peu visible et regard pensif
5	VIC	qu'est-ce que c'est quoi/ (0,6)		
6	LIM	des liens (0,3)		
7	VIC	ah des liens/ c'est hum: (0,9) euh: une relation en- entre une chose		Début d'un geste iconique bien encadré
		et une autre euh on dit		La main droite se sépare de la main gauche pour illustrer le concept d'un lien.
		qu'il y a un lien		Victor refait son geste illustrant un lien. Les deux mains sont mises ensemble au milieu de l'écran et se séparent.

8	LIM	(0,5) ah oui oui oui		Victor maintient la fin de son geste pendant 800 ms.
---	-----	--------------------------------	---	--

- 37 Victor utilise le mot « lien » pour la première fois à la ligne 1 sans faire de geste. Lors de ce tour de parole, il n'y a pas focalisation sur la langue car rien dans le discours de l'apprenant n'a déclenché une séquence de focalisation sur la langue, et Victor n'a pas anticipé une incompréhension. Liam indique à la ligne 2 la présence d'une incompréhension et demande une répétition sans préciser l'élément problématique. Victor, en ignorant quel élément de son tour précédent a déclenché l'incompréhension, reformule sa question en faisant des gestes à peine visibles qui illustrent le mot « lien » et l'expression « c'est pas du tout pareil » (ligne 3). Puis, Le mot « lien » passe d'un *déclencheur potentiel* à un *déclencheur* (Varonis & Gass, 1985) aux lignes 4 et 6 quand Liam précise quel mot n'a pas été compris. Victor est donc conscient à la ligne 7 que c'est le mot « lien » qui est non compris et fait des gestes amples et bien visibles lors du noyau explicatif. Il produit deux gestes iconiques illustrant le sens du mot « lien » et maintient son geste pendant 800 ms jusqu'au début du tour de parole suivant de Liam (ligne 8). Ce type de geste maintenu, nommé *arrêt sur image* par Kida (2014 : 115), est fortement communicatif et s'emploie lorsque l'intercompréhension devient une priorité.
- 38 Cet exemple montre un apprenti-enseignant qui semble considérer que les gestes constituent une ressource utile lors de cette séquence d'explication lexicale, et qui semble être conscient de leur visibilité. Il ne gestualise pas du tout avant que l'incompréhension soit manifestée, un peu plus lorsqu'il sait qu'il y a une incompréhension quelque part, et rend ses gestes pleinement visibles lorsqu'il se focalise sur un item lexical en particulier. On assiste donc à un exemple de *recipient design* (Sacks, Schegloff, & Jefferson, 1974) où les gestes sont ajoutés progressivement afin de rendre intelligible la parole de l'enseignant. La conscientisation de cet apprenti-enseignant est confirmée lors d'une séance de débriefing dans laquelle il indique que :
- La technologie, pour moi, depuis le début je la prends comme... C'est pas, c'est pas naturel, c'est pas authentique. Après, c'est pas, c'est pas désastreux, ça fait partie de cette expérience et puis c'est tout ! C'est pas un problème. Mais, oui je sens que dans la façon de s'exprimer ou dans la gestuelle, quand tu es obligé de t'aligner sur la caméra pour faire ton geste...
- 39 Tout porte à croire que cet apprenti-enseignant est conscient du fait que les gestes nécessitent une certaine conscientisation afin d'être rendus visibles par visioconférence, et qu'ils constituent un outil de valeur dans ce contexte. Dans les sections suivantes, à l'aide d'exemples, nous allons montrer les différentes fonctions remplies par les gestes pour mener à bien les séquences d'explication lexicale.

Mettre en exergue la forme d'un item lexical

- 40 Pour qu'il y ait apprentissage, on doit d'abord être conscient de ce que l'on perçoit (Schmidt, 1990). Concernant le lexique, la forme d'un item lexical doit être repérée par l'apprenant avant d'être apprise (Groot, 2000). Les apprenants doivent clairement

comprendre qu'il existe une focalisation sur un item langagier car on apprend ce sur quoi on se focalise (Barcroft, 2003 ; Hill & Laufer, 2003). Le repérage, ou le *noticing* est fondamental, surtout lors de l'ouverture d'une séquence d'explication lexicale quand l'objet de la séquence est repéré, décontextualisé et problématisé (Fasel Lauzon, 2014). Une séquence d'explication lexicale ne peut donc pas avoir lieu sans focalisation sur la forme. Puisque *la mémorisation de l'information lexicale dépend en partie de sa saillance dans l'interaction* (Lüdi, 1991 : 199), l'enseignement est plus efficace si l'enseignant fait un effort pour rendre saillante la forme de l'item lexical. Un enseignant a à sa disposition plusieurs techniques linguistiques et multimodales par lesquelles il peut décontextualiser un item lexical et mettre en exergue sa forme. Les gestes constituent l'une de ses techniques.

- 41 Si l'on compte le nombre de fois que la webcam a été utilisée par les apprentis-enseignants pour se focaliser sur la forme (expressions faciales d'incertitude, haussement de sourcils, regards vers la caméra, mouvements de tête, mouvements de bouche exagérés, et gestes manuels tous confondus), ces ressources visuelles ont été utilisées dans 53 séquences d'explication lexicale sur 295 pour se focaliser sur la forme. Les gestes manuels ont été utilisés dans 12 séquences.

Surligner un item lexical par gestes

- 42 Une façon d'objectiver un mot à l'écrit est d'utiliser des guillemets (Rey-Debove, 1997), mais l'on peut aussi utiliser cette stratégie gestuellement. Dans l'exemple suivant, qui fait partie d'une analyse plus exhaustive de l'explication du mot-valise « bobo » dans Holt & Tellier (2017), Adèle utilise le logonyme « on dit » et produit des gestes pour mettre l'emphasis sur l'item lexical qu'elle est en train d'expliquer.

Exemple 2 : Surlignage d'un item lexical par guillemets gestuels

Parole orale			Image d'Adèle	Phén. observés
1	ADL	mais en france on dit/ hum\ (0,3) pour dire hipsters/ (0,4) on dit/ bobo\		Adèle fait des guillemets gestuels.

- 43 En produisant des guillemets gestuels en même temps qu'elle dit « bobo », l'apprentie-enseignante rend saillant le fait que c'est le mot « bobo » qui est en train d'être expliqué, et assure qu'il y a focalisation sur celui-ci.

Travailler la prononciation

- 44 La représentation phonologique mentale d'un item lexical est primordiale pour l'apprentissage, même à la lecture (García, 2008 ; Marijanović, Billières, & Panissal, 2008 ; Nation, 2013 ; Schmitt, 2008 ; Tréville, 2000 ; Tréville & Duquette, 1996). Tréville (2000 : 63) affirme qu'il existe une forte corrélation entre l'aptitude à construire des représentations phonologiques de mots nouveaux dans la mémoire et l'apprentissage du vocabulaire de la L2. Une technique mnémonique qui renforce la mémorisation d'un item lexical en reliant sa forme phonologique à son sens est celle du mot-clé (Folse, 2004 ; Nation, 2013 ;

Pressley, Levin, & McDaniel, 1987). Elle consiste à penser à un mot en L1 qui ressemble phonologiquement au mot en L2. Puis, on construit une image mentale qui relie les deux choses. Dans l'exemple suivant, Adèle emploie cette technique un tant soit peu différemment en utilisant des gestes pour aider l'apprenante à se construire une représentation mentale de chaque syllabe du mot « prononciation ». Faute d'espace, uniquement la deuxième partie de la séquence est représentée ci-dessous. Dans la première moitié de l'explication, Adèle énonce toutes les syllabes une par une et dit « non comme non » pour la deuxième syllabe, en secouant la tête. Elle relie donc le son /nɔ̃/ de la deuxième syllabe du mot « prononciation » au mot « non ». Dans la suite de la séquence, Adèle utilise ses gestes pour aider l'apprenante à se construire une représentation phonologique du mot.

