

Affordances et positionnements d'expertise culturelle dans un tandem par visioconférence

Cappellini, Marco

Laboratoire *Savoirs Textes Langage* UMR 8163 CNRS
Université Lille 3 SHS
marco.cappellini@univ-lille3.fr

1	Introduction.....	29
2	Contexte	29
3	Cadre théorique.....	29
3.1	Pour une perspective interactionniste dans l'étude des positionnements d'expertise culturelle.....	29
3.1.1	Analyse conversationnelle et système des tours de parole	29
3.1.2	Procédés et co-construction des positionnements d'expertise	30
3.2	Pour une perspective écologique dans l'étude des interactions par visioconférence	30
3.2.1	Affordances et affordances communicatives	31
4	Cadre méthodologique.....	31
4.1	Recueil des données et constitution du corpus d'analyse	32
	<i>Tableau 1 : Corpus.....</i>	<i>32</i>
4.2	Approche analytique	32
5	Analyses et discussion	33
5.1	Expert-novice question-explication	33
	<i>Vidéo 1 : Cliquer sur l'image pour jouer l'extrait vidéo 1.....</i>	<i>33</i>
	<i>Extrait 1 : Transcription de l'extrait vidéo 1.....</i>	<i>34</i>
	<i>Image 1 : Organisation de l'écran</i>	<i>35</i>
5.2	Expert-novice « inversé »	36
	<i>Vidéo 2 : Cliquer sur l'image pour jouer l'extrait vidéo 2.....</i>	<i>36</i>
	<i>Extrait 2 : Transcription de l'extrait vidéo 2.....</i>	<i>37</i>
5.3	Discussion.....	38
6	Conclusion.....	39
	Références bibliographiques	39

1 Introduction

Depuis une dizaine d'années, la diffusion de logiciels de visioconférence poste à poste a stimulé différentes recherches sur la manière de gérer la communication à l'écran par la mobilisation de plusieurs régimes sémiotiques (Develotte *et al.* 2011). Entre autres, les chercheurs se sont intéressé(e)s à comment la dimension sociale de l'interaction est co-construite à partir des caractéristiques des environnements. Dans notre étude, nous nous intéressons à cette question dans le cadre d'un dispositif d'apprentissage des langues française et chinoise en tandem par visioconférence poste à poste, ou télé-tandem (Telles 2009). Plus particulièrement, nous étudions comment, à travers l'environnement de visioconférence, les interlocuteurs se positionnent en tant qu'expert et novice (Reichert et Liebscher 2012) concernant leurs cultures. Pour ce faire, nous analysons les captures d'écran dynamiques de 15 sessions télé-tandem (20 heures environ) de quatre binômes, en adoptant une approche écologique (Bronfenbrenner 1979) et interactionniste (Dausendschon-Gay 2010).

Dans notre article, nous commençons par présenter brièvement le contexte dans lequel les interactions ont eu lieu. Ensuite nous abordons nos cadres théorique et méthodologique. Le cœur de l'article est constitué de l'analyse de deux exemples et une discussion en termes plus généraux. Dans la conclusion nous revenons sur nos résultats et nous ouvrons des perspectives pour des études à venir.

2 Contexte

Nous nous intéressons à un dispositif d'apprentissage des langues en télé-tandem (Telles 2009). Dans un télé-tandem, deux étudiants de langues maternelles différentes et apprenant chacun la langue de l'autre interagissent par visioconférence poste à poste (dorénavant « visioconférence ») pour converser la moitié du temps dans une langue et l'autre moitié dans l'autre afin de s'entraider dans leurs apprentissages.

Plus particulièrement, nous nous concentrons sur un télé-tandem pour l'apprentissage des langues française et chinoise¹. Du côté chinois, les étudiants ont un âge compris entre 20 et 22 ans et sont inscrits à la troisième et avant-dernière année de licence en langue française à l'Université de Langues Etrangères de Dalian (Chine). Du côté français, les étudiants ont un âge compris entre 21 et 23 ans et sont inscrits en première année de master à l'UFR de Langues Etrangères Appliquées à l'Université Lille3. Les étudiants ont la liberté de choisir quand et où faire leurs sessions télé-tandem, ce qui a amené certains d'entre eux à les faire à leurs domiciles, alors que d'autres ont préféré se rendre dans un cybercafé pour bénéficier d'une meilleure connexion. Tous les binômes ont utilisé le logiciel *Skype*.

3 Cadre théorique

Dans notre étude de la co-construction des expertises culturelles, nous combinons une perspective interactionniste et une perspective écologique.

3.1 Pour une perspective interactionniste dans l'étude des positionnements d'expertise culturelle

3.1.1 Analyse conversationnelle et système des tours de parole

Dans leur article fondateur, Sacks, Schegloff et Jefferson (1974) se proposent d'identifier l'« appareil formel » qui permet à des interlocuteurs de gérer leur conversation et formulent quatorze règles de l'alternance des tours de parole. Le tour de parole est défini comme une

¹ Pour plus de précisions concernant ce dispositif, se référer à Cappellini et Rivens Mompean (2013).

contribution verbale d'un interlocuteur à un moment donné de l'interaction. Un tour de parole se termine soit avec un chevauchement amenant à une interruption, soit avec un point de transition. Ces points, dont la nature est toujours potentielle, sont produits et identifiés par les interlocuteurs à l'aide des structures langagières employées par le locuteur et surtout par des indices phonétiques et prosodiques.

