

## L'analyse des interactions multimodales par écran: questionnements méthodologiques

Ibnelkaid, Samira

Laboratoire ICAR - Université Lumière Lyon 2  
samiraibnelkaid@gmail.com

La troisième Révolution Industrielle est en marche. Comme les précédentes, elle repose sur la découverte et l'exploitation de nouvelles formes d'énergie et de communication. En effet, « la jonction de la communication par Internet et des énergies renouvelables engendre une troisième révolution industrielle [qui] va changer radicalement tous les aspects de notre façon de travailler et de vivre » (Rifkin, 2012 : 57). La communication est si centrale pour l'Homme qu'elle est la première et principale innovation visée par les nouvelles technologies numériques. Et le développement croissant des modes de communication disponibles à celui que d'aucuns nomment désormais l'« Homo-numericus » (Mounier, 2002) s'associe à l'évolution de ses compétences linguistiques.

En effet, nos sociétés contemporaines font face à une Révolution Numérique (Vial, 2012) au cours de laquelle l'émergence des nouveaux réseaux et interfaces de communication bouleversent l'organisation rituelle des interactions. Le vaste spectre de formats d'interactions rendus possibles par les écrans, de la communication asynchrone écrite (mails, forums,...) à la communication synchrone audiovisuelle (visio,...), en passant par le quasi-synchrone écrit (tchat,...), met à l'épreuve les théories linguistiques fondatrices de l'analyse des interactions. En outre, Develotte, Kern et Lamy affirment que « la complexité plurisémiotique de ce genre de conversation mérite d'être analysée dès ses premières années d'utilisation sociale courante » (2011 : 12).

Ainsi, nous partons du postulat que les situations de communication médiatisées par les écrans viennent revisiter les théorisations linguistiques précédemment établies à partir de situations d'interaction en présentiel, s'agissant notamment de l'étude de la spatio-temporalité et des modalités de production langagière. Il s'agit donc de repenser et de reconceptualiser les pratiques de communication contemporaines en s'appuyant sur les travaux existants. Néanmoins de nouveaux questionnements méthodologiques se posent. Comment appréhender l'impact de l'écran sur les nouveaux comportements communicatifs ? Quelles contraintes méthodologiques pèsent sur l'étude des interactions multimodales par écran ? À quels outils de récolte et d'analyse des données le chercheur en interactions numériques peut-il recourir ?

Nous proposons d'aborder ces questions au moyen d'une approche phénoménologique préconisant un « retour aux choses-mêmes » (Husserl, 1929), évitant ainsi le hiatus entre rejet et engouement pour les technologies, crainte et fantasme de l'« homo-connexus » (Louart, 2014). La démarche phénoménologique nous invite à « porter sur les choses un jugement rationnel et scientifique, se régler sur les choses mêmes, revenir des discours et des opinions aux choses mêmes, les interroger en tant qu'elles se donnent elles-mêmes et repousser tous les préjugés étrangers à la chose même » (Husserl, 1913). Cette démarche appliquée à l'analyse conversationnelle induit une définition de l'interaction comme une mise en espace-temps de l'intersubjectivité via des modalités. Et sa transposition à l'écran infère une modification de l'essence de ces quatre pôles : intersubjectivité, espace, temps, modalité.

L'intersubjectivité, en présentiel, se définit en opposition à la subjectivité —une conception autoréférentielle du sujet— comme une nécessité pour le Soi de se construire par rapport à un Autre (Mead, 1934). L'interaction entre Soi et l'Autre constitue alors le lieu où « un étant rencontre un étant, donc chaque étant se rencontre lui-même » (Marcel, 1936 : 9). Le numérique engage un niveau supplémentaire de subjectivité, celui de la technique. Technique qui a toujours