Exemple 3 : Utilisation de gestes pour construire une représentation phonologique d'un mot

Parole orale			Image d'Adèle	Phén. observés
31	ADL	dis (.) dis une par une\ (0,4)		Geste métaphorique qu'elle dirige vers le bas en disant « une par une »
32	ALN	pro/ (0,5) pro/ (1,5)		Pouce levé et maintenu
33	ADL	non/		Secouement de tête
34	ALN	(0,5) non/ (0,2)		Index tendu et maintenu
35	ADL	cia/ (0,5)		Hochement de tête puis doigts tendus et maintenus lors de la prononciation de la troisième syllabe

36	ALN	cia/ (.) non\ (1,5) [tion\]		
37	ADL	<((clavardage)) tion>		
		[MH\] (0,2)		
		MH\ (0,3)	[TION\]	
38	ALN		[pronon] (0,3)	
		prononation\ (1,3) prononc-		
		[((rire))] "oh: g[o:d"]		
39	ADL	[bien/] (1,2) [essaye de le dire] doucement		Haussement de sourcils et pouce levé
40	ALN	((rire))		
41	ADL	et dis-le tout doucement\ (0,3) PRO/ (.) NON/ (0,2) CIA/ (0,2) TION\ (0,2)		
42	ALN	pronona:tion\ (1,4) ah		
43	ADL	oui/ (0,5) c'est MIEUX qu'au début déjà c'est bien		Changement de position de la tête par syllabe, de droite à gauche de l'écran
44	ALN	[j'espère (.) ((rire))]		
45	ADL	[((rire)) il est dur à] dire hein/ (1,2) [donc ça ça t'aide/ ça vous aide/]		
46	ALN	[ah je suis f-/ ah je suis] fatiguée (c'est) (1,0)		
47	ADL	oui: je comprends\ vous travaillez beaucoup\ (0,4) ((rire))		

- 45 Dans cette séquence, Adèle utilise une variété de ressources multimodales à sa disposition pour aider l'apprenante à se former une représentation phonologique de l'item lexical « prononciation ». Elle commence par répéter la forme correcte lentement avec le menton levé et la bouche centrée devant la webcam. Après un mot

d'empathie, elle répète le mot syllabe par syllabe. Quand l'apprenante est toujours incapable de le prononcer correctement, Adèle répète sa stratégie de diviser le mot en syllabes mais augmente la saillance multimodale en redoublant par écrit chacune des syllabes. En plus de ce renforcement par écrit, Adèle crée un lien mental entre « non » (deuxième syllabe du mot) et « non comme non » accompagné d'un geste emblématique voulant dire « non ». Elle maintient sa bouche ouverte lors de la prononciation de la troisième syllabe « cia » pour la rendre encore plus saillante. Ensuite, elle verbalise sa stratégie de séparation syllabique en disant « dis une par une » en faisant un geste métaphorique, ses deux doigts tenant des syllabes imaginaires (ligne 31). Elle répète encore une fois le mot « prononciation » par syllabes, en maintenant devant la webcam les gestes indiquant le numéro de syllabe (lignes 32–37). Elle félicite l'apprenante de son effort et répète une dernière fois le mot, cette fois-ci en changeant la position de sa tête pour marquer les quatre syllabes (ligne 41). Avec patience et savoir-faire multimodal, Adèle utilise sa voix, le clavardage, son visage, ses mains, et la position de sa tête pour illustrer la forme orale du mot.

Focaliser sur le sens d'un item lexical

- 46 Il existe plusieurs façons d'expliquer le sens d'un item lexical : différents types de définitions dans la langue cible, la traduction, et l'explication ostensive/visuelle (Fasel Lauzon, 2014 ; Flowerdew, 1992 ; Nation, 2013). Ce qui nous préoccupe ici, c'est l'utilisation de gestes pour la focalisation sur le sens. En effet, on retrouve dans le corpus l'utilisation de gestes, mouvements de tête et mimiques faciales par les apprentis-enseignants pour se focaliser sur le sens dans 48 séquences d'explication lexicale sur 295. Ces gestes sont utilisés pour illustrer des concepts concrets et abstraits, et pour faciliter la recherche lexicale collaborative.
- 47 Dans notre corpus, nous avons repéré plusieurs gestes qui décrivent des mots aux référents concrets : un mouvement circulaire au-dessus de la tête pour décrire le mot « couronne », la main en forme de récipient pour décrire le mot « boisson », etc. Une méta-analyse de Hostetter (2011) a montré que les gestes sont particulièrement aptes à aider la compréhension de choses concrètes. Nous souhaiterions nous attarder ici sur les sens métaphoriques. Hostetter (2011) rapporte que les gestes métaphoriques ont moins d'effet sur la compréhension que les gestes iconiques. Dans une étude expérimentale, Macedonia & Knösche (2011) ont trouvé aussi que les mots concrets étaient mieux mémorisés que les mots abstraits, mais que les gestes avaient plus d'impact sur la mémorisation de mots abstraits que sur la mémorisation de mots concrets. Ils postulent que c'est parce que les mots concrets ont déjà une connotation incarnée dans le cerveau et pourraient faire l'économie d'un geste. Un mot abstrait, en revanche, pourrait bénéficier d'un geste qui devient le composant incarné du mot. Souvent, un item lexical comporte un sens littéral et un sens figuré (comme le mot « lien » que nous avons vu dans le premier exemple). Dans l'exemple suivant, Emilie explique les sens littéral et figuré du mot « arrondir » dans l'expression « arrondir ses fins de mois ».

Exemple 4 : Mise en lien des sens littéral et figuré

Transcription de la parole			Image d'Emilie	Phén. observés
1	EMI	alors\ arrondir:/ euh: c'est quand tu prends quelque chose: un objet/ (0,5) et/ tu lui donne une forme ronde\ (1,5) MAIS/ (0,2) euh:\ là/ c'est pas: hm: c'est au sens figuré si tu veux donc tu le trouveras pas si je te donne le sens euh\ (0,7) le sens littéral (0,5) le sens euh de base\ (.) du verbe arrondir\ (0,7) en fait/ (0,4) arrondir les fins de mois/ (0,6) ça veut dire gagner un peu plus d'argent/ que ce que tu gagnes déjà\ (1,2)		Geste iconique représentant un objet rond
		donc\ si par exemple tu touches une bourse/ tu comprends/ vous comprenez bourse/ (0,6)		Visage d'incertitude
2	FIN	oui (0,7)		
3	EMI	voilà\ (0,7) vous touchez une bourse/ par exemple vos parents vous aident/ (0,4) mais vous voulez (.) [ARROND]IR les fins de mois (0,2)&		
4	FIN	[ah: oui\]		
5	EMI	&donc/ vous allez travailler\ (0,4) ça veut dire/ que vous voulez gagner/ (.) PLUS d'argent\ (1,9) parce que en fait/ en france/ (.) euh: je je ne sais pas comme c'est/ en irlande/ (0,5) mais en france/ (.) on reçoit le salaire et le paiement/ (.) à la fin du mois\ (0,5) donc/ (.) quand on ARRONDIT la fin du mois/ (0,6) on est sensé rajouter UN ZÉRO (.) d'accord/		Geste emblématique représentant un zéro
		don[c avoir] &		
6	ADN	[hm\]		
7	EMI	&une paye un peu plus grosse\ (.) un peu plus im[portante\]		
8	ADN	[oui oui oui]		

		(0,5)	
9	EMI	c'est pour ça/ qu'on dit (.) arrondir les fins de mois en fait/ (0,5) on arrondit les chiffres	
		(0,6) sur le sur	[le]&
10	ADN		[hm\]
11	EMI	&chèque ou/ euh sur la carte bancaire\ (.) vous voyez/ (1,5) c'est une expression (0,3) voilà\ assez courante\ (0,3) voilà\ assez courante\	

- 48 Emilie commence son explication en donnant une définition lexicographique et décontextualisée du verbe « arrondir ». Afin de renforcer le transfert du sens, elle maintient un geste iconique qui représente un objet rond devant la webcam et appuie sa définition d'une pause longue de 1 500 millisecondes. Le geste nous renseigne sur la forme de l'objet avant la parole. Après, Emilie dit que le sens littéral ne va pas permettre aux apprenants de deviner le sens figuré.⁸ Pour expliquer le sens figuré, Emilie utilise un nouveau geste, cette fois-ci un geste qui représente un zéro qui serait ajouté au nombre indiqué sur une fiche de paie. Le geste emblématique est maintenu et sert de guide visuel qui renforce la transmission du sens figuré. Ce geste est assez remarquable chez cette apprentie-enseignante qui se positionne d'habitude devant la webcam de façon à ce que ses gestes soient invisibles. L'effort qu'elle fait pour rendre son geste visible et de le maintenir est donc digne d'intérêt. Malgré quelques critiques didactiques possibles (utilisation d'un vocabulaire quelque peu complexe, introduction de trop de nouveaux items lexicaux, acceptation de réponses minimales des apprenants de type « oui » sans vérification de compréhension), les gestes produits dans cette séquence sont représentatifs de leur potentiel sémiotique par visioconférence.