Deux caractéristiques de ce modèle sont importantes pour nous : il s'agit d'une part d'un système géré de manière interactionnelle (*interactionally managed system*) et d'autre part d'un système de gestion locale (*local management system*) (*ibid.* 725). Si ce « principe de détermination locale » concernait initialement seulement le système de l'alternance des tours de parole, par la suite plusieurs chercheurs, dont Dausendschon-Gay (2010), ont souligné que toute construction dans l'interaction a un caractère local. En particulier, la co-construction des catégories assignées aux interlocuteurs, dont la catégorie d'expert (Reichert et Liebscher 2012), est sujette à ce principe (Gajo et Mondada 2000 : 114). Enfin, un dernier principe régissant les interactions est celui de séquentialité (Dausendschon-Gay 2010), d'après lequel la valeur d'un tour de parole ne peut être établie qu'en fonction de ce qui le précède et de ce qui le suit dans l'interaction.

3.1.2 Procédés et co-construction des positionnements d'expertise

Par « procédé » (Gajo et Mondada 2000 : 23), on indique l'ensemble des actions mises en place par un interlocuteur afin d'accomplir une action, pour nous le traitement interactif de l'expertise culturelle dans les interactions télétandem. Nous abordons cet objet sous l'angle du positionnement discursif (Reichert et Liebscher 2012) dans l'accomplissement de ce que les ethnométhodologues appellent la *catégorisation* (Sacks 1972). Pendant l'interaction, les participants assignent des catégories aux différents interlocuteurs et manifestent « *qui* ils sont, afin d'organiser de façon intelligible leur action » (Gajo et Mondada 2000 : 107). Notons que les procédés de positionnement d'expertise ainsi que les catégories produites relèvent d'une perspective éémique (Dausendschon-Gay 2010 : 101).

Tout en sachant que les catégories assignées à un même interlocuteur peuvent être multiples, la catégorie à laquelle nous nous intéressons est celle d'« expert culturel ». La catégorie d'« expert » (Reichert et Liebscher 2012) caractérise la personne qui, dans l'interaction, se présente comme celle qui connaît le sujet de conversation et peut se permettre de l'expliquer à un interlocuteur qui ne le connaîtrait pas, le novice. Par l'adjectif « culturel », nous indiquons que nous focalisons notre attention sur les catégorisations d'expert et/ou de novice concernant des sujets des cultures françaises et chinoises (Cappellini et Rivens Mompean 2013).

3.2 Pour une perspective écologique dans l'étude des interactions par visioconférence

Bronfenbrenner (1979) indiquait par l'adjectif « écologique » une approche en psychologie qui s'oppose à la psychologie « expérimentale », caractérisée, cette dernière, par le découpage des situations à étudier en variables et leur manipulation en laboratoire. Au contraire, une perspective écologique s'intéresse à l'étude des sujets dans leurs environnements « naturels », caractérisés notamment par le fait que les sujets étudiés peuvent en manipuler les éléments (Bronfenbrenner 1979). En visioconférence, cela revient à affirmer que les sujets de notre étude ont une certaine marge de manœuvre sur leurs environnements physique et numérique (Develotte et Drissi 2013 : 56), par exemple par la possibilité d'organiser la distribution des différentes fenêtres à l'écran (Lamy et Hampel 2007 : 37).

De plus, une perspective écologique vise une validité phénoménologique de l'analyse (Bronfenbrenner 1979), qui fait écho à la perspective éémique des interactionnistes. En ce sens, nous suivons Le Bray (2002), qui note que dans l'étude des interactions avec l'ordinateur – et *a*

fortiori des interactions humaines *via* l'écran d'ordinateur – il est fondamental de s'intéresser non seulement aux manipulations de l'artefact, mais surtout aux sens que les acteurs, pour nous les interlocuteurs, leur attribuent.

Concernant les interactions par visioconférence, Drissi (2011 : 134 et sq.) distingue quatre perspectives pour envisager les moyens de la communication : l'objet technique, le mode, la modalité et le canal. Par le premier, elle indique l'artefact technique matériel qui permet l'interaction, dans notre cas l'ordinateur. La modalité correspond à une entrée par le logiciel de visioconférence. Ainsi, dans *Skype* les interlocuteurs disposent des modalités audio, vidéo et du clavier. Le mode correspond aux ressources sémiotiques mobilisées, pour nous verbales, paraverbales et non-verbales. Le canal enfin correspond au niveau de l'humain ; en visioconférence surtout le visuel et l'auditif.

Dans notre perspective écologique, l'environnement est envisagé comme un réservoir de ressources sémiotiques (Van Lier 2004), autrement dit d'« éléments » qui peuvent être introduits dans la communication pour co-construire du sens. En visioconférence, l'introduction de ressources sémiotiques venant de l'environnement (physique ou numérique), se fait par un investissement des modalités vidéo et audio (Vincent 2012) et/ou par la manipulation d'objets visuels à l'écran (Lamy et Flewitt 2011). Enfin, il nous faut noter que même si pour les besoins ponctuels de l'analyse nous pouvons isoler une modalité des autres, la communication se faisant de manière transmodale (Hauck 2010), son étude doit se faire par ce que des chercheurs comme Baldry et Thibault (2006) appellent le « principe d'intégration des ressources sémiotiques ».