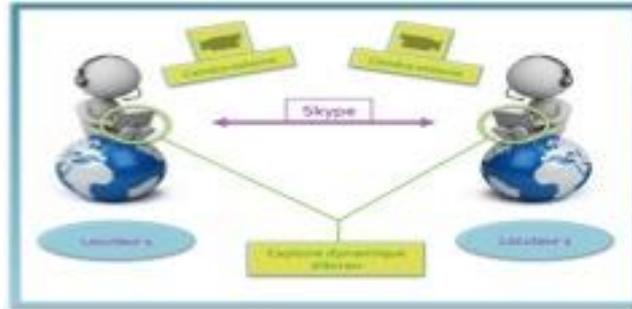
existé car pour qu'un Autre apparaisse à un Soi une médiation technique est nécessaire quelle qu'en soit la forme. Voirol (2013) distingue alors trois niveaux : usager-technique, usager-technique-usager, collectivité d'utilisateurs. L'écran fait ainsi varier le degré d'aura phénoménologique — à distinguer du degré d'existence — de l'utilisateur en ligne (Vial, 2013) en fonction des modalités utilisées. L'intersubjectivité ne fonctionne donc pas ex nihilo, elle nécessite la présence de supports tant symboliques que physiques (Coutant et Stenger, 2010).

L'espace, dans l'interaction, a jusqu'alors été perçu comme immédiat et commun aux interactants pour la bonne conduite de l'échange (Goffman, 1973) car il n'est de discours ni pratique social qui ne se produisent hors espace physique, historique et social (De Saint Georges, 2004). Or l'écran ouvre un « espace intermédiaire » (Jauréguiberry, 2000) qui a autant vocation de protection (l'écran fait écran et protège) qu'il a vocation d'ouverture (fenêtre sur le monde) (Chabert, 2012). Denouël (2008) parle alors de « coprésence à distance » à savoir « un accomplissement pratique lié à l'organisation située de ressources sociotechniques » favorisant ainsi l'échange à distance.

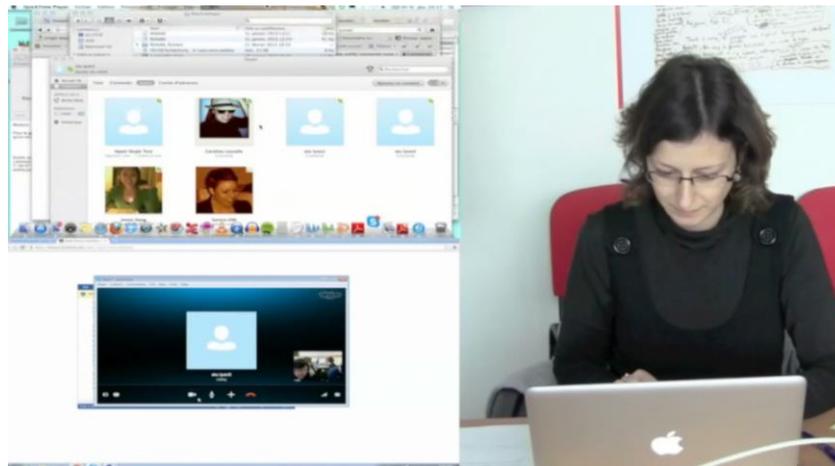
Le temps de l'interaction est qualifié, par les chercheurs en analyse conversationnelle, de « modifiable mais sans rupture » (Kerbrat, 1990), et constitué de séquences ordonnées (Sacks et al., 1974 : 727). La temporalité des échanges en ligne s'avère plus morcelé de la synchronie (visio) à la quasi-synchronie (tchat) et l'asynchronie (forum). L'outil technique prend part à l'allocation des tours de parole et provoque un décalage plus ou moins significatif entre la production du message et sa réception effective par l'interlocuteur réduisant ainsi les possibilités de copilotage de l'interaction (Garcia et Jacob, 1999). Par ailleurs plusieurs interactions simultanées sont rendues possibles par les divers plans de l'écran (fenêtres), la perception (canal visuel, auditif) et la temporalité, donnant lieu à une polyfocalisation. L'accès à plusieurs systèmes sémiotiques simultanés peut libérer ou contraindre l'utilisateur qui répondra, en fonction de ses compétences, par de nouvelles adaptations (Lamy, 2008).