Focaliser sur l'usage d'un item lexical

- 49 La facette « usage » d'un item lexical comprend les fonctions grammaticales, les collocations, et les contraintes liées à l'usage comme le registre, la fréquence, et la connotation (Nation, 2013). L'on peut aussi y inclure la charge culturelle partagée (Galisson, 1991). L'usage des items lexicaux est assez peu traité dans notre corpus (seulement 43 séquences d'explication lexicale sur 295 abordent l'usage), et les gestes ne sont utilisés pour aborder cette facette des items lexicaux que dans 5 séquences sur 295. Dans l'exemple 4 ci-dessus, Emilie parle de la fréquence de l'expression « arrondir ses fins de mois » en disant que c'est une expression « assez courante » (ligne 11), et dans l'exemple suivant, Adèle emploie un geste pour aborder la fréquence du mot « bobo ».

Exemple 5 : Explication de l'usage du mot « bobo »

Transcription de la parole	Image d'Adèle	Phén. observés
----------------------------	---------------	----------------

33	ADL	ça vous servira beaucoup en france/ parce qu[e: les gens utilisent]&			Haussement de sourcils et geste vers le haut
34	ALN		[okay le x marché/]		
35	ADL	&beaucoup (1,0)			
		ils disent beaucoup ah c'est un bobo euh\ (2,2) ça ça ça			Geste emblématique et regard hors champ

- 50 Adèle aborde l'usage du mot-valise en parlant de sa fréquence. Son geste métaphorique à la ligne 35 correspond temporellement et sémantiquement au mot « beaucoup ». Par ce geste et haussement de sourcils, elle renforce le mot « beaucoup » afin d'insister sur l'utilité et l'importance de l'item lexical qui fait l'objet de la séquence d'explication. L'insistance sur l'utilité du mot est susceptible d'encourager sa mémorisation, et éventuellement son emploi par les apprenants.

Impliquer les apprenants dans une séquence d'explication

- 51 La discussion et la participation active aident l'apprentissage du vocabulaire car quand on discute d'un mot, on le traite plus profondément que quand on le reconnaît comme la bonne réponse sur une liste (Stahl & Vancil, 1986). Selon Laufer & Hulstijn (2001), la *charge d'implication* serait facteur de trois choses : le besoin, la recherche et l'évaluation de l'item lexical. Quand c'est l'apprenant qui demande une explication, le *besoin* de le connaître est plus fort, favorisant l'acquisition. Quand l'apprenant recherche la forme ou le sens d'un mot, le côté *recherche* est privilégié, et quand l'apprenant doit juger de l'acceptabilité d'un mot ou d'une définition, le côté *évaluation* est soutenu. L'acquisition est favorisée quand les participants sont invités à participer activement dans la recherche lexicale (Brouwer, 2003), et quand ce sont les apprenants qui expliquent ou demandent des explications.
- 52 Comme le rappelle Fasel Lauzon (2014 : 30), l'explication est *non seulement un lieu de construction de connaissances, mais également un lieu de vérification de sa propre compréhension (pour la personne qui explique)*. Quand un apprenant demande la forme d'un item lexical, l'apprenti-enseignant doit être capable dans un premier temps d'avoir une compréhension adéquate de ce qui est recherché. Si l'apprenant n'est pas à même de faire du bricolage lexical (Lüdi, 1994) ou de décrire l'item lexical recherché en français, et si une traduction est inutile faute de compétences en langues communes, un geste est alors susceptible d'être la meilleure façon de signaler une panne lexicale ou de décrire ce que l'on cherche à dire (Kupetz, 2011). Les gestes de l'apprenant peuvent contribuer à signaler à l'enseignant ce qui est compris et ce qui ne l'est pas (Goldin-Meadow, 2003). L'enseignant pourra alors diagnostiquer son état de compréhension.

- 53 Dans notre corpus, les apprentis-enseignants impliquent les apprenants de plusieurs façons en effectuant des rappels productifs et réceptifs, en leur faisant remarquer des éléments problématiques, en leur faisant évaluer des items lexicaux, et en leur faisant expliquer, parmi d'autres. Dans l'extrait qui suit, Mélissa est en train de clôturer la séance quand Ana lui demande la traduction d'un mot inconnu. Les gestes permettent à l'apprenante de demander un *mot juste* et à l'apprentie-enseignante de vérifier sa propre compréhension de ce qui est demandé.

Exemple 6 : Utilisation de gestes pour trouver un mot juste

Transcription de la parole				Image de Mélissa ou d'Ana	Phén. observés
1	MEL	euh bon	[ben voilà/ ben&		Regards caméra
2	ANA		[oh et:]		
3	MEL	&écoutez (.) oui/ (0,4)			
4	ANA	je peux te poser une question/			
5	MEL	oui/ (0,8)			
6	ANA	oui euh comment est- [ce qu'on] d[it "YAW"]NING"\			
7	MEL	[oui/] (0,3) [°((rires))°]			
		(2,6) comment dit quoi/ (0,8)			Visage d'incertitude
8	ANA	"yawning"/ (1,9)			
9	MEL	tu tu peux m'écrire le mot/ peut-être\ (4,3)			
10	ANA	<((clavardage)) yawning> (0,6) "I don't even know if it's spelled like that" (1,7) hum (2,4)			
11	MEL	c'est un mot (0,2) anglais/ (0,9)			Ana mime l'acte de bâiller.

TIPA. Travaux interdisciplinaires sur la parole et le langage, 36 | 2020

12	ANA	ouais\ (0,3)		
13	MEL	AH		Mélissa mime l'acte de bâiller.
14	ANA	x euh		
15	MEL	bâiller (0,7)		
16	ANA	bâiller\ [comme ça/]		
17	MEL	qu- c'est quand tu fais (1,1)		
		[comme ça/]		
18	ANA	[ouais ouais] (0,6)		
19	MEL	c'est quand tu fais (1,1) okay (.)[((rires))]&		Mélissa répète son premier geste et fait des étirements tout de suite après.
20	ANA	[ouais ouais ouais\]		
21	AJL	[AH:\]		
22	MEL	<((en riant)) (alors)> [((rires))] [c'est le mot/]&		
23	ALJ	[((rires))] [merci] ((rires))		
24	MEL	&ah ben peut- tu le connais/ alejandra/ (1,3)		
25	ALJ	[non je ne connais]&		
26	MEL	[tu connais le mot alejandra/]		
27	ALJ	&mais mais ["you"] expliques&		
28	MEL	[non]		
29	ALJ	&((rires)) (0,7) "you": (0,5) e[uh:]		
30	MEL	[ah (.)] allô/		

31	ALJ	tu m'expliques euh (0,3) ((rires)) (0,5)		
32	MEL	°d'accord° (0,4) c'est <((clavardage)) bailler> c'est le mot/ (.) bâiller\		