3.2.1 Affordances et affordances communicatives

Dans une perspective écologique, l'ensemble des « moyens » de communication et leurs contraintes peut être envisagé par la notion d'affordance. Cette notion, formulée par Gibson (1979) dans le cadre de la psychologie de la perception, indiquait initialement la forte corrélation (au sens philosophique, non statistique, du terme) entre les éléments de l'environnement et leur perception par un sujet engagé dans une certaine action. En ce qui nous concerne, cela signifie affirmer la priorité de la perception des sujets sur une perspective technocentrée en termes de potentialités des artefacts et logiciels. La notion d'affordance a été reprise entre autres par Van Lier (2004), qui en a spécifié la relation avec l'action. Ainsi, ce chercheur fait des affordances des ressources sémio-pragmatiques, c'est-à-dire des éléments pouvant entrer en jeu non seulement dans l'action, mais aussi dans la construction de sens typique de l'interaction.

Pour sa part, Hutchby (2001), en étudiant comment la communication verbale se développe *via* différentes technologies, élabore la notion d'affordance communicative, autrement dit les « éléments » de l'artefact (et pour nous également du logiciel) qui permettent et contraignent la communication. En particulier, ce chercheur étudie comment les procédés ethnométhodologiques de gestion de l'interaction peuvent entrer en conflit avec les affordances communicatives. Dans la même lignée, Lamy (2008) affirme que dans l'étude de la CMO il faut réadapter les outils de l'analyse conversationnelle – notamment le principe de séquentialité – en fonction des nouvelles affordances, dont la coexistence de la modalité audio et du clavier.

4 Cadre méthodologique

Notre principale question de recherche est : comment les interlocuteurs exploitent les affordances et les affordances communicatives de l'environnement de visioconférence pour se positionner en tant qu'experts et novices concernant les cultures chinoises et françaises ? Une deuxième question, d'ordre plus méthodologique, est : quelles catégories permettent de mieux analyser les positionnements d'expertise culturelle ?

4.1 Recueil des données et constitution du corpus d'analyse

Pour répondre à ces questions, nous avons recueilli un corpus d'interactions télé-tandem. Parmi les étudiants participant au dispositif de formation, certains ont signé un document d'information et de cession des droits à l'image. Ensuite, ces étudiants ont reçu des instructions pour l'enregistrement des sessions télé-tandem à l'aide de différents logiciels de capture d'écran dynamique et audio.

Quatre binômes ont réussi à enregistrer une partie ou l'intégralité de leurs sessions télé-tandem. Tous les sujets sont des étudiantes avec un niveau de compétence dans leur langue étrangère correspondant au niveau B2 du *Cadre Européen Commun de Référence pour les Langues* (Conseil de l'Europe, 2001). Au moment des enregistrements, toutes les étudiantes françaises avaient effectué un échange en Chine pendant 6 à 10 mois. D'autre part, aucune des étudiantes chinoises n'avait été hors de la Chine. Le tableau suivant présente brièvement le corpus recueilli. Chaque binôme est identifié avec les lettres initiales des noms des étudiantes.

Binôme	Nombre de sessions enregistrées	Nombre d'heures enregistrées
LS	4	7h30
AH	3	3h20
CS	5	6h10
CW	3	3h20

Tableau 1 : Corpus

L'ensemble des interactions a fait l'objet d'une transcription des modes verbal et paraverbal et d'un alignement avec les enregistrements par le logiciel ELAN².

4.2 Approche analytique

Dans nos analyses, nous avons mobilisé les principes du cadre théorique. Ainsi, premièrement nous nous sommes intéressé principalement aux modes verbal et paraverbal et nous avons identifié, en fonction des principes de séquentialité et de détermination locale, les positionnements d'expertise culturelle possibles. Deuxièmement, en nous appuyant sur le principe d'intégration des ressources sémiotiques, nous avons visionné l'ensemble des séquences relevant de chaque positionnement et nous avons cherché des régularités dans la manipulation et l'investissement de la multimodalité, en identifiant des procédés récurrents.