Les modalités d'échange se font aussi bien « par le langage que par des expressions du corps, des sentiments devinés, une observation du comportement de l'autre » (Husserl, 1980). Ainsi les interactions sont considérées comme étant par nature des interactions de corps à corps (Cosnier, 2004). Et les gestes du corps remplissent des fonctions énoncives, énonciatives et de copilotage de l'échange (*Ibid.*). De même le visage du sujet, son épiphanie, est déjà discours, il s'exprime (Levinas, 1936). Tout comme le regard des interactants qui marque l'engagement dans l'interaction. D'où la crainte de la disparition du corps en ligne. Or les adaptations possibles de l'image de soi projetée à l'écran (photo, avatar, émoticône, séquence animée,...) participent d'une « présence corporelle assistée par ordinateur » (Casilli, 2012). Si bien que l'interaction par écran met en présence trois substances : le technique, l'intellectuel et le sensible ; l'artefact, le conceptuel, le corporel (Frias, 2004). Un nouveau régime de métaphores corporelles se met en place et la corporalité d'aujourd'hui se vit dans l'aller-retour entre technologie et chair (Casilli, 2009).

L'analyse conversationnelle ne peut donc se suffire d'une étude verbale des interactions. L'appréhension de l'hybridité des échanges par écran nécessite une approche multimodale et spatio-temporelle, la technologie ne rompant pas avec le présentiel mais le renouvelant, le transfigurant. Aussi des outils spécifiques sont-ils essentiels à la capture des traces linguistico-sémio-pragmatiques issues de l'adaptation des interactants aux modifications que nous venons d'aborder. En l'occurrence nous recourons pour nos recherches à un dispositif technique consistant en une capture dynamique d'écran des utilisateurs en ligne (via le logiciel QuickTime Player sur le système d'exploitation iOS) ainsi qu'à un enregistrement vidéo du contexte spatial présentiel (via des caméras externes). L'assemblage des vues ainsi que l'incrustation des transcriptions verbales (par le logiciel de montage Final Cut Pro X sur iOS) permet d'analyser simultanément les comportements communicatifs de tous les participants sur l'écran et hors champ.



Dispositif technique



Extrait vidéo avec montage des vues et incrustation des transcriptions verbales  
(sur ce corpus, nous n'avons pas pu placer une caméra externe sur le locuteur 2)

**Cliquer sur l'image pour jouer la vidéo**

Nous proposerons une analyse multimodale d'une séquence étendue de cet extrait lors de notre présentation orale au cours du colloque IMPEC 2014.

Nous montrerons comment pour les mêmes fins, les interactants recourent tantôt à l'outil tantôt au corps, indiquant une complémentarité et un ajustement interface-corps plus ou moins maîtrisée. L'étude spatio-temporelle et multimodale des interactions par écran nous évite ainsi de tomber dans le dualisme réel-virtuel et nous permet de décrire la co-construction de la présence corporelle assistée par le numérique.

### Références bibliographiques

- Chabert, G. 2012. « Les espaces de l'écran » In *Écrans et Médias* n°34, pp. 203-215.
- Cosnier, J. 1992. « Synchronisation et copilotage de l'interaction conversationnelle » In *Protée*, pp. 33-39.
- Denouël, J. 2008. *Les interactions médiatisées en messagerie instantanée*. Thèse en sciences du langage : Université Paul Valéry Montpellier III.
- Develotte, C., Kern R. & Lamy, M-N. 2011. *Décrire la conversation en ligne, le face à face distanciel*. Lyon : ENS Éditions.
- Herring, S. 1999. « Interactional coherence in CMC » In *Hawaii international conference on system sciences*, n°32, pp. 1-13.
- Rifkin, J. 2012. *La troisième révolution industrielle, comment le pouvoir latéral va transformer l'énergie, l'économie et le monde*. Paris : Les liens qui libèrent.
- Vial, S. 2012. *La structure de la révolution numérique*. Thèse en philosophie de la technologie : Université Paris Descartes.