- 54 Ana demande la traduction d'un mot anglais que Mélissa ne comprend pas (ligne 6). Mélissa lui demande de l'écrire (ligne 9), mais cette stratégie de clarification ne permet toujours pas à Mélissa d'atteindre la compréhension nécessaire pour donner le mot juste. A la ligne 11, Ana fait un geste pour illustrer le sens du mot recherché, et Mélissa, aussi par un geste, vérifie sa propre compréhension (ligne 17). Après avoir vérifié sa propre compréhension, Mélissa donne le mot juste à l'oral et à l'écrit (ligne 32).
- 55 Ce qui est digne d'intérêt ici, c'est la façon dont le geste est employé par Mélissa, car son geste à la ligne 19 remplace totalement la parole orale. On peut le considérer comme un autonome sémantique à l'instar de Rey-Debove (1997 : 84) qui considère qu'on peut imaginer un discours oral interrompu par un geste qui cite un item d'une gestuelle, et qu'on essaiera de transcrire artificiellement ainsi : /Il a fait (haussement d'épaules) sans relever l'allusion/. Alors que l'usage de gestes constitue une interruption pour Rey-Debove, nous rejoignons Sindoni (2013) lorsqu'elle argumente que toutes les ressources sémiotiques ont *a priori* le même statut et la même importance. Nous avons donc affaire au *mode-switching* (Sindoni, 2013) (alternance de modes ou alternance modale), plutôt qu'à une interruption.
- 56 Aussi digne d'intérêt est la densité modale qu'obtiennent les gestes de Mélissa aux lignes 17 et 19. Son début de phrase « c'est quand » suit au pied de la lettre le *template* pour une mise en situation (Fasel Lauzon, 2014), mais ce qui suit est un geste au lieu d'une explication linguistique. La communication sonore s'arrête complètement lors du geste, ce qui fait que sans ce geste, la communication n'existerait point. Voilà la définition du degré maximal de densité modale (Norris, 2004). Mélissa semble être bien consciente de son usage de gestes car son regard vers l'écran indique qu'elle vérifie le cadrage de ses gestes. Les gestes de Mélissa sont indissociables de sa parole, ce qui ne constitue pas une interruption, loin s'en faut. On assiste à un usage tout à fait naturel du corps de l'enseignante pour compléter sa parole, et la webcam facilite la transmission du sens. L'usage de gestes permet à Ana de se faire aider et à Mélissa de vérifier qu'elle donne l'item lexical recherché. La webcam permet aux interactants de changer de mode facilement et de manière fluide, quitte à produire des gestes autonymiques.
- 57 Enfin, un geste peut être utilisé pour encourager l'apprenante à prendre en charge l'explication et pour gérer l'utilisation des ressources sémiotiques. Adèle utilise ainsi un geste pour encourager Catriona à donner une explication du mot recherché à l'oral plutôt qu'à l'écrit.

Exemple 7 : Gestion des ressources multimodales à l'aide d'un geste

Transcription de la parole	Image d'Adèle	Phén. observés
----------------------------	---------------	----------------

1	CTR	et mon père il euh (0,2) il travaille dans l'école mais avec le (0,4) le travaux (oui oui) (0,6) je ne sais pas le mot en français/ (.) mais hum\		Hochement de tête
2	ADL	alors\ (0,4) décris-moi/ et je vais t'aider à trouver le mot\ (1,1)		Rapprochement de l'écran et sourire
3	CTR	hum (5,0)		
4	ADL	ou décris en parlant (0,5) en parlant\ (0,5)		Geste circulaire représentant l'acte de parler
5	CTR	"oh ok" (0,4) hum: il (0,5) il fait des répa- les les réparations dans l'école hum si: (0,8) si: quelque chose est cassée: (.) il le (0,3)		Rapprochement de l'écran et sourire
6	ADL	OUI:\ (.)		Regard pensif
		[euh:]		
7	CTR	[il le répare:]		
		(0,3)		
8	ADL	alors euh peut-être on appellerait ça:/ hu:m\ (1,0) agent/ de maintenance\ (1,1)		Sourire
9	CTR	ah oui de maintenance oui\ (0,3)		
10	ADL	<((clavardage)) agent de maintenance> un (0,6)		Sourire
11	CTR	oui		

12	ADL	agent de maintenance\ (0,3) [[((rires))]]		
13	CTR	[oui mais il travaille] dans un école (0,2) oui (0,5)		
14	ADL	très bien/ (0,3) donc euh tes deux parents sont à l'école\		

- 58 Avant même la panne lexicale, Adèle encourage son apprenante à parler au moyen de hochements de tête servant de régulateurs d'écoute (ligne 1). Quand Catriona commente la panne lexicale qu'elle rencontre, Adèle lui demande de décrire le mot recherché (ligne 2), ce qui la met en position haute dans l'interaction. A la suite de sa demande, Adèle sourit et se rapproche à l'écran, ce qui implique Catriona en lui donnant l'impression qu'elle est écoutée (lignes 2-3). S'ensuit alors une longue période de silence pendant laquelle Catriona regarde son clavier. Puis Adèle l'invite à reprendre la parole. Cette consigne méta-explicative et méta-sémiotique se réalise à l'aide d'un geste (ligne 4). Quand Catriona décrit le travail de son père, Adèle se penche à nouveau vers l'écran et sourit (ligne 5). Après l'explication de Catriona, Adèle réfléchit (lignes 6-8) et donne le mot oralement et par écrit (lignes 8, 10, 12). Elle utilise le marqueur de mitigation « peut-être » (ligne 8) quand elle donne le mot au lieu de le donner sur un ton péremptoire. Adèle termine la séquence par un commentaire qui témoigne de l'attention qu'elle porte à la parole de Catriona.
- 59 L'implication effectuée par Adèle se fait à trois niveaux. Au niveau cognitif, Adèle implique son apprenante en la mettant en position d'explicatrice. Elle ouvre un espace de collaboration dans lequel l'apprenante peut expliquer le mot recherché. Cette ouverture se fait oralement (« décris-moi », ligne 2) et visuellement, avec un rapprochement à l'écran et un sourire. Le microphone et la webcam sont utilisés ici de manière complémentaire. Puis, au niveau sémiotique, Adèle implique l'apprenante en lui demandant d'utiliser le microphone plutôt que le clavier pour s'exprimer (« décris en parlant », ligne 4). Cette demande se fait à l'aide d'un geste. Là aussi, deux modes sémiotiques sont combinés pour réaliser une action. L'implication permet à l'apprenante à la fois de s'entraîner oralement, en expliquant le travail que fait son père, et de trouver le mot juste recherché. Enfin, au niveau socioaffectif, Adèle encourage son apprenante en se rapprochant à l'écran et en souriant quand cette dernière explique ce que fait son père.
- 60 Ces deux derniers exemples montrent que lors des séquences d'explication lexicale, les gestes servent non seulement à expliquer, mais aussi à impliquer les apprenants.

Discussion et conclusion

- 61 Cet article s'est donné pour objectif d'analyser l'apport des gestes à des séquences d'explication lexicale par visioconférence. Bien que ces mêmes apprentis-enseignants ne gestualisent pas plus lors des séquences d'incompréhension (Holt, Tellier, & Guichon, 2015), et bien que la plupart des gestes soient produits hors champ (Guichon & Wigham, 2016), ces apprentis-enseignants semblent toutefois être conscients de l'utilité

de leurs gestes lors des séquences d'explication lexicale (cf. exemple 1). Toute analyse quantitative mérite donc d'être complétée par des analyses qualitatives comme celles-ci. Ce contexte particulier fait que tout geste rendu visible devant la webcam est digne d'intérêt, surtout ceux qui atteignent un statut autonymique comme nous l'avons vu dans l'exemple 6.