² <https://tla.mpi.nl/tools/tla-tools/elan/>

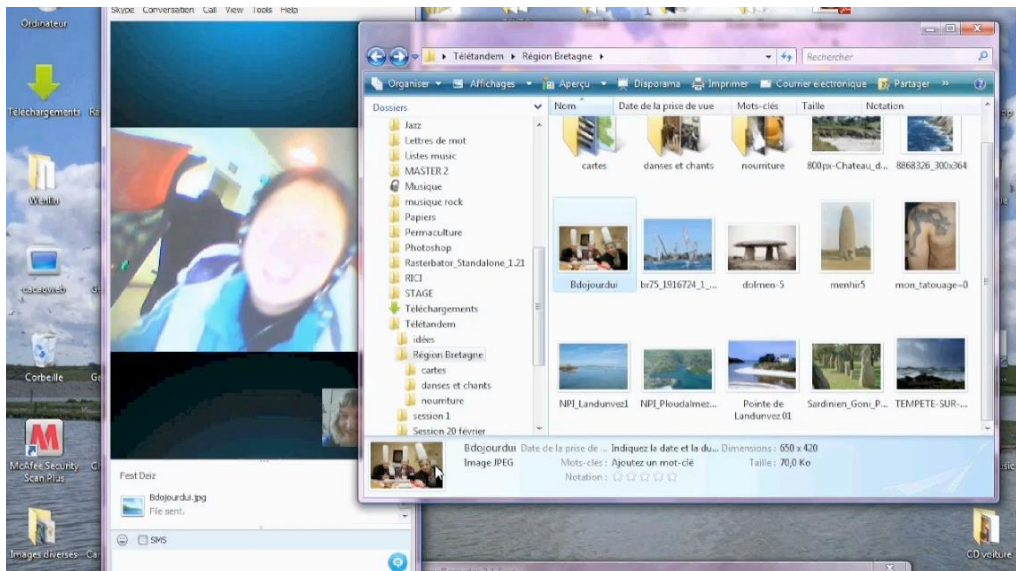
5 Analyses et discussion

D'un point de vue quantitatif, les analyses ont montré que sur les 20 heures d'enregistrements, 10 heures et 30 minutes relevaient des positionnements d'expertise culturelle. Trois catégories de positionnement ont émergé : les positionnements expert-novice, dont nous avons relevé 390 occurrences couvrant 9 heures et 39 minutes ; ceux expert-expert, avec 76 occurrences sur 44 minutes ; les positionnements novice-novice, avec 12 occurrences sur 6 minutes. Une étude exhaustive n'étant pas possible faute d'espace, nous nous concentrons sur deux procédés (sur sept) des positionnements expert-novice. Le premier a été choisi car il est le plus courant, le second pour son caractère d'exception.

5.1 Expert-novice question-explication

Le procédé le plus fréquent dans les positionnements d'expertise culturelle expert-novice est celui que nous nommerons « question-explication », visible dans l'exemple suivant. L'extrait est tiré d'une séquence où l'étudiante française (A) présente sa région natale à son interlocutrice (H) (AH2 15min57-17min38) :

Extrait 1 :



Vidéo 1 : Cliquer sur l'image pour jouer l'extrait vidéo 1

A verbal fr oui c'est un peu drôle et c'est très très drôle parce que euh ouais donc ça doit être dans les années mille-huit-cent-cinquante un peu des choses comme ça ils portaient encore beaucoup et euh quand il y a eu l'arrivée des premières voitures et ben en fait tu ne pouvais pas quand tu étais une femme bretonne tu ne pouvais pas euh 'fin de cette région là de la bretagne tu ne pouvais pas sortir de chez toi sans le chapeau si non c'était comme si tu sortais toute nue , il y avait pas le droit il y avait pas le droit de montrer ses cheveux sans ça tu vois si non si non c'était vraiment pas bien

TC 00:15:57.008 - 00:16:40.038

H verbal fr ok

TC 00:16:40.798 - 00:16:41.168

A verbal fr et du coup

TC 00:16:41.638 - 00:16:42.218

H verbal fr mais pourquoi les les chapeaux sont si hauts ?

TC 00:16:42.013 - 00:16:45.026

A verbal fr euh c'est pour c'est parce que partout en bretagne il y avait différents chapeaux et au début c'était petit comme ça et euh et en fait cette région de la bretagne c'est le pays bigouden ça s'appelle bi-gouden bigouden et c'est un une petite région de la bretagne euh dans le finistère et cette région là les gens étaient très fières tu comprends ? oui et ils voulaient ils voulaient toujours être plus beaux que les autres donc leurs costumes ils mettaient des petits boutons en or ils mettaient de l'or et ils faisaient les coiffes les plus hautes possible, comme ça c'était les plus beaux tu vois ça voulait dire ça et donc et donc ces gens là il y a que dans cette partie là de la bretagne que c'était aussi haut si non là où j'habite moi c'est plus petit c'est comme ça un peu tu vois et donc eux et ben ils étaient embêtés parce qu'ils devaient tout le temps porter ça et quand il y a eu les premières voitures ils devaient conduire quand elles conduisaient les femmes elles conduisaient la tête comme ça pour que ça sorte de la voiture

TC 00:16:45.414 - 00:17:50.842

A clavardage pays bigouden

TC 00:16:58.122 - 00:17:02.039

H paraverbal ah

TC 00:17:02.464 - 00:17:02.814

Extrait 1 : Transcription de l'extrait vidéo 1

Du point de vue discursif, l'étudiante française apporte des informations concernant la culture bretonne. Ce faisant, elle se positionne comme la personne qui connaît et explique, autrement dit en tant qu'experte. De son côté, l'étudiante chinoise non seulement ne remet pas en question ce positionnement, mais même le renforce en posant une question. Par cette question, elle positionne son interlocutrice en tant qu'experte et se positionne elle-même comme novice. Suite à la question de H, A répond en élaborant une explication. Ce faisant, elle se positionne à nouveau en experte. Avec ce procédé, les tours de parole de l'experte peuvent être très longs (environ une minute), alors que ceux de la novice ne sont que très courts et constitués principalement de questions ou de petites remarques. Enfin, les procédés mis en place par les interlocuteurs projettent des positionnements non seulement concernant la personne qui parle, mais aussi son interlocuteur.

En ce qui concerne la multimodalité, une première observation porte sur l'organisation de l'environnement à l'écran de la part de A.

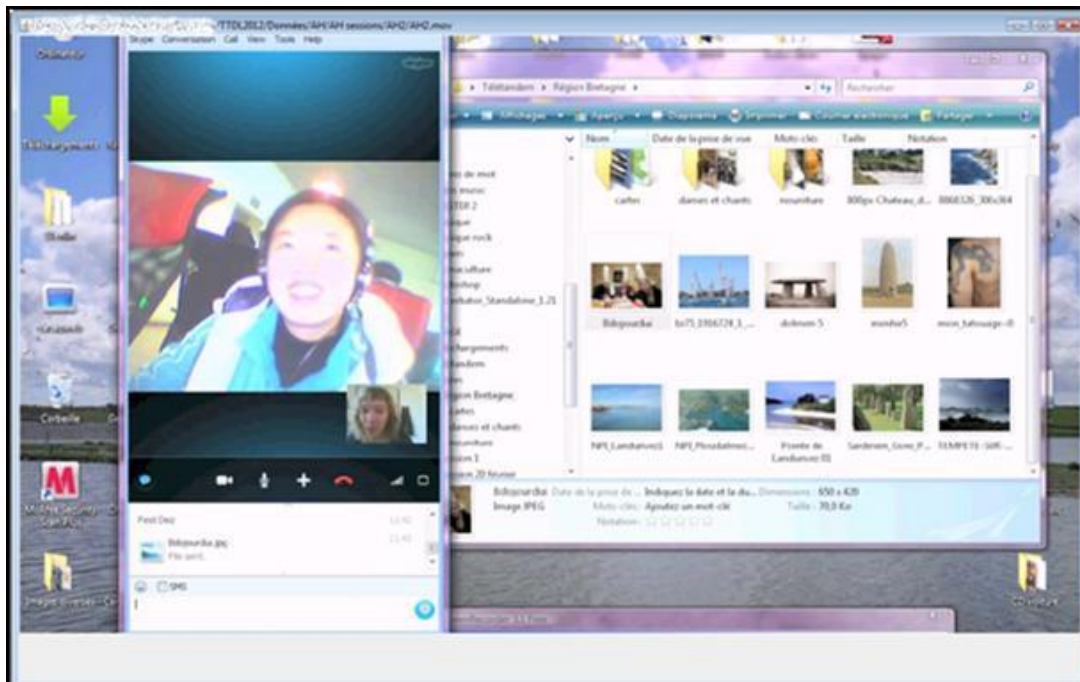


Image 1 : Organisation de l'écran

Dans ce cas, l'étudiante se positionnant en tant qu'experte a préparé une série d'images pour illustrer ses propos, images réunies dans un dossier. L'introduction de ce dossier à l'écran amène l'étudiante à réduire l'interface du logiciel de visioconférence, qui, au contraire, lors de la session télé-tandem précédente occupait l'entièreté de l'écran. Cette disposition des fenêtres facilite l'envoi des images par un *drag&drop* dans l'espace de clavardage de l'interface de Skype.

Le clavardage est utilisé lors de l'échange analysé également pour envoyer le nom d'un lieu (pays bigouden). A ce propos, il faut noter que l'utilisation du clavardage amène à une interruption temporaire de la conversation sur le mode verbal oral : A arrête son discours et semble passer à un langage autocentré focalisant son attention sur l'écriture. D'autre part, l'étudiante chinoise semble comprendre qu'il ne s'agit pas d'un point de transition et ne s'empare pas du tour de parole. Cette utilisation du clavardage nous semble un renforcement du positionnement d'expertise culturelle. En effet, l'étudiante française part de l'hypothèse que son interlocutrice non seulement ne connaît pas le (nom du) territoire en question, mais aussi elle a besoin d'en voir le nom pour pouvoir le comprendre.

Une autre série d'observations porte sur l'investissement de la modalité vidéo. En premier lieu, on peut noter parfois la présence de gestes autocentrés, même s'ils ne sont pas toujours visibles. Beaucoup plus intéressante est la présence de gestes illustratifs co-verbaux (Cosnier et Develotte 2011). Ces gestes, toujours produits par la personne ayant le tour de parole, sont présents à plusieurs reprises : pour illustrer « petit », pour mimer la conduite d'une voiture, et à plusieurs reprises pour les coiffes/chapeaux. Dans le cas de ces derniers, le même geste est décliné en plusieurs formats, en fonction de la hauteur des coiffes décrites. Notons que quand A veut produire un geste qui est le plus ample possible, elle recule avec sa chaise pour que le geste soit visible dans le champ de la caméra. Cela indique qu'A est consciente des limites du champ de la caméra et qu'elle adapte sa distance de celle-ci pour profiter au mieux de la modalité vidéo, ce qui confirme les observations de Cosnier et Develotte (2011). Enfin, l'illustration du

verbal ne se fait pas que par les gestes, mais aussi par la posture. C'est le cas de l'illustration de « fiers », qu'A produit en dressant son dos et en portant ses poings à la poitrine.