- 62 Les quelques exemples sélectionnés pour cet article ont montré certaines fonctions remplies par les gestes pédagogiques lors des séquences d'explication lexicale par visioconférence. Nous avons démontré notamment que les gestes peuvent contribuer à l'explication de toutes les grandes facettes de la connaissance d'un item lexical : la forme, le sens, et l'usage. Nous avons aussi montré des exemples où les gestes ont servi à impliquer l'apprenant dans la séquence explicative. Parmi les 295 séquences d'explication lexicale que nous avons repérées, la webcam a été utilisée dans 53 séquences pour se focaliser sur la forme, dans 48 séquences pour se focaliser sur le sens, dans 5 séquences pour se focaliser sur l'usage, et dans 24 séquences pour impliquer les apprenants. Ces gestes sont intéressants non seulement pour les fonctions didactiques qu'ils remplissent, mais aussi pour la façon dont ils sont employés.
- 63 Toute catégorisation n'est qu'un dispositif temporaire pour des études locales (Kendon, 2004), et les gestes doivent être annotés par rapport à la tâche en cours (Goldin-Meadow, 2003). La classification fonctionnelle des gestes par rapport à la facette de la connaissance d'un item lexical en cours d'explication a donc été adaptée au présent corpus car les données ont été abordées par cette entrée didactique spécifique. Nous espérons avoir élargi la description des gestes pédagogiques dans ce contexte particulier.
- 64 Alors que nous nous sommes focalisés sur les gestes, il faut se rappeler que les gestes font partie d'une palette plus large de ressources sémiotiques disponibles en contexte visiophonique. Le clavier, l'envoi d'images, les explications purement verbales, et les expressions faciales ont aussi contribué aux explications lexicales, même si leurs contributions n'ont pas été élaborées. Les exemples que nous avons choisis pour cet article montrent la nécessité de recourir à des approches multimodales pour de futures recherches.
- 65 Nos données ne nous ont pas permis de mesurer *via* des post tests la mémorisation des items lexicaux expliqués. Des études futures avec des données plus contrôlées permettraient de cerner les pratiques multimodales et gestuelles les plus propices à l'acquisition du vocabulaire par visioconférence poste à poste. Une autre limite concerne la reproduction des gestes par les apprenants. S'il est important de sensibiliser les futurs enseignants à l'importance de leurs gestes (Tellier & Cadet, 2014), surtout dans ce contexte particulier, il est aussi important de les sensibiliser à l'importance de la reproduction des gestes par les apprenants (Macedonia *et al.*, 2019). Enfin, nous n'avons pas pu mesurer le regard des apprenants par l'oculométrie. Cette technologie mérite d'être exploitée par de futures études afin de savoir avec certitude si et à quels moments les apprenants regardent les gestes produits par les enseignants lors des échanges pédagogiques par visioconférence poste à poste.

Remerciements

- 66 Nous remercions le LABEX ASLAN (ANR-10-LABX-0081) de l'Université de Lyon pour son soutien financier dans le cadre du programme "Investissements d'Avenir" (ANR-11-IDEX-0007) de l'Etat Français géré par l'Agence Nationale de la Recherche (ANR).

BIBLIOGRAPHIE

- Alibali, M. W., Heath, D. C., & H. J. Myers (2001) Effects of Visibility between Speaker and Listener on Gesture Production: Some Gestures Are Meant to Be Seen, *Journal of Memory and Language*, 44, p. 169–188.
- Azaoui, B. (2017) Faire face aux imprévus techniques, in Guichon, N. & M. Tellier (eds) *Enseigner l'oral en ligne : Une approche multimodale*, Paris: Didier, p. 184–214.
- Azaoui, B. (2019) Ce que les élèves voient et disent du corps de leur enseignant: Analyse multimodale de leur discours, in Rivière, V. & N. Blanc (eds) *Observer l'activité multimodale en situations éducatives: Circulations entre recherche et formation*, Lyon: ENS Editions.
- Baldry, A. & P. Thibault (2006) *Multimodal Transcription and Text Analysis*, London; Oakville, CT: Equinox Publishing.
- Bange, P. (1992) À propos de la communication et de l'apprentissage de L2 (notamment dans ses formes institutionnelles), *Acquisition et Interaction En Langue Étrangère*, 1, p. 53–85.
- Barcroft, J. (2003) Effects of questions about word meaning during L2 Spanish lexical learning, *The Modern Language Journal*, 87, p. 546–561.
- Beattie, G. & H. Shovelton (1999) Do iconic hand gestures really contribute anything to the semantic information conveyed by speech? An experimental investigation, *Semiotica*, 123, p. 1–30.
- Blake, R. (2000) Computer mediated communication: A window on L2 Spanish interlanguage, *Language Learning & Technology*, 4, p. 111–125.
- Brouwer, C. E. (2003) Word Searches in NNS: NS Interaction: Opportunities for Language Learning?, *Modern Language Journal*, 87, p. 534–545.
- Cappellini, M., & C. Combe (2017) Analyser des compétences techno-sémio-pédagogiques d'apprentis tuteurs dans différents environnements numériques: Résultats d'une étude exploratoire, *Alsic*, 20.
- Chanier, T. & M. Ciekanski (2010) Utilité du partage des corpus pour l'analyse des interactions en ligne en situation d'apprentissage: Un exemple d'approche méthodologique autour d'une base de corpus d'apprentissage, *Alsic*, 13.
- Cicurel, F. (2011) *Les interactions dans l'enseignement des langues*, Paris: Didier.
- Cohen, R. L. (1981) On the generality of some memory laws, *Scandinavian Journal of Psychology*, 22, p. 267–281.

Conseil de l'Europe (2001) *Cadre européen commun de référence pour les langues: Apprendre, enseigner, évaluer*, Conseil de l'Europe/Didier.

Cosnier, J. & C. Develotte (2011) Le face à face en ligne, approche éthologique, in Develotte, C., Kern, R. & M.-N. Lamy (eds) *Décrire la conversation en ligne, Le face à face distanciel*, Lyon: ENS Éditions, p. 27–50.

De Chanay, H. (2011) La construction de l'éthos dans les conversations en ligne, in *Décrire la conversation en ligne, Le face à face distanciel*, Lyon: ENS Éditions, p. 145–172.

De Gaulmyn, M.-M. (1991) Expliquer des explications, in Dausendschön-Gay, U., Gülich, E. & U. Krafft (eds) *Linguistische Interaktionsanalysen*, Tübingen: Max Niemeyer, p. 279–314.

De Pietro, J.-F., Matthey, M. & B. Py (1989) Acquisition et contrat didactique: Les séquences potentiellement acquisitionnelles dans la conversation exolingue, *Actes Du Troisième Colloque Régional de Linguistique*, p. 99–119.

Develotte, C. & S. Drissi (2013) Face à face distanciel et didactique des langues, *Le Français Dans Le Monde. Recherches et Applications*, p. 54–63.

Develotte, C., Guichon, N. & R. Kern (2008) “Allo Berkeley? Ici Lyon... Vous nous voyez bien ?” Étude d'un dispositif de formation en ligne synchrone franco-américain à travers les discours de ses usagers, *Alsic*, 11, p. 129–156.

Develotte, C., Guichon, N. & C. Vincent (2010) The use of webcam for teaching a foreign language in a desktop videoconferencing environment, *ReCALL*, 22, p. 293–312.

Develotte, C., Kern, R. & M.-N. Lamy (eds) (2011) *Décrire la conversation en ligne: La face à face distanciel*, Lyon: ENS.

Develotte, C. & F. Mangelot (2007) Discontinuités didactiques et langagières au sein d'un dispositif pédagogique en ligne, *Glottopol*, 10, p. 127–144.

Ekman, P. & W. V. Friesen (1969) The repertoire of nonverbal behavior: Categories, origins, and coding, *Semiotica*, 1, p. 49–98.

Ellis, R., Tanaka, Y. & A. Yamazaki (1994) Classroom interaction, comprehension, and the acquisition of L2 word meanings, *Language Learning*, 44, p. 449–491.

Fasel Lauzon, V. (2008) Interactions et apprentissages dans des séquences d'explication de vocabulaire, *Travaux Neuchâtelois de Linguistique*, 48, p. 83–104.

Fasel Lauzon, V. (2014) *Comprendre et apprendre dans l'interaction: Les séquences d'explication en classe de français langue seconde*, Bern: Peter Lang.

Flowerdew, J. (1992) Definitions in science lectures, *Applied Linguistics*, 13, p. 202–221.