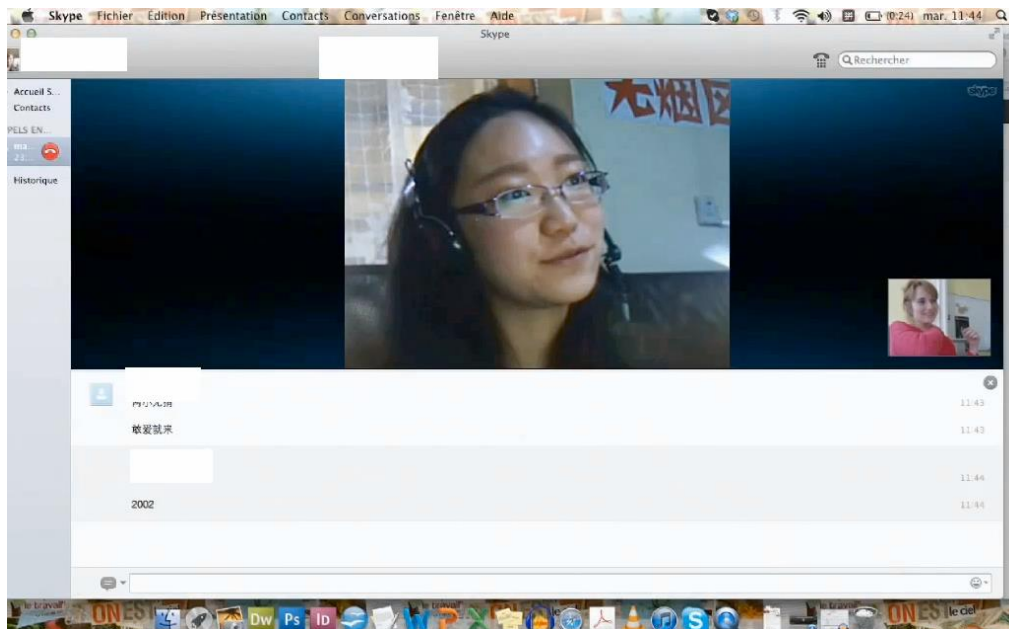
Un autre ensemble de gestes sont les gestes de pointage, produits par A soit quand elle se réfère à elle-même, soit à l'occasion de la référence au pays bigouden. Dans le premier cas, elle pointe avec les index des deux mains en direction de sa poitrine. Dans le deuxième cas, elle pointe avec l'index de la main droite en avant à droite. Dans ce deuxième cas, nous formulons deux hypothèses. La première serait que le pointage se fait vers l'ailleurs, en illustrant le fait qu'A ne vient pas de cette région – interprétation soutenue par le pronom « ils » et la désignation « ces gens là ». Une autre hypothèse serait que le pointage est dirigé vers une image présente à l'écran de A (et visible dans l'image 1) : celle de trois personnes avec la coiffe du pays bigouden. Bien que cette hypothèse soit intéressante pour les questions qu'elle ouvrirait concernant le destinataire du geste et la prise en compte du dispositif technique, nous pensons ne pas avoir suffisamment d'éléments pour la supporter.

Enfin, un dernier élément est le fait que la modalité vidéo est utilisée par l'étudiante chinoise lors des procédés régulateurs (Kerbrat-Orecchioni 1998 : 18). Ainsi, H produit des gestes faisant oui de la tête pour assurer à A qu'elle comprend l'explication, éventuellement suite à des procédés phatiques (*ibid.*) d'A.

5.2 Expert-novice « inversé »

Le positionnement expert-novice « inversé » est un positionnement rare (8 occurrences pour des sujets français et 5 pour des sujets chinois). Si dans les autres procédés de positionnement expert-novice nous avons observé que l'étudiante positionnée en tant qu'experte le faisait concernant sa culture, dans ce procédé c'est l'inverse. L'extrait suivant donne un exemple (CW4 23min53-24min21) :

Extrait 2³ :



Vidéo 2 : Cliquer sur l'image pour jouer l'extrait vidéo 2

³ Nous mettons entre crochets une traduction des tours de parole en chinois.

W verbal chn 还有比如 你要闭嘴 那个叫 叫法 语名应该 tais-toi
[il y a aussi par exemple « tu dois te taire »
celui-là s'appelle, en français il doit s'appeler
tais-toi]
TC 00:23:53.230 - 00:24:01.910

[C ne bouge pas]

W verbal fr tais-toi
TC 00:24:03.830 - 00:24:04.370

C verbal fr tais-toi ?
TC 00:24:05.070 - 00:24:05.530

W verbal chn oui 对 有 一 个 电 影 tais-toi
[oui oui il y a un film qui s'appelle tais-toi]
TC 00:24:05.690 - 00:24:08.582

C clarvardage (taituan)
TC 00:24:06.503 - 00:24:10.403

C verbal fr tais-toi mh
TC 00:24:09.020 - 00:24:10.760

W verbal chn 是 哪 也 是 很 老 的 了
[c'est aussi très vieux]
TC 00:24:13.807 - 00:24:17.949

C paraverbal rire
TC 00:24:18.060 - 00:24:18.670

C verbal chn 啊 对 对 对
[ah oui oui oui]
TC 00:24:19.740 - 00:24:21.000

Extrait 2 : Transcription de l'extrait vidéo 2

L'étudiante chinoise (W) dit avoir vu un film français (*Tais-toi*). Ce faisant, elle se positionne en tant qu'experte concernant un élément relevant des cultures françaises. Ensuite, après une demande de confirmation et la réponse qui suit, C fait comprendre avec différentes actions qu'elle ne connaît pas le film, en se positionnant ainsi en tant que novice. Après ce positionnement, W se repositionne en experte en apportant des nouvelles informations. Enfin C a recours à Internet, ce qui montre qu'il s'agit d'une affordance de l'environnement, et confirme à W qu'elle a compris.

Cet extrait est exemplaire des procédés conversationnels de ce qu'il nous semble possible d'appeler une « expertise inversée ». Avant tout, une étudiante introduit une référence à un sujet concernant un aspect de la culture de l'interlocutrice. Ensuite, un échange a lieu pour s'assurer du sujet de conversation. Enfin, l'experte donne des informations concernant le sujet de conversation.

La vidéo joue un rôle important dans les procédés de cogestion de l'interaction. En premier lieu, les interlocutrices produisent des mimiques de *doing thinking* (Reichert et Liebscher 2012), autrement dit, elles se mettent en scène dans l'acte de penser en dirigeant le regard en haut et/ou vers le côté. En second lieu, après le premier tour de parole de W, lors du point de transition, on peut voir que C ne bouge pas. Cette immobilité amène W à répéter le titre du film. Autrement dit, elle a le même effet de la question de confirmation formulée ensuite.