Folse, K. S. (2004) *Vocabulary myths: Applying second language research to classroom teaching*, Ann Arbor: The University of Michigan Press.

Galisson, R. (1991) *De la langue à la culture par les mots*, Paris: CLE International.

García, M. M. (2008) De l'importance des éléments suprasegmentaux dans l'apprentissage du lexique en L2, *Travaux de Didactique Du FLE*, 59, p. 13–23.

Goldin-Meadow, S. (2003) *How our hands help us think*, Cambridge, MA: Harvard University Press.

Groot, P. (2000) Computer assisted second language vocabulary acquisition, *Language Learning & Technology*, 4, p. 60–81.

- Grossmann, F. (2011) Didactique du lexique: État des lieux et nouvelles orientations, *Pratiques. Linguistique, Littérature, Didactique*, 149/150, p. 163–183.
- Groupe ICOR (2007) *Convention ICOR* [en ligne], Lyon: Université de Lyon. Disponible sur : http://icar.cnrs.fr/projets/corinte/documents/2013_Conv_ICOR_250313.pdf (consulté le 9 novembre 2020).
- Guichon, N. (2013) Une approche sémio-didactique de l'activité de l'enseignement de langue en ligne: Réflexions méthodologiques, *Éducation & Didactique*, 7, p. 101–116.
- Guichon, N. (2017a) Se construire une présence pédagogique en ligne, in Guichon, N. & M. Tellier (eds) *Enseigner l'oral en ligne: Une approche multimodale*, Paris: Didier, p. 29–58.
- Guichon, N. (2017b) Sharing a multimodal corpus to study webcam-mediated language teaching, *Language Learning & Technology*, 21, p. 55–74.
- Guichon, N., Bétrancourt, M. & Y. Prié (2012) Managing written and oral negative feedback in a synchronous online teaching situation, *Computer-Assisted Language Learning*, 25, p. 181–197.
- Guichon, N., & S. Drissi (2008) Tutorat de langue par visioconférence: Comment former aux régulations pédagogiques ?, *Les Cahiers de l'Acedle*, 5, p. 1–32.
- Guichon, N. & M. Tellier (eds) (2017) *Enseigner l'oral en ligne: Une approche multimodale*, Paris: Didier.
- Guichon, N. & C. R. Wigham (2016) A semiotic perspective on webconferencing-supported language teaching, *ReCALL*, 28, p. 62–82.
- Gülich, E. (1990) L'approche ethnométhodologique dans l'analyse du français parlé, in Charolles, M., Fisher, S. & J. Jayez (eds) *Le discours: Représentations et interprétations*, Nancy: Presses Universitaires de Nancy, p. 71–109.
- Gülich, E. (1991) Pour une ethnométhodologie linguistique: Description de séquences conversationnelles explicatives, in Dausendschön-Gay, U, Gülich, E. & U. Krafft (eds) *Linguistische Interaktionsanalysen*, Tübingen: Max Niemeyer, p. 325–364.
- Hatch, E. M. (1978) Discourse analysis and second language acquisition, in Hatch, E. M. (ed.) *Second language acquisition: A book of readings*, Rowley, MA: Newbury House, p. 401–435.
- Hill, M. & B. Laufer (2003) Type of task, time-on-task and electronic dictionaries in incidental vocabulary acquisition, *IRAL - International Review of Applied Linguistics in Language Teaching*, 41, p. 87–106.
- Holt, B. & M. Tellier (2017) Conduire des explications lexicales, in Guichon, N. & M. Tellier (eds) *Enseigner l'oral en ligne*, Paris: Didier, p. 59–90.
- Holt, B., Tellier, M. & N. Guichon (2015) The use of teaching gestures in an online multimodal environment: The case of incomprehension sequences, *Gesture and Speech in Interaction 4th Edition*, Nantes, p. 149–154.
- Hostetter, A. B. (2011) When do gestures communicate? A meta-analysis, *Psychological Bulletin*, 137, p. 297–315.
- Huckin, T. & J. Coady (1999) Incidental vocabulary acquisition in a second language: A review, *Studies in Second Language Acquisition*, 21, p. 181–193.
- Jewitt, C. (2014) An introduction to multimodality, in Jewitt, C. (ed.) *The Routledge Handbook of Multimodal Analysis*, 2nd ed., London: Routledge, p. 15–30.

- Joe, A. (1995) Text-based tasks and incidental vocabulary learning, *Second Language Research*, 11, p. 149–158.
- Joe, A. (1998) What effects do text-based tasks promoting generation have on incidental vocabulary acquisition?, *Applied Linguistics*, 19, p. 357–377.
- Jones, R. H. (2004) The Problem of Context in Computer-Mediated Communication, in Le Vine, P. & R. Scollon (eds) *Discourse and technology: Multimodal discourse analysis*, Washington, D.C.: Georgetown University Press, p. 20–33.
- Kelly, S. D., Özyürek, A. & E. Maris (2010) Two Sides of the Same Coin: Speech and Gesture Mutually Interact to Enhance Comprehension, *Psychological Science*, 21, p. 260–267.
- Kelly, S. D., McDevitt, T. & M. Esch (2009) Brief training with co-speech gesture lends a hand to word learning in a foreign language, *Language and Cognitive Processes*, 24, p. 313–334.
- Kendon, A. (2004) *Gesture: Visible action as utterance*, Cambridge: Cambridge University Press.
- Kida, T. (2014) *Geste et appropriation: Acculturation non verbale des étrangers*, Aix-en-Provence: Presses Universitaires de Provence.
- Kupetz, M. (2011) Multimodal resources in students' explanations in CLIL interaction, *Novitas-Royal*, 5, p. 121–142.
- Laufer, B. (1994) Appropriation du vocabulaire: Mots faciles, mots difficiles, mots impossibles, *Acquisition et Interaction En Langue Etrangère*, 3, p. 97–113.
- Laufer, B., & J. Hulstijn (2001) Incidental vocabulary acquisition in a second language: The construct of task-induced involvement, *Applied Linguistics*, 21, p. 1–26.
- Lazaraton, A. (2004) Gesture and speech in the vocabulary explanations of one ESL teacher: A microanalytic inquiry, *Language Learning*, 54, p. 79–117.
- Long, M. (1981) Input, interaction, and second-language acquisition, *Native Language and Foreign Language Acquisition*, 379, p. 259–278.
- Long, M. (1983) Native speaker/non-native speaker conversation and the negotiation of comprehensible input, *Applied Linguistics*, 4, p. 126–141.
- Long, M. H. (1991) Focus on Form: A Design Feature in Language Teaching Methodology, in De Bot, K., Ginsberg, R. B. & C. Kramsch (eds), *Foreign Language Research in Cross-Cultural Perspective*, Amsterdam/Philadelphia: John Benjamins Publishing Company, p. 39–52.
- Lüdi, G. (1991) Construire ensemble les mots pour le dire. A propos de l'origine discursive des connaissances lexicales, in Dausendschön-Gay, U., Gülich, E. & U. Krafft (eds), *Linguistische Interaktionsanalysen*, Tübingen: Max Niemeyer, p. 193–224.
- Lüdi, G. (1994) Dénomination médiate et bricolage lexical en situation exolingue, *Acquisition et Interaction En Langue Étrangère*, 3, p. 115–146.
- Macedonia, M. (2014) Bringing back the body into the mind: Gestures enhance word learning in foreign language, *Frontiers in Psychology*, 5.
- Macedonia, M. & W. Klimesch (2014) Long-Term Effects of Gestures on Memory for Foreign Language Words Trained in the Classroom: Long-Term Effects of Gestures on Memory for Words, *Mind, Brain, and Education*, 8, p. 74–88.
- Macedonia, M. & T. R. Knösche (2011) Body in Mind: How Gestures Empower Foreign Language Learning, *Mind, Brain, and Education*, 5, p. 196–211.

- Macedonia, M., Müller, K. & A. Friederici (2011) The impact of iconic gestures on foreign language word learning and its neural substrate, *Human Brain Mapping*, 32, p. 982–998.
- Macedonia, M., Repetto, C., Ischebeck, A. & K. Mueller (2019) Depth of Encoding Through Observed Gestures in Foreign Language Word Learning, *Frontiers in Psychology*, 10.
- Macedonia, M. & K. von Kriegstein (2012) Gestures enhance foreign language learning. *Biolinguistics*, 6, p. 393–416.
- Marijanović, V., Billières, M. & N. Panissal (2008) L'identification du mot écrit en FLE auprès des apprenants débutants, *Travaux de Didactique Du FLE*, 59, p. 25–33.
- McDonough, K. & W. Sunitham (2009) Collaborative dialogue between Thai EFL learners during self-access computer activities, *TESOL Quarterly*, 43, p. 231–254.
- McNeill, D. (1992) *Hand and Mind*, Chicago: The University of Chicago Press.
- Mol, L., Krahmer, E., Maes, A. & M. Swerts (2011) Seeing and Being Seen: The Effects on Gesture Production, *Journal of Computer-Mediated Communication*, 17, p. 77–100.
- Nation, I. S. P. (2013) *Learning Vocabulary in Another Language*, 2nd ed., Cambridge: Cambridge University Press.
- Newton, J. (1995) Task-based interaction and incidental vocabulary learning: A case study, *Second Language Research*, 11, p. 159–176.
- Nicolaev, V. (2010) Les négociations de sens dans un dispositif d'apprentissage des langues en ligne synchrone par visioconférence, *Les Cahiers de l'Acedle*, 7, p. 169–198.
- Norris, S. (2004) *Analyzing Multimodal Interaction: A Methodological Framework*, New York: Routledge.
- O'Dowd, R. (2000) Intercultural learning via videoconferencing: A pilot exchange project, *ReCALL*, 12, p. 49–61.
- O'Dowd, R. (2016) Emerging Trends and New Directions in Telecollaborative Learning, *CALICO Journal*, 33.3, p. 291–310.
- Pica, T. (1994) Research on Negotiation: What Does It Reveal About Second-Language Learning Conditions, Processes, and Outcomes?, *Language Learning*, 44, p. 493–527.
- Pica, T., Young, R. & C. Doughty (1987) The impact of interaction on comprehension, *Tesol Quarterly*, 21, p. 737–758.
- Porquier, R. & B. Py (2008) *Apprentissage d'une langue étrangère: Contextes et discours*, Mesnil-sur-l'Estrée: Imprimerie Nouvelle Firmin Didot.
- Pressley, M., Levin, J. & M. McDaniel (1987) Remembering versus inferring what a word means: Mnemonic and contextual approaches, in McGeown, M. & M. E. Curtis (eds), *The nature of vocabulary acquisition*, Hillsdale, N.J.: Erlbaum.
- Quinn-Allen, L. (1995) The effects of emblematic gestures on the development and access of mental representations of French expressions, *The Modern Language Journal*, 79, p. 521–529.
- Renner, J. (2017) Negotiation of meaning and language-related episodes in synchronous, audio-based Chinese-German eTandem, *Language Learning in Higher Education*, 7, p. 137–159.
- Rey-Debove, J. (1997) *Le métalangage: Étude linguistique du discours sur le langage*, Paris: Armand Colin.
- Rivens Mompean, A. & N. Guichon (2009) Assessing the use of aids for a computer-mediated task: Taking notes while listening, *JALT CALL Journal*, 5, p. 45–60.

- Sacks, H., Schegloff, E. A. & G. Jefferson (1974) A Simplest Systematics for the Organization of Turn-Taking for Conversation, *Language*, 50, p. 696–735.
- Sarré, C. (2011) Computer-mediated negotiated interactions: How is meaning negotiated in discussion boards, text chat and videoconferencing?, in Thouësny, S. & L. Bradley (eds) *Second language teaching and learning with technology: Views of emergent researchers*. Dublin: Research-publishing.net, p. 189–210.
- Schegloff, E. A., Jefferson, G., & H. Sacks (1977) The Preference for Self-Correction in the Organization of Repair in Conversation, *Language*, 53, p. 361–382.
- Schmidt, R. W. (1990) The role of consciousness in second language learning¹, *Applied Linguistics*, 11, p. 129–158.
- Schmitt, N. (2008) *Teaching vocabulary*, Pearson Education.
- Sime, D. (2006) What do learners make of teachers' gestures in the language classroom?, *IRAL - International Review of Applied Linguistics in Language Teaching*, 44, p. 211–230.
- Sime, D. (2008) "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 (eds) *Gesture: Second Language Acquisition and Classroom Research*, New York: Routledge, p. 259–279.
- Sindoni, M. G. (2013) *Spoken and Written Discourse in Online Interactions: A Multimodal Approach*, New York, NY: Routledge.
- Smith, B. (2003) Computer-Mediated Negotiated Interaction: An Expanded Model, *The Modern Language Journal*, 87, p. 38–57.
- Stahl, S. A. & S. J. Vancil (1986) Discussion is what makes semantic maps work in vocabulary instruction, *The Reading Teacher*, 40, p. 62–67.
- Stam, G. & S. McCafferty (2008) Gesture Studies and Second Language Acquisition: A Review, in McCafferty, S. & G. Stam (eds) *Gesture: Second Language Acquisition and Classroom Research*, New York: Routledge, p. 3–24.
- Swain, M. (1985) Communicative competence: Some roles of comprehensible input and comprehensible output in its development, in Gass, S. M. & C. G. Madden (eds) *Input in second language acquisition*, Rowley, MA: Newbury House, p. 235–252.
- Swain, M. (1995) Three functions of output in second language learning, in *Principles and Practice in Applied Linguistics: Studies in Honour of H.G. Widdowson*, Oxford: Oxford University Press, p. 125–144.
- Swain, M. (2000) The output hypothesis and beyond: Mediating acquisition through collaborative dialogue, in Lantolf, J. P. (ed.) *Sociocultural theory and L2 learning*, Oxford: Oxford University Press, p. 97–114.
- Tellier, M. (2008a) Dire avec des gestes, *Le Français Dans Le Monde: Recherche et Application*, p. 40–50.
- Tellier, M. (2008b) The effect of gestures on second language memorisation by young children, *Gesture*, 8, p. 219–235.
- Tellier, M. (2013) De l'usage du corpus semi-contrôlé dans la recherche en didactique des langues, *Les Cahiers de l'ASDIFLE, FLE : L'instant et l'histoire, Actes des 49e et 50e rencontres*, p. 39–47.
- Tellier, M. & L. Cadet (2014) *Le corps et la voix de l'enseignant: Théorie et pratique*, Paris: Éditions Maison des Langues.

- Tréville, M.-C. (2000) *Vocabulaire et apprentissage d'une langue seconde: Recherches et théories*, Québec: Les Éditions Logiques.
- Tréville, M.-C. & L. Duquette (1996) *Enseigner le vocabulaire en classe de langue*, Paris: Hachette F.L.E.
- Van der Linden, E. (2006) Lexique mental et apprentissage des mots, *Revue Française de Linguistique Appliquée*, 11, p. 33–44.
- Van der Zwaard, R. & A. Bannink (2019) Toward a comprehensive model of negotiated interaction in computer-mediated communication, *Language Learning & Technology*, 23, p. 116–135.
- Van Leeuwen, T. (2005) *Introducing social semiotics*, London: Routledge.
- Varonis, E. M. & S. Gass (1985) Non-native/Non-native Conversations: A Model for Negotiation of Meaning, *Applied Linguistics*, 6, p. 71–90.
- Waring, R. (2002) Scales of vocabulary knowledge in second language vocabulary assessment, *Kiyo. Occasional Papers of Notre Dame Seishin University, Okayama*, 26, p. 40–54.
- Wigham, C. R. (2017) A multimodal analysis of lexical explanation sequences in webconferencing-supported language teaching, *Language Learning in Higher Education*, 7, p. 81–108.
- Wittenburg, P., Brugman, H., Russel, A., Klassmann, A. & H. Sloetjes (2006) Elan: A professional framework for multimodality research, *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006)*, p. 1556–1559.