Ce type de procédé soulève des questions concernant les catégories d'analyse de la conversation par visioconférence. Avant tout, il prouve (s'il en était besoin) la pertinence du principe de séquentialité (Dausendschon-Gay 2010), car l'immobilité de C ne prend une signification que parce qu'elle se place dans un point de transition, un moment où au contraire il devrait y avoir une réaction, une prise de parole de la part de C. Si la valeur communicative de l'immobilité en fonction du principe de séquentialité ne nous semble pas faire des doutes, ce qui est plus déstabilisant est que C ne fait ni de manipulation – il n'y a pas de contact avec les périphériques de l'ordinateur – ni d'investissement de l'environnement à proprement parler. En effet, un investissement présuppose une action. Or, rester immobile est justement le contraire d'une action. Il nous semble alors pouvoir parler dans ce type de procédé d'une sorte d'« investissement 0 », un non-investissement des modalités de communication qui, en vertu du principe de séquentialité, peut constituer un comportement communicatif. Par conséquent, c'est une catégorie d'analyse de la multimodalité pertinente seulement par rapport au système des tours de parole, et en particulier des points de transition. Enfin, notons qu'une telle catégorie remet partiellement en question des analyses se basant exclusivement sur le repérage des actions des interlocuteurs, et non, aussi, sur leurs non-actions.

5.3 Discussion

Dans cette section, nous répondons à notre question de recherche principale en systématisant les remarques faites lors des analyses et en les complétant avec d'autres exemples, avant tout en abordant l'utilisation spécifique des différentes affordances, enfin en insistant sur l'intégration des ressources sémiotiques dans la production du sens et la cogestion des interactions.

Dans les procédés de positionnement d'expertise culturelle, les modalités audio et vidéo sont investies en continu. La modalité audio devient une affordance en mobilisant principalement les modes verbal et paraverbal, avec des procédés similaires à ceux du face à face présentiel. Autrement dit, lors des positionnements d'expertise culturelle expert-novice, les interlocutrices utilisent un ton et un débit de voix comme dans les conversations quotidiennes, avec parfois des phénomènes d'adaptation rappelant le xénolecte (Bange 1992). De plus, la modalité audio est employée dans certains cas pour faire écouter à son interlocuteur des chansons ou des musiques. Cela arrive quand il n'est pas possible d'envoyer un lien hypertexte vers la chanson en question, soit parce qu'il s'agit d'un fichier stocké sur la machine d'une des interlocutrices, soit parce que la musique se trouve sur un site Internet non accessible par l'autre interlocutrice. La modalité vidéo est investie principalement avec la production d'expressions faciales pour la gestion de l'interaction (par exemple de *doing thinking*) et avec des productions posturo-mimo-gestuelles visant à faciliter l'intercompréhension. La production de gestes – exception faite pour les gestes autocentrés – est toujours faite en s'adaptant au champ de la caméra et les gestes sont principalement illustratifs.

Nous avons pu observer également des investissements plus ponctuels de la modalité vidéo. En effet, l'image peut être utilisée aussi pour porter des objets dans le champ de la caméra afin de les rendre visibles à l'interlocutrice, ce qui rappelle des observations de Vincent (2012). Par exemple une étudiante, pendant la présentation d'un auteur chinois, porte un de ses livres face à la caméra. D'ailleurs, cela montre l'importance de l'environnement physique où les interactions ont lieu, dans la mesure où seulement les interlocutrices choisissant leur domicile ont la possibilité de prendre des objets de leur environnement quotidien.

Une autre modalité qui est utilisée de manière ponctuelle est le clavardage. Ainsi, l'experte peut se servir du clavardage pour envoyer un terme ou une expression afin de faciliter la compréhension de la novice ou des images préparées avant la conversation. En outre, le clavardage peut être le moyen d'envoyer des informations recherchées sur Internet, comme par exemple le texte d'une chanson ou une image.

Bien qu'Internet ne soit pas une modalité de la communication, il représente une affordance importante de l'environnement. En effet, quand les aléas de l'interaction le rendent nécessaire, l'étudiante se positionnant en tant qu'experte – mais parfois aussi la novice, comme dans l'exemple 2 – peut ouvrir un navigateur Internet pour se rendre sur des sites, dont le contenu sera ensuite utilisé dans la communication par les modalités du clavardage ou de l'audio. L'utilisation d'un navigateur Internet pendant la vidéoconférence, ainsi que l'ouverture de dossiers, pose des questions d'organisation de l'espace des fenêtres à l'écran.

En conclusion, notre recherche confirme les affirmations de Cosnier et Develotte (2011), d'après qui les éléments de la communication « sont étroitement unis pour réaliser l'interaction conversationnelle dans ses aspects d'échanges informationnels et de dispositif de co-pilotage régulateur de ces échanges ». Autrement dit, en adoptant le principe méthodologique de l'intégration des ressources (Baldry et Thibault 2006), les analyses des positionnements d'expertise culturelle expert-novice montrent bien que la co-construction du sens se fait de manière transmodale (Hauck 2010).