NOTES

1. « Les sens des mots de vocabulaire est l'un des types d'informations qui sont souvent véhiculés dans la classe de langue, et c'est peut-être la situation langagière prototypique qui requiert le déploiement explicite de gestes à des fins pédagogiques » (notre traduction).
2. La visioconférence poste à poste réunit deux personnes, chacun à son « poste » devant un écran. Il peut y avoir plusieurs personnes par poste s'ils sont tous devant un même écran. Il existe de nombreux logiciels de visioconférence comme Skype qui mettent des utilisateurs en contact via une webcam, un microphone, et un clavier. Develotte, Kern & Lamy (2011 : 12) estiment que la visioconférence « s'apparente le plus à la conversation de face à face traditionnelle ».
3. Définies comme les actions, matériels et artefacts que l'on utilise pour communiquer (Jewitt, 2014 ; Van Leeuwen, 2005).
4. Le Cadre européen commun de référence pour les langues est disponible en ligne sur <https://rm.coe.int/16802fc3a8>
5. Voir <http://icar.cnrs.fr/ccf/>
6. Les vidéos de notre corpus ont un taux de rafraichissement d'environ trois images par seconde.
7. Sur la colonne de gauche, nous mettons la transcription de la parole orale en suivant la convention ICOR. Nous mettons en caractères gras la partie de la transcription qui correspond temporellement à la capture d'écran que nous mettons dans colonne du milieu. Pour annoter le mouvement des mains et de la tête, nous mettons une flèche par mouvement détecté lors du noyau de la phrase gestuelle. Sur la colonne de droite, nous décrivons les phénomènes dont la capture d'écran est l'illustration. Cette représentation à trois colonnes nous semble être un bon compromis entre exhaustivité et lisibilité.

8. Ici on assiste à une opportunité ratée pour encourager l'usage créatif (Joe, 1995, 1998). Emilie aurait pu dévoiler le sens littéral du verbe « arrondir » et faire deviner aux apprenants le sens figuré afin d'entraîner leur capacité à faire de l'usage créatif, c'est-à-dire accéder aux acceptions d'un item lexical qui ne sont pas apprises explicitement. La capacité à faire de l'usage créatif réduit le temps nécessaire pour apprendre du vocabulaire, car il est coûteux en temps d'apprendre individuellement toutes les acceptions de chaque item lexical. Le développement de cette compétence devrait être visée et favorisée par tous les participants.

RÉSUMÉS

Cet article vise à explorer les fonctions des gestes pédagogiques produits par de futurs enseignants de français lorsqu'ils enseignent à des apprenants de niveau intermédiaire par visioconférence poste à poste. Plus particulièrement, l'apport des gestes est analysé lors des séquences d'explication lexicale. À partir d'un corpus audiovisuel de quinze heures et demie d'interaction visiophonique entre sept apprentis-enseignants et douze apprenants, nous avons repéré environ trois cents séquences d'explication lexicale. Par une approche multimodale, nous analysons l'apport des gestes produits par les apprentis-enseignants pour de se focaliser sur la forme, sur le sens, et sur l'usage des items lexicaux, et pour impliquer les apprenants dans les séquences explicatives. À partir d'exemples tirés du corpus, nous mettons à jour les fonctions des gestes pédagogiques dans ce contexte particulier.

Coverbal gestures are tightly linked to spoken language and play a vital role in communication. Speakers have a tendency to produce more gestures when their interlocutor is visible, such as during face-to-face communication, or when they know that they can be seen, such as in a videoconferencing environment. Gestures are especially useful to teachers, who tend to be more aware of the gestures that they produce and who use them to fulfill various pedagogical functions. These so-called teaching gestures are particularly useful for lexical explanations, where a gesture can be associated with a word. Gestures have been shown to help enhance word memorization, especially if the learner reproduces the gesture.

Lexical explanations occur frequently during foreign language teaching, both in physical classrooms and in online videoconferencing environments. Empirical evidence has shown that negotiation sequences, which are conducive to language acquisition, are primarily triggered by lexical items. It is therefore interesting to analyze the role that gestures play during these videoconferenced lexical explanation sequences.

Explicit lexical explanation sequences are important for beginner and intermediate learners because their vocabularies are not yet rich enough to learn new words incidentally by reading or by listening. During a lexical explanation sequence, teachers use various techniques to focus on different aspects of the lexical item. The classic model of a lexical explanation sequence is comprised of three steps: the opening phase, during which the unknown lexical item is identified and highlighted; the core, during which the lexical item is explained; and the closing phase, during which the learner's comprehension is ratified.

The objective of a lexical explanation sequence is to help the learner to know the lexical item in

question. Lexical knowledge can be broken down in many different ways, but for our purposes and in most general terms, knowing a lexical item involves knowing its oral and written forms, its meanings, and its usage constraints. The form includes written form (spelling) and spoken form (pronunciation). Meaning involves what is signified by the word and what it refers to, as well as the other words that are associated with it. Usage refers to grammatical functions, collocations, and constraints on usage such as frequency and register. Some or all of these three main facets of lexical knowledge are dealt with during lexical explanation sequences.

According to the multimodal perspective that we have adopted, interlocutors make use of a variety of semiotic resources during lexical explanation sequences such as oral and written language, prosody, gestures, facial expressions, posture, and others. No single resource is always more important than another, and choices are made depending on the interactional context and on the semiotic environment. The videoconferencing platform allows for a wide range of modes to be used including gestures, as long as the interlocutors make their gestures visible in front of the webcam. Gestures produced in a pedagogical environment are known to fulfill cognitive, emotional, organizational, informational, and evaluative roles, among others. This study aims to describe the roles that gestures play during lexical explanation sequences that occur during videoconferencing sessions.

Our data come from a semester-long videoconferencing project that took place during the fall of 2013 between 12 master's students in France who were training to become French teachers and 18 undergraduate business students in Ireland who were learning French for business purposes in preparation for an internship the following year in France. The students had an intermediate level of French and the weekly interactions lasted between 25 and 40 minutes. Consent form responses meant that 7 French master's students and 12 Irish undergraduates were retained for the study.

After aligning all of the various audio and video recordings using video editing software, all spoken and written language was transcribed using multimodal annotation software. From 15.5 hours of video, 295 lexical explanation sequences were identified and selected for study. About half of the lexical explanation sequences were triggered by comprehension problems, and the other half were initiated by production problems. The teacher trainees initiated the lexical explanations slightly more frequently than the students. Each lexical explanation sequence was divided into the three phases listed above (opening, core, and closing). Aside from the written and spoken language, multimodal phenomena such as posture, facial expressions, eyebrow movements, hand gestures, and document sharing were annotated.

Although a large number of gestures were produced outside the field of view of the webcam, we were able to analyze the role that visible gestures played in the lexical explanation sequences. We postulate that the gestures that were produced in the limited space in front of the webcam were done so with communicative and pedagogical intent. The teacher trainees seem to think that gestures constitute a valuable explanation strategy because they are often used during lexical explanation sequences. Our examples show that teacher trainees used their gestures to explain all of the main facets of word comprehension: form, meaning, and usage. In addition to this, they used their gestures to help involve the learners in the explicative process, which facilitates the acquisition of lexical knowledge.

Our examples show the utility of gestures in explaining different facets of lexical items in this particular multimodal environment, but due to the lack of a posttest, we are not able to make any claims about acquisition or draw any conclusions about which gestures or types of explanations

work better than others. Furthermore, we are not able to ascertain whether or not the learners pay attention to the gestures that teachers produce in front of the webcam. Future research should address this.

INDEX

Mots-clés : didactique du français langue étrangère, visioconférence, gestes pédagogiques, explication lexicale, multimodalité

Keywords : teaching French as a foreign language, videoconferencing, teaching gestures, lexical explanation, multimodality

AUTEUR

BENJAMIN HOLT

Aix Marseille Univ, CNRS, LPL, Aix-en-Provence, France
benjamin.holt@univ-amu.fr