6 Conclusion

Dans notre étude, nous avons montré comment les différentes modalités présentes en visioconférences deviennent des affordances communicatives lors des positionnements d'expertise culturelle. En particulier, nous nous sommes arrêté sur deux exemples de positionnements expert-novice qui nous ont permis d'observer les différents types d'investissement des modalités ainsi que la transmodalité des positionnements. A partir de ces exemples, nous avons formulé une série d'observations et les avons enrichies d'autres remarques venant de notre corpus. Si dans cette étude nous nous sommes concentré spécifiquement sur les positionnements expert-novice, à l'avenir il nous semble profitable de prendre en compte d'autres types de positionnements (Cappellini et Rivens Mompean 2013) pour dégager des éventuelles différences et/ou similitudes dans les procédés. Enfin, en ce qui concerne les catégories d'analyse, notre étude, en se basant sur le principe de séquentialité, a mis en relief l'existence d'un « investissement 0 » des modalités de visioconférence.

Références bibliographiques

- Baldry, A. & Thibault, P. J. (2006). *Multimodal transcription and text analysis*. Sheffield : Equinox Publishing.
- Bange, P. (1992). A propos de la communication et de l'apprentissage en L2. *Aile*, 1, 53-86.
- Bronfenbrenner, U. (1979). *The Ecology of Human Development. Experiments by nature and design*. Cambridge : Harvard University Press.
- Cappellini, M. & Rivens Mompean, A. (2013). Positionnements culturels dans un tandem sino-fançais par visioconférence. *Sinergies Chine*, 8, 137-149.
- Conseil de l'Europe (2001). *Un Cadre Européen Commun de Référence pour les Langues*. Paris : Didier.
- Cosnier, J. & Develotte, C. (2011). Le face à face en ligne. Approche éthologique. In Develotte *et al.*, *op. cit.*, 27-50.
- Dausendschon-Gay, U. (2010). La construction interactive d'objets. In Macaire, D., Narcy-Combes, J.-P. & Portine, H. (éd.) *Interrogations épistémologiques en didactique des langues*. Paris : Clé, 98-110.
- Develotte, C. & Drissi, S. (2013). Face à face distanciel et didactique des langues. In Ollivier, C. & Puren, L. (éd.) *Mutations technologiques, nouvelles pratiques sociales et didactique des langues*. Paris : Clé, 54-63.
- Develotte, C., Kern, R. & Lamy, M.-N. (ed.) (2011). *Décrire la conversation en ligne. Le face à face distanciel*. Lyon : ENS Editions.
- Drissi, S. (2011). *Apprendre à enseigner par visioconférence : étude d'interactions pédagogiques entre futurs enseignants et apprenants de FLE*. Thèse soutenue le 2 décembre 2011 à l'ENS Lyon, inédite.
- Gajo, L. & Mondada, L. (2000). *Interactions et acquisitions en contexte*. Fribourg : Editions Universitaires Fribourg Suisse.

- Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. London : Lawrence Erlbaum Associates.
- Hauck, M (2010). Telecollaboration: At the interface between multimodal and intercultural communicative competence. In Guth, S. & Helm, F. (éd.) *Telecollaboration 2.0*. Oxford : Peter Lang, 219-244.
- Hutchby, I. (2001). *Conversation and technology: from the telephone to the internet*. Cambridge, Polity.
- Kerbrat-Orecchioni, C. 1998 [1990]. *Les interactions verbales. 1. Approche interactionnelle et structure des conversations*. Paris : Armand Colin.
- Kerbrat-Orecchioni, C. (2011). Conversation en présentiel et conversations en ligne : bilan comparatif. In Develotte *et al.*, *op. cit.*, 173-196.
- Lamy, M.-N. (2008). Ce que cyber-parler veut dire : quels cadres théoriques pour l'analyse des conversations multimodales en réseau ? In Gerbault, J. (éd.). *La langue du cyber-espace. De la diversité aux normes*. Paris : L'Harmattan, 91-104.
- Lamy, M.-N. & Flewitt, R. (2011). Describing online conversations : insights from a multimodal approach. In Develotte *et al.*, *op. cit.*, 71-94.
- Lamy, M.-N. & Hampel, R. (2007). *Online communication in language learning and teaching*. New York : Palgrave MacMillan.
- Le Bray, J.-E. (2002). Description de l'usage des nouvelles technologies dans l'apprentissage des langues. In Barbot, M.-J. & Pugibet, V. (éd.), *Apprentissage des langues et technologies : usages en émergence*. Paris : Clé, 37-44.
- Reichert, T. & Liebscher, G. (2012). Positioning the expert: Word searches expertise and learning opportunities in peer interactions. *The Modern Language Journal*, 96 (4), 599-619.
- Sacks, H. (1972) An Initial Investigation of the Usability of Conversational Data for Doing Sociology. In Sudnow, D. (éd.), *Studies in Social Interaction*. New York : The Free Press, pp. 31-74
- Sacks, H., Schegloff, E. & Jefferson, G. 1974. A Simplest Systematics for the Organization of Turn-Taking in Conversation. *Language*, 50(4), 696-735.
- Telles, J. A. (éd.) (2009). *Teletandem. Um contexto virtual, autônomo, colaborativo para aprendizagem das linguas estrangeiras no século XXI*. Campinas, SP : Pontes Editores.
- Van Lier, L. (2004). *The Ecology and Semiotics of Language Learning: a Sociocultural Perspective*. Dordrecht : Kluwer Academic Publishers.
- Vincent, C. (2012). Interactions pédagogiques « fortement multimodales » en ligne : le cas de tuteurs en formation. Thèse soutenue le 26 octobre 2012 à l'ENS Lyon, inédite.